

An accumulation of gender inequalities in old age?

Exploring life course- and gender-sensitive
approaches for analysing Gender Pension Gaps



Carla Rowold

Nuffield College, Department of Sociology
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy in Sociology

Michaelmas 2023

Word count (including Tables, Figures, notes): 95,283

To my grandmother, Lore Bethge (*1950 - †2008),

and my mother, Maja Rietz (*1970).

Acknowledgement

Growing up in what I would now describe as a ‘complex’ family environment, I witnessed numerous challenges and injustices faced by the women close to me. This sparked an early interest in inequalities, especially those related to gender, and laid the foundation for this thesis many years ago. At the same time, I could not have wished for a more supportive upbringing, thanks to the unwavering love, care, and courage of my mother, Maja, and my grandmother, Lore. I’m deeply grateful to them for teaching me how to navigate and overcome obstacles while keeping sight of life's true priorities. The skills and values they imparted have played a crucial role in motivating my work. It is to them that I dedicate this thesis.

This thesis is rooted in my previous studies of social sciences at the Humboldt University of Berlin and was written at the Nuffield College which provides a unique environment for the study of social inequalities. I would like to appreciate the work of the staff at Nuffield College, especially of the housekeepers, the porters, and the buttery staff. Throughout my academic journey, I have been fortunate to meet many people who supported me and this thesis in a multitude of ways.

First, I want to thank my thesis advisor Christiaan Monden for his guidance, advice, and patience with me despite my many side projects. I’m grateful to Ridhi Kashyap, Katja Möhring, Selçuk Berdük, and Brian Nolan, for reading different versions of this thesis and providing helpful feedback. I want to thank Kène Henkens and Hannah van Solinge for their feedback during my research stay at the NIDI.

I am indebted to Emanuela Struffolino and Anette Eva Fasang, not only for being great co-authors, but also for their invaluable mentorship over the years, their inspiration, and their continuous support. Without them, I would not have started this PhD in the first place.

A big thank you goes to Zachary Van Winkle and Nicole Kapelle for their helpful feedback at various stages of this thesis, and for being great and patient co-authors. I also want to thank them and Mine Kühn for all the fun at the conferences. I am grateful to Anna Hammerschmid for encouraging me to start a PhD and for embarking on the journey of exploring gendered inequalities in old age alongside me.

A huge shout out to Conor, Fernando, Gabriel, Joan, Kasimir, Kat, Maksim, Said, and many others for sharing this journey, offices, and many beers. Having such great friends around made the time in Oxford always enjoyable.

A special thanks to Alicia García Sierra for being the best PhD companion I could have wished for. Having someone so warm, empathetic, and fun beyond the academic brilliance to share this experience with, including all its ups and downs, has kept me going several times.

To all my friends in Berlin, particularly Timo, thanks for always being there for me and being patient with me and my absences during the last years.

I am glad for my family, and my dad, for all that they have taught me. My mum, Maja, Frank, and my brother Mathis for their unconditional support, love, and always providing a home for me.

Finally, I am incredibly grateful to Kent Troutman for his exceptional support and for encouraging me in every way possible in everything I do, including this PhD. This thesis would not have been started or completed without him.

Abstract

Women in Europe receive much less pension income compared to men. Based on the life course perspective and the pertinent gender literature I argue that previous approaches risk concealing complex inequality (re)producing mechanisms that arise over the life course and shape these Gender Pension Gaps (GPG). Particularly, the existing empirical literature on GPGs lacks a multidimensional, gender- and life-course-sensitive analysis by focussing predominantly on the impact of employment. This thesis aims to implement novel methodological approaches to analyse gendered pension income inequalities considering gendered life course complexities and contribute empirically to our understanding of how GPGs unfold.

I apply Sequence analysis, decomposition, and feature selection techniques to assess and quantify the relation between gendered work-family life courses, or specific life course aspects, with gendered pension income inequality. Comparing the analyses across country contexts or pension types reveals how these associations are channelled through different pension designs.

The results reveal that, independently of the pension system and methods applied, the overarching driver of the GPGs is the large amount of unpaid care work that only women performed over their life courses and that is not equivalently rewarded in pension systems compared to other activities. Particularly, typical life courses with strong interdependences between family and work are only experienced by mothers and are highly associated with GPGs. In other words: gendered pension inequality emerges due to an interaction of i) welfare state contexts of the 20th century which incentivised a traditional gendered division of labour and gender inequalities arising therefrom, with ii) pension policies rewarding gendered life courses emerging from it highly unequally nowadays. I conclude that pension policymakers must consider this intertemporal interaction of present pension designs and past welfare state policies for pension reforms if they aim to prevent a severe reproduction of accumulated gender inequalities in old age.

Table of content

1	Introduction.....	1
1.1	Motivation and policy context	1
1.2	Outline.....	4
1.3	Theoretical frame: The life course paradigm through a gender and old-age lens	5
1.4	Previous research on gendered pension income inequality.....	11
1.5	Aim of the thesis	23
1.6	Chapter overview and contributions	24
2	Conceptualizing analytical decisions in studies on the elderly: a gender- and life-course-sensitive approach	33
2.1	Perspective and definition of pension income	35
2.1.1	Unit of analysis: Individual versus couple-level perspective	35
2.1.2	Pension income types.....	44
2.1.3	Considering access to pensions and the Gender Coverage Gap	48
2.2	Sample definition: Housewives never retire!? Gender biases in popular sample definitions for studies on the elderly	52
2.2.1	Introduction.....	53
2.2.2	Gendered life courses and self-reported retirement	55
2.2.3	Gendered life courses and pension receipt	60
2.2.4	Work trajectories of individuals not identifying as retirees or without pension receipt	62
2.2.5	Discussion.....	66
2.3	Retrospective life course data	72
2.3.1	Work life course data in SHARE: How to solve gendered missing data in the Job Episodes Panel.....	73
2.3.2	Further limitations of the retrospective life course data and implications for this dissertation.....	81
3	Life-course-sensitive analysis of group inequalities: Combining Sequence Analysis with the Kitagawa-Oaxaca-Blinder decomposition.....	85
3.1	Introduction	86
3.2	Illustrative application: Life courses and Gender Pension Gaps in Germany and Italy.....	90

3.2.1	Gendered life courses and retirement systems.....	90
3.2.2	Empirical evidence on Gender Pension Gaps.....	91
3.2.3	Italy and West Germany	95
3.2.4	Data, sample, and variables	97
3.3	Analytical strategy	99
3.3.1	Step I: Sequence Analysis.....	99
3.3.2	Step II: Kitagawa-Oaxaca-Blinder Decomposition (KOB).....	106
3.4	Results of the illustrative application on GPGs in Germany and Italy	124
3.4.1	Standard KOB using life course summary measures.....	124
3.4.2	Gendered work-family life courses.....	128
3.4.3	SA-KOB decomposition of Gender Pension Gaps	131
3.4.4	Comparison of SA-KOB with standard KOB using life course summary measures.....	138
3.4.5	Extension: Single-channel or multichannel sequence analysis.....	142
3.5	Discussion and conclusions.....	151
4	Differences in Gender Pension Gaps in public and private pensions in West Germany: What role do work-family life courses play?.....	157
4.1	Introduction	159
4.2	Theoretical and institutional background.....	164
4.2.1	Empirical evidence	164
4.2.2	Gender, life courses, and the pension system	167
4.2.3	The German context.....	169
4.2.4	Expectations.....	173
4.3	Data and Methods	175
4.3.1	Data	175
4.3.2	Methods	179
4.4	Results	182
4.4.1	Step I: Work-family life course patterns.....	182
4.4.2	Step II: Decomposition of the Gender Pension Gaps	189
4.5	Discussion	210
5	Full-time employment is all that matters? Decomposing Gender Pension Gaps based on relevant life course features in Germany & the Netherlands.....	218
5.1	Introduction	220

5.2	Background	224
5.2.1	Life courses and pension outcomes from a gender perspective.....	224
5.2.2	Empirical evidence on Gender Pension Gaps.....	231
5.2.3	Welfare state context & country comparison design	234
5.2.4	Expectations.....	241
5.3	Data & Methods	242
5.3.1	Data	242
5.3.2	Analytical strategy	243
5.4	Results.....	251
5.4.1	Step I: Life Course Feature Selection	251
5.4.2	Step II: Decomposition of Gender Pension Gaps	265
5.5	Discussion	279
6	Discussion & Conclusion.....	290
6.1	Aims and Chapters overview	290
6.2	Discussion of findings and contributions	291
6.2.1	Conceptual contribution.....	291
6.2.2	Methodological contribution.....	293
6.2.3	Substantive contribution and discussion.....	299
6.3	Social policy implications	318
6.4	Limitations & avenues for future research.....	325
6.4.1	Limitations	325
6.4.2	Future research.....	327
6.5	Concluding remarks and outlook	331
7	Bibliography	334
8	Data, software, and funding	374
8.1	Data	374
8.2	Software and packages	376
8.3	Funding	376
A2	Appendix for Chapter 2	377
A3	Appendix for Chapter 3	382
A4	Appendix for Chapter 4	425
A5	Appendix for Chapter 5	453

List of Figures

Figure 1.1 Example work life courses	18
Figure 2.1 Self-assessed employment status at age 65 or older by country and gender....	58
Figure 2.2 Cohort trend of self-assessed female retirees for selected countries.	59
Figure 2.3 Pension income receipt at age 65 or older by country and gender.....	61
Figure 2.4 Self-defined employment status of female pension recipients.....	63
Figure 2.5 Distribution of work states from age 18-65 by self-definition of retirement ...	64
Figure 2.6 Distribution of work states from age 18-65 by pension receipt	65
Figure 2.7 Most frequent work trajectories of individuals not self-defining as retirees ...	66
Figure 2.8 Missingness before adjustment	75
Figure 2.9 Missingness after adjustment	76
Figure 2.10 Distribution of filled working states for all countries by gender	77
Figure 2.11 Distribution of filled working states for Germany by gender	79
Figure 2.12 Distribution of filled working states for Italy by gender.....	79
Figure 2.13 Distribution of filled working states for the Netherlands by gender.....	79
Figure 3.1 Work and family life courses: relative frequency sequence plots, distribution by gender and across countries.....	130
Figure 3.2 Decomposition results – explained shares by work-family life courses and country	133
Figure 3.3 Decomposition results – shares from returns component of the unexplained part by work-family life courses and country.....	138
Figure 3.4 Sequence distribution for work life-course clusters from single-channel SA	143
Figure 3.5 Gender-specific distribution over work life-course clusters (single-channel SA) by country.....	144
Figure 3.6 Sequence distribution for family life-course clusters from single-channel SA	147
Figure 3.7 Gender-specific distribution over family life-course clusters (single-channel SA) by country.....	148
Figure 4.1 Observed outcomes along the life course, institutional impacts on the GPG and data.....	174
Figure 4.2 Family and work life courses by gender.	179

Figure 4.3 Gender-specific distribution on work-family life courses, relative frequency plots.....	186
Figure 4.4 Shares of the GPG due to the gender-specific distribution on work-family life courses (explained part).....	191
Figure 4.5 Shares of the GPG due to gender-specific returns for the same work-family life course pattern (returns component)	195
Figure 4.6 Gender gap in public pensions over the distribution.....	198
Figure 4.7 Decomposition over the distribution: selected explained shares	199
Figure 4.8 Decomposition over the distribution: returns component	201
Figure 4.9 Public and private pension distribution by gender	204
Figure 5.1 Example work life courses	227
Figure 5.2 Work life courses across genders in West Germany and the Netherlands.....	237
Figure 5.3 Mean importance of selected variables by country	252
Figure 5.4 Mean importance of selected variables by country and life course state	253
Figure 5.5 Ñopo decomposition of Gender Pension Gap in West Germany.....	266
Figure 5.6 Ñopo decomposition of Gender Pension Gap in the Netherlands.....	267
Figure 5.7 Contributions of years spent in care work (left panel) and age at first care work spell (right panel) to GPG (DA/DB), West Germany (model 3).....	274
Figure 5.8 Work life courses of women by matching status and model, Netherlands	276
Figure 5.9 Work life courses of women by matching status and model, West Germany	277
Figure A2.1 Share of missingness for all person-year spells before filling missing information	381
Figure A2.2 Share of missingness for all person-year spells after filling missing information	381
Figure A3.1 Average silhouette width (ASW) by cluster, country & gender	400
Figure A3.2 Dissimilarity from the medoid for the sequence plots in Figure 3.1.....	401
Figure A3.3 Sequence distribution plots, post-reassignment.	402
Figure A3.4 Distribution of silhouette values and cut-off values by cluster	403
Figure A3.5 Share of individuals with no pension income by country and gender.....	403
Figure A3.6 Confidence intervals (95%) for the decomposition of the GPG in Italy	404
Figure A3.7 Confidence intervals (95%) for the decomposition of the GPG in West Germany	404

Figure A3.8 Average years spent in work life-course states from age 18–65.....	413
Figure A3.9 Average years spent in family life-course states from age 18–65.....	414
Figure A3.10 Distribution on work-family life courses for men by country, before and after reassignment (frequencies).....	423
Figure A4.1 Earning point quartiles	429
Figure A4.2 Distribution on family states by gender, person-year spells	429
Figure A4.3 Distribution on working states by gender, person-year spells	429
Figure A4.4 Sequence distribution plots by cluster.....	430
Figure A4.5 Original distribution on clusters by gender	431
Figure A4.6 Final distribution on clusters by gender used for decomposition.....	431
Figure A4.7 Decomposition results with confidence intervals, public pension income .	432
Figure A4.8 Decomposition results with confidence intervals, private pension income	432
Figure A4.9 Gender-specific distribution on work-family life course clusters by private pension receipt	433
Figure A4.10 Gender gap in public pensions over the distribution, quantiles	434
Figure A4.11 Decomposition over the deciles: selected explained shares.....	435
Figure A4.12 Decomposition over the deciles: selected shares due to returns component.....	436
Figure A4.13 Share with any independent public or private pension income overall and by gender.....	437
Figure A4.14 Distribution of public pension income by gender.	438
Figure A4.15 Distribution of personal pension income by gender.....	439
Figure A4.16. States replacing non-contributory periods in administrative records.....	443
Figure A4.17 Distribution on states (based on survey data only) if civil servants missing in register data	444
Figure A4.18 Distribution on states (based on survey data only) if self-employee missing in register data	444
Figure A4.19 Frequency of states from register data imputed with civil service reported in the survey.....	445
Figure A4.20 Frequency of states from register data imputed with self-employment reported in the survey	445
Figure A4.21 Frequency of individuals in clusters before and after reassignment	450
Figure A5.1 Family life course states considered by country (person-year spells).....	455
Figure A5.2 Working life course states considered by country (person-year spells).....	455

Figure A5.3 Family life courses across genders in West Germany and the Netherlands	456
Figure A5.4 Contributions of years spent in care work (to GPG (DA/DB), Netherlands (model 2)).....	457
Figure A5.5 Contributions of years spent in civil service (to GPG (DA/DB), West Germany (model 3)).....	457
Figure A5.6 Contributions of age at transition from civil service to retired (to GPG (DA/DB), West Germany (model 3)).....	458
Figure A5.7 Contributions of working life complexity categories to GPG (DA/DB), Netherlands (model 2)	458
Figure A5.8 Gender-specific pension predictors West Germany	463
Figure A5.9 Association between age at transition from FT employment to retirement and pension income, West Germany	463
Figure A5.10 Association between age at transition from civil service to retirement and pension income, West Germany	464
Figure A5.11 Association between duration of retirement and pension income, West Germany	464
Figure A5.12 Gender-specific pension predictors the Netherlands.....	465
Figure A5.13 Ñopo decomposition of Gender Pension Gap in West Germany.....	468
Figure A5.14 Ñopo decomposition of Gender Pension Gap in the Netherlands.....	469
Figure A5.15 Ñopo decomposition of Gender Pension Gap in West Germany.....	483
Figure A5.16 Ñopo decomposition of Gender Pension Gap in the Netherlands.....	484
Figure A5.17 Ñopo decomposition with men as group A ('returns reference'), West Germany	486
Figure A5.18 Ñopo decomposition with men as group A, Netherlands.....	487
Figure A5.19 Ñopo decomposition using PS Matching, West Germany	489
Figure A5.20 Ñopo decomposition using PS Matching, the Netherlands.....	489
Figure A5.21 Example trajectories and life course variables.	490
Figure A5.22 Work life courses of men by matching status and model, Netherlands	493

List of Tables

Table 1.1 Main streams of literature on gendered pension income inequalities and their limitations, focused on conceptual and methodological components	17
Table 1.2 Chapter overview.....	26
Table 2.1 Overview of literature on gendered pensions concerning conceptual choices	38
Table 2.2 Pension income types included.....	47
Table 3.1 Overview of KOB studies on Gender Pension Gap	93
Table 3.2 Options for reference coefficients	114
Table 3.3 Cross table of gender with the clusters (illustrative application, for Italy).	117
Table 3.4 Results from standard KOB decompositions.....	127
Table 3.5 Results from SA-KOB decompositions.....	135
Table 3.6 Baseline coefficients from linear regression analysis used for SA-KOB, normalized	136
Table 3.7 Comparing benefits and limitations of the standard and the SA-KOB decomposition.....	139
Table 3.8 Single-channel SA.	146
Table 4.1 Summary characteristics of life course clusters by gender.....	187
Table 4.2 Decomposition results of the Gender Pension Gap in public and private pension.....	193
Table 4.3 Decomposition results of the gender gap in private pension receipt and private pension among recipients	206
Table 4.4 Decomposition results of the gender gap in public pension among public pension recipients	209
Table 5.1 Selected features for overall pension inequality, including gender as predictor and not, ranked.	258
Table 5.2 Selected features for overall, female- and male-specific pension inequality, ranked.	261
Table 5.3 Ñopo decomposition of Gender Pension Gap in West Germany	266
Table 5.4 Ñopo decomposition of Gender Pension Gap in the Netherlands	267
Table 5.5 Observed characteristics by gender and matching status for West Germany (model 3).....	272
Table 5.6 Observed characteristics by gender and matching status for the Netherlands (model 3).....	273

Table 6.1 Overview of substantive findings and derived policy implications	315
Table A2.1 Summary table for retirement identification.....	379
Table A2.2 Summary table for pension receipt	379
Table A3.1 Norms and family policies in Italy and Germany	382
Table A3.2 Pension systems in Italy and Germany	383
Table A3.3 Waves, survey years and number of observations.....	385
Table A3.4 Overview of sample selection and dropouts.	385
Table A3.5 Pension types included in the analysis.....	385
Table A3.6 Options for reference coefficients, interpretation, benefits, and specifics for SA-KOB decomposition	386
Table A3.7 Characteristics of life-course clusters for Italy	387
Table A3.8 Characteristics of life-course clusters for West Germany	389
Table A3.9 Gender-specific distribution on clusters and gender gaps therein, before and after reassigning	391
Table A3.10 Robustness checks, set I: different reference coefficients used for KOB...	392
Table A3.11 Robustness checks, set II: excluding individuals poorly assigned to cluster based on low silhouette widths.	393
Table A3.12 Robustness checks, set III: Including more controls.	394
Table A3.13 Robustness checks, set IV: using original cluster and cross-sectional weights.	396
Table A3.14 Comparing the SA-KOB results to standard KOB (summary measures)...	397
Table A3.15 Results from Kitagawa-Oaxaca-Blinder decompositions, full table.	398
Table A4.1 Family policies, gender norms and outcomes across West Germany	425
Table A4.2 Pension policies and indicators for West Germany	426
Table A4.3. Sample selection	427
Table A4.4 Waves, survey years and n of observations (final sample).....	427
Table A4.5 Pension types included and definition of dependent variables.....	427
Table A4.6 Baseline coefficients from linear regression models for public and private pension income	428
Table A5.1 Pension types included and definition of dependent variables.....	453
Table A5.2 Sample selection	453
Table A5.3. Full results of Boruta Life Course Feature Selection, main model	454
Table A5.4 KOB decomposition over the matched sample for West Germany, stepwise models.....	471
Table A5.5 KOB decomposition over the matched sample for the Netherlands.....	472
Table A5.6 Important pension predictors by identified by different algorithms, ranks ..	475
Table A5.7 Important pension predictors for different pension definitions, ranks	477
Table A5.8 Ñopo decomposition of Gender Pension Gap in West Germany	483
Table A5.9 Ñopo decomposition of Gender Pension Gap in the Netherlands	484

1 Introduction

1.1 Motivation and policy context

Women throughout Europe receive on average much lower pension income compared to men and the level of this ‘Gender Pension Gap’ largely exceeds the level of the much more popular Gender Wage Gap in most countries (Bettio et al., 2013; Hammerschmid & Rowold, 2019a; Jędrzychowska et al., 2020). The Gender Pension Gap (GPG) documents the difference between the mean pension income of women compared to the mean pension income of men. For example, in West Germany, women receive 61% less pension income compared to men, in Italy 53% and in the Netherlands 46%. In monetary terms, this means that women have on average €13,449, €6,651, and €12,489 less in pension income per year.¹

Despite the significantly higher levels of Gender Pension Gaps in most European states, GPGs have received much less attention compared to gender inequalities focused on outcomes in pre-retirement life course stages. However, studying gendered pension inequalities is important because women are at higher risk of poverty in old age (Haitz, 2015; Smeeding & Sandstrom, 2005) and (public) pensions are the most important income source in old age and consequently particularly important to prevent poverty (Kuitto et al., 2023; OECD, 2019). Recent studies call for more policy-relevant evidence on how to diminish gendered inequality in pension income (Kuitto et al., 2021).

Studying gender inequalities in pensions becomes increasingly pressing for a multitude of reasons. First, a higher number of individuals will have to face such inequalities in the future due to population ageing. Second, and partly because of population ageing, pension

¹ Own estimation, see Chapters 3 and 5.

policy trends are characterized by privatisation, individualisation, and deinstitutionalization of pensions (Cooke, 2006; Ebbinghaus, 2011b; Frericks et al., 2007; Hinrichs, 2021; Kuitto et al., 2021; Pedersen, 2004). A consequence is, and most likely will be, cuts in public pensions. These trends thus particularly threaten the financial well-being of individuals in old age who have to rely on redistributive elements of public pensions, such as those with unstable careers and low earnings (e.g., Kuitto et al., 2021). Since these individuals are disproportionately women, pension privatisation is expected to be particularly disadvantageous for women (Foster & Ginn, 2018).

Third, family biographies and structures are becoming more complex across countries, i.e., less stable, over cohorts, and this trend is more pronounced for women (Rowold & Van Winkle, 2023; Van Winkle, 2018; Van Winkle & Fasang, 2021). Particularly union dissolutions are becoming more prevalent (Van Winkle, 2018). This trend is also observable among the elderly: the ‘grey divorce revolution’ describes the increase in union dissolution among the elderly (Brown & Wright, 2017). Moreover, elderly women are less likely to remarry after divorce but more likely to face poverty upon divorce (Lin & Brown, 2021; Schimmele & Wu, 2016). Elderly women are thus decreasingly financially secured by the household context and will have to rely increasingly on their own income sources.

As a consequence of these trends, I take an individual perspective and consider individual, independent pension income for the empirical analyses in this dissertation. I argue that this conceptual decision allows me to derive conclusions that are relevant for the future since independent individual pensions will most likely become more important (see for a detailed discussion Chapter 2.1.1). Some argue that the financial well-being of women with low independent pensions is often secured by relatively high-income partners and that, therefore, there is no need to be concerned about Gender Pension Gaps (Faik & Köhler-Rama, 2012). However, for women with low own pension claims, it becomes increasingly

risky to rely on the high income of a partner or redistributive elements of the pension system or survivor pensions. The financial dependences of these women are likely to make them more vulnerable to poverty risk in old age in the future given more unstable partnerships and cuts in redistributive elements of pensions. Consequently, the focus on financial independence in old age is more sustainable for women's financial well-being in the future and in line with the demographic change as well as the individualisation of pensions.

Fourth, the concept of normatively pursued life courses, primarily characterized by stable full-time employment upon which pension systems are traditionally built (Allmendinger et al., 1993; Fasang, 2010; Foster & Ginn, 2018; Grady, 2015; Kuitto et al., 2021; Leitner, 2001), is becoming increasingly uncommon (e.g., Guillemard & Rein, 1993). At the same time, pension policies have been criticized for ignoring life course complexities and heterogeneities in general as well as between different life course domains, such as work and family lives (Cooke, 2006). This is particularly important from a gender perspective given that literature has shown across countries, that women's life courses are more complex compared to men's (Rowold & Van Winkle, 2023; Van Winkle, 2018; Van Winkle & Fasang, 2017, 2021).

Consequently, researchers have highlighted the need for more life-course-sensitive approaches to pension income inequalities, particularly the Gender Pension Gap (most recently: Foster & Ginn, 2018). However, I argue that this has not been empirically and methodologically implemented yet, at least when it comes to the literature on gendered pension inequality (see Chapter 1.4). Instead, the scarce literature on gender disparities in pension income has primarily focused on the link between life course summary measures of employment types (e.g., the years spent in full-time employment in an individual's life) and the Gender Pension Gaps.

Contributing to the current literature, the main research question of this thesis is:

How are gendered life courses related to Gender Pension Gaps?

To answer this question and create new substantive evidence focussing on the role of life courses, novel methods are required that capture such life course complexities while still being gender-sensitive. Addressing the challenges in the literature, I introduce gender- and life-course-sensitive methodological approaches to explore gendered inequalities in pensions and create new empirical evidence. Applying these methodological approaches sheds light on the mechanisms through which Gender Pension Gaps are related to work-family life course patterns or specific aspects of work and family lives. The contributions of this dissertation are of methodological, conceptual, and substantive nature (see Chapters 1.6 & 6).

1.2 Outline

The dissertation consists of six Chapters. In the introductory Chapter, I give an overview of the theoretical frame applied (Chapter 1.3) and the previous research on gendered pension inequality including its challenges (Chapter 1.4), before I present the aim and contribution of the thesis (Chapter 1.5) as well as an overview of the Chapters (Chapter 1.6).

In Chapter 2, I discuss the conceptualisation of the main components of empirical analyses regarding GPGs beyond methodological approaches from a gender- and life-course-sensitive perspective. This lays the conceptual foundation for the empirical Chapters 3-5.

Chapter 3 suggests combining Sequence analysis with the Kitagawa-Oaxaca-Blinder decomposition (SA-KOB) to achieve a more gender- and life-course-sensitive analysis of group inequalities in outcomes from a holistic perspective. This Chapter introduces the application of the SA-KOB decomposition, placing particular emphasis on the discussion

of methodological and technical key considerations when integrating both methods. It further provides researchers with empirical evidence on the decomposition of Gender Pension Gaps based on work-family life courses in West Germany and Italy.

In Chapter 4, I apply the SA-KOB decomposition suggested in Chapter 3 to richer data and focus on the differences in the gender gaps in public and private pensions to derive policy-relevant conclusions in light of pension privatisation. Linked admin-survey data allows me to consider earnings of the life course for the categorisation of work-family life course and thus contributes to the literature in revealing how gendered earnings inequalities over the life course are associated with the gender gaps in public compared to private pensions.

In Chapter 5, I discover which aspects of work and family life courses (duration, timing, ordering) are relevant pension predictors and assess whether the years spent in full-time employment are indeed the most important predictor of pension inequality in West Germany and the Netherlands. I then decompose the Gender Pension Gap based on these most important pension predictors contributing by quantifying which parts of the gap are due to gender-exclusive combinations of these life course analyses.

Finally, in Chapter 6 I summarize the main findings and conclude with a discussion of the overarching results, contributions and policy recommendations, limitations, and avenues for future research.

1.3 Theoretical frame: The life course paradigm through a gender and old-age lens

The theoretical framework relies primarily on the life course paradigm, linked to perspectives on old age and gender. The life course paradigm generally builds up on the idea that life courses develop based on individual constraints and opportunities that result from, among others, the socio-historic context, the interlinkage with other individuals as well as the timing of events along an individual's life (Elder et al., 2003). The life course

perspective adopts a longitudinal approach to time, i.e., the approach usually observes a longer period, and considers multiple domains of life courses as well as several levels of analysis such as individual and supra-individual levels (Bernardi et al., 2019).

Life courses are usually referred to as whole entities considering all their complexities, e.g., the ordering and timing of events over the whole life course or the studied period. However, research can also explore the various aspects that constitute life courses, that are the duration, timing, and ordering of events or transitions throughout one's life (Halpern-Manners et al., 2015; Liao et al., 2022; Studer & Ritschard, 2016), as well as the overall complexity of life courses (Elzinga, 2010) (for a detailed description of these aspects, see Chapter 5.2.1).²

Foster and Ginn (2018) recommend adopting a life course perspective to study *gender* inequalities in pensions for a better understanding of how different life course transitions and domains are linked to the situation of women in old age. The life course perspective is a powerful tool for studying Gender Pension Gaps for several reasons.

First, taking a life course perspective is straightforward for studying any inequalities in pension income because *pension systems reward life courses* (Ginn et al., 2001; Madero-Cabib & Fasang, 2016; Möhring, 2021). Consequently, outcomes in old age, such as income, are increasingly determined by life course inequalities (Kuitto et al., 2021). As such, pension systems are life-course-sensitive (e.g., Möhring, 2021). However, pension systems are built based on the notion of continuous full-time employment (Allmendinger et al., 1993; Fasang, 2010; Foster & Ginn, 2018; Grady, 2015; Kuitto et al., 2021; Leitner,

² The overall complexity of life courses is defined as “internal heterogeneity of a life course in terms of position occupied” (Pelletier et al., 2020, S. 1129) and measures often incorporate the number of transition between life course states, the entropy (i.e., the unpredictability) and the number of life course states experienced in a life course (see for an overview Pelletier et al., 2020).

2001), and consequently, continuous full-time employment is often rewarded with generous pensions, while other forms of work, most importantly “family work is insufficiently, or not at all, protected.” (Leitner 2001: 202). The pension system’s focus on employment has been criticised for its negative implications for women: “it is the gender blindness that is woven into the framework of the pension system that creates and sustains disadvantage for women” (Grady 2015: 452).

Second, the life course perspective enhances our understanding of the complex dynamics in play for shaping gender inequalities over the life course. Two concepts of the life course paradigm are particularly relevant.

First, the ***cumulative advantage or disadvantage (CAD)*** concept emphasizes the process of accumulating individual or group-based inequalities over time (Dannefer, 2003; DiPrete & Eirich, 2006; O’Rand, 1996). Inequalities in pre-retirement life course stages are often gendered. Prominent examples are the differences in the labour force and working hours, occupational segregation as well as the gender gap in wage and lifetime earnings (Anker, 1998; Blau & Kahn, 2017; Budig et al., 2016; de Castro Galvao, 2023; Gangl & Ziefle, 2009; Jalovaara & Fasang, 2020; Joshi & Davies, 2002; Kunze, 2018). Applying the concept of CAD, these gender inequalities are likely to accumulate over the life course and lead to larger, accumulated gender inequalities: “aged inequality is significantly affected by women's cumulative disadvantage [... which ...] is anchored in earlier opportunity structures and the individual behaviours constrained by them” (O’Rand, 1996, S. 233). Thus, following CAD, the Gender Pension Gap can be understood as “the sum of gender inequalities over a lifetime, including differences in the life course (motherhood penalty), segregated labour market and gendered social norms and stereotypes more generally.” (European Institute for Gender Equality, 2015, S. 4).

The second important concept for the emergence of gender inequalities is that of ***linked lives***. The concept of linked lives highlights that one individual's life decisions have implications for the opportunities of another's individuals' life (Elder et al., 2003). Examples are the implication of parental decisions on their children's lives (e.g., García-Sierra, 2023), or, how the partner's decisions and preferences impact each other. In this thesis, the linked lives concept is employed to analyze the historically prevalent gendered division of labour, in many Western European countries since the mid-20th century (e.g. Lewis, 1992; Meyer & Pfau-Effinger, 2006; Pfau-Effinger, 2005), i.e., the 'linked lives' between individuals' careers and their family members' lives and needs. The division of labour was a key gender arrangement during the working lives and family formation phases of the current elderly cohorts. In its most traditional form, it assigns men the role of the sole breadwinner and women the role of homemaker and caretaker (e.g. Meyer & Pfau-Effinger, 2006; Rosenfeld et al., 2004). Even in the more common "male breadwinner/female part-time carer model" (Pfau-Effinger 2005: 332), women bear the majority of care and family work (OECD 2020). As a result of these gender arrangements, "[w]omen's linked lives with other people have an enormous influence on their careers" (Hofmeister et al., 2006, S. 401), while men's employment life courses are less affected by family events (Allmendinger et al., 1992; Hofmeister et al., 2006; Krüger & Levy, 2001). This indicates how gender inequalities across the life course arise from "sharing of unpaid [and paid] work over the life course" (Lewis, 1992, S. 170) and thus from care responsibilities overwhelmingly taken by women.

For gender inequalities, the two most important life domains that are 'linked' through interactions with other individuals are consequently work and family. Work life courses are important determinants of pensions since earnings often contribute to pension systems and work outcomes directly determine pension outcomes. Scholars have highlighted the

relevance of considering family life courses additional to the work domain for achieving a more gender-sensitive analysis of old-age inequalities (Allmendinger et al., 1992, 1993; Fasang, 2009, 2010; Frericks, 2020; Ginn et al., 2001; Meyer & Pfau-Effinger, 2006).

On one hand, family life courses can determine pension outcomes directly, for example through policies aiming directly at family life course aspects, such as childcare benefits or couple-level pension claim splitting after divorce (Fasang et al., 2013; Kreyenfeld et al., 2023). Moreover, women's typical pathways to pensions are more strongly linked to their family lives compared to men's (Fasang, 2009, 2010). Family life courses are also more complex for women and become increasingly complex over cohorts, especially for women (Rowold & Van Winkle, 2023; Van Winkle, 2018; Van Winkle & Fasang, 2021). Consequently, it is important to include family life courses for considering these gendered trends and dynamics and assess whether they might have implications for gender inequalities in pension in the long run.

On the other hand, family life courses can be linked to pension income through interactions with work careers. Strong interdependences between family and work life courses exist for women, but are low or do not exist at all for men (Fasang & Aisenbrey, 2022; Madero-Cabib & Fasang, 2016; Rowold et al., 2023). As such, it is important to consider family life courses next to work life courses as well as their interactions to assess whether the gendered level of work-family life course interdependences translates into gendered pension income inequality.

However, the work-family life course interactions, or the reconciliation between family duties and career, are not only determined on the individual level. The opportunity structures are strongly shaped by the broader context.

Third, the life course perspective is meaningful to study the GPG as it emphasizes the *role of societal and welfare contexts in framing the opportunities and obstacles* in which life

courses unfold (Bernardi et al., 2019; Elder et al., 2003). In this tradition, gendered life courses are the result of welfare state institutions and norms, which enhance or prevent certain gender arrangements and thus facilitate life course opportunities of men and women differently (Krüger & Levy, 2001; Pfau-Effinger, 1998). Considering the context is firstly relevant to recognize the “institutionalization of the work-family conflict” (Möhring & Weiland, 2022, S. 2) for shaping life courses of men and women equally or distinctly. Empirical comparative life course literature shows clear country and more broadly, welfare state system differences, in the degree of gender differences in work(-family) life courses (Komp-Leukkunen, 2019; Rowold, 2022; Uccheddu et al., 2022).

The emphasis on the welfare context is also important to understand that pension systems do not operate in isolation, but in interaction with other welfare state policies, and, most importantly, across several decades. Life courses that are rewarded by the pension system once the individual reaches a certain age were shaped by welfare state policies and norms that were in place along the past life course. Given the expected role the interdependences between work and family life courses play for gendered pension income, the inclusion of family life courses in the analysis is also relevant to reveal the role of family policies for women’s pension income (Fasang et al., 2013), and thus the GPG.

Consequently, to disentangle the emergence of gender gaps in pension income, it is crucial to assess whether and to what extent pension systems reward the gendered work-family life courses differently: “[w]elfare states both construct and reward particular uses of time through the rules of public pension schemes.” (Ginn et al., 2001, S. 7). In sum, the interrelation of all pension institutions with gendered life courses is an ideal example of “[w]hat is valued, protected, or rewarded over the life course” (O’Rand 1996: 230).

To adequately address the intricate intersections across different policy domains and delve into detailed discussions, comparative life course sociology narrows its focus to small sets

of countries (Fasang & Ulrich Mayer, 2020). This dissertation follows the comparative life course sociology in their focus on a few countries for comparison. All countries studied in this dissertation (West Germany, Italy, and the Netherlands) can be classified as male breadwinner societies with intermediate to high gender inequalities supported by conservative or rudimentary welfare policies for the studied cohorts (primarily birth cohorts 1930-1950) (e.g., Komp-Leukkunen, 2019, 2021; Pfau-Effinger, 2005) but have implemented different pension systems. The research design assesses how the different welfare and pension regimes interact with each other concerning the formation of Gender Pension Gaps.

The thesis aims to implement the life course perspective methodologically to enhance our understanding of the emergence of Gender Pension Gaps in a more life-course- and gender-sensitive way.

1.4 Previous research on gendered pension income inequality³

I identify three main streams in the literature on gender inequalities in pension income that all apply micro-level analyses but different methodological approaches that address varying research questions (Table 1.1).⁴ The first one uses decomposition techniques to assess the gender gap in pensions directly. The second one focuses on women's pension income mostly and applies linear regressions. As covariates of interest, studies of both approaches use either cross-sectional information (for example, the marital status at the time of the interview but also the occupation in a last job) or some form of proxy for different life course elements, what I refer to as (life course) summary measures. Both approaches use

³ This section focuses on discussing the different methodological approaches and covariates used, and the limitations thereof as well as a short summary of findings. Chapter 2 discusses other conceptual decisions, such as the pension income definition and Table 2.1 gives an overview of the conceptual choices the studies discussed have made.

⁴ Another stream in the literature takes a multi-country comparison approach to disentangle patterns across countries and policy impacts (Bettio et al., 2013; Butts, 2020; Hammerschmid & Rowold, 2019a; Lis & Bonthuis, 2019; Samek Lodovici et al., 2016; Tinios et al., 2015). However, due to the focus of this dissertation, I consider only micro-level analyses for the literature review.

selective sets of summary measures for the work or (more rarely) the family biography. These can include the age at first childbirth or marriage, but most commonly these research streams use the sum of years spent in different life course states such as the years being (full-time) employed (i.e., the *duration* in such states). The 3rd literature stream models life courses from a holistic perspective.

Literature stream I: Decomposition studies. The decomposition studies usually apply the Kitagawa-Oxaca-Blinder (KOB) decomposition based on linear regressions (Bardasi & Jenkins, 2010; Bettio et al., 2013; Bonnet et al., 2020; Cordova et al., 2022; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Hänisch & Klos, 2014; Johnson et al., 1999; König et al., 2019; Levine et al., 1999; Nolan et al., 2019; OECD, 2021; Veremchuk, 2020; Zhao & Zhao, 2018). The KOB allows researchers to differentiate between parts of the Gender Pension Gap that are due to differences in endowments with characteristics (e.g., the years spent in full-time employment) and a part that is due to different returns for the same characteristic (e.g., different pension income for the same years in full-time employment). The first component is often referred to as the explained or compositional component and the second as the returns or effect component (for more methodological details on KOB, see Chapter 3).

Among the summary measures, almost all these studies assess the role of the sum of years spent in specific employment types, most often full-time employment (see for an overview of decomposition studies Table 3.1). Fewer studies consider the duration of part-time employment or unemployment and only one study (Frommert & Strauß, 2013) includes the years engaged in unpaid care work. Only the current family status is considered, if at all, but no proxies for the family life courses are included (except for Bardasi & Jenkins, 2010).

Most studies find that women's lack of years in full-time employment is the predominant or one of the main reasons for the GPG (compositional component of the KOB, Bonnet et

al., 2020; Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014; Johnson et al., 1999; Levine et al., 1999; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018, see last column of Table 3.1 for an overview). The duration of part-time employment is explicitly assessed only in Germany and the UK. In Germany, a large share of the gap is reduced due to the more years women have spent in part-time employment while in the UK it drives the gap in public and private pensions (Cordova et al., 2022; Ezeyi & Vujic, 2017; Frommert & Strauß, 2013). Other studies highlight the important impact of differences in wages (Bonnet et al., 2020; Even & Macpherson, 2004; Levine et al., 1999; Johnson et al., 1999; König et al., 2019) or occupational segregation between genders (Levine et al., 1999; Veremchuk, 2020). Cordova et al. (2022) also find evidence suggesting that parts of the gap in Germany are due to gender segregation over industries and company size and that the gender pension wealth gap is higher among married couples.

If family life indicators are assessed at all, results show that gender differences in the prevalence of being married drive the GPG, whereas gender differences in widowhood rather decrease⁵ the Gender Pension Gap in most countries (e.g. Hänisch & Klos, 2014; Veremchuk, 2020) – but these findings are not further discussed. The interlinkage between the work and family life course dimensions is not captured in any of these studies.

Except for two studies on the UK, the returns components of the decomposition are not discussed in detail. However, in the UK, large shares of the GPG are due to the gender *differences in returns* for marital status and fertility (Bardasi & Jenkins, 2010; Ezeyi & Vujic, 2017), suggesting that family characteristics play a role as well.

Literature stream II: Regression analyses. Other studies do not assess the GPG directly but analyze determinants of women’s pensions or gender inequality in pension income in

⁵ This is the case as the prevalence of widowhood is much larger among women (Bettio et al., 2013) and the studies include survivor pensions.

regression frameworks (e.g. Fasang et al., 2013; Kuivalainen et al., 2020; Möhring, 2015; Sefton, Evandrou, & Falkingham, 2011; Kreyenfeld et al., 2023; Gardiner et al., 2016; Evandrou et al., 2009; Allmendinger et al., 1993) or old-age income more broadly (e.g., Sefton, Evandrou, Falkingham, et al., 2011; Sefton, Evandrou, & Falkingham, 2011). In line with literature stream I, they find a strong positive association between women's length of both, full-time and part-time employment careers, and their income in old age in West Germany (Fasang et al., 2013; Sefton, Evandrou, & Falkingham, 2011). Some studies operationalise work life course complexities as career volatility or career types and find that such career patterns matter beyond the duration of full-time employment (Möhring, 2015; Sefton, Evandrou, & Falkingham, 2011; Sefton, Evandrou, Falkingham, et al., 2011). Overall, this stream of literature puts slightly more emphasis on family characteristics compared to the decomposition studies. Going beyond the employment lives only, Fasang et al. (2013) found a negative impact of the duration of marriage and a positive association of being never married, divorced, or widowed on West German women's retirement income. This holds while controlling for the employment history. Kreyenfeld et al. (2023) find similar associations of partnership status with women's pension income and highlight the impact of splitting pension rights upon divorce which results in divorce being a 'Gender Equalizer' in public pension income inequality. Allmendinger et al. (1993) additionally find that the age of childbirth matters for women's own pension income.

A smaller group of studies focuses on the motherhood pension penalty, i.e., the differences between mothers and childless women's pension income (Dotti Sani & Luppi, 2021; Mika & Czaplicki, 2017; Möhring, 2018; Stegmann, 2007). Comparing the development of public pension claims of mothers and childless women reveals that West German mothers are far behind childless women and this gap remains over the whole life course (Stegmann, 2007). Mika and Czaplicki (2017) find large unadjusted motherhood penalties in public

pension claims increasing with the number of children in West Germany. When considering public pension claims that include childcare benefits, they find a more modest motherhood penalty and conclude that the childcare benefits compensate for parts of the employment interruptions in the public pension system. However, Möhring (2018) observes that while the pension system's overall redistributive elements tend to mitigate motherhood pension penalties, pension care benefits tend to exacerbate them. Further analyses show that care benefits are only positive for mothers' pension income if they are very generous, but rather counterproductive if their generosity is low to medium-high. Overall Europe, the motherhood pension penalty vanishes when controlling for employment characteristics but years of full-time employment increase mothers' pension income to a lesser extent compared to childless women (Möhring, 2018). Additionally, the longer mothers stay home after childbirth, the larger the motherhood penalty (Dotti Sani & Luppi, 2021).

Literature stream III: Sequence analysis and regressions. The third literature stream applies regression analyses together with sequence analysis. Sequence analysis allows researchers to consider many different work and family life course categories without multicollinearity issues as well as interactions between work and family life courses (e.g., a transition in work life follows a transition in family life; in the case of Multichannel sequence analysis). It visually reveals in which order life course categories unfold over time. Life course complexities become visible and are categorised into similar groups. Madero-Cabib and Fasang (2016) were the first to introduce such a holistic perspective of how typical life course patterns impact men's and women's pension income. They first apply Multichannel Sequence analysis to categorize typical patterns of work-family life courses as clusters and then use the life course clusters as covariates in regressions. Compared to the male standard life course of stable married parenthood and full-time employment their results document the highest pension penalties for typical female work-

family life courses characterized by care work and motherhood followed by patterns in which women reconcile their motherhood with part-time employment in both, West Germany and Switzerland.

Applying Sequence analysis to the employment sphere only, Tophoven & Tisch (2016) find large public pension claim penalties for female life courses characterized by care work, where the penalty is smallest if the care work started at later ages. While both, female full- and part-time employment careers contain small shares of care work and generate pension penalties compared to male full-time time careers, the pension penalty is larger for part-time careers. Möhring and Weiland (2022) were the first to take a couple's perspective. Focussing on employment trajectories as well, they show the highest pension incomes for women in dual-earner and the lowest in male-breadwinner couples.

Table 1.1 Main streams of literature on gendered pension income inequalities and their limitations, focused on conceptual and methodological components

	1. Decomposition of GPG	2. Regression analyses	3. Sequence analyses + Regression	
Outcome	Gender Pension Gap	(Women's) pension income	(Women's) pension income	
Covariates of interest	Summary measures	Summary measures	Life course typologies (clusters)	
Method	Decomposition analyses, usually KOB (based on OLS)	Usually OLS	Sequence analysis (SA) + OLS	
Examples	Bardasi & Jenkins, 2010; Bonnet et al., 2020; Cordova et al., 2022; Even & Macpherson, 2004; König et al., 2019; Nolan et al., 2019; Zhao & Zhao, 2018	Fasang et al. 2013; Sefton, Evandrou, and Falkingham 2011; Kreyenfeld et al 2022, Möhring 2018; Dotti Sani & Luppi 2020	Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2021; Tophoven & Tisch, 2016; Overweg, 2023	
Main limitations				Addressed by
Concealing life course complexities	x	x		Chapters 3 + 4
Risk of multicollinearity	x	x		Chapters 3 + 4 + 5
Focus on a selective set of life course states	x	x		Chapters 3 + 4 + 5
Not considering work-family interdependences	x	x	(x) ⁶	Chapters 3 + 4
Focus on employment life course categories (primarily full-time empl.)	x	(x)	(x)	Chapters 3 + 4 + 5
Risk concealing importance of single life course aspects	x	x	x	Chapter 5
No counterfactual analysis, i.e., no quantification of shares of GPG due to different mechanisms (e.g., compositional and returns effects)		x	x	Chapters 3 + 4 + 5

Note: x= applies. (x)= applies partly.

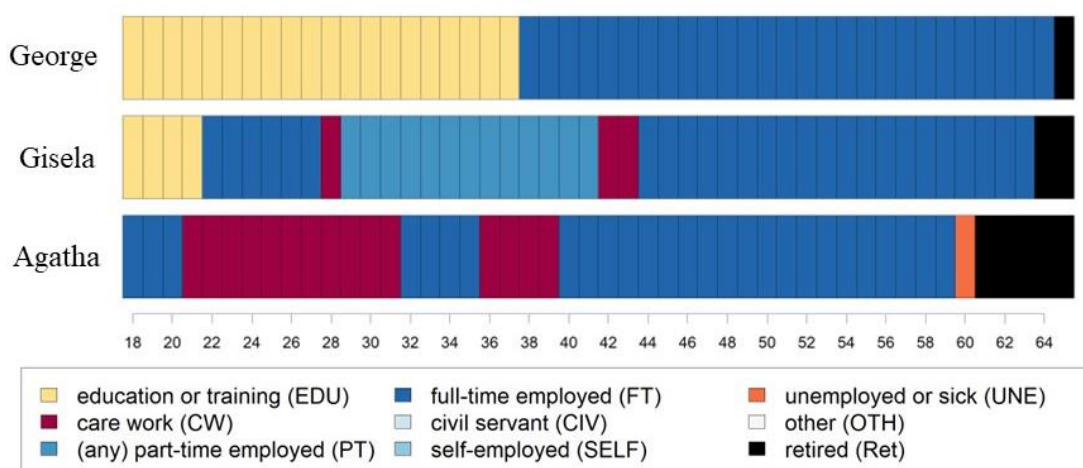
⁶ If single-channel SA applied, e.g., Tophoven & Tisch 2016 and Möhring & Weiland 2021.

Gaps in previous literature

Table 1.1 gives an overview of the main methodological and conceptual shortcomings of the different literature streams. The literature streams I (decomposition analyses) and II (regression analyses) have several limitations in common.

First, they rely on life course summary measures in their models, usually the years in different employment (or, less often, family) states, if they use life course data at all. However, modelling the working and family biographies based on summary measures, these approaches do not capture the complexity of life courses as pointed out more generally by Halpern-Manners et al. (2015) and Nutz & Lersch (Nutz & Lersch, 2021). Next to the notion of duration, life course complexities include the aspects of timing, ordering and overall complexity (see Chapter 1.2). Relying on the duration of states is likely to simplify complex life courses. This becomes visible in Figure 1.1. George, Gisela, and Agatha all spent almost the same duration in full-time employment (26-27 years). As such, based on the summary measure ‘duration’ their life courses are almost identical.

Figure 1.1 Example work life courses



Note: Based on Figure 5.1. Example working life courses selected from the analysis sample of Chapter 5. The colour code shows which working life states the individuals experienced for each year, spanning ages 18-65.

However, this is not the case. While George has a continuous full-time employment career until retirement after a long educational period, Gisela, and Agatha both experience several

care-related employment interruptions. Such employment interruptions are likely to matter for the accumulation of pension claims. Similarly, the timing of care-related employment interruptions (very early for Agatha, six years later for Gisela) might have consequences for the individuals' subsequent careers and, consequentially, for their pension claims (for more examples of life course complexities, see section 5.2.1).

As shown by Halpern-Manners et al. (2015) more sophisticated ways of measuring life courses reveal “something about the temporal characteristics of people’s work and family experiences” (2015, S. 1387) and therefore (statistically and substantively) improve the understanding of old age well-being. Given that the complexity of life courses is higher among women (Rowold & Van Winkle, 2023; Van Winkle & Fasang, 2017, 2021), using simplifying life course summary measures covering only one dimension of life course complexity, namely the duration, seems especially inadvisable for analysing gender inequalities in outcomes. Consequently, for the Gender Pension Gap, considering gendered life course complexities is highly relevant. However, this has been largely neglected by literature streams I and II.

Second, considering the duration in certain categories of one life course domain generates a risk of multicollinearity, because the categories of one domain are naturally correlated. For example, the years spent in full-time employment, part-time employment, unemployment, and care work will be somewhat collinear. Consequently, the duration in all categories of one life course domain cannot be included in the regression-based models. Some of the studies indeed mention such multicollinearity issues (Fasang et al., 2013; Frommert & Strauß, 2013)⁷

⁷ For instance, Frommert and Strauß (2013) note, that they did not include the wage in full-time employment due to the high correlation with education and sum of year working full-time. Similarly, Fasang et al. (2013) discuss that they cannot include all summary indicators on family histories due to collinearity.

Third, if multicollinearity is an issue, the researchers have to decide which variables to keep in the model and thus create selective sets of life course variables considered. Even though this selection will be theoretically driven, it might be biased. At least it is clear, that the previous literature focuses on the straightforward duration in states, and usually the duration in (full-time) employment (at least the studies applying decompositions, Table 3.1). If studies (have to) rely on a selective set of covariates, they risk concealing important associations with life course variables that were included in the model.

Fourth, life course complexities can be multi-dimensional (Bernardi et al., 2019). This means, for example, that complexity might not only unfold within one isolated life course domain (e.g., work) but transitions in one domain might be related to transitions in another domain. The most prominent, and gendered, example would be family and work interdependencies. For example, a mother withdraws from full-time employment (work domain) upon childbirth (family domain), but the father remains in the labour force. Indeed, research has shown across different countries that women's labour force participation, working hours and earnings drop dramatically for years after childbirth, while those of fathers tend to increase (de la Vega, 2022; e.g., Harkness et al., 2019; Musick et al., 2020; Van Winkle & Fasang, 2020). As such, women's careers are much more interrelated with their family lives than those of men. Consequently, and as was argued by literature very early on, it is important to consider the family life course next to the work life courses and the potential interdependences between them (Allmendinger et al., 1992, 1993; Fasang, 2009, 2010; Frericks, 2020; Ginn et al., 2001; Meyer & Pfau-Effinger, 2006). However, the decomposition and regression studies (literature streams I and II) have not taken into account work-family interdependences. For example, they could have included interactions between certain work and family characteristics in their modelling approaches. This, however, would likely have worsened the multicollinearity issues.

While addressing challenges 1-4 of the other two literature streams, studies combining regressions with sequence analysis (stream III) have two challenges in common with the other two approaches. First, they all risk concealing the association of single life course aspects (i.e., duration, timing, ordering) with the Gender Pension Gap. Streams I and II do not cover the link between single life course aspects and the GPG because of the selective sets of life course aspects that are considered due to multicollinearity issues or they only consider the duration in states. Literature stream III reveals the complexities of life courses. Similar life courses are grouped by SA and the life course groups are characterized by dominant patterns, e.g., out-of-labour & part-time employed or continuous full-time employment. However, the single trajectories within these groups of life courses will still differ, for example with regard to the age at the first full-time or part-time employment spell or the order in which periods of part-time and out-of-labour force periods occurred (for a more detailed discussion and examples, see Chapter 5.2.1). Thus, it remains unclear which of the different aspects clustered together in the life course groups drives the association with the outcome: the duration, ordering or timing of which categories?

Second, all literature streams focus on work-life-course aspects. The selective perspective on work biographies is particularly pronounced for the decomposition studies (stream I), while literature streams II and III provide few exceptions of studies that consider the family sphere (Evandrou et al., 2009; e.g., Fasang et al., 2013; Kreyenfeld et al., 2023; Madero-Cabib & Fasang, 2016). Beyond these exceptions, family characteristics are often only included as control and only a few times as life course variables in all three literature streams. This ongoing focus on the employment sphere, particularly among decomposition studies, is surprising given that empirical literature has confirmed the relevance of the family situation for women's or gendered retirement pathways and timing (Damman et al., 2014; Fasang, 2010; Riekhoff & Järnefelt, 2017), old age income (Evandrou et al., 2009;

Fasang et al., 2013; Madero-Cabib & Fasang, 2016) and private pension coverage (Yabiku, 2000).

Lastly, both regression analyses not using sequence analysis (literature stream II) as well as studies combining regressions with sequence analysis (literature stream III) only assess how life course patterns or aspects are associated with pension income. Unlike the decomposition studies, they cannot distinguish and quantify the parts of the Gender Pension Gap that are due to compositional differences between men and women and the parts that arise due to differences in returns for the same characteristics. However, this allows researchers to disentangle which characteristics are most, and through which mechanisms, related to the Gender Pension Gap and what should be targeted by policymakers if they wish to reduce the Gender Pension Gap.

In conclusion, I argue that the existing empirical literature on GPGs largely misses multidimensional and gender- and life-course-sensitive analyses. Despite the theoretical emphasis on the significance of the family sphere for women's pension income and empirical evidence that has lent support to this perspective, the majority of the literature has primarily concentrated on employment characteristics. Furthermore, the literature so far does not effectively model life course complexities (streams I and II) and their associations with Gender Pension Gaps (stream III), even though life course complexities become more important across cohorts, and are largely gendered (Rowold & Van Winkle, 2023; Van Winkle, 2018; Van Winkle & Fasang, 2017, 2021). The disadvantages of simplifying life courses become more severe, the more complexities (e.g., transitions) individuals experience. As a result, previous studies either insufficiently cover the complexity of life courses or focus on selective associations.

Given the focus on employment and the disregard of life course complexities, the “steady worker bias” (O’Rand, 1996, S. 235) applies to the conceptual and methodological choices

the literature has made: the empirical approaches centre around the notion of a continuous worker.⁸ These approaches thus risk concealing inequalities that arise beyond the axis of continuous work which apply disproportionately to women who are more likely to deviate from continuous work careers and experience more interdependences between work and family lives. This steady worker bias of the literature is in line with the reference of pension systems to continuous employment as well as the overall notion of retirement (see Chapter 2.2) but risks concealing the complex mechanisms in play for shaping GPGs.

Lastly, we do not know yet which specific elements of life courses are the most important drivers of the (gendered) pension income inequality and how a less selective set of empirically relevant life course covariates is related to GPGs. While the literature has largely focussed on the duration of (full-time) employment, we do not know whether it actually is the most important predictor of pensions or whether other important mechanisms have been concealed so far.

1.5 Aim of the thesis

The overall aim of this thesis is to enhance our understanding of how Gender Pension Gaps (GPG) are related to previous inequalities over the life course and draw policy implications therefrom. It intends to highlight and effectively address the necessity to more systematically realize the gender- and life-course-sensitive approach that I argue is needed for a more suitable analysis of gendered inequalities in old age. For this purpose, I identify and integrate crucial components of empirical analysis, encompassing the entire spectrum

⁸ O’Rand primarily applies the steady worker bias to the collection of data, but I argue that her definition also applies to studies that are conceptualised with a focus on continuous (full-time) employment: “This bias is derived from a lifetime employment model of the work career that conceptualizes the career as a relatively sustained sequence of full-time jobs of longer as opposed to shorter durations over the adult life span. Of course, this model has never been applicable to the entire workforce, especially to women’s more interrupted, mobile, and truncated work career patterns. Yet, this bias permeates the way we think” (O’Rand, 1996, S. 235).

from sample definition to methodological applications, and design, employing a gender- and life-course-sensitive framework.

Substantively, and in line with the notion of Cumulative Advantages and Disadvantages and the pertinent life course and gender literature, the aim is to study to what extent gendered work and family life course inequalities are reproduced in old age and provide an in-depth picture of the associations between life course complexities and GPGs. Methodologically, I aim to equip research in the field with analytical tools to study Gender Pension Gaps in a more life-course- and gender-sensible way. To address these aims I develop and enhance a combination of research designs and methods that effectively implement multidimensional gender- and life-course-sensitive approaches, which provide a comprehensive description of the complex accumulation of gender inequalities over the life course and their relation to outcomes in old age. Based on the established theoretical considerations, I consider family life courses beside work biographies as well as interactions between them.

Jointly, analysing the association of life course patterns, or relevant aspects thereof, with the GPGs, Chapters 3-5 address the challenges of the empirical literature as identified in Chapter 1.3 (see Table 1.1). The presented conceptual choices, methodological innovations, and substantive applications aim to actively prevent the gendered implications of the “steady worker bias” (O’Rand, 1996, S. 235) that applies largely to previous research.

1.6 Chapter overview and contributions

Chapters 2-5 address different limitations of previous literature and thus contribute in different ways to the field of Gender Pension Gaps, and, partly, inequalities in old age more broadly. Table 1.2 gives an overview of the main objectives for each of the Chapters and Table 1.1 depicts which Chapter addresses which limitations in previous literature.

In Chapter 2, more broadly, I discuss the main conceptual decisions which I identify as highly important when analysing gender inequalities in old age such as the Gender Pension Gap. I argue that the implementation of life-course- and gender-sensitive analyses do not only start with methodological modelling, but that the research design, data, and any measures used should be chosen and discussed carefully as well. I reflect on different options to conceptualize and measure pension income, and the analysis sample in more gender- and life-course-sensitive manners. Such conceptual decisions are the prerequisite for any robust and thoughtful empirical analyses but their detailed description and justification often come short in studies on old-age inequalities, particularly compared to the methodological approaches. Thus, I carefully discuss and partly descriptively underline the reasoning for the conceptualisation of the dependent variable, pension income, the covariates of interests (work, and family life courses), as well as the sample definition as it is applied in Chapters 3-5. This includes a descriptive analysis of different sample definitions across countries and genders that reveals severe selection biases in popular sample definitions for studies on the elderly (section 2.2).

Additionally, I show how the processing of retrospective life course data can involuntarily be subject to gendered biases by uncovering and solving highly gendered patterns of missing information for the employment life courses in the generated Job Episodes Panel (Brugiavini et al., 2013, 2019) (section 2.3).

Chapter 2 provides the background for the conceptual and empirical decisions in Chapters 3-5. It contributes to the literature by guiding future research in key conceptual choices and preventing gendered sample and data biases.

Table 1.2 Chapter overview

Chapter	Authors	Data	Countries	Focus	Perspective	Methods	Main contribution
2	Rowold	SHARE	All countries covered in SHARE	Conceptual: gender- and life-course-sensitive research design		Theoretical discussion, literature review, descriptive analyses	- Identifies and discusses main conceptual components of empirical research relevant to gender- and life-course-sensitive research designs (Chapter 2.1) - Reveals gendered sample bias of popular sample definitions used in studies on old age (Chapter 2.2) - Reveals and solves highly gendered bias in popular generated dataset on work life courses (Chapter 2.3)
3	Rowold, Struffolino, Fasang	SHARE	West Germany, Italy	Methodological: Introducing SA-KOB, Guidelines for applied research Substantive: Illustration Gender Pension Gap	Holistic life course perspective	Multichannel sequence analysis, SA-KOB	- Introduces SA-KOB for a more life course and group-sensitive decomposition of group inequalities in outcomes - Derives guidelines for the empirical implementation of SA-KOB - Substantial: reveals and quantifies how gendered work-family life courses are related to the GPG
4	Rowold	SHARE, SHARE-RV	West Germany	Substantive: pension privatisation and role of redistribution in public pensions, earnings over the life course	Holistic life course perspective	Multichannel sequence analysis, SA-KOB	- Decomposition based on relative earnings over the life course (included as sequence state) - Differentiation between gender gap in public versus private pensions and their relationship with work-family life courses - Richer and more reliable data for retrospective work life courses (admin records) - Decomposition along the distribution and of the Gender Coverage Gap
5	Rowold	SHARE	West Germany, Netherlands	Substantive: life course aspects, important pension predictors Methodological: combination of Life Course Feature Selection with Ñopo decomposition; Matching based on most important life course pension predictors	Focus on a multitude of single life course aspects	Life course feature selection + Ñopo and KOB decomposition	- Reveals most important life course predictors of pension income (among hundreds of proxies for the duration, ordering, timing, and complexity of life courses) - Reveals and quantifies how most important life course aspects are related to the GPG - Ñopo decomposition quantifies which parts of GPG due to gender-exclusive combinations of life course experiences

Chapters 3-5 apply decomposition techniques and complement each other. Even though Chapters 3-5 are closely related, they are written as stand-alone journal articles.⁹

Chapters 3 and 4 take a holistic life course perspective by decomposing the GPG based on groups with similar life courses. I apply a novel two-step approach combining two established methods that developed and were so far applied independently from each other: sequence analysis (SA) and the Kitagawa-Oaxaca-Blinder (KOB) decomposition. Sequence analysis is a fast-growing toolbox that allows to investigate patterns of trajectories based on categorical states (Liao et al., 2022), such as family or work categories. In combination with KOB, SA is used in its classical form with cluster analysis to categorise individuals with similar trajectories into the same cluster (Abbott & Forrest, 1986). Specifically, I apply Multichannel SA (Gauthier et al., 2010; Pollock, 2007) which allows to consider work and family life courses at the same time and results in clusters of individuals with similar work *and* family trajectories. This is to cover empirically the theoretically relevant interdependences between work and family lives, particularly for women's pension income (Allmendinger et al., 1992; Fasang, 2009; Ginn et al., 2001). Benefitting from the powerful visualization tools of SA (Fasang & Liao, 2014), the clusters then reveal life course complexities such as ordering, timing, and duration of life course states (Liao et al., 2022) and allow for a thick description of gendered work-family life courses, including the interdependences of work and family events.

The KOB decomposition (Blinder, 1973; Kitagawa, 1955; R. Oaxaca, 1973) is a popular approach to investigate the underlying mechanisms of group inequalities in outcomes, such as the Gender Pay, Wealth or Pension Gap (Bardasi & Jenkins, 2010; Blau & Kahn, 2017; Cordova et al., 2022; Kunze, 2018; Waitkus & Minkus, 2021). It reveals the part of the

⁹ Abbreviated versions of the Chapters are accepted for publication (Chapter 3), under review at international scientific journals (Chapters 2.2 & 4) or being prepared for submission (Chapter 5).

group inequalities due to differences in characteristics (composition component) and the part arising from differences in returns for the same characteristics (returns component).

While Chapters 3 and 4 both apply the SA-KOB decomposition, Chapter 3 focuses on the methodological contributions and Chapter 4 adds more on the substantive contributions that can be derived from the SA-KOB decomposition.

In *Chapter 3*, together with my co-authors, I methodological and technically developed and introduced the combination of sequence analysis with the Kitagawa-Blinder-Oaxaca (KOB) decomposition to analyse gender gaps in outcomes, such as pension income. The SA-KOB decomposition allows for the differentiation between the part of the Gender Pension Gap due to compositional life course differences and the differences in returns for similar life courses.

Chapter 3 contributes to the literature by introducing the SA-KOB decomposition, describing the steps of both methods – the sequence analysis as well as the KOB decomposition – in detail and developing specific guidelines that will help researchers implement the SA-KOB decomposition. While this Chapter's contributions are primarily methodological, we apply the SA-KOB to GPGs in Italy and West Germany to illustrate the analytical steps and interpretation of the results. Our results prove that the combination of MSA and KOB provides new substantive insights into the study of Gender Pension Gaps, such as the relevance of the gender-specific interrelations of work and family life courses. By enabling a more life-course- and group-sensitive decomposition of group inequalities, the SA-KOB decomposition provides a promising analytical tool for assessing inequalities in other life-course-sensitive outcomes as well as across other groups.

Chapter 4 builds upon Chapter 3 and contributes to the literature primarily substantively by applying the SA-KOB to richer and more reliable life course data. It reveals i) how earning differences are reproduced in pension income and ii) to what extent work-family

life courses are related differently to gender inequalities in public and private pensions. Linked survey-register data allows the inclusion of annual earning levels beyond different employment states in the sequence analysis. I thus decompose gendered pension income inequalities the first time based on the relative earnings positions over the life course. Differentiating between public and private pensions allows disentangling more precisely the potential role of redistribution in public pensions as well as of the pension privatisation on the GPG. Furthermore, it reveals which pension pillar reproduces gendered earnings inequalities over the life course to a larger extent.

The methodological approach in Chapters 3 and 4, the SA-KOB decomposition, addresses multiple limitations of the previous literature (see Table 1.1). By reducing dimensions through clustering, sequence analysis largely prevents multicollinearity issues and allows for considering all categories within life course domains (up to a feasibility maximum). Consequently, researchers do not have to rely on selective sets of life course aspects. Life course complexities are modelled and visualised, and, in the case of Multichannel sequence analysis, work, and family interdependences are considered and their roles in shaping Gender Pension Gaps are quantified.

Chapter 5 shifts away from the holistic perspectives of Chapters 3 and 4 and focuses on specific, single life-course aspects underlying the whole work and family trajectories. GPGs are so far either analysed based on one life course dimension, mostly the duration in (full-time) employment (previous literature), or based on life course typologies from sequence analysis (as in Chapters 3 and 4). I parse work and family life courses into the ordering, timing, and duration of the life course categories individuals experience as well as their overall life course complexity. Applying the Life Course Feature Selection (Bolano & Studer, 2020) I disentangle which of the resulting hundreds of life course proxies are the most important pension income predictors. In the second step, I apply the Ñopo

decomposition which expands on the logics of the KOB decomposition by using matching techniques. In a nutshell, the explained and unexplained shares of the KOB are estimated over a sample of matchable men and women concerning the set of observed characteristics. Combining these advanced quantitative methods allows for an in-depth analysis of the underlying mechanisms of the Gender Pension Gap with regard to work and family life course aspects.

Chapter 5 contributes to the previous literature as well as to Chapters 3 and 4 by uncovering the most important single life course aspects (duration, timing, order, and complexity) of family and work life course characteristics for pensions. Thus, the decomposition of the Gender Pension Gap in the second analytical step does not rely on a selective set of life course aspects but on those that are most important for predicting pensions. Compared to previous literature, it assesses specifically whether research has focussed so far on the correct life course aspect, that is the duration in specific states and particularly the years spent in full-time employment. Compared to the holistic approaches of Chapters 3 and 4, it disentangles which of manifold life course aspects within typical life course patterns are most relevant. Furthermore, introducing the Ñopo decomposition to the field of gendered pension inequalities and combining it with life course covariates proves very promising for revealing new insights into how gender-specific life course experiences are related to pension inequalities in old age.

Country context design & data. I apply country comparisons (Chapters 2, 3, & 5) and stratify the analyses by pension types (Chapter 4) to discuss the role of pension systems in shaping gendered pension income inequalities. By comparing countries that represented gender-conservative welfare contexts in the 20th century but have different pension systems, I highlight the added value of the analytical approaches for comparative welfare state studies. The country comparisons further add to the literature on gendered pension

inequalities which has mostly focussed on the comparison of Germany with the UK (Fasang, 2010, 2012; Fasang et al., 2013; Meyer & Pfau-Effinger, 2006; Sefton, Evandrou, Falkingham, et al., 2011) or multicountry comparisons (Bettio et al., 2013; Dotti Sani & Luppi, 2021; Möhring, 2018; Veremchuk, 2020). The multicountry studies do not allow to discuss in detail the contexts and results for all countries considered. Most studies, however, have focussed on one country (Table 2.1, Table 3.1). In sections 2.2 and 2.3 of Chapter 2, however, I apply multicountry comparisons to show to what extent the gender biases I uncover are stratified across different country contexts.

This thesis explicitly focuses on conservative country contexts, which are likely to have shaped similarly highly gendered life course experiences, especially for today's birth cohorts in old age. The aim is to keep the level of gender differences in the life course more or less constant in order to better identify the role of different pension systems for the Gender Pension Gaps. A preliminary application of SA-KOB has shown that Gender Pension Gaps are not related to gender-specific life courses in contexts that shape more gender-equal life courses such as East Germany and the Czech Republic (Rowold, 2022). As such, welfare policies that help to shape more equal life courses lead to the lower GPGs in both former socialist countries. While this is an important finding that should be tested for other countries, focussing on different contexts with high levels of gender segregation on different life courses helps to disentangle to what extent different pension policies channel the relation between life courses and the GPG. Furthermore, a certain degree of gender segregation on life courses is necessary to introduce the two combinations of methods aimed at quantifying the relationship between gendered life courses and GPGs. Otherwise, it would not have been possible to discuss exemplary interpretations, advantages, and limitations of these approaches. Finally, the inclusion of additional country contexts to cover different gender arrangements and pension systems may have added to

the complexity of chapters 3, 4, and 5 given the space required to introduce the advanced methodological approaches. However, it would be important for further policy-relevant conclusions to extend the analyses to other country contexts and comparisons in the future (see section 6.4.2).

Rich survey data from the Survey of Health, Ageing and Retirement in Europe (SHARE) enables these country comparisons as it provides harmonized data for most European countries. The SHARE survey focuses on the elderly by surveying individuals aged 50 and older and their partners. The birth cohorts of the analysis samples used in this thesis were born primarily between 1930 and 1950. This thesis thus assesses the gender inequalities of individuals who are at old age nowadays. I derive the dependent variable pension income as well as main sociodemographic data such as gender and age and information on controls from SHARE. For the operationalization of life courses, rich data on the work and family biographies is essential. SHARE surveyed two waves of retrospective life course data (SHARELIFE) which provide information on work and family lives for each year in the past of the observed individuals (see Chapter 2.3 for more details).

Jointly, Chapters 2-5 raise awareness for the need and usefulness of more life-course- and gender-sensitive analyses and provide the field on gender inequalities in old age with new methods and empirical evidence. Applying novel methodological approaches, Chapters 3-5 reveal how Gender Pension Gaps are shaped through the interaction of welfare state policies in the 20th century, incentivising highly gender-segregated life courses, and pension policies nowadays, rewarding these gender-specific life courses unequally. Highlighting the relevance of specific work-family life courses or specific life course aspects allows policymakers to employ more targeted policies to prevent the reproduction of gender inequalities over the life course in old age. The contributions and implications derived from the analyses are discussed more in detail in Chapter 6.

2 Conceptualizing analytical decisions in studies on the elderly: a gender- and life-course-sensitive approach

Researchers need to make informed decisions on how to define their sample and measure key variables for empirical analyses. However, the conceptualization of such core decisions is more complex if indicators and outcomes of interests are measured at older ages because they are shaped by an individual's prior life courses. As such, indicators measured at later life states might have been impacted by inequalities as they emerged and accumulated over the life course. For example, how to deal with individuals with no pension claims? Which pension pillars and specific pension income types should be considered? Or, how to define an old-age sample without excluding systematic groups of individuals? Thus, particularly when studying inequalities in old age, a careful conceptualization of analytical core decisions and measurements is necessary, specifically regarding prior life course inequalities.

In Chapters 3-5 of this dissertation, the sample is the old-age population in different countries, the outcome is pension income and the covariates of interest model work-family life courses. All these elements can be defined and operationalized in various ways. While Chapters 3-5 are transparent about these definitions, in this Chapter I discuss them more in detail and partly ground the decisions on empirical investigations. Clarifying these choices and their implications for the results is highly important given that they are likely to impact the results substantially. Still, such conceptual decisions often do not receive the same attention and careful investigation as the methods applied. Given my focus on gender inequalities in pension income, I argue that it is particularly important to consider that the measures and data used are conceptualized in a gender and life-course-sensitive way. Otherwise, research risks systematically excluding specific subgroups of the elderly and

therefore unconsciously bias the results. Discussing such conceptual decisions beyond the modelling approach satisfies the aim of the thesis to systematically establish elements of empirical analysis needed for a gender and life-course-sensitive analysis of gender inequalities in old age.

Chapter 2.1 systematically identifies and discusses three important parameters for the conceptualization of (gendered) pension income: the unit of analysis (individual or couple), the definition of pension income, and whether to take into account access to pensions in the first place. Table 2.1 gives a literature overview of how previous studies have dealt with these conceptual choices that I argue are the most important and influential ones for gendered pension income inequality. Empirical literature needs to be transparent about these.

In Chapters 3-5, I take an individual perspective and focus on individuals' own, independent pensions. This conceptualization focuses on financial independence and sustainability and is, as I argue, most relevant for the future of gendered pension income inequalities given the recent socio-demographic and policy-related developments (e.g., pension privatization). Due to this focus, I do not consider survivor pensions for the measurement of pensions. At the same time, I include individuals with no access to own pensions ('Gender Coverage Gap') in my operationalization of pension income to model the full extent of gendered pension income inequality.

Chapter 2.2 discusses and shows empirically how popular sample definitions used in studies on the elderly population produce severe gendered sample biases in many countries. To not exclude selective groups of individuals, I apply an age cutoff as the main sample criteria in Chapters 3-5.

Chapter 2.3 discusses how the retrospective data on life courses is surveyed, how biases can also apply to data generation (here: the popular Job Episode Panel used for work trajectories in SHARE) how to solve it, as well as which limitations of the retrospective data remain.

After each subchapter, I describe the implementation of the discussion or the empirical analysis for Chapters 3-5.

2.1 Perspective and definition of pension income

Pension income is a complex outcome and can be defined in various ways. There are several dimensions to consider and the most relevant parameters for the aims of this dissertation are the unit of analysis, the pension income pillars and types considered, and how to deal with individuals without any pension claims., I discuss the benefits and limitations of different choices from a gendered perspective.

2.1.1 Unit of analysis: Individual versus couple-level perspective

The *couple* (or household) *perspective* considers the income of both partners in a heterosexual couple, e.g., sums up all pension income of both partners. Particularly in old age, the couple's perspective is likely to often overlap with the household perspective (i.e., how many individuals live in the household) given that children have usually moved out. The household perspective is important since it is the better proxy for individuals' material well-being (Sefton, Evandrou, & Falkingham, 2011), and social security indicators are measured based on it, such as poverty risks in old age (e.g., Smeeding & Sandstrom, 2005). As such, the couple's perspective might reveal who requires higher pensions from a welfare and social security perspective.

From the couple's perspective, women in a traditional male breadwinner setting, i.e., the man's main responsibility is paid work and the women's is unpaid care work, are secured

against poverty by their partner's (high) pensions, even if they receive no or only a minor own pension. Indeed, in Germany, women with a retirement pathway characterized by no own personal income still have a medium-high household income (Fasang, 2010). It is argued that women with low pensions are not subject to old-age poverty but rather live in household contexts with relatively high incomes (Faik & Köhler-Rama, 2012). The authors conclude that the GPG is an inappropriate indicator because women's financial well-being is often secured by relatively high-income households. However, the couple's perspective implicitly assumes that both partners have access to the household income and share their finances equally (Sefton, Evandrou, & Falkingham, 2011), ignoring that women's financial well-being depends fully on their partners' income (Fasang, 2010) and in fact, that wives are only "a husband away from poverty" (Orloff, 1993, S. 319). Similarly, Sefton et al. (2011; 2011) argue that the couple's perspective obscures the financial dependence of women and the direct association of women's own work and family life courses on their pension income. While some public pension systems secure women financially in case of the death of a partner (widow or survivor pensions), such pension types seem not sustainable for the future as they do not provide financial independence. Additionally, pension privatization threatens the existence and generosity of (public) survivor pensions (Bonnet & Geraci, 2009; Frericks et al., 2007). Moreover, the couple perspective assumes that the gendered division of labor is desired in the couple context (Faik & Köhler-Rama, 2012) and thus ignores the impact of welfare state policies for shaping gendered opportunities, particularly concerning the reconciliation of paid work and family responsibilities (e.g., Ferragina, 2019; Gangl & Ziefle, 2015; Hegewisch & Gornick, 2011).

On the other hand, the *individual perspective* considers the income of each individual separately, independent of household context and income available in a household. From this perspective, women with low or no own pension in traditional male breadwinner

couples are doing financially worse than in the couple perspective, where a partner's pension buffers their own low pensions.

The two different perspectives are related to three different pathways through which women can attain pension income that the literature on gender inequalities in pensions distinguishes (Ginn et al., 2001). The first one takes the couple's perspective and considers the sharing of the partner's (mostly the husband's) pensions. The second pathway is based on own pension claims that are derived from the (former) husband's pension contributions, e.g., survivor pension (see Chapter 2.1.2), and can be seen as dependent individual level. The third pathway is own pension derived from women's own contributions only.

While only the third pathway, the independent individual level, provides women with an independent and secured source of income (Ginn et al., 2001), the first two pathways are usually the main pathways to pension income for women in traditional male breadwinner models (Leitner, 2001). The independent individual perspective has been highlighted as particularly relevant for women's income from the theoretical and feminist perspectives. It provides financial security and independence and "reduced vulnerability" (on personal earning; Muller et al. 2020: 2). Conversely, other options such as sharing a partner's pension or pension claims that are derived from family members, predominantly the husband, are tied to a loss of autonomy and increased dependence (Ginn et al., 2001).

Table 2.1 Overview of literature on gendered pensions concerning conceptual choices

					Pension type			Sample	
Author	Year	Literature stream*	Country	Unit of analysis	Pillars	Survivor pension	Form	Criteria	Gender coverage gaps
Bardasi & Jenkins	2004	1	UK	individual	occ. + pers.	excluded	log	pension receipt for age KOB, cutoff	two-part model; 1. decomposition of gender coverage gap, 2. decomposition of GPG among pension recipients with selection correction
Bettio	2013	1	multi-country	individual	unclear, probably sum of all three pillars	included	log	pension receipt, age cutoff, and ever paid work	0 income excluded
Bonnet et al	2020	1	France	individual	public; occ. + pers.	included	log	pension receipt	0 income excluded
Cordova	2022	1	Germany	individual	pension wealth (accrual), sum of all three pillars	unclear, not addressed	absolute	age cutoff, non-retired population	unclear, not addressed
Dotti Sani & Lupi	2021	2	multi-country	Individual, household	sum of all three pillars	included	log	age cutoff, not in active employment	0 income included
Evandrou et al Even & Macpherson	2009	3	UK, US, and West Germany	individual	sum of all three pension pillars + public transfers + other income (incl. earnings)	included	log	pension receipt, age cutoff, only women	0 income excluded
	2004	1	US	individual	unclear, probably sum of all three pillars	unclear, not addressed	log	pension receipt for KOB, otherwise unclear	two-part model; 1. decomposition of gender coverage gap, 2. decomposition of GPG among pension recipients with selection correction

					Pension type			Sample	
Author	Year	Literature stream*	Country	Unit of analysis	Pillars	Survivor pension	Form	Criteria	Gender coverage gaps
Ezeyi & Vujčić	2017	1	UK	individual	public; occ. + pers.	unclear, not addressed	log	self-reported retirement status, age above state pension age, not working	two-part model; 1. decomposition of gender coverage gap, 2. decomposition of GPG among pension recipients
Fasang et al.	2013	2	UK, West Germany	individual	sum of all three pillars, + social assistance benefits + employment income	included	absolute	age cutoff, only women	0 income included
Frommert & Strauß	2013	1	West Germany	individual	public; sum of all three pillars	unclear, not addressed	absolute	public pension receipt, age cutoff	unclear, not addressed
Gardiner et al.	2016	2	UK	individual	occ. + pers.	included	absolute	employees aged 50 to 59	two-part model; 1. model of gender coverage gap, 2. model of GPG among pension recipients with selection correction
Hänisch & Klos	2014	1	Germany	individual	sum of all three pillars	included	log	retirees, age cutoff	unclear, not addressed
Johnson et al.	1999	1	US	individual	occ. + pers.	unclear, not addressed	log	full-time employed worker with pension plan, aged 51-61	two-part model; 1. decomposition of gender coverage gap, 2. decomposition of GPG among pension recipients with selection correction
König et al.	2019	1	Sweden	individual	sum of all three pillars	unclear, not addressed	log	retirees, pension receipt	0 income excluded
Kreyenfeld et al.	2022	2	West Germany	individual	public	excluded	absolute	age cutoff	0 income included
Levine	1999	1	US	individual	household pension wealth (anticipated), sum of all three pillars	unclear, not addressed	log	only currently non-married, age cutoff	0 income included

Pension type								Sample	
Author	Year	Literature stream*	Country	Unit of analysis	Pillars	Survivor pension	Form	Criteria	Gender coverage gaps
Madero-Cabib & Fasang	2016	3	Germany, Switzerland	Individual, household	sum of all three pillars (first independent pension income, (2. outcome: household income)	included	absolute	self-reported retirement status, age cutoff	included for subset (never worked and married)
Möhring	2015	2	multi-country	individual	sum of all three pillars, and other market income	excluded	log	age cutoff + not in active employment	0 income included
Möhring	2018	2	multi-country	individual	sum of all three pillars, including social assistance	included	log	age cutoff + not in active employment	0 income included
Möhring & Weiland	2021	3	Germany	couple	sum of all three pillars	does not apply	absolute	married couples, age cutoff	0 income included
Nolan et al	2019	1	Ireland	individual	occ. + pers.	unclear, not addressed	log	retirees, age cutoff	unclear, not addressed
Overweg	2023	3	West Germany, Netherlands	Individual, household	sum of all three pillars, including social assistance, (2. outcome: household income)	excluded	log	age cutoff, only women	0 income included
Sefton et al.	2011	2	UK	individual	total person income: sum of all three pension pillars, public transfers, and other income	included	log	age cutoff	unclear, not addressed

Pension type								Sample	
Author	Year	Literature stream*	Country	Unit of analysis	Pillars	Survivor pension	Form	Criteria	Gender coverage gaps
Sefton et al.	2011	2	UK, US, and West Germany	individual	total person income: sum of all three pension pillars, public transfers, and other income	unclear, not addressed	log	age cutoff	unclear, not addressed
Tophoven & Tisch	2016	3	Germany	individual	public	does not apply	absolute	age cutoff	unclear, not addressed
Veremchuk	2020	1	multi-country	individual	sum of all three pillars	included	log	self-reported retirement status, age cutoff	unclear, not addressed
Zhao & Zhao	2018	1	China	individual	occ.	unclear, not addressed	log	age cutoff	two-part model; 1. decomposition of gender coverage gap, 2. decomposition of GPG among pension recipients

Note: Literature streams refer to the three categories of empirical literature discussed in Chapter 1.4 and depicted in Table 1.1 [stream 1: Decomposition of GPG, stream 2: regressions-type analyses with women's or gendered pension income inequality as focus, stream 3: sequence analysis and regression with women's or gendered pension income inequality as focus]. The studies selected for the Table do not cover all studies in the field but provide a selection of studies relevant to this dissertation. Occ.= occupational pensions, pers.=personal pension.

Möhring and Weiland (2022) were the first to fully empirically implement the couple's perspective by assessing the relationship between couples' life courses with the share of women's income within the elderly couple. Other studies in the field have used both, individual and household pension income (or equivalents) as outcomes in their analysis (Dotti Sani & Luppi, 2021; Fasang, 2010; Möhring, 2021; Overweg, 2023; Riekhoff et al., 2023). Two studies for European countries find no large differences in the association of life course summary variables with individual income compared to household income for both, men and women (Möhring, 2021; Riekhoff et al., 2023). Fasang (2010) finds in Britain and West Germany that part-time employment trajectories of women before retirement are relatively similarly associated with individual compared to household level income, whereas the trajectories characterized by no income are secured against poverty through a male breadwinner in the household. Overweg (2023) suggests in West Germany and the Netherlands that strong associations between women's typical work-family life course patterns and women's individual pension income do not hold for household income. The differentiation between individual and household pension income still needs more attention. The literature on gendered pension income inequalities focuses on the individual perspective so far (see Table 2.1). The current evidence is mixed but the majority of studies rather suggest fewer differences in results across using individual and household income than one would expect.

Implications for this dissertation

In this dissertation, I take the individual perspective and thus define pension income as independent individual income based on an individual's own achievements. I follow the argumentation that this type of pension income provides financial security and independence.

I argue further that the individual, independent perspective is the most relevant for the aim of this dissertation and for drawing conclusions about women's financial risks in the future given recent demographic trends. First, individual, independent pensions are more directly linked with individual work-family life course experiences (similarly argued by Sefton, Evandrou, and Falkingham 2011). Consequently, individual pensions are best suited to reveal the associations between Gender Pension Gaps and life course complexities – the aim of this dissertation. The empirical Chapters 3-5 thereby focus on the financial independence of women. Second, while the individual perspective cannot draw concrete implications for the overall financial well-being of current cohorts of women in old age, it most likely becomes more relevant in the future because of different demographic and social policy trends. Family life courses become more complex across cohorts, and union dissolutions increase not only in general (Van Winkle, 2018) but also the divorce rates among the elderly rise (gray divorce revolution, Brown & Wright, 2017). Women in old age are much less likely to repartner than men and face a much larger decline in the standard of living after divorce in old age (Lin & Brown, 2021; Schimmele & Wu, 2016). Less stable marriages make it riskier to rely on the household component or survivor pensions. Additionally, the division of labor following the traditional male breadwinner concept is increasingly less incentivized by family policies compared to decades ago (e.g., increasing public child care, parental leave reserved for fathers), and leads slowly to a more equal division of labour realized in couples across cohorts (e.g., Sullivan, 2021; Zoch & Heyne, 2023). And lastly, in most pension systems we observe a trend towards the individualization and privatization of pensions (Cooke, 2006; Ebbinghaus, 2011b; Frericks et al., 2007; Hinrichs, 2021; Kuitto et al., 2021; Pedersen, 2004). This trend highlights the increasing relevance of own, independent pension claims, as the privatization of pensions leads to cuts in redistributive elements of the pension systems (e.g., Frericks et al., 2007).

As a result, I assume that individual independent pensions are more relevant for overall financial well-being in old age in the future (similarly argued by Frommert & Strauß, 2013), where male breadwinner models and stable marriages will be less common, related pension elements such as survivor pensions as well as overall redistributive elements might be less *generous*, and individuals must increasingly rely on their own, independent pensions. Financial independence is more sustainable for the future as it provides more financial security than financial dependence on partners' incomes or certain partnership dynamics (e.g., stable marriage or widowhood without remarriage).

The couple perspective is still important for social policy reasons, particularly for current cohorts in old age, for which unions are more stable than for younger cohorts. The presented methodological innovations and research designs can be applied in future research to the couple or household perspective, for example for decomposing within-couple GPGs.

2.1.2 Pension income types

Another important choice is the definition of pension income. More broadly, pension income categories or ‘pillars’ usually distinguish public, occupational, and personal pensions. The first pillar, public (or state) pensions, covers any pensions paid by the state, such as public old-age pensions. The second pillar, occupational pensions, refers to pension schemes on the firm level sponsored by the employer and the third pillar, personal pensions, to personal saving plans (Ebbinghaus, 2011a). In the theoretical and empirical literature, there is a consensus that women benefit more from the public compared to occupational and personal pensions because public pensions are more accessible and contain more redistributive elements (e.g. Ginn, 2004; Jefferson, 2009; Möhring, 2018). As a result, few studies focus on private pensions (here: sum of occupational and personal pensions) or differentiate between public and private pensions. Most studies assessed the sum of income

from all pension pillars and do not differentiate between pillars, while few do not specify in detail which pillars are included in their analyses (Table 2.1).

The types of pensions vary from the usual old-age pensions, and early retirement pensions to survivor pensions, and war pensions. While most studies on the Gender Pension Gap apply the individual perspective (see Chapter 2.1.1), there is more variance and often a lack of transparency regarding which pension income types are used. However, such decisions have important implications for pension inequalities. For gender inequalities in pensions, it is highly relevant how to treat survivor pensions. Whether to include survivor pensions largely impacts the GPG (Bonnet & Geraci, 2009; Tinios et al., 2015) given that survivor pensions particularly benefit women (e.g., Corsi & D'Ippoliti, 2009). While a slight majority of studies on gendered pension income inequality include survivor pensions, almost as many studies do not address how they deal with survivor pensions, and a minority exclude survivor pensions from the analysis (see Table 2.1).

Widow pensions were implemented in pension systems to financially secure specifically women in traditional breadwinner couples relying on their husbands (Bonnet, Hourriez, & Reeve, 2012). Given the changing family forms over cohorts with less stable marriages and more union dissolutions (Van Winkle, 2018) as well as the grey divorce revolution (e.g., Brown & Wright, 2017) strongly questions “[b]oth the principle and the fitness for purpose of [...] survivor’s pension[s]” (Bonnet, Hourriez, & Depledge, 2012, S. 147). It follows that fewer individuals might be eligible for survivor pensions in the future. Moreover, relying on survivor pensions derived from a former partner is a form of financial dependence (Fasang, 2010) and thus not financially sustainable. The decline in spending on survivor pensions in several European countries limits the reliance on this form of pension further (Bonnet & Geraci, 2009; Frericks et al., 2007). Lastly, it is not a durable

pension entitlement women can rely on as survivor pensions expire upon remarriage and thus depend on women's marriage status (e.g., in Germany Fasang, 2009).

Another parameter is whether to transform the pension distribution. Some studies use the natural logarithm of absolute pension income to transform the distribution towards a normal distribution while others rely on pension income in absolute terms (see Table 2.1).

Implications for this dissertation

In Chapters 3 and 5, I use the sum of all pension pillars to analyze gender inequalities in overall pensions. Doing this, the pension income definition also considers country differences in the pension design with regard to different relevancies of the different pension pillars which is highly important for the country comparisons. Additionally, considering all pension pillars allows to model the overall pension income inequality. However, it does not allow to reveal whether findings differ across pension pillars. In Chapter 4, focussing on West Germany, I differentiate between public and private pensions (sum of occupational and personal pensions) to assess whether the gender inequality in the different pillars is differently associated with gendered work-family life course patterns.

In Chapters 3-5, I do not consider income sources that are derived from other individuals. Consequently, I exclude survivor pensions. This is in line with the overall focus on independent pension income and financial independence (see Chapter 2.1.1) as has been argued previously (e.g., Kuivalainen et al., 2020; Möhring, 2015) but rather increases the level of the GPG (Bonnet & Geraci, 2009; Tinios et al., 2015). Table 2.2 shows the pension income types considered per pension pillar in the main analyses.

Table 2.2 Pension income types included.

Public pension	Occupational pension	Personal Pension
1. Old age pension	1. Old age pension from the last, second, or third job	1. Average payments of regular life insurance
2. Old age supplementary pension or public old-age second pension	2. Early retirement pension	2. Regular private annuity or private personal pension
3. Early retirement or pre-retirement pension	3. Disability or invalidity insurance	
4. Main or secondary disability insurance pension, or sickness benefits		

In Chapters 3-5, I use absolute pension income. While some argue that using the logarithm eases the interpretation since the coefficients reflect the percentage change (e.g., Nolan et al., 2019), the approximation works less well for large gaps in outcomes, such as the Gender Pension Gap. I argue instead that using the absolute income allows for a more straightforward and tangible interpretation. I also report the relative contribution of different decomposition components to the GPG (calculated based on absolute pensions), which is important for the country and pension pillar comparisons, as the absolute pension levels differ.

2.1.3 Considering access to pensions and the Gender Coverage Gap

Having access to pension claims in the first place is an important dimension of pension inequality. Depending on the aim of the analysis, it should be considered in modeling pension income.

Access to pension is determined by pension regulations. In some pension systems, mostly Beveridgean types, flat-rate public pensions guarantee access to pensions for almost all individuals of a certain age (e.g., Netherlands). In other pension systems, access to pension entitlements is limited by a minimum contribution requirement, often defined by a minimum of years individuals have contributed. The minimum contribution requirement varies largely across pension systems. For example, only 5 years were required in Germany, while 20 years were required in Italy in 2013. Whether individuals reach such criteria largely depends on their work and family life courses since contribution years are mostly defined by pension-relevant employment but also care work periods can count as a contribution. Access to pension is thus determined by an interaction of pension policies and life course inequalities, making it a relevant dimension for exploration within this dissertation.

Furthermore, in many countries, there is a substantial Gender Coverage Gap, i.e., more women than men have no access to pension income (Bettio et al., 2013). However, there is no consistent approach in the literature on how to consider Gender Coverage Gaps when analyzing the GPG. While some studies include individuals without pension receipt with a pension income of 0 (e.g., Dotti Sani & Luppi, 2021; Fasang et al., 2013; Möhring, 2015, 2018), others just focus on individuals who receive pension income and thus exclude those without any pension claims (e.g., Bettio et al., 2013; Bonnet & Rapoport, 2020; Evandrou et al., 2009; König et al., 2019). Excluding individuals without own pension receipt can lead to a large gendered sample selectivity depending on the country (see Chapter 2.2) and

consequently affects the level of the Gender Pension Gap in many countries (Amarante et al., 2017; Hammerschmid & Rowold, 2019b). Consequently, whether individuals have access to any pension entitlements is an important dimension of the overall pension inequality. Excluding individuals without own pension claims particularly excludes individuals with work life courses that deviate from continuous employment and these are most frequently women with many years engaging in unpaid care work (empirically demonstrated in Chapter 2.2). However, literature that excludes individuals with no pension income while still referring to Gender Pension Gaps likely underestimates the full level of gendered pension income inequality. Such studies should be more transparent about the fact that they analyze pension inequality among *pensioners* but not among the elderly and likely exclude groups of individuals, especially women (see Chapter 2.2). Albeit the relevance of the conceptual decision how to deal with gender difference in pension receipt for the level of gendered pension income inequality, some studies do not address the issue at all (see Table 2.1).

Few studies consider the Gender Coverage Gap, but there are different approaches to how to do so. The first one is to include individuals with no own pension claims with a pension income of zero. This is a popular approach in the literature (Dekkers et al., 2022; Dotti Sani & Luppi, 2021; Fasang et al., 2013; Möhring, 2015, 2018), but risks producing zero-inflated pension distributions depending on the number of individuals with no pensions. Including individuals with no pension income in the calculation of the GPG is sometimes referred to as the ‘Elderly pension gap’ (Amarante et al., 2017; Bettio et al., 2013). The benefit is, that this indicator considers the gender coverage and the pension income gap among pensioners simultaneously (Bettio et al., 2013) and consequently provides an indicator for full gender pension inequality. Another approach is to apply a two-part model. Herefore, the outcome is split into two variables, i) a dummy for the coverage (receipt of

any pension income) and ii) a continuous variable for pension income including only nonzero values (i.e., only these individuals with pension receipt). Subsequently, separate models for both outcomes are estimated with the latter one either unconditional (Boulton & Williford, 2018) or conditional on the presence of a nonzero outcome (Belotti et al., 2015). Six studies on the Gender Pension Gap applied this approach and first decomposed i) gendered access to pensions and then ii) the gender gap in pensions among those receiving pensions (Bardasi & Jenkins, 2010; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Gardiner et al., 2016; Johnson et al., 1999; Zhao & Zhao, 2018). For the latter, the gender gap in pension benefits among pension recipients, some also used Heckman-corrected decompositions of the GPG to correct for gender selectivity in pension receipt (Bardasi & Jenkins, 2010; Ezeyi & Vujic, 2017; Gardiner et al., 2016; Johnson et al., 1999).¹⁰ However, satisfying the exclusion criteria, i.e., finding covariates that affect pension receipt but not the pension income amount, is particularly difficult for pensions and across genders. Most covariates available will impact both, pension receipt and pension income, and will be correlated with life course patterns.

Implications for this dissertation

Given that access to pensions is determined by pension institutions and often in interrelation with life course inequalities, it is in the field of interest of this dissertation that aims to uncover associations between gendered life course and pension income inequalities. Consequently, in Chapters 3-5 I consider the gender gap in coverage and the pension benefits simultaneously by including individuals without any pension income with a pension income of zero. I argue that only considering differences in access to pensions additionally to pension income inequality of pension recipients considers the complete

¹⁰ Belotti et al. (2015) differentiate between two-part models and models applying the Heckman selection approach.

inequality in pension income. This is particularly important when taking a life course perspective since access to pensions is life-course-sensitive – just like the level of pension entitlements. Thereby, this dissertation addresses the overall pension income inequality and not the inequality among pension recipients only.

The consequences of these decisions are that the Gender Pension Gaps are larger compared to studies that exclude individuals without any pension receipt. Moreover, as Chapter 2.2 shows, the exclusion would disproportionately affect women with many years engaged in unpaid care work and thus risks concealing the role that unpaid care work plays for gendered inequalities in pensions.

However, in Chapter 4 I include extensive additional analyses to assess the sensitivity of the results concerning the treatment of zero pension income and how this varies across public and private pensions. For Chapter 5, I conduct sensitivity analysis for the first step (Life Course Feature Selection) with regard to excluding individuals without pension receipt.

2.2 Sample definition: Housewives never retire!? Gender biases in popular sample definitions for studies on the elderly

Abstract

Even though research has emphasized the risk of a gendered sample selection bias decades ago, the empirical literature on old-age inequalities up to date seems to be largely unaware of it. This Chapter raises awareness of this issue by revealing the gendered sample selectivity for the population aged 65 or older for two popular sample criteria, the self-reported retirement status and pension receipt, for the countries covered in the Survey of Health, Ageing, and Retirement in Europe (SHARE).

The results reveal that up to more than half of elderly women are excluded when using these sample measures. The gender selection bias varies largely across countries and is less pronounced in post-socialist or liberal welfare states. Visualizing the work trajectories by sample status indicates that particularly women with long trajectories of unpaid care work and men with large shares of self-employment, unemployment, and longer educational trajectories do not declare themselves retired or do not receive pensions.

Studies using such sample criteria to capture the population of the elderly likely underestimate the level of inequalities. The selectivity arising from these two sample criteria can be particularly severe in large country comparison studies for which the sample bias more easily remains undetected. Such studies across many countries become more popular due to the availability of harmonized surveys such as SHARE or the Health and Retirement Study. While this Chapter cannot provide a universal recommendation for a sample definition it aims to facilitate less biased sample definitions in studies on the old-age population.

Key words: old age inequalities, retirement, sample bias, gender bias, life course

2.2.1 Introduction

Studies on old-age inequalities become more and more pressing given that more individuals will be affected by these inequalities due to population aging in many countries. Important outcomes under study are pension income, health, or well-being. At the same time, large survey data harmonizing projects focussing on the old age population such as the Survey of Health, Ageing, and Retirement in Europe (SHARE) and the gateway to global aging data covering SHARE and other Health and Retirement Study (HRS) family surveys emerged. They facilitate country comparisons focussing on the elderly and drawing policy implications therefrom. However, defining and measuring the elderly or old-age population for such studies, especially across countries, is not always straightforward given that many different measurements can be applied. For example, Denton et al. (2009) discuss nine different measurements for retirement alone and various combinations thereof are applied in empirical studies. The “convoluted” (Fasang, 2012, S. 686) and “fuzzy concept[s]” (Fasang, 2009, S. 277) of retirement or related constructs particularly unfold their implications when applied in the empirical literature, for example for the sample definition. Which sample definition is most meaningful for the aims of the study likely differs across the country and the life course context of the studied population. Thus, particularly for country comparisons it can become challenging to find a sample definition that proves meaningful across contexts and subgroups.

Luckily, the harmonized surveys contain some items that can be applied across countries easily and facilitate an accessible starting point for defining the elderly. A popular sample definition is the self-reported retirement status which is applied as sample criteria for analysing a range of outcomes such as well-being, social security wealth, and pension income (e.g., Belloni et al., 2020; Bridgen et al., 2022; Madero-Cabib & Fasang, 2016; Nutz & Lersch, 2021; Palomäki, 2017; Ponomarenko, 2016; Tambellini, 2023). Another

criterion used to adjust the sample is the receipt of pension income (e.g., Dingemans & Henkens, 2019; König et al., 2019).¹¹ However, both measures risk not covering the old-age population broadly but rather excluding a selective proportion of the elderly, and as I show, this applies over proportionally to women. Self-reported retirement is likely to be gendered because the broad notion of retirement is more applicable to male-typical biographies (Allmendinger et al., 1992; Fasang, 2009; O’Rand, 1996). Sample selection due to pension receipt depends on which types of pension income are considered and how pension systems regulate access, which often depends on previous contributions based on the work biography. Pension receipt and self-identification with retirement thus both likely depend on previous life courses. And these are often gendered, particularly in conservative and liberal welfare contexts (e.g., Madero-Cabib & Fasang, 2016; Rowold et al., 2023; Uccheddu et al., 2022). As a result, the meaning, and implications of these two sample definitions likely differ across countries and gender and thus lead to country-specific, gendered sample biases.

To descriptively reveal the sample bias for the elderly I apply the two popular sample criteria to the countries covered in SHARE. The sample consists of all respondents aged 65+ to demonstrate the implications of the sample criteria based on a broad sample of the elderly. To maximize the number of observations, I pool waves 2-6 (see Appendix A2.1 for details). For both sample definitions, I first show the shares of men and women separately by country that would be included and excluded. Second, I visualize the work life-courses by sample status to reveal the selectivity arising from applying the sample criteria.

¹¹ This is a highly selective and by far not fully exhausted list of studies applying these two measurements. These studies are transparent about their sample selection criteria. Many other studies in the field do not describe their sample selection in such detail which is highly problematic.

The results confirm that both definitions can lead to dramatic gendered sample biases which differ largely across countries. In some countries, 30-50% of women aged 65 or older are erased from the sample when applying these criteria. Studies in the field seem not to be utterly aware of these gender selection biases pointed out decades ago (Allmendinger et al. 1992) and more generally do often not discuss their sample definitions. Thus, sample biases are likely to remain unnoticed, particularly in large-scale country comparisons. Whereas the definition of the sample is always subject to the individual research question, a more careful investigation of sample issues is essential when using self-reported retirement status or pension receipt.

2.2.2 Gendered life courses and self-reported retirement

Pension systems operate based on the notion of idealized male-typical biographies and retirement pathways characterized by continuous full-time employment as ‘prototype life courses’ linking public pension receipt to stable paid employment (Allmendinger et al., 1993; Fasang, 2010; Grady, 2015; Kuitto et al., 2021; Leitner, 2001). The resulting common understanding of retirement “as the withdrawal of older workers from paid working life” (Denton & Spencer, 2009, S. 65) does not capture the reality of the whole population in old age, such as individuals with weak or interrupted labor market experiences, and thus particularly women (Allmendinger et al., 1992; Fasang, 2009; O’Rand, 1996). For these women, there is less of a strict age range or transition of states that determine retirement (Allmendinger et al. 1992). Indeed, current studies show highly gendered retirement pathways across different country contexts, where female pathways to retirement are often characterized by part-time or out-of-labor-force periods (e.g., Fasang, 2010; Madero-Cabib et al., 2021; Riekhoff & Järnefelt, 2017). These gendered retirement pathways demonstrate that the meaning and experience of retirement strongly differ across genders (Fasang, 2010). Ambiguity about self-identification with retirement is likely

greater the more individual experiences diverge from the institutionalized transition from employment to pensions (Fasang, 2009).

The assignment of retirement to typical male life courses terminating continuous employment parallel to replacing earned income with pension income leads to a “Steady Worker Bias in Retirement Research” (O’Rand, 1996, S. 235). The bias is not only applicable to data collection as highlighted by O’Rand but also risks creating a gendered sample definition (Allmendinger et al. 1992). Consequently, men are more likely to report themselves being retired in surveys, while selective groups of individuals that deviate from ideal-typical male pathways to retirement, above all women with more interrupted careers, are less likely to identify with the common notion of retirement.

The reporting and “sample selection bias” (Allmendinger et al. 1992: 92) can be further reinforced by survey design which forces respondents to choose between being ‘retired’ and other states, such as ‘homemaker’ when reporting their employment status (e.g., in HRS and SHARE, see A2.2). Such design is particularly likely to exacerbate the gender bias given gendered work trajectories and the gendered reality of retirement views qualitative research has put forward: “Housewives never retire!” (Loretto & Vickerstaff 2013: 77). Women who engage in unpaid care work, for which there exists no discrete institutional pathway for retirement equivalent to that of paid work (Fasang, 2010), are likely to describe themselves as homemakers and thus the structure of the question likely levels down the share of women self-reporting themselves as retired. Thus, using self-reported retirement as sample criteria (and more generally as a covariate in any model) risks excluding large shares of women and producing gendered selectivity issues.

However, the magnitude of such bias likely depends on how gendered normative and welfare contexts are in which life courses and pathways to retirement are shaped given that

the level of gender segregation in work(-family) life courses differs across countries (Komp-Leukkunen, 2019; Rowold, 2022; Uccheddu et al., 2022). It follows that the closer women's work life courses to the ideal type of continuous full-time employment, the more likely they are to experience a smooth transition from employment to retirement and declare themselves as retired. Additionally, social desirability likely affects subjective retirement classifications (Fasang, 2009). Social desirability and the overall notion of retirement might vary across countries as well.

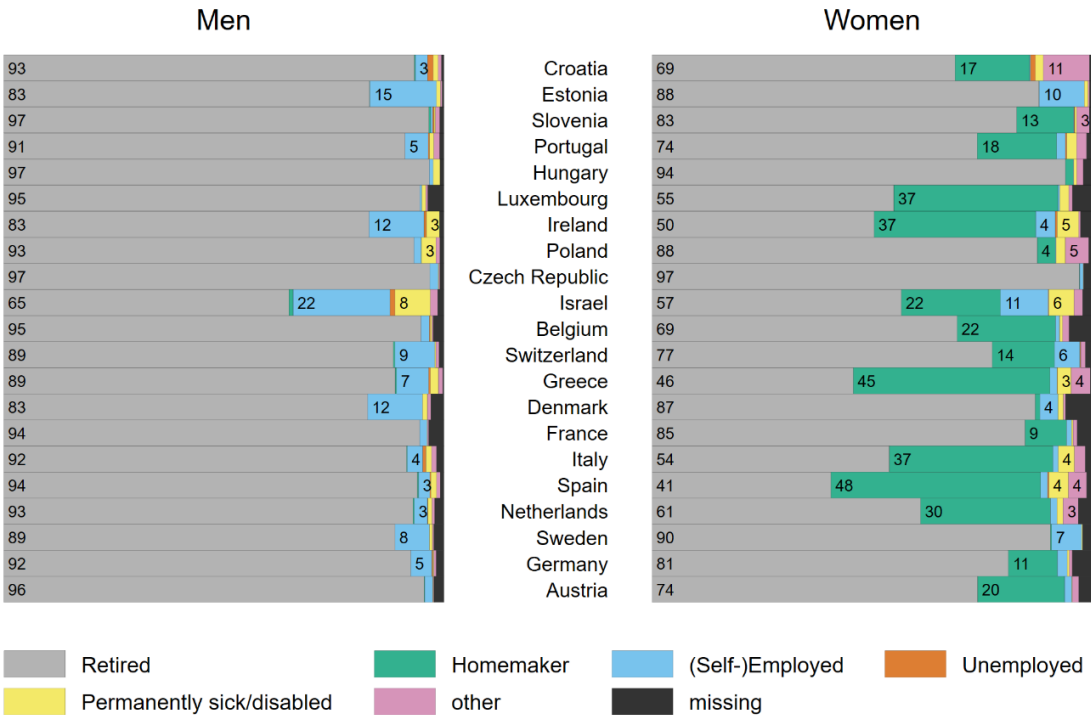
To the best of my knowledge, this issue has not been documented empirically previously. Closest comes a study on the recall bias in the retirement age which compares the self-reported retirement age in SHARE with the retirement age captured based on the administrative records of the German Public Pension Insurance (Korbmacher, 2014). A much higher share of 'retired' women (as captured in the administrative records) compared to men did not self-identify as retirees (83.2% vs. 92.6%): 8% of them report in the SHARE survey being homemakers instead. The implication of this gendered reporting bias is not discussed further. Since the retirement age in SHARE is only surveyed for respondents who declare themselves as retirees, those who do not report retirement were excluded from the rest of the analysis. As a result, the recall bias of retirement age did not differ across genders for the remaining sample.

Some studies on gender inequalities in pension use the criteria 'retiree' as one of several sample criteria, but only a few studies state clearly that retirement for the sample selection is measured as self-assessment (Ezeyi & Vujic, 2017; Madero-Cabib & Fasang, 2016; Veremchuk, 2020, see Table 2.1 for an overview).

The implications of the gendered retirement notion for the sample of the elderly population are visible in Figure 2.1 which depicts the self-assessed employment status of individuals

aged 65 or older by gender and country. Indeed, in many countries, large gender differences appear in the share of self-identification as a retiree with lower shares of women reporting retirement (Figure 2.1). For example, using self-defined retirement as a measurement excludes more than 50% of women compared to only 6-11% of men in Spain and Greece.

Figure 2.1 Self-assessed employment status at age 65 or older by country and gender



Notes: Shares higher than 2.5% are displayed as relative frequencies. Own calculations based on SHARE sample including all respondents aged 65 or older, waves 2-6, release 7.1.0. Not weighted.

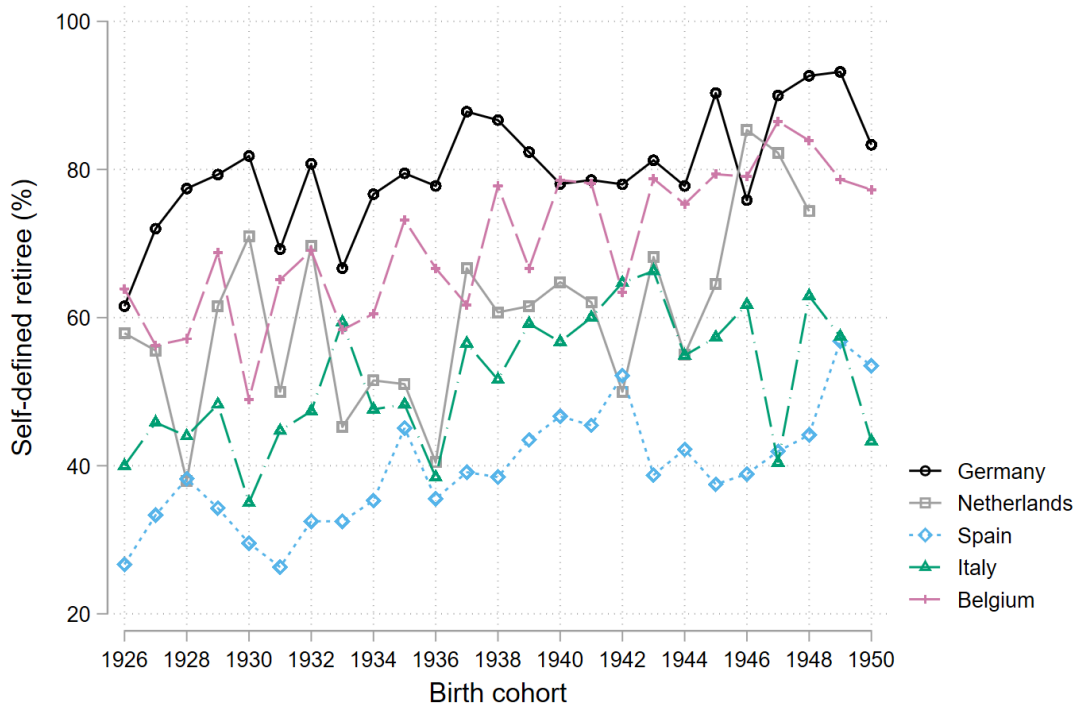
However, as expected, the level of the gender gap varies strongly across countries. It stands out that the gender bias in self-assessed retirement is largely in line with the level of gendered work trajectories across countries which broadly follows the overall welfare states regime typology (e.g., Komp-Leukkunen, 2019). Except for France and Germany¹², large shares of women aged 65 or older, ranging from 14.2% in Switzerland to 47.6% in Spain, would rather identify with homemaking than with retirement or any other

¹² In Germany, the relatively higher share of women identifying with retirement is likely driven by women from East Germany, whose life courses were shaped in a socialist context and are much more dominated by labour force compared to West Germany (Rowold, 2022).

employment state in liberal, conservative, and Southern European welfare states. At the same time, a large majority of men (in most cases close to or more than 90%) in these countries report themselves as being retired. On the other hand, I find much less gendered self-identification with retirement in post-socialist and social-democratic welfare states except for Slovenia and Croatia.

The self-defined retirement might become less gendered for younger cohorts given that the increasing female labor market participation likely converges women's employment biographies with men's and thus with the common understanding of retirement.

Figure 2.2 Cohort trend of self-assessed female retirees for selected countries.



Notes: Own calculations based on SHARE sample including all respondents aged 65 or older, waves 2-6, release 7.1.0. Not weighted.

Indeed, the shares of women who identify as retirees slightly increase across cohorts for a selected set of countries (Figure 2.2). In Germany and the Netherlands for example, the share increases from about 60% for the 1926 birth cohort to up to 80-90% for the youngest birth cohorts (1946-1950). In Italy, it increases from about 30% up to 55% roughly, in Spain from 25% up to almost 60%, and in Belgium from. This suggests that the sample

bias can also lead to cohort selectivity among women. Even though the magnitude of the gendered sample bias tends to decline over cohorts, it is still relevant for the youngest cohorts observed.

2.2.3 Gendered life courses and pension receipt

Similarly to the identification with retirement and given that pension policies reward life courses (Ginn et al., 2001; Madero-Cabib & Fasang, 2016; Möhring, 2021), pension receipt can be determined by life courses. However, this depends on the pension system. Beveridgean types of pension systems such as in the Netherlands for example guarantee relatively broad access to public pensions through flat-rate pensions, whereas high minimum contribution years such as 20 years in Italy constrain access particularly for individuals with interrupted careers (Leitner, 2001; Rowold et al., 2023). This applies disproportionately to women, depending on the country context, and can lead to large differences between men and women concerning access to pension income (Gender Coverage Gap, e.g., Tinios et al., 2015). Furthermore, the definition and operationalization of pension incomes considered matters, e.g., which pillars (public, occupational, and personal pensions) and specific pension types (e.g., survivor pension) are considered. For example, excluding survivor pensions risks excluding women particularly, as most women are more likely to outlive their partners. The narrower the definition of pension income, the more likely to exclude specific groups of individuals.

Applying pension receipt as one of the pension criteria is relatively common in the literature on gendered pension inequalities (e.g., Bettio et al., 2013; Bonnet & Rapoport, 2020; Evandrou et al., 2009; König et al., 2019, see Table 2.1).

For the empirical investigation, I apply pension receipt based on total imputed pension income, that is the sum of all pension income from public and occupational pensions (including survivor pensions but excluding social assistance) generated by SHARE, to use

an inclusive definition of pension receipt. As a result, the sample bias shown in Figure 2.3 is a conservative approximation.

Using pension receipt as sample criteria for the population aged 65 or older produces highly gendered sample biases in about half of the countries but often to a lower degree compared to using the self-reported retirement status. In many Southern European countries, however, the share of women that would be excluded remains relatively high (e.g., 42% in Spain and 31% in Greece). While in some countries a lower share of men receive pension income (Figure 2.3) than those who self-identify as retirees (Figure 2.1; e.g., Portugal, Spain, Italy, Greece, Croatia), a higher share of women receive pension income than those who report themselves being retired in all countries (particularly large differences in Netherlands, Spain, Israel, Ireland, Italy, Croatia, and Luxembourg).

Figure 2.3 Pension income receipt at age 65 or older by country and gender



Notes: The gender difference documents the Gender Coverage Gap in pensions. Own calculations based on SHARE sample including all respondents aged 65 or older, waves 2-6, release 7.1.0. Not weighted. SHARE provides five imputation alternatives. I display the results using imputation 1, but the results remain robust when using the other imputation outcomes. In the imputed pension income, no personal pensions are included.

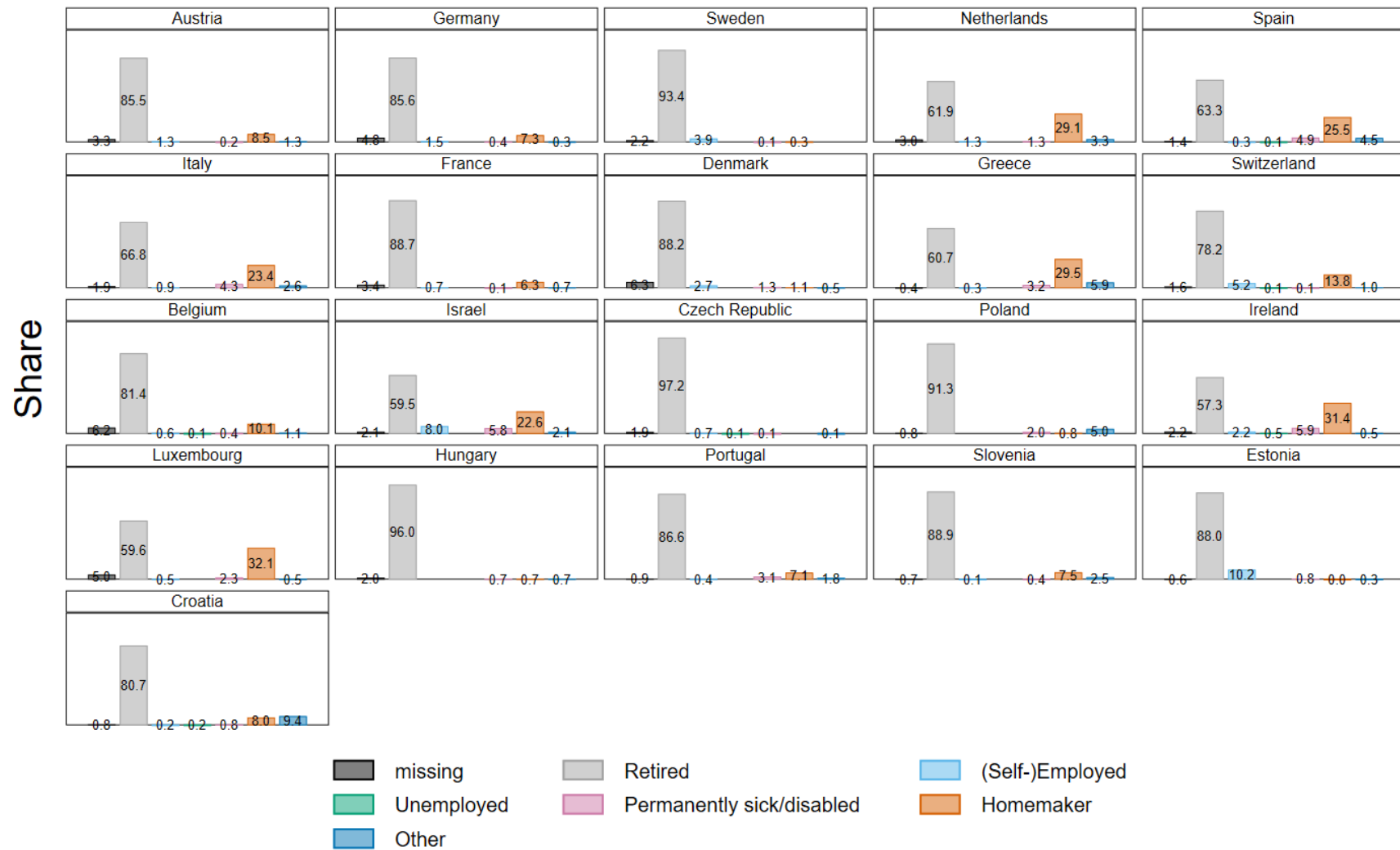
This sheds light on the complexity of retirement identification: even among women who receive pension income a substantial share (up to 32%) identify as housewives rather than as retirees (Figure 2.4). This is particularly the case in countries that are known for traditional gendered division of labor. At the same time, the share of women self-identifying as retirees is larger among pension recipients compared to the overall population of elderly, for example in Spain, Italy, Greece, and Belgium. As a result, using self-reported retirement as sample criteria excludes much more women compared to using pension receipt.

2.2.4 Work trajectories of individuals not identifying as retirees or without pension receipt

To explore the selectivity into self-reported retirement and no pension receipt with regard to work life-courses, Figure 2.5, Figure 2.6, and Figure 2.7 visualize the work trajectories from age 18 to age 65 by sample status respectively, and pooled over countries.¹³ The information on the work biographies is derived from the generated Job Episodes Panels (Brugiavini et al., 2019) based on the SHARELIFE data (waves 3 and 7 of SHARE). Roughly 40% of the sample have no retrospective data and must be excluded. This affects the distribution of and the gender bias in both sample criteria only very marginally (see A2.3). The Appendix describes the retrospective data sample, the handling of missing retrospective data, and the operationalization of the work categories.

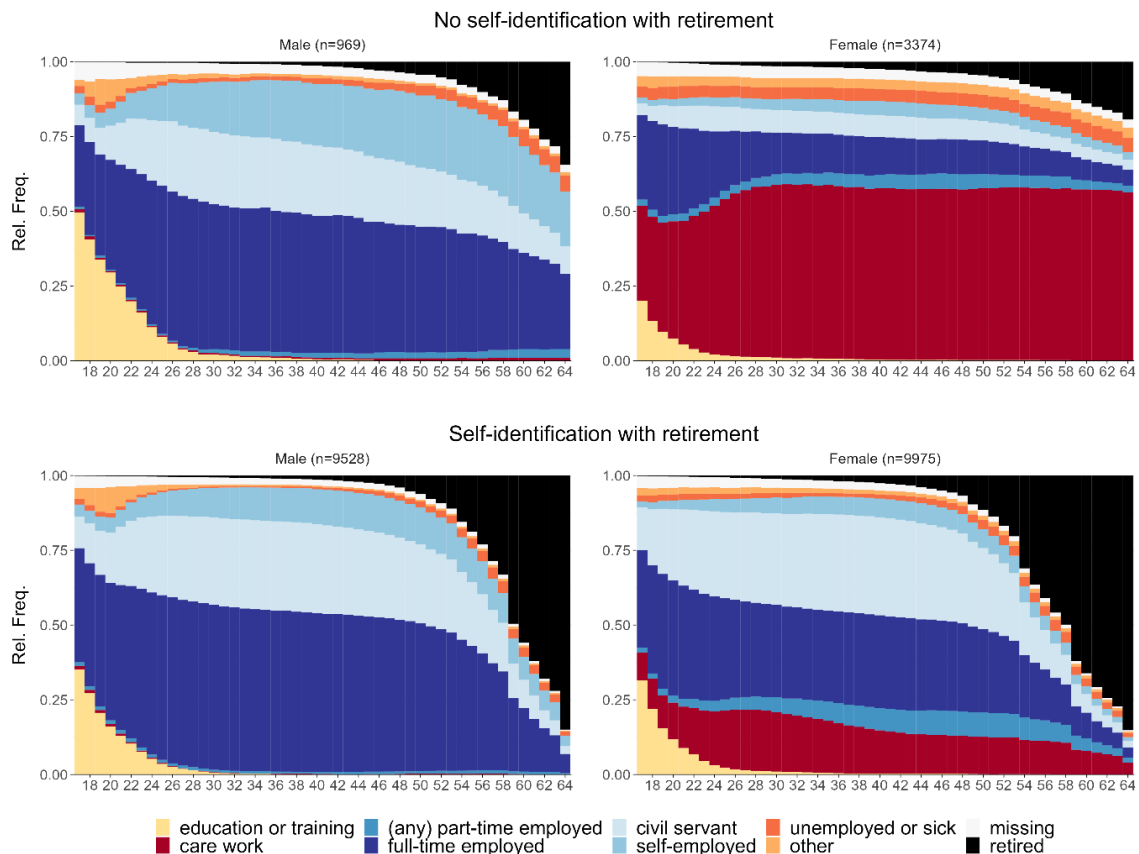
¹³ Pooling over countries conceals potential country differences. Moreover, countries with high shares of non-identification with retirement or no pension receipt drive the findings for these samples respectively. Still, the Figures give an indication of how selective the life courses of those who do not meet the sample criteria are and who would be excluded from the analysis.

Figure 2.4 Self-defined employment status of female pension recipients.



Notes. Own calculations based on SHARE sample including all respondents aged 65 or older with pension receipt, waves 2-6, release 7.1.0. Not weighted.

Figure 2.5 Distribution of work states from age 18-65 by self-definition of retirement

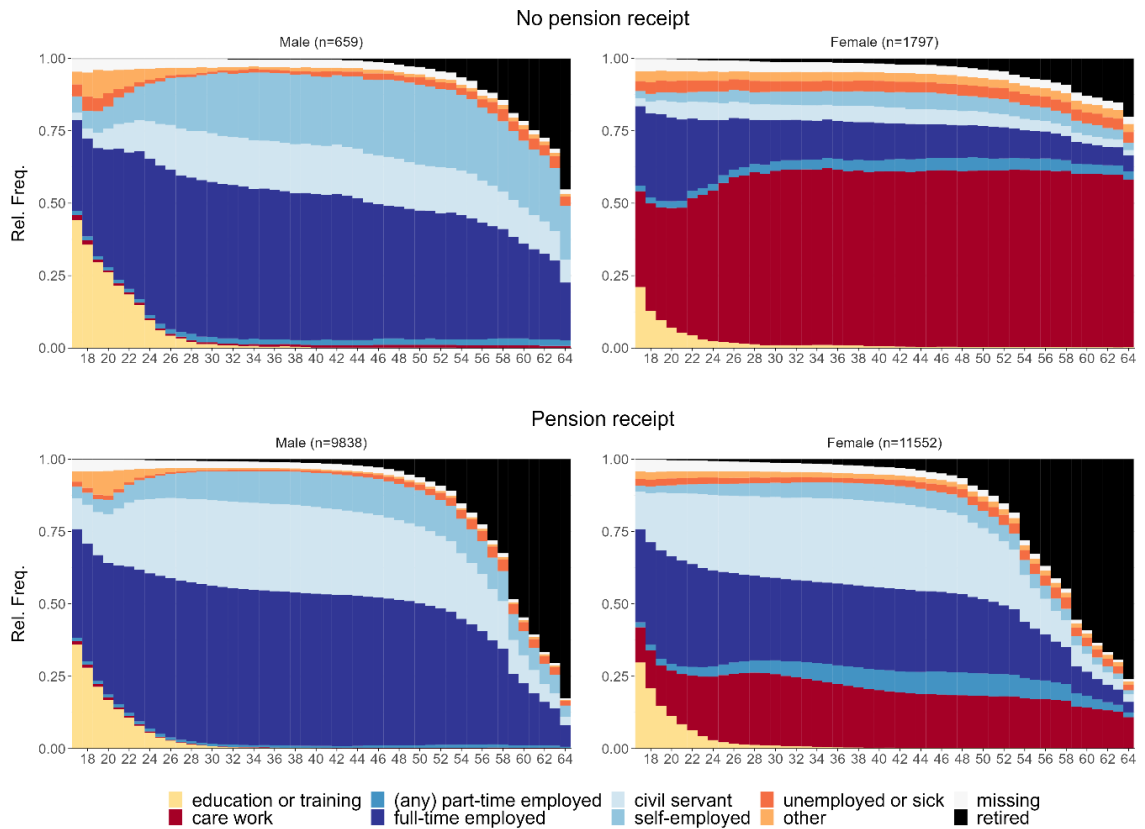


Notes: 'No self-identification with retirement' covers all categories of the self-assessed employment state at age 65 or older other than retired. Own calculations based on SHARE sample including all respondents aged 65 or older with retrospective data, waves 2-6, release 7.1.0. Not weighted.

Figure 2.5 and Figure 2.6 reveal for each age the distribution on the different work categories by gender and sample status (i.e., self-identification with retirement and pension receipt). The patterns are similar for both sample definitions. For men who do not identify as retirees (Figure 2.5) or receive pensions (Figure 2.6) a much higher share is self-employed from age 22 onwards compared to those who self-identify as retirees or receive pensions. Furthermore, a larger share of these men were in education up until their 30s and worked in part-time employment (particularly at older ages). For women, who would be excluded based on both sample criteria, the share of unpaid care work exceeds the one of those who do identify as retirees or receive pension income by at least three times for all ages. At the same time, the shares of any type of employment, particularly civil service and full-time employment in the private sector, are much lower, especially from age 20 on.

For both men and women who do not define themselves as retirees or do not receive pension income, larger shares experience unemployment or sickness, as well as ‘other’ work categories. Additionally, and straightforwardly, the shares of retirement are much lower.

Figure 2.6 Distribution of work states from age 18-65 by pension receipt

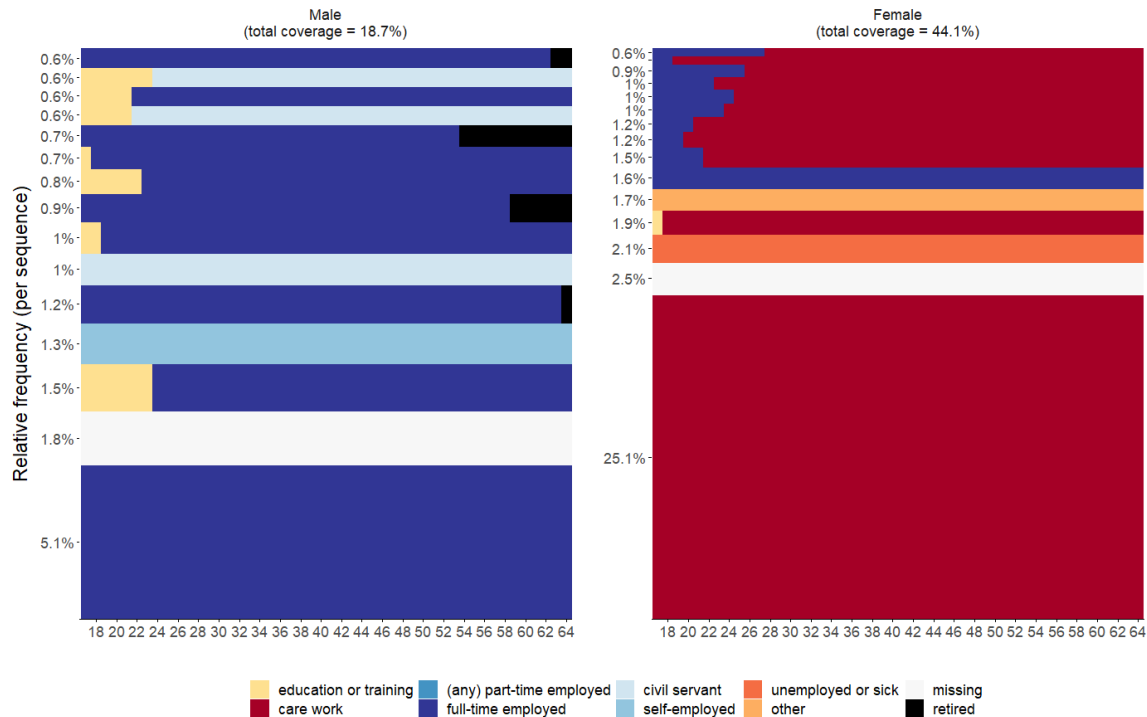


Notes: Own calculations based on SHARE sample including all respondents aged 65 or older with retrospective data, waves 2-6, release 7.1.0. Not weighted.

Figure 2.7 displays the fifteen most frequent work trajectories among men and women who do not self-identify as retirees and thereby reveals how the ordering and timing of working states most often unfold over the life course. One-quarter of women who do not report being retired continuously engaged in unpaid care work from age 18-65. Another frequent work life-course pattern is characterized by women transiting from full-time employment to continuous unpaid care work in their early twenties, which might be linked to family formation as suggested for Italy and West Germany in Chapters 3 and 4. For men, the work trajectories are more heterogeneous because the fifteen most frequent trajectories only cover 18.7% of those who do not self-identify as retirees. They are mostly characterized by

either continuous full-time employment or continuous full-time employment following educational spells until the early 20s.

Figure 2.7 Most frequent work trajectories of individuals not self-defining as retirees



Notes: Sequence Frequency Plot (N. S. Müller et al., 2008) displaying the 15 most common work trajectories for men and women separately. Own calculations based on SHARE sample including all respondents aged 65 or older with retrospective data, waves 2-6, release 7.1.0. Not weighted.

2.2.5 Discussion

This research note reveals descriptively that large, gendered sample biases can appear when relying on self-reported retirement status or pension income receipt as sample criteria for the elderly as was suggested decades ago (Allmendinger et al., 1992). However, in light of the increasing accessibility of harmonized multi-country data focussing on the old-age population, using such easily applicable sample definitions is popular. Especially for country comparisons, this is problematic given that the gender bias differs by country context. The steady worker bias (O’Rand, 1996) leads to gendered sample biases: relying on the self-reported retirement status assumes that the notion and identification with retirement, understood as the withdrawal from (continuously) paid work, is neutrally applicable to the whole old-age population. Relying on pension receipt assumes equal or

universal access to pension, though in most pension systems access is constrained by the number of contribution years, i.e., most often the employment years. This undermines life course experiences deviating from the norm of stable (full-time) employment and risks excluding systematic groups of individuals.

The results confirm that up to 59 percent of women aged 65 or older are excluded when applying the self-assessed retirement as sample criteria. Particularly in conservative, liberal, and Southern European welfare states, where previous research revealed strongly gendered work(-family) life courses (Komp-Leukkunen, 2019; Rowold, 2022; Uccheddu et al., 2022), the exclusion of high shares of women is likely. Using pension receipt as sample cutoff criteria produces a similar gendered sample bias but the level is lower (exclusion of up to 42 percent of women). Furthermore, the interaction between the two sample definitions highlights the complexity and gendered notion of retirement identification. While even among female pension recipients up to 40% do not classify themselves as retirees, particularly in countries with a traditional gendered division of labor, a higher share of men self-identifies as retirees than those who receive pension income. This highlights that retirement means “very different things for women and for men” (Loretto and Vickerstaff, 2013, p. 65) and suggests that the identification with retirement is more strongly shaped through the overall distance to the ideal-typical retirement transition from continuous paid work than through pension receipt.

Visualizing the employment trajectories by the sample status confirms that women and men with employment trajectories deviating from the ideal-typical continuous employment biography are excluded if any of both sample criteria are applied. Women not identifying as retired or not receiving pension income spent many more years doing unpaid care work at all ages from 18-65. Most of their work trajectories are characterized by continuous unpaid care work which descriptively confirms that “[h]ousewives never retire” (Loretto &

Vickerstaff 2013: 77). For men, the findings are more mixed. On the one hand, the shares of less privileged life course states such as unemployment or being sick are more common among those not captured as retirees or not receiving pension income. On the other hand, I observed longer educational trajectories in this sample of men.

In line with the life course concept of cumulative advantage or disadvantage (e.g., Dannefer, 2003), such differences in the work trajectories across samples are likely reproduced in old age (O’Rand, 1996). When analyzing inequalities in outcomes of the elderly, excluding individuals with less privileged trajectories (e.g., higher shares of unemployment and care work) or more privileged trajectories (e.g., higher educated men) is likely to make a difference. Excluding individuals with life course experiences deviating from continuous employment trajectories, research using such sample cutoffs is likely to conceal the extent of inequalities (Allmendinger et al., 1992; Fasang, 2009). For example, using pension receipt compared to an age cutoff as sample criteria dramatically changes the level of the Gender Pension Gap in many countries (Hammerschmid & Rowold, 2019b) and partly alters the association between financial hardship along the life course and pension income (Riekhoff et al., 2023). There is an overall need to systematically analyze how selective old age sample definitions skew different outcomes to get a more profound understanding of the implications of the gendered sample selectivity. Furthermore, whereas this analysis focuses on gender, other marginalized groups such as non-natives or individuals with low SES might be affected by sample selectivity as well. Systematically revealing how these characteristics differ by sample criteria would be relevant as well.

The results more generally speak to the need for an “appropriate conceptualization of retirement for the research question at hand” (Fasang, 2009, S. 290). The same bias applies to the use of self-reported retirement and pension receipt in empirical models more generally as well as measures that are based on it. For example, the retirement age or

reasons are only surveyed if respondents self-identify as retirees in SHARE and therefore automatically replicate the gendered sample bias.

Researchers should be aware of it and its implications. For example, the literature on retirement transitions or age uses the self-reported retirement status straightforwardly as a measure of being retired (e.g., Radl, 2013; Tambellini et al., 2023). While some studies are overall transparent about their “very specific sample” (Ponomarenko et al., 2019, S. 261) or discuss that individuals who identify as homemakers instead of as retirees are excluded (e.g., Sohler et al., 2021; Tambellini, 2023), they do not indicate or discuss the (gendered) implications of the sample selection for the results. In line with the results presented in this study, Fasang (2009) shows that the determinants for retirement timing differ more strongly for women than for men depending on the measurement of retirement timing (self-reported retirement status or first pension receipt). A more inclusive retirement definition might be to additionally consider the first pension receipt or first experiences of economic inactivity after a certain age as proxies for retirement (Komp-Leukkunen, 2021) rather than solely relying on self-reported retirement status. More generally, using an age-based metric for the elderly population attenuates the sample selection bias but might require follow-up adjustments depending on the research questions such as excluding individuals still in the workforce.

This Chapter raises awareness for gender selectivity deriving from popular sample criteria in studies on the elderly and calls for more caution and transparency regarding sample definitions and their consequences. The documentation of the gender bias in sample criteria across countries aims to give researchers an indication of whether the bias applies to their research and to facilitate better-informed decisions for the sample definition when studying inequalities in old age. Otherwise, research risks systematically excluding selective subgroups of the elderly and therefore skewing the results.

Acknowledgement

The author is grateful to Alicia Garcia Sierra and Christiaan Monden for their helpful comments. She further wants to thank Hanna van Solingen and Kène Henkens for motivating this research note and Maik Hamjediers for his advice for the visualization.

Implications for this dissertation

As a result, Chapters 3-5 use a broad sample definition which is popular in old-age studies as well: a simple age cutoff (e.g., Dotti Sani & Luppi, 2021; Fasang et al., 2013; Kreyenfeld et al., 2022; Möhring, 2015, 2018; Sefton, Evandrou, Falkingham, et al., 2011; Zhao & Zhao, 2018, see Table 2.1 for an overview). Respondents aged 65 or older are included in the baseline sample. However, to make sure I do not include individuals whose main activity is still paid work and who might only yet receive part of their own, independent pension income, I include some additional conditions for the analysis sample. I excluded individuals who were still in the active labor force and received payments from employment or unemployment benefits in the respective year or had any missing information on this information. However, I retain non-employed respondents who report work which could be classified as spare time work to keep those retirees in the sample who require additional income. I further excluded all respondents with any missing information on any variable needed for the calculation of the annual pension income, controls, or the differentiation between East and West Germany.

I provide for each of the Chapters 3-5 tables which document for each step of the sample selection how many individuals were dropped at each step of the sample definition.

2.3 Retrospective life course data

This dissertation relies heavily on retrospective data collected for the work and family life courses. In the SHARE dataset, the annual work and family situation is surveyed retrospectively in waves 3 and 7 (SHARELIFE). SHARELIFE contains valuable and detailed retrospective data about individual work, family, and health situations over the life course. Retrospective data risk suffering from recall biases, i.e., respondents cannot recall what happened at what time particularly if the events are way in the past (e.g., Korbmacher, 2014; Kreyenfeld & Bastin, 2016). In SHARE, different survey techniques are used to minimize the recall bias and ensure an accurate collection of retrospective information. The SHARE team uses graphically displayed Event History Calendars with landmarks and anchoring (e.g., using the birth of the respondent's child or the date J. F. Kennedy was shot as an anchor to trigger the recall of events happening at the same time) and utilizing the life history approach to reconstruct and visualize the whole retrospective life course (Börsch-Supan & Schröder, 2011). Comparing the cross-sectional information from SHARE waves 1 and 2 with the SHARELIFE data, research shows that the "SHARELIFE data is overall strongly consistent with the information reported at the time of occurrence of the events" (Garrouste & Paccagnella, 2011, S. 70). However, linking the SHARELIFE data with precise and reliable administrative records for Germany (SHARE-RV) enables more recent research to reveal that particularly respondents with complex life courses tend to simplify their biographies in the survey data: only 14% of the respondents "do not have a single inconsistent year in the entire time observed" (Bühler et al., 2022, S. 12). The former is particularly problematic from a gender-sensitive perspective because women experience more complex life courses than men. In line with this, unpaid care work, which is mostly reported by women, is the most underreported work category.

In Chapter 4, I make use of the same database linked to German administrative records for the pension-relevant work biographies (SHARE-RV) to minimize this gender-specific bias. However, due to the country comparison in Chapters 3 and 5, I must rely on the retrospective survey data of SHARELIFE as no link to similar administrative records covering the whole life course is available in Italy and the Netherlands.

Processing such retrospective data can be complex. To minimize the researchers' degrees of freedom and make Chapters 3-5 more comparable to other studies in the field, I use a generated dataset for the SHARELIFE data instead of the raw data: the Job Episodes Panel (Brugiavini et al., 2019). However, the Job Episodes Panel contains a large amount of person-year spells with missing information occurring disproportionately for women. This is particularly problematic given that women's more complex life courses are more likely to be underreported in survey data anyhow (Bühler et al., 2022, S. 12).

In this Chapter I document the missingness, discuss the reason for it, and provide a solution on how to fill the missing spells with the raw data from SHARELIFE that was not used to generate the Job Episodes Panel.

2.3.1 Work life course data in SHARE: How to solve gendered missing data in the Job Episodes Panel

Raw retrospective data can be difficult to handle and requires often heavy data management. However, for the SHARELIFE data, Brugiavini et al. (2013, 2019) have and most literature using SHARELIFE for work trajectories relies on their generated dataset, the Job Episodes Panel (JEP). The JEP is valuable for the academic community for research as it helps harmonize research because it decreases the variation in the data used stemming from different data management decisions and thus improves the comparability of research. Thus, it is a popular data source for almost one decade and was used in many studies on several outcomes and across different disciplines (e.g., Antonova et al., 2017; Bravo &

Herce, 2022; Cohn-Schwartz & Naegele, 2023; Ice et al., 2020; Mousteri et al., 2018; Ponomarenko, 2016; Riekhoff & Vaalavuo, 2021; Tambellini, 2023; Uccheddu et al., 2022; van Hedel et al., 2016; Van Winkle & Fasang, 2017).

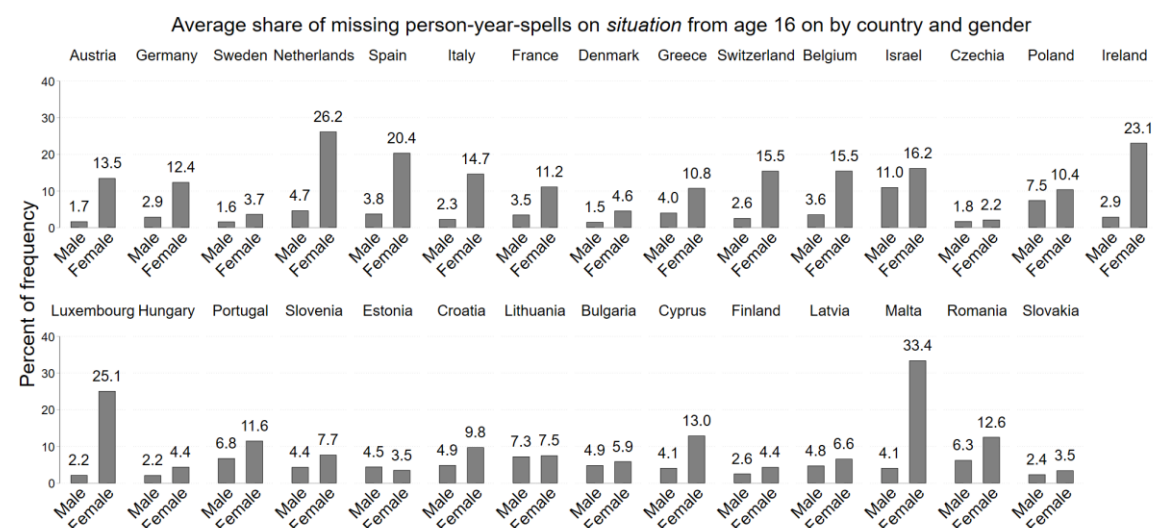
One of the particularly rich variables in the JEP is the *situation* indicator containing the self-reported situation or activity individuals were in for each year, covering multiple employment lives but also categories out of the labor force. This information is key in this thesis as I use it to operationalize the work life courses. However, the *situation* variable contains missing information for sizeable shares of person-year spells. This is problematic because missingness seems to be not random when looking at the two dimensions country and gender.

Using the person-year-spells for ages above 15 of the whole JEP sample, the average share of missing observations of all person-year spells per person on the *situation* variable ranges from 1.5% in Denmark to 7.5% in Poland for men and from 3.5% in Estonia and Slovakia to 33.4% in Malta for women (Figure 2.1). In many countries, there is a considerable gender difference in relative missingness, while the level varies across countries. This is also the case for the countries studied in the following Chapters. The missingness is largest among women in the Netherlands (26.2% compared to 4.7% for men) but also considerable in Italy (14.7% for women, 2.3% for men) and Germany (12.4% for women, 2.9% for men).

Due to this sizeable missingness, individual studies have to make decisions on how to proceed with such missing data (e.g., listwise deletion or imputations) if they are aware of them in the first place. These different approaches are likely to produce biased results since specific countries as well as specific groups of men and particularly women (see below) will have been more affected by large shares of missing information on their working lives.

When focussing on all person-year-spells the gender gaps become narrower (Figure A2.1). However, in most empirical applications, the *situation* variable might be applied for employment biographies which is why I focus on the person-year spells for ages above 15 in the main Figures.

Figure 2.8 Missingness before adjustment



Notes: Own calculations based on the whole sample of JEP release 7.1.0.

The Source of Missingness and a Potential Solution

The Job Episodes Panel (up until release 8.0.0) does not consider variables of the raw data containing information on the situations people were in after their last job (variables re035_*, re039_*, re039a_*). They are then asked whether the situation changed again and to report the second situation they were in after the change etc. until their situation does not change anymore. Most of the spells with missing information can be filled using this information of the raw data.

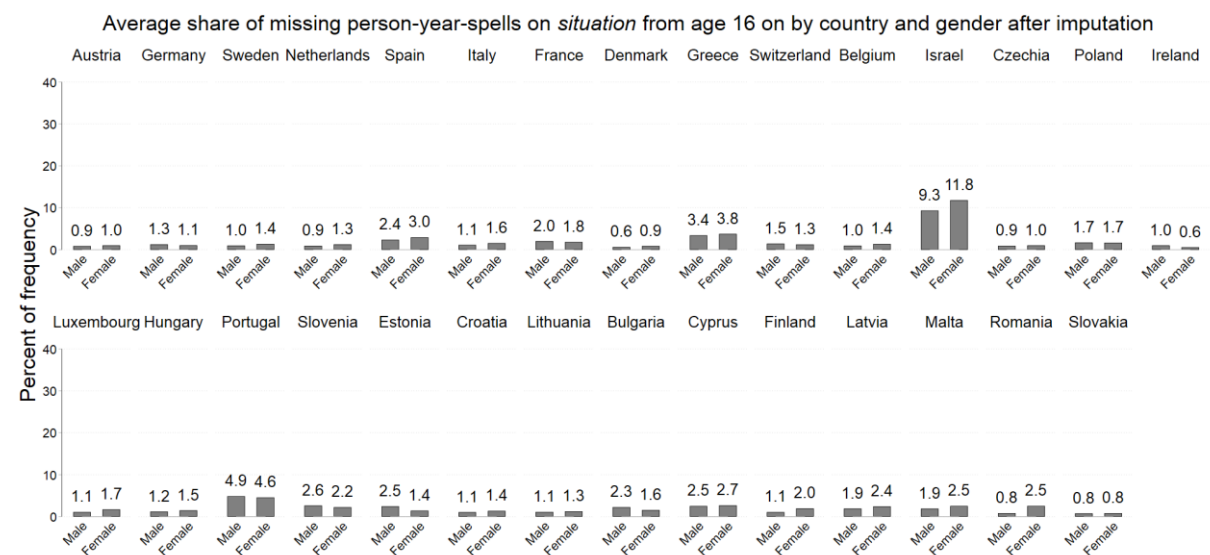
I impute only person-year-spells that are missing on the *situation* variable and that occur in the years after the last reported job (re035). Spells with missing information are filled for each year until the first year with non-missing information on the *situation* variable or until the situation changes again (re039a). If the situation changes, the same procedure is done

for the n^{th} situation after the last job n^{th} times. The year in which the situation after the last job change is filled with the new situation.¹⁴

Reduction in Missingness & country- and gender-specific patterns

Figure 2.9 shows the share of missingness by country and gender after filling in the missing information as described above. The level of relative missingness drops down to on average below 2% of all person-year spells for most countries. The improvement differs across countries and genders, but overall, the gender gaps in missingness are reduced to a minor level.

Figure 2.9 Missingness after adjustment



Notes: Own calculations based on the whole sample of JEP and SHARELIFE release 7.1.0.

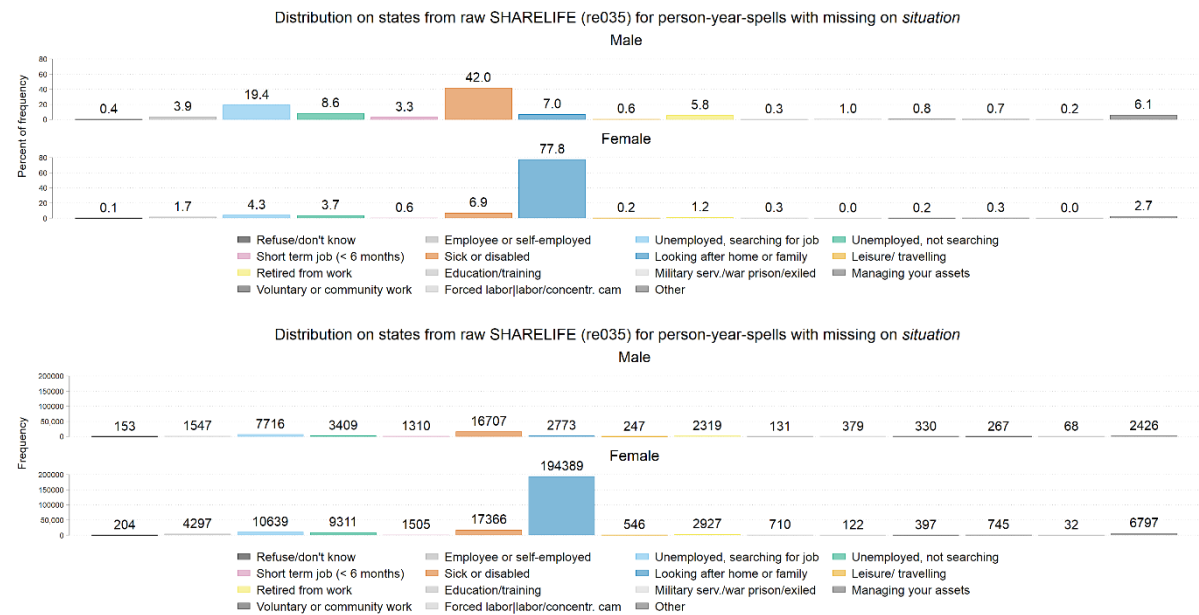
When considering all person-year spells and not only those for ages 16 onwards the share of missing data drops down to an average of 7-9% for most countries (Figure A2.2). This is because most respondents have missing information in their first 4 years of life which cannot be filled by the applied procedure – 81% of the remaining missing information on the *situation* variable occurs during the first 15 years of age. A large share of the remaining

¹⁴ The code for filling the missing information is [available online](#).

missing information is due to gaps between education and the first situation after the end of the education phase.

For the countries studied in Chapters 3-5 the share of missing person-year spells is reduced to 0.9-1.6%. While the gender gap in missingness has fully diminished and even slightly reversed in Germany, it is still observable in Italy and the Netherlands but at a minor level (e.g., 1.6% for women compared to 1.1% for men in Italy).

Figure 2.10 Distribution of filled working states for all countries by gender



Notes: Own calculations based on the whole sample of JEP and SHARELIFE release 7.1.0. Some small categories merged.

Figure 2.10 shows the categories of the *situation* variable that have been filled in relative and absolute terms by gender and pooled over all countries. 77.8% of the missing data (a total of 194,383 person-year spells) for women is due to unpaid care work (*Looking after home and family*) they engaged in after their last job. Thus, not considering the situation respondents were in after their last job produces particularly large shares of missing spells for women engaged in care work. This is most likely due to the traditionally gendered division of labor in most of the observed countries, in which women are more likely to drop out of the labor market and retire earlier due to care duties (e.g. Carr et al., 2016; Zagel &

Van Winkle, 2022). Thus, a large share of unpaid care work is obscured in the original situation variable. For women, small shares of the previously missing data represent being *sick or disabled* as well as unemployment periods – spells that were underestimated in the JEP data so far too.

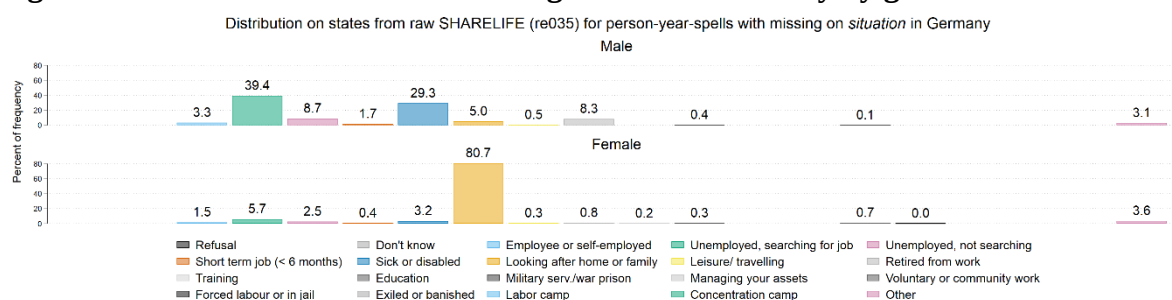
For men on the other hand, a majority of previously missing data is due to being *sick or disabled* (42%), followed by *unemployed and searching for a job* (19.4%) as well as *unemployed but not searching for a job* (8.6%) and *looking after home and family* (7%). Thus, even for men, specific working spells that mostly deviate from employment were concealed in the original situation variable.

However, for both, men and women, the situation that is filled and thus was previously neglected differs across countries (see Rowold, 2023b for the distribution of working states that have been filled for each country by gender). For women, in all countries, the highest share of missing spells was filled with care work. However, the magnitude differs. Whereas in Spain 91% of all missing spells of women are due to *looking after home and family*, in Poland it is 44%. For example, for men in Italy, roughly as many spells of *unemployed and searching for a job* as well as *being sick or disabled* were filled (27% each), whereas in the Netherlands 67% of spells were filled with *being sick or disabled* alone. For men in Slovenia, only 9% of the previously missing social situation spells are due to being sick or disabled, while 35% are due to *unemployed and searching for a job* and 27% are due to *unemployed and not searching for a job*.

In all three countries used in Chapters 3-5, a very similar and large share of missing person-year-spells of women is filled with *looking after home and family*, what is referred to as ‘unpaid care work’ in this dissertation (80.7-83.3%, see bottom panels Figure 2.11, Figure 2.12, Figure 2.13). This is a slightly higher share compared to the average over all countries in the SHARE data (77.8%) and underlines that the empirical studies in Chapters 3-5 would

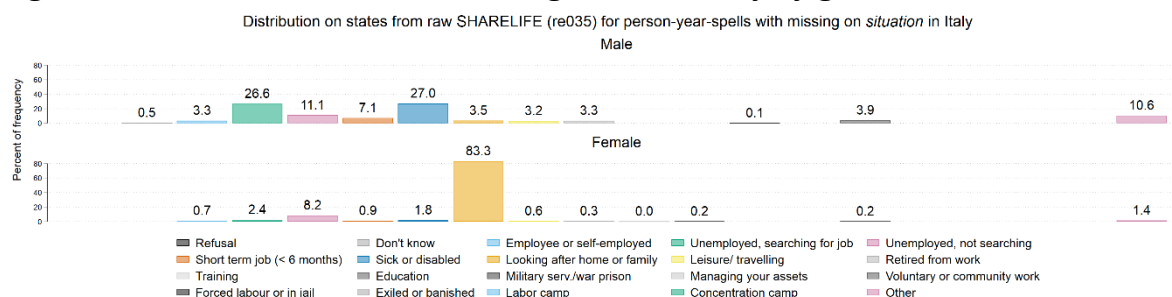
have underestimated the share of unpaid care work done by women and the emerging trajectories severely. In Italy, the second largest share of filled person-years spells that would have been concealed otherwise refers to *unemployed not searching for a job* (8.2%). In the Netherlands and Germany, substantial shares are filled with *other* as well as *sick and disabled* in the Netherlands and *unemployed, searching for job*.

Figure 2.11 Distribution of filled working states for Germany by gender



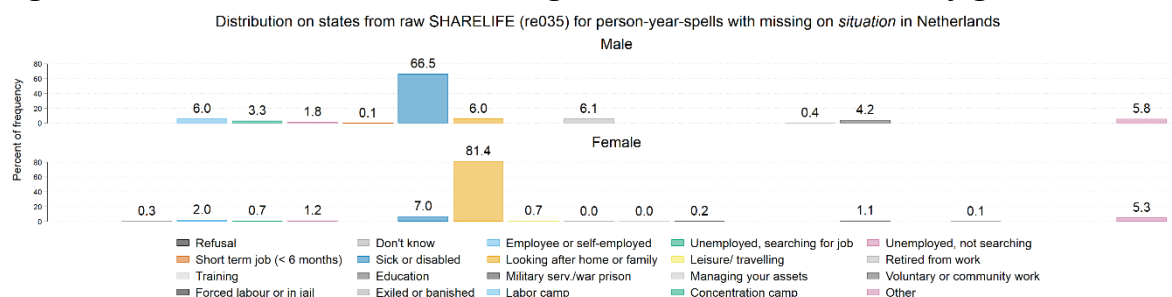
Notes: Own calculations based on the whole sample of JEP and SHARELIFE release 7.1.0. Some small categories merged.

Figure 2.12 Distribution of filled working states for Italy by gender



Notes: Own calculations based on the whole sample of JEP and SHARELIFE release 7.1.0. Some small categories merged.

Figure 2.13 Distribution of filled working states for the Netherlands by gender



Notes: Own calculations based on the whole sample of JEP and SHARELIFE release 7.1.0. Some small categories merged.

For men, the work states with which the missing person-year spells are filled are more diverse in general and differ more strongly across countries. In the Netherlands, 67% of

spells were filled with *being sick or disabled*, in Germany 29.3% and in Italy 27% (see top panels Figures 2.11-2.13). In Italy and Germany, *unemployed, searching for job* (26.6% and 39.4%) and *unemployed not searching for a job* (11.1% and 8.7%) are the other two work categories that were underestimated before the filling of missing person-year-spells.

As a result, for both men and women in the countries assessed in this dissertation, specific work categories that deviate from employment and could be characterized as precarious were concealed in the original situation variable of the JEP. At the same time, gender differences in the work items that are concealed by the missingness are strongly pronounced in Italy, the Netherlands, and Germany. Thus, in light of the gender focus of this dissertation, uncovering the source of the missingness and filling the missing person-year spells is highly relevant.

Summary and use

The missingness of the *situation* variable is not random, but very gendered. Furthermore, the situations concealed by the missing information in the original *situation* variable are highly skewed as well, mostly covering rather precarious situations beyond employment. The most concealed work state is by far unpaid care work, which is particularly problematic because it is the most underreported item in SHARELIFE data before introducing the gender bias in the data generation revealed here (Bühler et al., 2022). Consequently, using the JEP without filling in the missing data largely conceals the amount of unpaid care work that individuals provided over their life courses. Given that the bias was due to overlooking the situation of individuals after their last job until retirement, this example shows that the “steady worker bias” (O’Rand, 1996, S. 235) can also apply to data management.

Depending on how researchers have dealt with this missingness in their analysis, they might have created skewed data, if they, for example, excluded respondents with large shares of missing data. Particularly, they might have excluded selective individuals such as

sick/disabled or unemployed men and female unpaid caregivers. Thus, the analyses are likely to underestimate the share of working states deviating from employment as well as the level of inequalities. This is particularly problematic since most of the studies using the Job Episodes Panel do not discuss how they deal with the large shares of missing person-year spells. The results aim to give researchers a first idea of whether the missingness issue might be problematic for their studies.

Other variables provided by the Job Episodes Panel that are originally based on the same variables as the *situation* variable (Brugiavini et al., 2013) might be updated based on the new *situation* variable as well. This is particularly the case for the *unemployed* and *retired* dummies. The *working* dummy on the other hand should not be adjusted based on the new *situation* variable as it was constructed from another set of variables.

The team of the Job Episodes Panel will address this issue, among others, and publish a new version of the Job Episodes Panel including some other improvements with the next SHARE release. Until this release is publicly available, scholars who used or currently use the Job Episodes Panel and the *situation* variable might still want to use the (almost) full retrospective working/activity lives. This might be relevant for ongoing projects, but also for replications of already finished studies for which the *situation* variable was relevant. The code provided can be applied to fill most of the missing person-year spells on the *situation* variable and is meant to be used only as a transition solution until the fully revised Job Episodes Panel is available.

2.3.2 Further limitations of the retrospective life course data and implications for this dissertation

Chapters 3-5 apply the retrospective data for work states with the filled missing information based on the described procedure to prevent concealing life course states that deviate from employment.

There are still some limitations in the retrospective data which should be kept in mind. First, employment and retirement spells tend to be overestimated in the data processing in the Job Episodes Panel (Brugiavini et al., 2019).¹⁵ Second, due to the organization of the SHARELIFE questionnaire, the actual degree of unpaid work per year is likely to be underestimated. Full-time care work that was done for less than 6 months per year is not considered if respondents were in the labor force for the remainder of the year. More generally, unpaid care work that was done simultaneously with employment or other activities, is not captured, as just one item of the item battery could be selected. Consequently, unpaid care work as surveyed in SHARELIFE exclusively captures *full-time* care work that was done for at least 6 months a year.

Furthermore, as literature has also shown that women are likely to reduce their working hours due to caring responsibilities for the studied cohorts (e.g., Carr et al., 2016; Spiess & Schneider, 2003), it is very likely that the state of part-time work captures unpaid care work as well, most likely to reconcile care responsibilities with paid work.

Other limitations refer to the family life course data. First, for the number of children, there is no information on when the children lived in the same household as the adult respondent. This would be important to know, particularly from a gender perspective for understanding work-family interdependences, as women are more likely to be involved in childcare which limits their career opportunities. For example, in Italy, the US, Sweden, and France, women, and to a lesser extent men, provide significantly more hours of unpaid care when there are (young) children in the household compared to empty nest couples (Anxo et al., 2011). As such, individuals who have children, but do not share the household with them likely experience lower (child) care needs, and thus an overall level of care work that is more comparable to that of people without children. However, the data do not allow to

¹⁵ In years in which working and other activities were reported, working spells were used. Retirement was used for years in which working and retirement was surveyed.

distinguish between sharing the household with children or not. Such information may be particularly relevant after union dissolution when children usually live with only one parent.

A rough approximation of the parents who might be affected that is feasible with the data is the proportion of parents who are divorced or not living together since in these cases one of the parents is most likely not living with their children. For example, 12.2% (9.5% for spells if the youngest child was below 17 years old) of years Dutch parents were not cohabiting, and 6.6% (4.4%) of parents in West Germany. In both, the Netherlands and West Germany 3.4% (2.1% and 2.3% for spells if the youngest child was below 17) of person-year spells parents reported being divorced. The proportions were slightly higher for mothers than for fathers. As a result, a small but significant proportion of parents may not have lived in the same household as their children and may therefore have had a lower care need. Future research might want to consider different data that allow a distinction to be made between parents living with their children or in separate households.

In addition, I do not differentiate between cohabitation and marriage for partnership status due to the small number of cohabitation periods. More generally, the retrospective data is surveyed yearly which means it does not survey multiple transitions within a year and therefore simplifies life courses.

Lastly, the limitation remains that the work state 'retired' as it is surveyed and measured in the retrospective data similarly risks underestimating the retirement of women as shown for the self-reported employment status of the survey years as shown in Chapter 2.2. As a result, the retrospective data likely observes fewer retirement spells for women which might have implications for the life course cluster and aspects used in Chapters 3-5. However, these should not have a large impact on the results for Chapters 3-4 at least, since retirement is not among the most dominating characteristics of typical life course patterns.

3 Life-course-sensitive analysis of group inequalities: Combining Sequence Analysis with the Kitagawa-Oaxaca- Blinder decomposition

Carla Rowold, Emanuela Struffolino, Anette Eva Fasang

Abstract

Processes that unfold over individuals' life courses are often associated with inequalities later in life. The literature lacks methodological approaches to analyze inequalities in outcomes between groups, for example, between women and men, in a life-course-sensitive manner. We propose a combination of methods—of sequence analysis, which enables us to study the multidimensional complexity of life courses with Kitagawa-Oaxaca-Blinder decomposition. This approach allows us to distinguish the share of inequalities between groups that is due to group-specific life courses from the share that is due to group-specific returns to similar life courses. We illustrate the combination of the two methods by analyzing work-family life courses and Gender Pension Gaps in Italy and Germany. Our contribution is to systematically compare possible core analytical choices when combining typologies derived using sequence analysis with the Kitagawa-Oaxaca-Blinder decomposition. For future applications, we propose a set of practical guidelines for sequence analysis–Kitagawa-Oaxaca-Blinder decomposition.

Keywords: sequence analysis, Kitagawa-Oaxaca-Blinder decomposition, gender, pensions, life course, SHARE data, Gender Pension Gap

3.1 Introduction

What processes link earlier life courses to group-specific inequalities later in life? Are there differences across welfare states? To address these and related questions, we propose a stepwise combination of sequence analysis (SA) and Kitagawa-Oaxaca-Blinder decomposition (KOB) and systematically compare different possibilities for core analytical choices. In the process, we highlight advantages and limitations and derive a set of practical guidelines for combining the two methods.

Sequence analysis has become the standard method in the social sciences for analyzing trajectories of several categorical states —i.e., how these states develop over a certain time (Liao et al., 2022). The most prominent of these states are work and family life courses. SA is used to identify, categorize and visualize similar life-course patterns in cluster-based typologies. Multichannel sequence analysis (MSA) is a special form of SA that is used to describe how events in multiple domains (e.g., employment and family life) unfold jointly (Gauthier et al., 2010; Pollock, 2007). MSA's advantages include its capacity to simultaneously consider the occurrence and duration of certain life-course states, their timing, and their sequencing in more than one life domain. A typical example is the association between parenthood (family domain) and employment (work domain): parents often leave the labor force to care for their children. To date, life-course typologies emerging from SA and cluster analysis have been used as dependent or independent variables in regression models (e.g., Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022; Riekhoff & Järnefelt, 2017). Yet, these existing SA designs have not assessed whether group inequalities in outcomes, such as the Gender Pension Gap (GPG), arise from group-specific compositional differences in life courses or group-specific returns to similar life courses. For example, does the GPG mainly result from men and women experiencing

entirely different life courses that are unequally rewarded in pension systems or from men and women receiving differential pension returns to similar life courses?

KOB decomposition is the method of choice to estimate group-specific inequalities in outcomes, for example, gender wage gaps (Kunze, 2018). Going beyond regression analyses, KOB decomposition reveals what part of the gap is due to differences in mean characteristics between groups and what part is due to group differences in returns to the same characteristics. However, previous studies have heavily relied on summary indicators (e.g., years of full-time work) or “point-in-time” indicators to account for individual experiences over the life course. Such measures simplify life-course complexities, such as the timing, ordering, or co-occurrence of events over the life course (Halpern-Manners et al., 2015). We propose combining the strengths of SA and KOB to provide a life-course-sensitive analysis of group inequalities. Specifically, a combination of SA and KOB captures the extent to which gender gaps are related to (a) gender-specific life-course clusters (i.e., compositional differences), or (b) unequal rewards for men and women with similar life courses who are classified in the same life-course cluster (i.e., differential returns).

We illustrate the SA-KOB decomposition by analyzing the Gender Pension Gap in Italy and West Germany for the years 2006–2015 and the birth cohorts 1911–1950 using data from the Survey of Health, Ageing and Retirement in Europe (SHARE). By comparing Italy and (West) Germany, we highlight the added value of the proposed procedure for comparative welfare state studies. Germany and Italy share characteristics typical of conservative welfare states but have different pension systems. So far, empirical studies have typically decomposed the GPG based on one-dimensional summary measures and focused on the employment sphere (e.g., Even & Macpherson, 2004; Zhao & Zhao, 2018). However, life-course research shows that work and family life courses and their

combinations are highly gender specific and vary across welfare states (e.g., Fasang & Aisenbrey, 2022; Komp-Leukkunen, 2019; Madero-Cabib & Fasang, 2016; Tosi & Grundy, 2021). The literature currently lacks an account of work-family life course relations that uncovers gender-specific inequalities in later life outcomes, such as the GPG (Allmendinger et al., 1992; Ginn et al., 2001). To address this in the context of the SA-KOB application, we make use of multichannel sequence analysis in our illustration and provide guidelines that can be used to decide whether to cluster different life-course domains separately based on single-channel SA or account for their relationship by clustering them jointly using MSA. Only a few studies have used SA and regression analysis to identify typologies of life courses and relate them to women's pension incomes (Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022; Tophoven & Tisch, 2016). To date, no study has combined SA with decomposition methods to identify the share of GPGs due to compositional differences in life courses between men and women or differential returns to similar life courses.¹⁶ Combining SA and decomposition can provide new insights on a broad range of related questions regarding how trajectories of categorical states are associated with subsequent group inequalities. To combine the two methods, we first run sequence and cluster analyses to identify a typology of prevalent life courses. We then use this typology as a covariate in a KOB decomposition of group differences in an outcome, here gender gaps in pensions.

Our contribution is therefore twofold. First, we introduce and systematically compare different possibilities in core analytical choices when combining SA and KOB to derive a set of general guidelines for applying SA-KOB decomposition. Second, our empirical

¹⁶ Van Winkle (2020) suggested combining sequence analysis metrics with KOB to link compositional changes in individual level characteristics to population level standardization of family life-courses. In his approach, the outcome of the SA is a single average metric representing family formation standardization, which serves as an outcome variable to be decomposed across cohorts. Our approach differs in that it constructs typologies with SA and uses them as independent variables to decompose group-specific inequalities in outcomes.

example generates new insights into the association between gendered life courses and the GPG, with implications for other fields of application. Specifically, we show that a substantial proportion of the GPG is due to the temporal interdependence between the occurrence of certain family transitions (to partnership and parenthood) and transitions from different types of labor-market participation to unpaid care work. However, these transitions only arise in life course clusters typical for women. MSA captures this link over time between life-course spheres overlooked by previous research, which has focused on the characteristics of labor-market participation. SA-KOB decomposition also overcomes limitations of previous GPG studies, which had to restrict the number of measures of working lives used to avoid multicollinearity.

Both SA and KOB are descriptive methods, and their combination does not allow any causal inference. Yet, the distinction between compositional and returns effects of life courses on later life outcomes has implications for life-course theory and social policy. The SA-KOB combination offers more detailed descriptive findings that can support or refute theoretical propositions on processes of accumulation over the life course. For instance, different mechanisms are likely to drive compositional or returns effects to life courses, as we elaborate below, and this has immediate policy implications. Specifically, reducing compositional differences in life courses would likely require policy measures directed at early and middle adulthood, whereas adjusting pension regulations might alleviate group differences in pension returns to similar life courses. The combination of SA and KOB can easily be applied to any outcome linked to processes of accumulation over the life-course that are group-specific, for example, wealth or health gaps by gender or race.

We first introduce our application and then discuss the necessary steps in SA and KOB in detail. As we move through the discussion of the various steps, we outline practical guidelines for applications. Before we discuss the SA-KOB results, we first replicate the

most popular way of analyzing GPGs so far, which is decomposing them based on life course summary measures. Last, we compare the results of KOB based on life course summary measures and SA-KOB and conclude with a discussion of the main contributions of SA-KOB.

3.2 Illustrative application: Life courses and Gender Pension Gaps in Germany and Italy

3.2.1 Gendered life courses and retirement systems

Family life courses are more strongly associated with women's pension income than with men's (e.g., Fasang, 2010; Fasang et al., 2013; Hofmeister et al., 2006; Krüger & Levy, 2001; Meyer & Pfau-Effinger, 2006; Muller et al., 2020). Family and care responsibilities are “key individual determinant[s] of women's employment” (Zagel & Van Winkle, 2022, S. 3), and the literature documents employment and earnings penalties of marriage and motherhood that vary across welfare states (e.g., Aisenbrey et al., 2009; Arntz et al., 2017; Boeckmann et al., 2015; Markus Gangl & Andrea Ziefle, 2009). Cumulative advantage or disadvantage (CAD) in life-course theory (Dannefer, 2003) expresses “that the advantage of one individual or group over another grows (i.e., accumulates) over time” (DiPrete & Eirich, 2006, S. 272), leading to increasing within-cohort inequality that culminates in old age (Dannefer, 1987). In the tradition of CAD “[t]he gender gap in pensions can be understood as the sum of gender inequalities over a lifetime, including differences in the life-course (motherhood penalty), segregated labour market and gendered social norms and stereotypes more generally.” (European Institute for Gender Equality, 2015, S. 4). Therefore, both employment and family life courses have to be considered when analyzing gender gaps in pension income. Institutional and cultural contexts foster or discourage certain gender arrangements and thus shape gendered life courses and outcomes (Elder et al., 2003; Krüger & Levy, 2001; Pfau-Effinger, 1998; Rosenfeld et al., 2004). Moreover,

pension systems reward certain life courses more than others (Ginn et al., 2001). Most pension systems assume continuous full-time employment as the norm (Leitner, 2001) and disregard the gendered division of labor.

A gender-sensitive analysis of pension inequalities requires researchers to consider all factors shaping *access* to and *amounts* of pension income (Ginn et al., 2001; Jefferson, 2009). Regarding *access* to pension claims, eligibility rules, age thresholds, and mechanisms for considering time spent doing unpaid care work are particularly relevant for women (Ginn et al., 2001). A usual qualifying condition is the minimum length of contribution, which systematically excludes individuals with discontinuous employment records and thus disproportionately affects women (Leitner, 2001). Regarding the *amount*, overall redistributive elements tend to augment women's pension incomes (Ginn, 2004; Vlachantoni, 2012). Conversely, earnings-related contributions and less progressive pension systems are disadvantageous for women (Grech, 2013; Horstmann et al., 2009; Samek Lodovici et al., 2011). Care benefits and the upgrading of part-time work in pension entitlement further elevate women's pension income (Möhring, 2018). Because women have more interrupted work lives, lower earnings, and limited access to occupations that guarantee generous occupational pension schemes, they often find that personal pensions¹⁷—which are gaining importance in most European countries—yield lower returns or are not viable in the first place (Ginn, 2003, 2004; Ginn & Arber, 1996; Jefferson, 2009; Möhring, 2018).

3.2.2 Empirical evidence on Gender Pension Gaps

Most studies have decomposed the GPG based on a set of individual-level characteristics, including retrospective information on labor-market attachment, operationalized as the sum

¹⁷ If not explicitly stated otherwise, “personal pensions” refers to the sum of both occupational and private pensions.

of years individuals spent in a certain employment status over their life course (Bardasi & Jenkins, 2010; Bettio et al., 2013; Bonnet et al., 2020; Cordova et al., 2022; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Hänisch & Klos, 2014; König et al., 2019; Levine et al., 1999; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018). Table 3.1 provides a literature overview of decomposition analyses of GPGs. Most of the studies on public pensions or total pension income have stressed that gender differences in labor-market attachment are among the main explanatory factors for GPGs (Hänisch & Klos, 2014; Even & Macpherson, 2004; Levine et al., 1999; Bonnet et al., 2020; Nolan et al., 2019; Frommert & Strauß, 2013). The few studies that have included family life-course indicators (Table 3.1) show that gender differences in marriage rates are related to higher GPGs, whereas the higher share of widowed women compared to men is related to lower GPGs in most countries (e.g., Hänisch & Klos, 2014; Veremchuk, 2020). Two studies support the idea that much of the GPG is due to gender differences in returns to marital status and fertility, i.e., suggesting that women experience pension penalties for marriage and parenthood but men enjoy pension premiums (Bardasi & Jenkins, 2010; Ezeyi & Vujic, 2017).

Other studies have focused on women's pensions or gender inequalities in pension incomes by applying OLS regressions instead of decomposition techniques. They have incorporated the life course perspective by including similar summary measures as the decomposition literature as well as composite indicators, for instance on "career volatility" (Möhring, 2015, S. 11) or career types (Sefton, Evandrou, Falkingham, et al., 2011). Evandrou et al. (2009) and Fasang et al. (2013) found strong associations between summary measures of family history and women's later life income, partly even after controlling for socioeconomic status and employment history.

Table 3.1 Overview of KOB studies on Gender Pension Gap

			Characteristics on employment sphere							Characteristics on family sphere					Pension income	GPG		
Author	Year	Country	ft empl	pt empl	unemployed	care work	occupation	wage	other	married	divorced	single	widow	n of children	Impact family discussed?	types	Total share explained	Compositional differences in (F/T) employment main reason?
Bardasi & Jenkins	2010	UK	yes, retro*	yes, retro*	yes, retro		yes, retro		share of time spent in inactivity	yes, retro*	yes, retro*		yes, retro*	yes, combined with marital status*	no	occ. + priv.	8-18%	no, differences in returns much more relevant
Bettio et al.	2013	multi-country	yes, retro*	na					share of pillar 3 (partly *)	yes*	yes	yes	yes*			occ. + priv.	too many	unclear
Bonnet et al.	2020	France	contribution period, unclear what covered, probably mostly employment years*						yes*					yes	no	public; occ. + priv.	86%	larger share due to wage than to contribution period, but both main contributors
Cordova et al.	2021	Germany	yes, retro*	yes, retro*	yes, retro*		yes		industry, company size, pension type coverage	yes	yes	yes	yes	yes + age*	marginally	pension wealth (accrual), sum of all three pillars	>100%	yes
Even & Macpherson	2004	US	yes, retro*	na					yes*	KOB by married and single women					no	unclear, probably sum of all three pillars	31-55%	yes, besides salary, but salary more important
Ezeyi & Vujić	2017	UK	yes, retro	yes, retro					starting wage last job before retirement*	yes	yes	yes			no	public; occ. + priv.	public: 25%, personal: 12%	yes, largest share for public, second largest for private, but unexplained shares are larger

Table 3.1. Overview of KOB studies on Gender Pension Gap (continued)

			Characteristics on employment sphere							Characteristics on family sphere					Pension income	GPG		
Author	Year	Country	ft empl	pt empl	unemployed	care work	occupation	wage	other	married	divorced	single	widow	n of children	Impact family discussed?	types	Total share explained	Compositional differences in (FT) employment main reason?
Frommert & Strauß	2013	West Germany	yes, retro*	yes, retro*	yes, retro	yes, retro			marginal employment						margi-nally	public, sum of all three pillars	54-56%	yes
Hänisch & Klos	2014	Germany	yes, retro*				yes		noncontinuous employment history, never been employed	yes*	yes*	yes	yes*	yes	no	sum of all three pillars	27%	yes, together with education
Johnsen et al.	1999	US	yes, retro*				yes*	yes*	education, race, industry, Union, firm size	yes*, combined					no	occ. + priv.	78%	yes
König et al	2019	Sweden					yes*	yes*	industry, pension type coverage						no	sum of all three pillars		not included
Levine et al.	1999	US	yes, retro*	na			yes*	yes*		Only non-married					no	pension wealth (anticipated), sum of all three pillars	85%	yes, besides earnings, but earnings more important
Nolan et al.	2019	Ireland	yes*	na					disability, nationality, living status, location etc.	yes, status*	marital			yes	margi-nally	occ. + priv.	2-44%	yes, but only at top deciles of distribution
Veremchuk	2020	multi-country	yes, retro*	na			yes*		migrant status, chronic illnesses	yes*	yes		yes*		no	sum of all three pillars	too many	yes
Zhao & Zhao	2018	China	yes, retro*					yes	hukou type, province, education						no	occ.	30%	yes

Notes: This table includes decomposition analyses of Gender Pension Gaps published 1990–2021. Significance for the explained share by characteristics indicated by *. Ft= full-time employment, pt= part-time employment, retro: retrospective/longitudinal measure used (mostly sum of years over life course). If not 'retro', cross-sectional data is used (or no information was available). If not otherwise indicated, individual annual or monthly pension income was used as dependent variable. An exception are the studies of Cordova et al. (2022) and Levine et al. (1999), who used pension wealth (as the expected pension entitlement from start of retirement until end of life). Note that the total share of the GPG that is due to the explained part differs largely by which reference coefficients are used.

Madero-Cabib & Fasang (2016) analyzed (gendered) old age income inequalities using MSA to identify typologies of work and family life courses. Findings indicated that, in West Germany and Switzerland, female-dominated life courses are much less well-rewarded than men's standard life courses (characterized by continuous full-time employment, being married, and having at least two children). More recently, Möhring & Weiland (2022) have taken a couple perspective and shown for Germany that women's old-age income is highest for women in dual-earner couple life courses and lowest for women in male-breadwinner partnerships, though this association weakens when controlling for the number of children and the share of childcare over women's life courses. Focusing on the association between employment life courses and public pension claims in Germany, Tophoven & Tisch (2016) found that women with unstable or interrupted work life courses had lower pension claims than those with relatively stable full-time or part-time trajectories. However, neither Möhring and Weiland (2021) nor Tophoven and Tisch (2016) included family life courses in the SA step.

3.2.3 Italy and West Germany

Italy and West Germany¹⁸ are two examples of gender-conservative welfare state systems and have among the highest GPGs in Europe (Hammerschmid & Rowold, 2019b). While the pension systems are different, both are described as Bismarckian models that are particularly disadvantageous for women due to the tight link between earnings and pension benefits (Corsi & D'Ippoliti, 2009; Fasang, 2010; Horstmann et al., 2009; Samek Lodovici et al., 2011).

Family policies and gender norms. In the 20th century, West Germany was an ideal-typical gender-conservative welfare regime (Trappe et al., 2015). A significant policy shift towards

¹⁸ Given the contrasting welfare state contexts in the formerly separated states of Germany (Rosenfeld et al., 2004; Trappe et al., 2015), we consider them separately and focus on West Germany here.

more gender equality started at the beginning of the 2000s; however, this was not relevant for our study cohorts, who had long completed their active family formation at this time. In the familial Italian welfare state, benefits have traditionally been structured around family units rather than individuals (Saraceno, 1994), and the state has relied on family members (namely women) to provide care work (Hofmeister et al., 2006). Family policies were similar in West Germany and Italy for our study cohorts. Table A3.1 in the Appendix displays a summary of family policies and gender norms in the two countries. For instance, childcare in both countries only recently started to be publicly funded and regulated for children aged three or older, and gender and family norms were similarly traditional in the 1980s and 1990s. In both countries, female labor force participation was very low in 1990, but the rate in Germany was higher due to female part-time employment. Unlike Germany, Italy has no legacy of part-time work (Hofmeister et al., 2006). Additionally, joint taxation of married spouses and limited daycare opening hours in West Germany encouraged women to perform part-time work (Rosenfeld et al., 2004). Apart from women's higher access to part-time work in Germany, we expected to find similar gender differences in typical life courses for women and men in these two traditional welfare states and normative contexts.

Pension systems. Table A3.2 summarizes the pension systems. Both pension systems are Bismarckian, but the Progressivity Index suggests that the Italian public pension system is overall less redistributive than the German one. To qualify for old age pensions, individuals in Italy must have a much higher contribution record (20 years vs. 5 years), which is more difficult for women to reach.¹⁹ This might explain why 35% of Italian women but only 9% of West German women aged 65+ do not receive any own pension income. Even though

¹⁹ Since the early '90s the Italian system has shifted from a pay-as-you-go defined-benefit system to a defined contribution system, with an increase of the age and the number of years of contribution. The years of contribution are partly age dependant (see Table A3.2).

both public systems provide benefits for child (and elderly) care, women with low incomes particularly benefit from child benefits in Germany. Unlike in Italy, in Germany, these claims are based on the overall average income for the respective year and not on previous individual income. Thus, care benefits in Germany do not reproduce previous (gendered) labor-market disadvantages (Horstmann et al., 2009). Moreover, upon divorce, partners split their pension entitlements in Germany, but they do not do so in Italy—that is, entitlements are transferred from the partner with the higher claim to the one with the lower claim, which generally benefits women (Kreyenfeld et al., 2023). Finally, women in Germany benefit slightly more from the redistributive elements in the pension system, whereas men in Italy tend to be favored (suggested by the gender gap in replacement rate). At the same time, regressive personal pensions that tend to disadvantage women are more widespread in Germany than in Italy.

Because public pensions are the main source of total pension income in both countries, we expect gender inequalities over the life course to be slightly more reproduced in the Italian pension system. In sum, we expect that compositional differences in life courses will account for sizeable shares of the GPG in both countries due to the high level of gendered segregation on different life courses that we expect and unequal pension rewards for these. However, we expect the share accounted for by a) compositional differences and b) returns effects to similar life courses in the pension system to be larger in Italy than in Germany. This is due to the more restrictive access to pensions, and due to the lower level of overall redistribution in the Italian pension system.

3.2.4 Data, sample, and variables

The data came from the Survey of Health, Ageing and Retirement in Europe (SHARE, Börsch-Supan et al., 2013). We used Wave 5 as a base sample of individuals aged 65 or older at the year of interview and merged respondents from Waves 2, 4, and 6 who were

not surveyed in Wave 5 to increase case numbers in the analytical sample. The survey years spanned 2006 to 2015 (see Table A3.3 in the Appendix). We combined the cross-sectional information in the analytical sample with retrospective annual data on family and work life courses from age 18–65 included in SHARELIFE, collected in 2008–2009 & 2017 (Brugiavini et al., 2019).²⁰

The analytical sample included individuals aged 65 and older and excluded respondents: (i) with missing information on controls and annual pension income; (ii) who were part of the labor force and in receipt of a salary or unemployment benefits at the time of the interview; or (iii) who had entirely missing retrospective data (see Table A3.4). We retained non-employed respondents who reported work that could be classified as spare-time work or similar to keep retirees in need of additional income in the sample.

Pension income was specified as individual income from public, occupational, and private pensions based on independent own achievements (see Table A3.5). Income sources that were derived from other individuals, most importantly survivor pensions were excluded, as they were associated with economic dependence and a loss of autonomy (e.g., Ginn et al., 2001).

Women were more likely than men to not receive pension income. We retained respondents who reported receiving no pension income and assigned them a pension income of 0. Our analysis thereby broadly covers the entire population 65+. We used the absolute annual pension incomes and adjusted for purchasing power. We only considered regular payments (i.e., no lump sum payments, which played a marginal role in Germany and Italy). We further top-coded the highest 1% annual pension income with the 99th percentile.

²⁰ See Chapter 8 for details on the data.

3.3 Analytical strategy

3.3.1 Step I: Sequence Analysis

3.3.1.1 Overview of Sequence Analysis

Over the last twenty years, sequence analysis (SA) has become a key analytical tool for analyzing trajectories of categorical states in the social sciences, particularly in the field of life-course research (Abbott, 1995; Liao et al., 2022; MacIndoe & Abbott, 2004). The standard SA workflow pursues an exploratory approach to uncover regularities in temporal processes by, first, operationalizing sequences of categorical states that capture a temporal process of interest (e.g., educational or employment trajectories) and then using data-reduction techniques such as cluster analysis to identify the most typical empirical realization of that process. Typically, sequence analysis is used to calculate pairwise distances between all individuals in a sample to determine sequence similarity. This pairwise dissimilarity matrix captures the extent to which each sequence is similar to any other one in the sample. Similarity is understood as the cost of the operations computed to transform one sequence into another. Operations include the insertion, deletion, or substitution of states along the sequences: this is the basic optimal matching strategy, as introduced by Abbott and Hrycak (1990; see Studer & Ritschard, 2016 for an overview of distance metrics). The sequence dissimilarity matrix is used in a cluster analysis to identify typologies of sequence clusters. The clusters can be used as dependent or independent variables in a regression framework to link them to baseline individual characteristics or to an outcome. For a detailed and step-wise introduction to SA, including instructions for coding in the TraMineR module in R (Gabadinho et al., 2011), see Raab & Struffolino (2022).

3.3.1.2 Single versus multichannel sequence analysis

To address life-course-related questions, researchers may need to account for parallel processes in different life domains—for example, employment and family formation. There are several strategies used to construct a joint multidomain typology. The most popular consists of the computation of a joint dissimilarity matrix based on costs derived additively from domain costs (Gauthier et al., 2010; Pollock, 2007). This is referred to as multichannel sequence analysis (MSA, see Piccarreta 2017 and Raab and Struffolino 2022 for an overview and Ritschard, Liao, and Struffolino 2023 for a critical review).

Both single-channel SA—i.e., the application of SA to one life-course domain (e.g., work or family)—and multichannel SA—i.e., the application of MSA to account for several domains simultaneously (e.g., work *and* family)—can be combined with KOB. The question of whether to apply single or multichannel SA depends on the research question and the data characteristics. There are three options for combining single-channel SA and KOB. The first applies to research questions that focus on one domain; in this case, clusters are generated for the single domain of interest and they are used in KOB as independent variables to decompose group inequalities in a given outcome. The second option is to generate clusters on the different domains separately and use the different cluster typologies as independent variables (potentially in a step-wise fashion) in KOB. In this case, for example, researchers might be interested in how typical work trajectories are related to GPG while controlling for typical family life courses. The third option draws on the second one but includes interactions between the cluster typologies. Including interactions will likely yield a large number of comparison groups and very small cell sizes in conventional survey data, which might result in imprecise and hard-to-interpret coefficients. This third strategy is the only one that promises to account for the relationships between trajectories in different domains using single-channel SA. However, due to the aforementioned

limitations, we prefer multichannel SA. In the version proposed by Gauthier et al. (2010) and Pollock (2007) it summarizes core patterns of parallel unfolding of different (life course) domains such that it accounts for the link between domains in the generation of the clusters.

Additionally, in most applications of SA-KOB decomposition, the assumption is that the relationship between trajectories in domains differs by the group of interests. For example, we are interested in gender differences, and it has been shown that family lives are more consequential for women's employment trajectories than for men's (see above). Therefore, the relationship between work and family domains is relevant for our illustrative empirical example. Piccarreta (2017) and Ritschard et al. (2023) have introduced criteria to examine the nature of the link between states across different domains and between trajectories. It is advisable to consider these depending on the research question. In our case—and potentially for other SA-KOB applications—group-specific associations between the life-course domains are expected and the link between the different domains should be assessed for each of the groups separately (e.g., men and women in our application) (Raab & Struffolino, 2022, S. 121). Even if the relationship between domains is empirically found to be driven by only one of the groups (e.g., women), we suggest proceeding with MSA rather than single-channel SA when combining SA with KOB. This is justified by the theory-based expectation that the process that generates the inequality between groups in the outcome can be related to group differences in the link between different domains. However, if life-course domains are, in empirical terms, not strongly linked in any of the groups, then two separate single channel cluster analyses might yield a more efficient clustering.

As we assumed that the relationship between work and family life courses was gender-specific and we were interested in how this relationship was related to GPGs, we applied

MSA in our main model.²¹ In section 3.4.5, we apply the single-channel options, describe the results, and elaborate on the advantages and disadvantages of single versus multichannel SA for the analysis of GPGs.

3.3.1.3 Definition of the work and family life-course states

In our illustrative application, we constructed family and work life courses from age 18 to 65. The observation window spanned from 1929, when the oldest respondents were 18, until 2015, when the youngest respondents were 65. The family trajectory accounted for partnership status and the number of children in six mutually exclusive states: single, no children; single, 1+ children; married, no children; married, 1 child; married, 2+ children; and divorced, with/without children. Married also included non-cohabiting couples (3% of person-year spells) and nonmarried cohabiting couples (0.93% of person-year spells). Single referred to individuals not cohabiting and not married or divorced. Divorce overwrote all other family states occurring in the same year, except for marriage. The number of children included any children alive, including adopted children, in the given year (Brugiavini et al., 2019). However, it is not surveyed whether the children lived in the same household as the respondent although this information could be important for approximating childcare needs. Thus, respondents could have two or more children without ever having lived with them in the same household.

The work trajectory accounted for eight mutually exclusive states: education or training; full-time care work/other/missing; part-time employment in the private sector; full-time employment in the private sector; civil servant; self-employment; unemployed; and retired.²² Care work was identified based on the response option “looking after home or

²¹ As suggested by Piccarreta (2017), we calculated the correlation between the work and family trajectories before proceeding with any other step. We find a positive correlation between the distance matrices of both domains for women only. This confirms our decision to choose a multichannel over single-channel SA.

²² There are other characteristics of working biographies that are also relevant for pension income, such as earnings, industries, and occupations. However, we lack detailed retrospective data on these characteristics. Heterogeneity regarding these characteristics is thus not visible in the life-course clusters.

family”. This measure exclusively captured full-time care work that was done for at least 6 months a year and therefore underestimated the actual degree of unpaid care work per year. We combined care work with a category capturing the rest of the activities, mostly the states “sick and disabled” and “other”—that is, when none of the other items were applicable. These can be expected to yield limited pension entitlements. “Part-time employment” combined part-time spells and any reported part-time work done as (short-term) employment. “Unemployment” captured both unemployment and inactivity.

Using the generated dataset for employment histories (Brugiavini et al., 2019), we found that 11.56% of all person-year spells contained missing information in our sample. Missing information along the sequences was filled following the procedure described in A3.3 and Chapter 2.3.1 more broadly.

3.3.1.4 From individual sequences to typologies

In general, we recommend grounding the decision on the final cluster solution—the outcome of SA—on statistical measures and on the plausibility and theoretical fit of the clusters for the research question (see Aisenbrey and Fasang 2010 for an extension of the concept of construct validity to this case) and applying an informed sensitivity analysis (see Studer, 2013). In line with the exploratory approach of SA, we first applied multiple options to the main parameters; we computed different dissimilarity matrices and clustering algorithms (results available upon request). We considered standard statistical measures to choose the number of clusters per specification (average silhouette width (ASW), Hubert’s Somers’ D, point biserial correlation). Next, we compared multiple cluster solutions from the different specifications regarding their ASW. The ASW is the most used cluster cut-off criterion following sequence analysis (Studer 2013). Low values suggest that many individual cases have ambiguous cluster membership and might as well have been classified in other clusters. High ASW values indicate a coherent and discriminant grouping

in which sequences in each cluster are internally homogeneous and distinct from the other clusters. Nonetheless, in SA, quantitative cluster quality criteria should always be combined with a visual inspection of the content of the typology in light of the research question and theoretical considerations. Thus, as a last step, we visually explored the best performing clusters for the work and family lives (see Fasang & Liao, 2014 for an overview of visualization techniques).

By employing optimal matching with indel costs of 1 and substitution costs based on transitions rates combined with Ward-linkage hierarchical clustering, we identified an eight-cluster solution that we considered theoretically meaningful in our case. Importantly, it included a theoretically expected cluster dominated by atypical work arrangements, such as part-time work, and by unemployment and early retirement. This cluster was relevant for our study since it incorporated life-courses that differed systematically from the norm of stable full-time employment that is highly rewarded by the pension systems. Particularly in West Germany, part-time employment enabled access to the labour force for many women and the part-time and full-time divide is a core feature of gender inequality in the labour market in Germany. All the other cluster solutions that performed high on at least one statistical measure suggested a very similar work-family life-course typology. This confirmed the validity of our typology. The final eight-cluster grouping had an ASW of 0.37, which was high for an MSA application and was the most parsimonious one while still including the theoretically important part-time cluster. Tables A5.1 and A5.2 in the Appendix display cluster characteristics by country and gender.

3.3.1.5 Guidelines for SA when combining it with KOB: pooling groups and visualization

When applying SA comparatively across countries or social groups, researchers have to decide whether to run SA for group-specific samples such as countries or gender or for

pooled samples of these groups. In the literature both group-specific and pooled analyses are common. The benefits of group-specific analyses—for example, separate life-course typologies for Black men, Black women, White men and White women (Fasang & Aisenbrey, 2022)—are that they can identify distinct small life-course clusters that are only relevant for one of the groups of interest. Group-specific SA typologies therefore are particularly useful for descriptive purposes to illuminate which life-course patterns signify the most prominent life-course experiences of social groups. In contrast, pooled analyses across groups—for example, including men and women or multiple countries in one pooled analysis—are generally preferable when combining SA typologies with regression-based methods in a second step (e.g., Kapelle & Vidal, 2021; Madero-Cabib & Fasang, 2016; Muller et al., 2020; Raab & Struffolino, 2020; Uccheddu et al., 2022). When clusters are used as independent variables, as in SA-KOB, pooled analyses allow researchers to interact life-course clusters with core group variables, such as gender or country, to assess whether the same life-course type yields different rewards for men and women or in different welfare states (Madero-Cabib & Fasang, 2016).

For the combination of SA with KOB decomposition, the groups of interest used in KOB (in our case, gender) must be pooled, so that they help generate the clusters jointly. This is necessary to have the same covariates of interests (i.e., life course typology as clusters) for each group (i.e., men and women) in the KOB. The main risk is that the SA on the pooled sample might conceal distinct group-specific life-course patterns, and this might be especially true if one of the groups is much smaller than the others. This can be monitored by an accurate visual inspection of different cluster solutions extracted from the pooled and the nonpooled samples. This enables researchers to interpret the results accurately.

In our illustrative application, we also pooled the SA across Germany and Italy, as we were confident that we had captured the theoretically most relevant life-course clusters in both

countries with our final typology (see above). A joint typology across countries reduced complexity and simplified the interpretation of the findings. However, it would have been technically possible to calculate separate SAs by country if the research question suggested it or there were empirically highly country-specific life-course typologies. Note that we chose our comparison countries based on the rationale of similar conservative work-family policies over the study cohorts' life courses but different pension systems. Calculating a joint SA across countries but conducting the decomposition of the Gender Pension Gap separately for countries thus corresponded with our comparative life-course design.

Below, we propose a decomposition of the Gender Pension Gap based on the gender distribution across clusters. The decomposition step should thus rely on a thorough description of how the groups of interest, in our case men and women, are distributed across the life-course clusters (see Section 3.4.2). We recommend visualizing the life-course clusters jointly with the group-specific distribution (here: of gender) over the clusters (see Figure 3.1).

3.3.2 Step II: Kitagawa-Oaxaca-Blinder Decomposition (KOB)

In the second step, we applied a KOB decomposition (Blinder, 1973; Kitagawa, 1955; R. Oaxaca, 1973)²³ to each of the two countries separately²⁴ to quantify the share of the gap that was explained by gender differences in average characteristics (i.e., the independent variables introduced in the regression models; *explained part*, *compositional*, or

²³ Recently there has been a discussion whether to refer to the method as Kitagawa-Oaxaca-Blinder or only Blinder-Oaxaca decomposition. While the latter has been the default for many years (e.g., Jann, 2008) based on the work of Blinder (1973) and Oaxaca (1973), more recently, researchers have started to use the former to acknowledge that Kitagawa's (1955) work on decompositions two decades previously might have pioneered this method. Oaxaca and Sierminska (2023) show that the Kitagawa and Oaxaca-Blinder decompositions are identical in one specific case, but that the Kitagawa decomposition was an early version of the more general decomposition suggested by Blinder and Oaxaca. Even though we do not use this specific case of the OB decomposition, and our application is closer to the one of Oaxaca and Blinder, we prefer to use the name Kitagawa-Oaxaca-Blinder to acknowledge Kitagawa's fundamental contribution to strongly related decomposition techniques in the 1950s.

²⁴ In line with our comparative analysis design, we run the KOB separately by country to identify country differences in the association between the GPG and work-family life-courses. This is an analytical choice that can be altered in other applications.

endowment effect) and the share of the overall gap that was explained by gender-specific differences in returns to these characteristics, that is, differences in the coefficients (*returns* or *group effect*; Kunze 2018).²⁵ For a general introduction to KOB decomposition, see Jann (2008) and Fortin et al. (2011). In this section, we will first present our application and then justify the analytical decisions by systematically comparing them with alternatives, especially with regard to the specification of the KOB reference coefficient.

The decomposition was based on separate linear regression models by group, here of men and women. We applied the twofold decomposition (Jann, 2008, S. 455), which is computed as:

$$\bar{y}_M - \bar{y}_W = \underbrace{\alpha_M - \alpha_W}_{\text{Intercept}} + \underbrace{\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W(\hat{\beta}^* - \hat{\beta}_W)}_{\text{Returns}} + \underbrace{(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*}_{\text{Explained}}$$

where $\bar{y}_M - \bar{y}_W$ stands for the average GPG. $\bar{X}_M$ and $\bar{X}_W$ denote vectors of mean values of the independent variables for men and women. α_M and α_W denote the intercepts and $\hat{\beta}_M$ and $\hat{\beta}_W$ the coefficient vectors of men (M) and women (W) from the gender-specific linear regression model.²⁶ $\hat{\beta}^*$ is a so-called nondiscriminatory or reference coefficient vector. In the two-fold decomposition applied here, the standard assumption is that the reference coefficient, $\hat{\beta}^*$, captures returns that are not subject to discrimination. To achieve this, the “non-discriminatory coefficient vector” (Jann, 2008, S. 455) must be estimated or defined. In our illustrative application, we wanted to apply a pension reward structure that did not discriminate with regard to the groups of interest. The literature suggested different options for defining a nondiscriminatory reward structure in terms of the outcome, here pension income, that we discuss below.

²⁵ The use of “explain” and “explained” does not refer to causal relations but to the “explained part” of the decomposition, a common wording in the application of the KOB decomposition. Overall, the methods used are descriptive and cannot claim causality.

²⁶ We use bootstrapped standard errors for the KOB.

Since we were interested in the contribution of each life-course cluster to the gap in outcomes, we focused on the detailed decomposition, which revealed the contribution of each individual predictor respectively (Jann, 2008). The total explained and unexplained shares were thus less relevant for our application.

The *explained or compositional part* weighted the mean gender difference in characteristics by the nondiscriminatory coefficient.

$$(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*$$

In combination with SA, the explained shares indicated what part of the gap in the outcome was associated with the group-specific distribution over different trajectory types (row percentages, see section 3.3.2.1). It thus quantifies which of the group-specific life course patterns contributes most to the group inequalities in the outcome of interest. In our application, the work-family life-course clusters were the main characteristics of interest. The explained part thus quantifies the share of the Gender Pension Gap that was driven by the gender segregation of work-family life-course clusters, that is, men and women experiencing different life courses. Note that such associations should only have appeared if a pension system rewarded typical male and female life-course patterns differently. Even if men and women had completely separate life-course clusters, GPGs could have been very small if the typical male and female life course trajectories yielded similar pension income. For example, the explained part indicates what part of the gap was due to the underrepresentation of women in a certain life-course cluster $(\bar{X}_M - \bar{X}_W)$, such as continuous full-time employment, that is highly rewarded by the pension system (indicated by a high return; $\hat{\beta}^*$). The interpretation of the explained part that refers to compositional differences is thus sensitive to the underlying reward structure of the outcome.

The *total unexplained part* consists of two components. First, the sum of the group differences in the intercept from the group-specific regression models (*intercept component*). Second, the sum of the differences between the respective group-specific coefficients from the reference coefficient (here: the male and female coefficients, see section 3.3.2.1; *returns component* or *effect*). Differences in the intercept reflect differences in returns for men and women unrelated to life-course-cluster membership. These emerge due to overall unobserved heterogeneities and within-cluster heterogeneity. Thus, given that the intercept component is not meaningful since it is comprised of unobserved heterogeneity, we focused on the component of the unexplained share arising from different returns for the same life-course patterns (*returns component*):

$$\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W(\hat{\beta}^* - \hat{\beta}_W)$$

In combination with SA, the *returns component* of the unexplained shares indicates to what extent the different groups *within* the same life-course cluster—that is, experiencing a similar trajectory—were rewarded differently in the outcome. Applied to our example, they show what part of the GPG was due to differences in the pension incomes of men and women with the same work-family life-course cluster. As such, the returns component of the SA-KOB provides a tool to quantify how much of a gap is reduced due to better returns for the same life course pattern of the disadvantaged group. If the outcome is shaped by a component of welfare systems that aims to buffer some degree of inequalities, such as pensions through pension systems, the returns component of SA-KOB can capture the share of the gap that is mitigated by policies producing group-differences in returns. Examples would be policies that target specific life course characteristics and lead to higher pensions for the disadvantaged group, such as child care benefits for the care work mostly done by women. The returns component can also highlight how group-inequalities are accelerated because disadvantaged groups receive lower returns for the same life course cluster. For

example, GPGs might have been very high even if life-course clusters were not gender segregated at all, but men and women received different pension incomes for the same work-family life-course cluster.²⁷ However, differences in returns can also be result of gender-specific unobserved heterogeneity, as well as within-cluster heterogeneity.

Due to the expected high level of gendered work-family life courses coupled with the overall regressive pension systems in both countries (see section 3.2.3), we expected to find large compositional parts of the GPGs. We expected overall lower returns compared to compositional effects of the GPG, given that the redistributive elements in both pension systems are likely not strong enough to result in higher returns for women compared to men for the same life course cluster that would help mitigating the GPG. On the other hand, redistribution should still mitigate the production of GPGs due to higher returns of men for the same life course cluster that might result due to unobserved gender inequalities, such as in earnings.

Three key modeling choices and interpretations in KOB require special attention in combination with SA: 1) the choice of the reference coefficient, 2) dealing with (group-specific) within-cluster heterogeneity and its implications for the interpretation of the returns component, and 3) the specification of the baseline categories for categorical variables (here: clusters).

3.3.2.1 Choice of reference coefficients

One of the most important parameters of KOB decomposition (albeit one that is often not discussed in empirical applications) is the choice of the reference coefficients, as it is

²⁷ In some applications of KOB, the total unexplained part is attributed to discrimination. However, the unexplained part might be driven by unobserved characteristics more generally and by differences in returns resulting from group-specific heterogeneity in categories of observed characteristics (i.e., within-cluster heterogeneity in the case of SA-KOB, see Section 3.2.2). The former might be particularly applicable to life courses because information on all relevant characteristics over the life course are hardly available in the sort of data necessary to reconstruct individual life courses. Therefore, we do not claim to identify discrimination in our application.

consequential for the results (Fortin et al., 2011, S. 47ff.) and has implications for interpreting the estimates. For the two-fold decomposition, we had to define the nondiscriminatory reference category, which we assumed was not subject to discrimination. There were multiple suggestions for estimating the reference coefficients (Jann, 2008; Rahimi & Hashemi Nazari, 2021). However, because there is no consensus in the literature and the choice of the reference coefficients also depends on theoretical and empirical considerations, we systematically compared the most prominent options and our proposal of cluster-specific reference coefficients with regard to their general and specific advantages and disadvantages in the SA-KOB combination. Table 3.2 and Table A3.6 give an overview of the benefits and shortcomings of using the different options of reference coefficients as well as their interpretation.

Option I: using the coefficients of one group for all covariates. Discrimination might only be directed towards one of the groups of interests, e.g., women; therefore the coefficients of the other group, e.g., men, might be a good proxy for a nondiscriminatory rewarding structure. These coefficients are obtained from regression models for the respective group and can then be applied as the reference coefficients for all covariates used in the decomposition. In the gender wage-gap literature, for example, the male coefficients (i.e., coefficients from regression model for men only) are usually used as the reference because it is assumed that wage discrimination is not directed against men (e.g., Blau & Kahn, 2017; Kunze, 2008). Female disadvantages in pension income are particularly likely because most pension systems, including those of Germany and Italy, are structured around a typical male life course by rewarding stable, full-time employment (see Section 2). Thus, in our application, the coefficients of the regression models of men come closest to a nondiscriminatory return for pension income in general terms (see Table A3.6 in the

Appendix for example interpretations of the KOB estimates). However, this might not apply to all clusters.

When combining KOB and SA, some life-course clusters might contain much higher shares of the group that is not chosen as the reference, in our case, women. This could be problematic for at least two reasons. First, the underrepresented group in that life-course cluster is more likely to deviate from the majority group in some characteristics of the life-course trajectories and the return for this smaller group may be more likely to be biased by outliers. Second, the reward structure (in this case, the pension system) might have some special benefits that only apply to the majority group. For example, many pension systems have specific unpaid care benefits that are (more or less explicitly) directed to typical female life courses dominated by care work. When using the coefficients for men (or the minority group more generally) as reference, such benefits are not reflected in the reward structure used in the decomposition. Furthermore, especially when groups are very unequally distributed across life-course clusters or when within-cluster heterogeneity varies across groups (see discussion of guidelines below), it might be more difficult to choose the reference coefficients for all clusters based on one group than in usual applications of KOB.

The researcher will have to decide on which clusters to compromise on, because most likely one group—here, either men or women—will not be the ideal reference for all clusters. Disregarding the group composition of clusters for the choice of references as in option I means that researchers will inevitably apply theoretically or empirically inappropriate reference coefficients for some of the clusters.

Option II: pooled reference coefficients. Another alternative is to use coefficients from a pooled regression model of both decomposition groups by merging the reward structures of both groups, so that the reward structure considers the group composition of the covariates (Neumark, 1988). Even though this is a widely used strategy, it leads to

ambiguous interpretations because the reference is an artificial quantification that does not apply to any of the groups in real life (see Table A3.6). As a result, it is unclear what the reward structure actually captures. For example, it is impossible to tell whether the pooled coefficients are driven by one of the groups (especially the majority group) or represent a well-balanced combination of the reward structures of both groups.

Table 3.2 Options for reference coefficients

Reference coefficient option	I.a) Group 1 (here: men)	I.b) Group 2 (here: women)	II) Pooled	III) Cluster-specific majority group
Concept	Regression coefficients from group 1: $\hat{\beta}_M$	Regression coefficients from group 2: $\hat{\beta}_W$	Regression coefficients from pooled model over both groups: $\hat{\beta}^*$	For group-specific clusters: coefficients of the majority group For group-neutral clusters: coefficients of the theoretically nondiscriminatory group
Explained part	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_M$	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_W$	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*$	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_M$ <u>or</u> $(\bar{X}_M - \bar{X}_W)' \hat{\beta}_W$
Returns component from unexplained part	$\bar{X}'_W (\hat{\beta}_M - \hat{\beta}_W)$	$\bar{X}'_M (\hat{\beta}_M - \hat{\beta}_W)$	$\bar{X}'_M (\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W (\hat{\beta}^* - \hat{\beta}_W)$	$\bar{X}'_W (\hat{\beta}_M - \hat{\beta}_W)$ <u>or</u> $\bar{X}'_M (\hat{\beta}_M - \hat{\beta}_W)$
Benefits	- Straightforward interpretation - Reflects assumption that one group is not discriminated against		- Sensitive to the group composition of the covariates and implications for the rewarding structure	- Straightforward interpretation - Sensitive to the group composition of the covariates and implications for the rewarding structure
Limitations	- Nondiscrimination assumption unlikely to hold (esp. for all categories of all covariates of interests) - Disregards group composition of the cluster		- Artificial rewarding structure - Unclear which group drives coefficient - No straightforward interpretation	- Necessary to define thresholds for group-specific and neutral clusters & choice of majority group; application-specific
Specifics for SA-KOB	- Concept theoretically and empirically challenged in case of high level of group-specific life-course clusters		- In case of highly group-specific life-course cluster, the less typical group might skew the nondiscriminatory coefficients in the pooled regression model	- Addresses group-specific life-course clusters and its theoretical and empirical implications - Addresses some issues of group-specific within-cluster heterogeneity as majority group more likely to be coherent with overall cluster characteristics

Additionally, the pooled reference coefficients might be skewed by less typical minority-group life courses in the cluster because they are based on regression models for the whole sample. Related strategies use the weighted average of the male and female coefficients or weight them based on group sizes (Cotton, 1988; Reimers, 1983). Yet these approaches generally suffer from similarly ambiguous interpretations as pooled reference coefficients.

Option III: cluster-specific reference coefficients using the majority group for each cluster.

Addressing the limitations of options I and II, we suggest using cluster-specific reference coefficients based on the group for which the life course is more typical. For example, if a life-course cluster is much more typical for men (i.e., men are the majority group), the male coefficient should be used as the reference for this cluster. But if another life-course cluster is more typical for women, we should use the female coefficient as the reference. We propose this cluster-specific specification of the reference coefficients as a particularly suitable approach when combining SA and KOB.

Compared to using the regression coefficients from one group (option I), our preferred alternative (option III) prevents researchers from using a reference based on a nontypical and small subgroup of a given life-course cluster (and therefore with a probably nontypical reward) as a nondiscriminatory coefficient. An example arises when a few men, who systematically engaged in much less care work compared to the female majority group in a cluster, are assigned to the same cluster characterized by care work. Additionally, gender-oriented policy interventions, such as childcare benefits in pensions, are likely to be picked up by women's reward structures but not by men's. Thus, for such a life-course cluster, it is substantively and theoretically meaningful to use women's reward structures as a reference.

Similarly, compared to using the pooled reference coefficients (option II, Table 3.2), using cluster-specific references based on the majority group (option III) prevents the reference

coefficients from being driven by outliers in the reference group. At the same time, option III is sensitive to the group composition of the covariates and the implications for the reward structure but offers a more straightforward interpretation than using the pooled coefficient. In most applications, cluster-specific reference coefficients from the majority group (option III) will perform best from the theoretical and empirical point of view. The majority group in the specific life-course cluster is less likely to be discriminated against and the group's returns might be closest to the "real" returns for the cluster. The estimation of the returns for the majority group will be more robust than those obtained by using the other options, because they are less likely to be affected by outliers.

Empirically, the results will tend to be similar to those obtained when using the pooled reference coefficients, because pooled coefficients are likely driven by the majority group in the clusters (confirmed in our illustrative application, A3.5.1). But using cluster-specific references (option III) allows for a clearer and more substantively relevant interpretation. Depending on which group is the majority in the cluster, the shares can be interpreted in the same way as they could be when using the coefficients of one group (option I; see interpretation examples in Table A3.6). To sum up, option III combines the benefits of the other two options (straightforward interpretation and consideration of group composition of clusters, see Table 3.2) while providing researchers with greater flexibility to adjust the reference based on specific empirical cases or theoretical or policy considerations.

Guidelines for cluster-specific reference coefficients using the majority group

We propose a set of guidelines for the application of cluster-specific reference coefficients (option III). To begin with, researchers must identify group-specific life course clusters and the majority groups. For this, we suggest applying 1) theoretical, 2) empirical, and 3) statistical considerations. First, theoretically, one has to consider whether the reward

structure of a given cluster contains group-specific benefits that are expected to be exclusive to a particular group's reward structure (e.g., pension benefits for care work will mostly be picked up in women's reward structures). If this is the case, there might be reasons to pick this group as the reference. Second, when seeking to identify the majority groups empirically, researchers can use a descriptive table that shows the interaction between groups and clusters (Table 3.3).

Table 3.3 Cross table of gender with the clusters (illustrative application, for Italy).

(a) row %									
	<u>Cluster</u> <u>1</u>	<u>Cluster</u> <u>2</u>	<u>Cluster</u> <u>3</u>	<u>Cluster</u> <u>4</u>	<u>Cluster</u> <u>5</u>	<u>Cluster</u> <u>6</u>	<u>Cluster</u> <u>7</u>	<u>Cluster</u> <u>8</u>	<u>Total</u>
<u>Men</u>	45.39	6.48	15.4	3.61	1.18	15.48	8.25	4.2	100
<u>Women</u>	13.78	3.25	4.97	13.78	36.85	14.84	7.42	5.1	100
<u>Mean</u>	28.75	4.78	9.91	8.97	19.96	15.14	7.82	4.68	100
<u>Abs. Difference Men/Women</u>	31.61	3.23	10.43	10.17	35.67	0.64	0.83	0.9	
<u>Rel. Difference Men/Women</u>	69.64	49.85	67.73	73.8	96.8	4.13	10.06	21.43	
(b) column %									
	<u>Cluster</u> <u>1</u>	<u>Cluster</u> <u>2</u>	<u>Cluster</u> <u>3</u>	<u>Cluster</u> <u>4</u>	<u>Cluster</u> <u>5</u>	<u>Cluster</u> <u>6</u>	<u>Cluster</u> <u>7</u>	<u>Cluster</u> <u>8</u>	<u>Mean</u>
<u>Men</u>	74.76	64.23	73.59	19.07	2.8	48.39	50	42.54	47.35
<u>Women</u>	25.24	35.77	26.41	80.93	97.2	51.61	50	57.46	52.65
<u>Total</u>	100	100	100	100	100	100	100	100	100
<u>Abs. Difference Men/Women</u>	49.52	28.46	47.18	61.86	94.4	3.22	0	14.92	
<u>Rel. Difference Men/Women</u>	66.24	44.31	64.11	76.44	97.12	6.65	0	35.07	

Notes: Relative difference: the respective majority group (group with higher row percentages) used for the denominator. Bold are the row percentages we use for defining the cluster majority groups. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Both column and row percentages can be considered.²⁸ However, the column percentages are sensitive to the group sizes. Since, unlike the fairly equal distribution of men and women in our application, group sizes may differ in other applications, we suggest considering the gender gap in the distribution over the clusters, that is, in the row percentages, which are equivalent to the $\bar{X}_A$ and $\bar{X}_B$ used in the decomposition. We categorized life-course clusters

²⁸ The column percentages show the gender distribution *within* the clusters (i.e., the total number of individuals per cluster equals 100%) and the row percentages reveal the gender distribution *over* the clusters (i.e., the percentage of all men [women] in the cluster, all men [women] equals 100%). For example, for cluster 5, the row percentages reflect that 1.18% of all men compared to 36.85% of all women are assigned to this cluster. The column percentages reflect that 2.8% of the cluster are men and 97.2% are women.

as group-specific if the relative gender gap in the distribution was greater than 50% ($50 < 100 * (\frac{\bar{X}_A - \bar{X}_B}{\bar{X}_A})$), with $\bar{X}_A > \bar{X}_B$).

The choice of the cut-off for defining majority groups depended on the overall level of the group distribution on clusters. For example, in scenarios with lower levels of group-specific segregation of life courses, researchers might choose thresholds of 30% or 40%. Furthermore, comparing the cluster characteristics across groups can help identify the group that shares the key characteristics of the cluster and thus guide the choice of the majority group (i.e., men or women in our illustration). Lastly, the categorization of the majority group can be statistically validated by comparing the ASW by group for each cluster. The ASW—that is, the average fit of the trajectories to the cluster—should be higher for the majority group.²⁹ For most applications, consulting theoretical, empirical, and statistical guidelines will roughly result in the same choices for the cluster-specific majority group.

In a second step, the researcher must decide which reference to use for group-neutral clusters, that is, life-course types with no clear majority group. For group-neutral life-course clusters, we suggest using the reference coefficients of the overall nondiscriminatory group. This should be guided by theoretical assumptions on which group is least discriminated against regarding the outcome of interest. For wages and pensions, for example, researchers typically expect that men are not discriminated against, suggesting the coefficients of the regression model for men should be used as the reference for gender-neutral clusters.

In our application, we illustrate the choice of cluster-specific references (option III) based on the outcome of the SA at the beginning of Section 4.2. For our main model, we specify

²⁹ This validation is likely to work best for homogenous clusters.

cluster-specific reference coefficients based on the coefficients from the linear regression models for women for the female majority work-family life courses (b_{4W}, b_{5W}); we specify cluster-specific reference coefficients based on men's coefficients for the male majority and gender-neutral clusters ($b_{1M}, b_{2M}, b_{3M}, b_{6M}, b_{7M}, b_{8M}$):

$$\bar{y}_M - \bar{y}_W = (\bar{X}_M - \bar{X}_W)' \hat{\beta}^* + \bar{X}'_M (\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W (\hat{\beta}^* - \hat{\beta}_W)$$

$$\text{with } \hat{\beta}^* = \begin{matrix} b_{1M} \\ b_{2M} \\ b_{3M} \\ b_{4W} \\ b_{5W} \\ b_{6M} \\ b_{7M} \\ b_{8M} \end{matrix}, \hat{\beta}_M = \begin{matrix} b_{1M} \\ b_{2M} \\ b_{3M} \\ b_{4M} \\ b_{5M} \\ b_{6M} \\ b_{7M} \\ b_{8M} \end{matrix}, \hat{\beta}_W = \begin{matrix} b_{1W} \\ b_{2W} \\ b_{3W} \\ b_{4W} \\ b_{5W} \\ b_{6W} \\ b_{7W} \\ b_{8W} \end{matrix}$$

We apply the different options for setting the reference coefficients to our empirical case and discuss the results in A3.5.1. Our main findings are robust when applying all discussed options and we illustrate how, in addition to being easier to interpret, option III has advantages in our case based on theoretical and empirical considerations.

In the case of highly group-specific covariates, researchers might want to apply our suggested option to standard KOB applications as well and choose the reference coefficients for each of the covariates or categories separately.

3.3.2.2 (Group-specific) within-cluster heterogeneity

Poorly classified cases in the cluster typology, or outlier life courses that do not clearly fit into any cluster might also affect reference coefficients. These life courses are not a good match to the main characteristics of the life-course type. This is a common issue, as cluster analysis will always allocate each case to a cluster (see Raab & Struffolino, 2022). What this means for SA-KOB is that the reference coefficients used to weight the mean differences between the groups when computing the *explained component* might not be representative and fail to reflect the nondiscriminatory returns for the life-course cluster if

it is skewed by poorly assigned individuals instead. The explained part of the KOB might thus be overestimated or underestimated.

Within-cluster heterogeneity might affect the *returns component* of the unexplained part of the decomposition as well and especially so if the within-cluster heterogeneity is group specific. The group differences within clusters means that there are some differences between groups with regard to their trajectories, here men and women, in endowment with the same characteristic, here work-family life-course cluster. If this is the case, the groups within the same cluster have *similar* life courses but not the very *same* ones—these are understood as different empirical realizations of a similar underlying life-course type. However, as the main idea of the returns component in the KOB is to show how the returns to the *same* characteristics (e.g., tertiary education) differ by groups by holding the endowment with this characteristic constant, this interpretation cannot be applied to characteristics that are highly heterogeneously endowed. As a result, the differences in returns by group might be driven by the group-specific heterogeneity within the same life-course clusters.

When combining SA and KOB, the returns component will likely never show the impact of different returns to the *same* endowment but rather to a *similar* trajectory or the same life-course *cluster*. That is, the returns component does not necessarily distinctly show how the different groups with the exact *same* life course are rewarded differently as intended in the standard KOB decomposition but could also pick up ways that the groups are rewarded differently because of *differences* within the same life-course cluster. Researchers should keep this in mind when interpreting the returns component in SA-KOB applications and generally only use sufficiently coherent cluster typologies to ensure meaningful results in SA-KOB.

More generally, within-cluster heterogeneity underlines the complexity of life courses and highlights an additional dimension of inequality.³⁰ As such, some within-cluster differences might only become visible through SA and can be accommodated in the modeling and interpretation (Aisenbrey & Fasang 2010); for instance, the sum of the years in a specific state used previously in decomposition analyses might conceal differences in the volatility, timing, or order of these states between groups. Such group-specific heterogeneity is not necessarily specific to the SA-KOB combination but might apply to any variable used for KOB, albeit it may be less visible. For instance, a common assumption in KOB would be that women and men with tertiary education have, overall, the same endowment of education. However, heterogeneity within tertiary education is plausible due to differences in the educational trajectory that leads to a certain degree, different fields of study, or different reputations of the university among individuals, which are likely to influence wage and pension returns. By including more (life-course-related) variables simultaneously than in previous decomposition techniques (e.g., differentiating between full-time and part-time employment, civil service, and self-employment) and considering the timing and order of these along life courses, the SA-KOB combination can reduce heterogeneity within categories. At the same time, visualization techniques make heterogeneity explicit and can inform the interpretation of the results.

SA provides tools to deal with within-cluster heterogeneity by identifying individuals who are poorly allocated to the clusters based on low silhouette widths and excluding them from further regression-based analyses (Jalovaara & Fasang, 2020). Estimating the KOB when excluding poorly classified sequences in the cluster typology shows how sensitive the results are to (group-specific) within-cluster heterogeneity: see A3.5.3 for a discussion of

³⁰ Overall, the association of the within-cluster heterogeneity with the group gap in the outcome is captured by the differences in the intercepts between groups, alongside any additional unobserved heterogeneity. However, it is difficult to disentangle which part is due to within-cluster heterogeneity and which part is due to further unobserved heterogeneity.

this sensitivity check. We recommend assessing the impact of within-cluster heterogeneity on the KOB results. When using relative cut-offs for silhouette widths sensitive to the cluster-specific silhouette distribution, we found that the returns component in our illustrative application remained roughly the same, while the explained shares tended to be larger (see Table A3.11). This is in line with the generally stronger effect sizes for sequence clusters in regression-based analyses when poorly classified cases are excluded and only individuals above a certain silhouette threshold are retained to ensure relatively “pure” types in the life-course clusters (Jalovaara & Fasang, 2020).

3.3.2.3 Baseline categories of categorical variables and choice of controls.

The decomposition results for categorical variables, including life-course clusters, differ depending on the baseline category used (Fortin et al., 2011; Jann, 2008). We normalized the categorical variables to calculate shares of the GPG arising from the returns component that were independent of the choice of the reference category: The coefficients in the OLS models underlying the KOB indicate deviations from the grand mean and not from the chosen reference category (Jann 2008; Yun 2005). Normalization is helpful when combining decomposition tools with SA. Especially in empirical scenarios of highly group-specific life-course clusters, it is more difficult to find a theoretically meaningful and empirically unproblematic baseline category for all groups than in standard applications of KOB. While researchers have to choose the same baseline category for all groups for KOB, the standard life course, which would be a suitable baseline category, will likely differ across groups if life courses are group specific. Adding more dimensions, such as country comparisons, will add more difficulties, since standard life courses are also likely to differ across countries.

A related issue is which control variables to include in addition to life-course clusters. To avoid multicollinearity issues and over controlling, additional control variables should

generally be temporally located prior to the start of the life-course sequences, they should not be components of the life-course sequences, and they should be largely unrelated to them. For example, researchers are not advised to control for education when it is a sequence state, but variables such as parental education or region are unproblematic. Accordingly, we only adjusted the main models for birth cohort and survey wave but undertook robustness checks with more extensive control scenarios as presented in A3.5.4.

Further robustness checks

Besides systematically comparing alternatives with regard to 1) the choice of reference coefficients and dealing with 2) within-cluster heterogeneity, we conducted a series of further robustness checks. We showed that the main association between the GPG and the life-course clusters was not driven by: (i) differences in education or migration background between men and women (Appendix A3.5.4, Table A3.12) or (ii) the use of weights (Table A3.13). Furthermore, we discuss our approach to dealing with the common support problem that might occur when life-course clusters from the first step of the SA-KOB decomposition are highly group-specific (Appendix A3.5.5).

Software and packages

We performed SA in R using the packages *TraMineR*, *TraMineRextras* (Gabadinho et al., 2011), and *WeightedCluster* (Studer, 2013) and used the *oaxaca* package (Jann, 2008) for decomposition in STATA. However, both SA and KOB can be performed in R and STATA. In STATA, the packages for sequence analysis are *SQ* (Brzinsky-Fay et al., 2006) and *SADI* (Halpin, 2017), and in R, *oaxaca* (Hlavac, 2014) can be used for KOB decomposition.

3.4 Results of the illustrative application on GPGs in Germany and Italy

In this section, we summarize the results of our application to illustrate the potential of SA-KOB to provide new insights into long-standing questions in life-course research, social demography, and stratification. For this, we first replicate the standard KOB approach with our data and briefly discuss the results (Section 4.1). We then present the results of the MSA (Section 4.2) and of the SA-KOB decomposition (Section 4.3). In Section 4.4 we compare the results, the benefits, and the limitations of the standard KOB to the SA-KOB decomposition.

3.4.1 Standard KOB using life course summary measures

To highlight the theoretical, empirical, and methodological contribution of SA-KOB decomposition compared to the standard decomposition approach in the literature, we replicated KOB decomposition using summary indicators as proxies for family and work life courses (hereafter referred to as “standard model”). For comparability, we aimed to use the same life-course states as in the SA but could not include all of them due to high multicollinearity levels. As a result, we opted to include the life-course states that are used most often as main covariates in standard decomposition applications to estimate the GPG (see Table 3.1) and therefore excluded the duration of care work, retirement, self-employment, years spent in marriage and having one child, and years spent in single parenthood. Please refer to Appendix A3.5.2 for more details on the multicollinearity issue and on the results; the latter are only summarized briefly in the following paragraphs.

We decomposed the GPG, which amounts to 53% in Italy (€6,651 in absolute terms) and 62% in West Germany (€13,089 in absolute terms).³¹ The overall lower GPG in Italy is in line with previous research (Bettio et al., 2013; Hammerschmid & Rowold, 2019a; Tinios

³¹ The relative GPG is estimated as $\left(\frac{\bar{y}_M - \bar{y}_W}{\bar{y}_M}\right) * 100$.

et al., 2015) but might seem counterintuitive since the Italian public pension system is more regressive than the German one (see Table A3.2) and thus more likely to reproduce gender inequalities. The lower GPG in Italy could have several explanations, such as higher gender inequalities in work-family life courses in Germany (not confirmed in section 3.4.2) or the higher coverage of private pensions in Germany leading to a higher reproduction of gender inequalities. Most likely, however, is the combination of a more pronounced shadow economy in Italy (Schneider, 2012)³² and the higher minimum contribution period of 20 years needed to qualify for old-age pensions. As informal employment is not registered with the state, it cannot be considered for the old age pension contribution period. Thus, individuals who work for many years in informal jobs risk not reaching the 20 years of contributions and thus not qualifying for an old-age pension. Summary statistics support this assumption, as much higher proportions of men in Italy than in Germany have no pension income, regardless of the life course type.³³ For example, even if Italian men have reported continuous full-time employment, 10.9% of them do not receive pension income compared to only 1.15% in West Germany (see Table A3.7, Table A3.8). The lack of pension income of these men decreases the average pension income of all men, which mitigates the Gender Pensions Gap in Italy.

For the standard KOB decomposition, we applied the same parameters as in our main analysis to ensure comparability. Table 3.4 shows the decomposition results. Consistent with previous literature, we found that the lower number of years women spent on average in full-time work compared to men (see Figure A3.8) was the main reason for the GPG in both countries (explained share). Women's pension would increase by €1,368 in Italy (21%

³² Such activities are defined for the analysis as 'productive economic activities that would normally be included in the national accounts but which remain underground due to tax or regulatory burdens' (Schneider, 2012)

³³ An alternative explanation for the higher proportion of those not receiving a pension could be that Italian men retire later than German men. However, this is unlikely to be the case as the effective exit age from the labour market is on average earlier for Italian men (see Table A3.2).

of GPG) and €3,136 in West Germany (24% of the GPG) if they worked the same number of years on average as men in full-time employment. The "missing" years women spent working in the civil service are another driver of the GPG but to a much higher extent in West Germany than in Italy (15% vs 1% of the GPG). While the higher number of years spent on average in education by men after age 17 explains 7% of the GPG in West Germany, 5% of the GPG in Italy is due to the greater number of years on average women have spent in unemployment. In both countries, the gender inequality in pensions would be even higher if women spent as few years as men in part-time employment (1% of GPG in Italy, and 10% in West Germany). We find that family life characteristics matter as well but to a lesser extent. In both countries, the GPG would be even larger if women were childless and single for as many years as men (see Figure A3.9), and in Italy a part of the GPG is due to the greater number of years women spent married and with two or more children.

The returns component reveals that substantial parts of the gap in both countries are due to lower pension returns for women compared to men for years spent in education. In Italy, the GPG would be even higher if women received the same returns for full-time employment and unemployment as men. In West Germany, in contrast, the GPG would be even higher if women received the same pension returns for years spent being divorced as men.

In line with previous decomposition literature (Table 3.1), the present analysis's main conclusion based on the standard KOB is thus that women have much lower pensions than men due to their fewer years spent in employment.

Table 3.4 Results from standard KOB decompositions

	Italy				West Germany			
Mean pension inc. men	12,660**				21,119**			
	(64.22)				-58.62			
Mean pension inc. women	6,009**				8,030**			
	(28.42)				-26.82			
Absolute GPG	6,651**				13,089**			
	(21.27)				-26.33			
Explained, total	1,682**				4,499**			
	(6.781)				-6.457			
Unexplained, total	4,970**				8,590**			
	(12.09)				-12.55			
	explained	shares	returns	shares	explained	shares	returns	shares
Years FT employment	1,368**	20.56%	-706.7**	-10.63%	3,136**	23.96%	206.7	1.58%
	(6.491)		(-3.644)		-5.452		-0.431	
Years PT employment	-68.74**	-1.03%	-116.2**	-1.75%	-1,282**	-9.79%	-34.23	-0.26%
	(-2.964)		(-3.268)		(-6.136)		(-0.973)	
Years civil service	79.66*	1.20%	-311.1**	-4.68%	1,939**	14.82%	-134.9	-1.03%
	(2.224)		(-4.926)		-8.523		(-1.340)	
Years education	109.7	1.65%	629.9**	9.47%	940.9**	7.19%	916.5**	7.00%
	(1.141)		(3.931)		-3.169		-2.879	
Years unemployment	322.6**	4.85%	-362.1*	-5.44%	156.3	1.19%	-280.2	-2.14%
	(2.798)		(-2.450)		-0.843		(-1.066)	
Years divorced	3.471	0.05%	30.66	0.46%	27.32	0.21%	-514.3**	-3.93%
	(0.299)		(1.078)		-0.621		(-3.764)	
Years married, 2+ch.	145.3**	2.19%	-776.8	-11.68%	55.69	0.43%	517.6	3.95%
	(2.782)		(-1.324)		-0.51		-0.541	
Years married, no children	-1.068	-0.02%	-60.36	-0.91%	44.46	0.34%	87.09	0.67%
	(-0.0634)		(-0.423)		-1.226		-0.386	
Years single, no children	-254.3**	-3.82%	-856.5**	-12.88%	-558.2**	-4.26%	-829.2**	-6.33%
	(-3.080)		(-3.685)		(-3.278)		(-2.840)	
Constant			7,648**				8,369**	
			(6.353)				-4.163	
n male	1357		1357		804		804	
n female	1509		1509		815		815	

Notes: Based on linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. The coefficient from the women's regression is used as reference coefficient for the duration in part-time employment, and the ones from men's otherwise. Z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Shares do not add up to 100% since the controls are not depicted. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

3.4.2 Gendered work-family life courses

Figure 3.1 depicts the eight work-family life-course clusters, with the work-life courses on the left and the family life courses at the right-hand side. The outer portions of the Figure show the gender distribution over these clusters by country.

In line with expectations, the work-family life-course profiles in the clusters mirror the traditional distribution of paid and unpaid work between men and women as incentivized by the conservative and family-oriented institutional and normative contexts in both welfare states in the second half of the 20th century.

Cluster 1 reflects a prototypical male breadwinner life course of early continuous full-time employment in the private sector. Most individuals in this group entered marriage and parenthood between the ages of 20 and 30. It is the most prevalent life-course type (29%). 45% of Italian and West German men experienced this life-course type, compared to only 14% of women (see Table A3.9), pointing to notable similarities in the unequal gender distribution for this life-course type across countries.

Cluster 2 displays a similar prototypical male breadwinner life course with continuous careers as civil servants after longer periods of education and correspondingly a somewhat later onset of family formation compared to cluster 1. A greater proportion of people in this group have only one child. This pattern is not quite as gendered as the private sector male breadwinner life course, but it is still more exclusively experienced by men, particularly in Germany.

Cluster 3 shows employment lives that transition into self-employment, mostly from full-time private sector work by age 40 and then remain in stable self-employment. Family lives are characterized by having two or more children in stable relationships with a slightly later

onset on average compared to clusters 1 and 2. In both countries, this cluster is more common among men than women, especially in Italy.

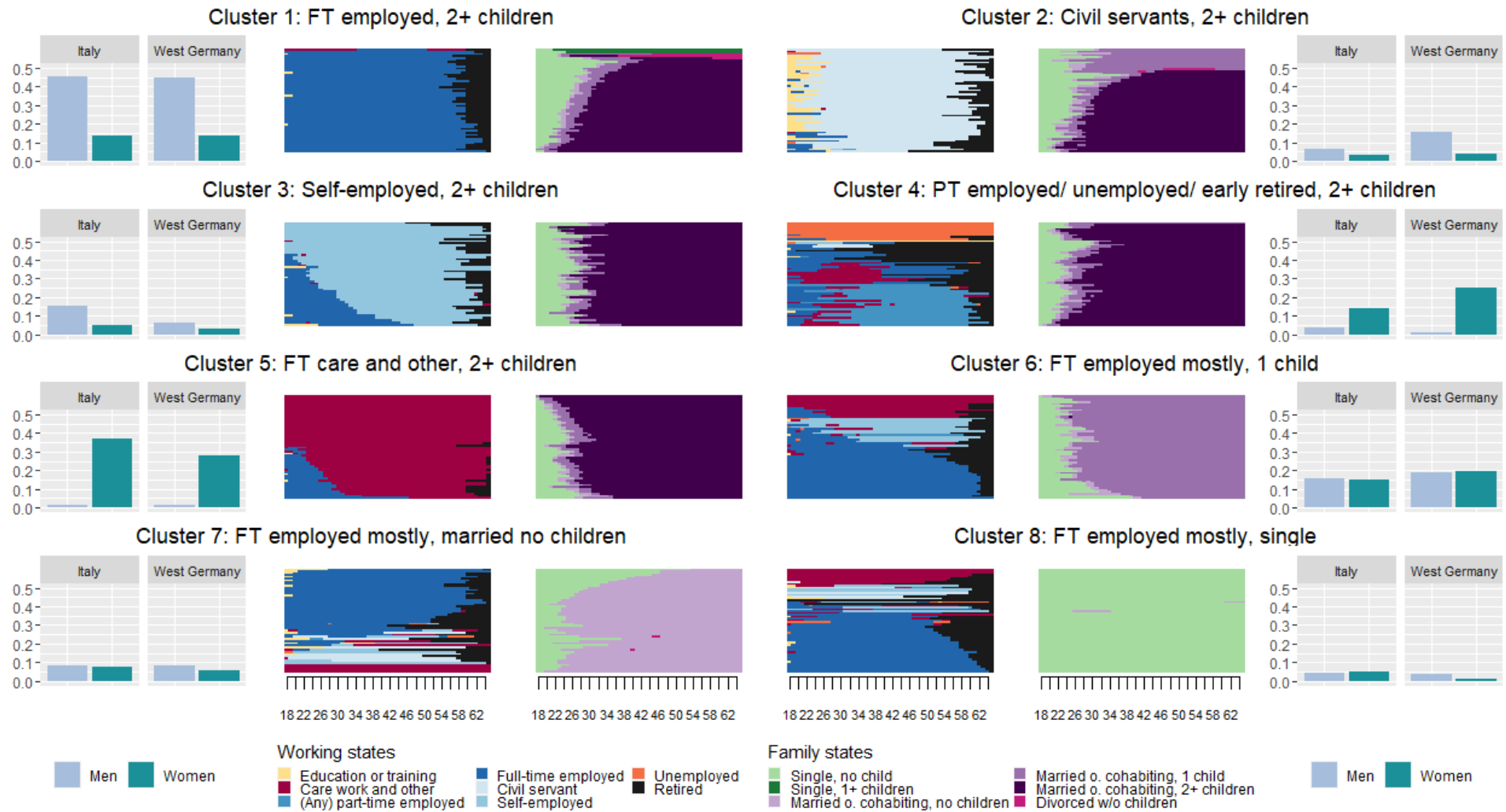
Together, clusters 1–3 thus distinctly signify typical male breadwinner life courses. This provides empirical support for a strong gender segregation in life courses in the gender-conservative environments, as also found by Madero-Cabib and Fasang (2016).

A tight linkage between work and family life only becomes visible for clusters 4 and 5: changes in working life (mostly from full-time employment to part-time or care work) often parallel family transitions. Both clusters are characterized by early family formation, being married, and having at least two children and are highly dominated by women. For the life courses in clusters 4 and 5, with transitions between employment and care work, the start of the full-time care period corresponds almost exactly to a shift from being single to being married or childbirth, highlighting the interdependence of women's family and working lives.

Cluster 4 is characterized by the absence of stable and paid full-time work. Most individuals experience volatile work trajectories, transitioning from full-time employment to longer periods of care work and finally to fairly continuous part-time employment, and to a lesser extent full-time employment, after mid-life. A considerable proportion of respondents have retired early after a short period of full-time employment or are continuously unemployed. Among West German women, 25% used (interrupted) part-time work to reconcile work and family life; the figure was only 14% in Italy.

Cluster 5 is characterized by care work throughout the whole life course or stable care work after a short period of full-time employment in early adulthood. Care-dominated life courses are most prevalent for women in both countries: they apply to 37% of women in Italy and 28% of women in Germany.

Figure 3.1 Work and family life courses: relative frequency sequence plots, distribution by gender and across countries



Notes: (i) Gender-specific distribution on clusters depicted as relative shares per gender; (ii) Clusters depicted as relative frequency sequence plot at the center (Fasang & Liao, 2014), with depicting representative life courses per cluster from age 18–65. The dissimilarity from the medoid for each plot in this Figure is displayed in Figure A3.2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Life-course clusters 6–8 show very similar work lives but different family life courses and are gender-neutral. All three are dominated by continuous full-time employment but contain a substantial share of individuals with continuous care work or self-employment as well. Cluster 6 comprises individuals who are married and have one child. Cluster 7 is characterized by childlessness and relatively late marriages and cluster 8 is characterized by lifelong singlehood.

The highly gender-specific clusters raise challenges for the KOB decomposition in the second step of our analysis, given that there are almost no men in clusters 4 and 5. This empirical finding is specific to our illustrative application and does not appear for SA-KOB more generally. We discuss this issue and a way of dealing with it in detail in A3.5.5. To reduce the potential impact of lacking common support for men in the female majority clusters 4 and 5, we identified male life courses that are similar to the characteristics of these clusters based on the fit with the cluster (ASW) and life course summary measures. We then reassigned these male trajectories to clusters 4 and 5 for the decomposition to create counterfactuals of men for these highly women-majority clusters manually.

3.4.3 SA-KOB decomposition of Gender Pension Gaps

We used the MSA cluster typology to decompose the GPG in Italy and West Germany.

For clusters 4 and 5, we used the coefficients from the regression model for women as reference coefficients in the KOB; for all other clusters we use the coefficients for men. First, based on empirical considerations, the share of women more than doubles that of men's for clusters 4 and 5 in both countries (relative difference higher than 50%, see Table A3.9). Comparing the cluster characteristics across groups confirms that women's characteristics are much more in line with the overall cluster characteristics than men's (see Table A3.7, Table A3.8 in the Appendix). Second, based on theoretical considerations, we prefer to use the women's coefficients as a reference for clusters 4 and 5, as these are more

likely to capture pension-related benefits that reward care work and part-time employment, which are more common for women in these clusters. Lastly, the group-specific ASW by cluster and gender validates using women's coefficients as references for clusters 4 and 5: women's life courses are on average a better fit to both clusters (see Figure A3.1).

For similar reasons, we use men's coefficients as references for clusters 1–3. The clusters are more common among men, and men's life courses in these groups mirror the cluster characteristics. Men's coefficients best capture the nondiscriminatory reward structure, given that most pension systems are “shaped around an idealized male worker“ (Grady, 2015, S. 454). This is also why we use men's coefficients for the gender-neutral clusters 6–8, which are characterized by full-time employment careers too. Moreover, men's life courses share the overall working life-course characteristics of these clusters (see Table A3.7, Table A3.8, and a detailed discussion in A3.5.1) and fit the cluster better in most cases (see Figure A3.1).

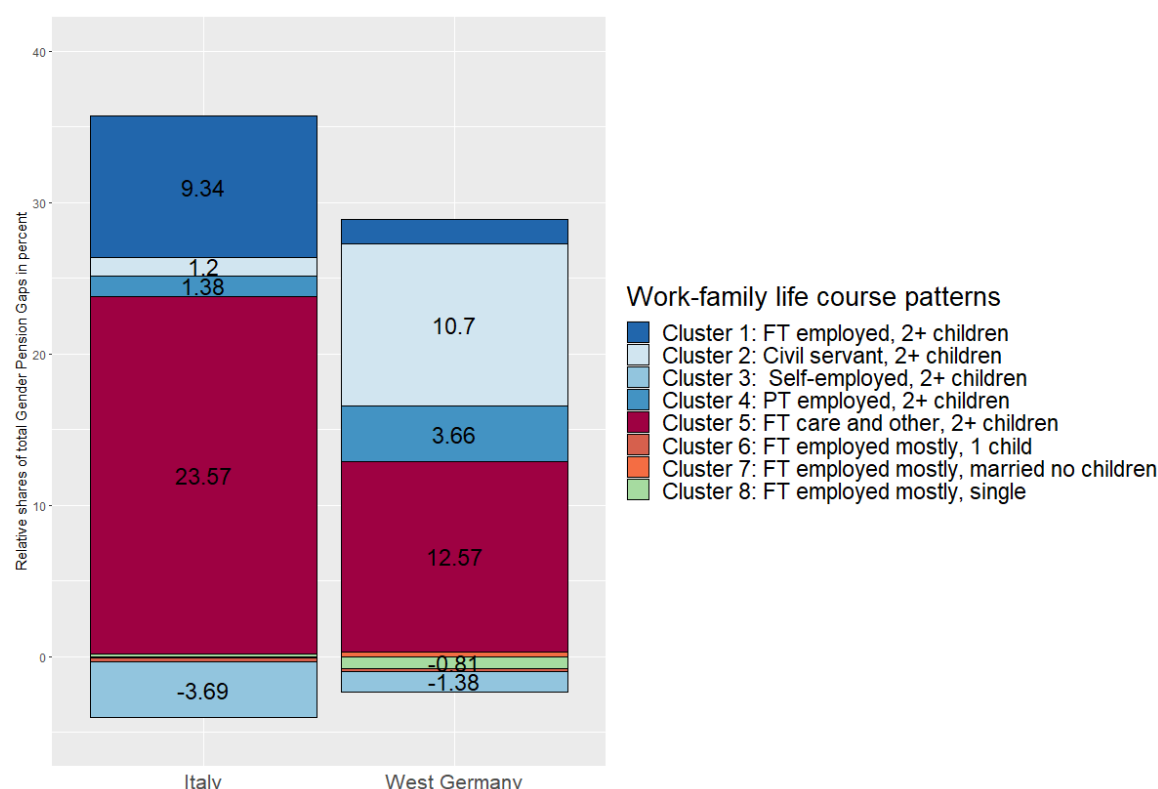
In line with our expectations of relatively large compositional life-course effects on GPGs in both countries, Table 3.5 shows that 32% of the GPG in Italy and 27% in Germany is explained by the gender-segregation of work-family life courses, that is, men and women experiencing different life courses.³⁴ Further, in line with expectations, both compositional and returns effects account for a larger share of GPGs in Italy than in Germany, albeit the country differences in the returns component are relatively small.

In both countries, the majority female life-course cluster characterized by continuous unpaid care work (cluster 5) is the main driver of the GPG. 24% (€1,568) of the gap in Italy and 13% (€1,646) of the gap in West Germany are due to the absence of almost any men with life courses that are poorly rewarded in the pension system (Table 3.5, Table 3.6). The

³⁴ Confidence intervals depicted in Figure A3.6 and Figure A3.7.

much higher minimum contribution years in Italy means more Italian women with this life course have no pension income (56% of women compared to 23% of men in the cluster, Table A3.7)—less than half of the women in this life course reach the required 20 years of paid work. In Germany, lower qualification criteria give women with employment periods before family formation access to independent pension income (14% of German women in this cluster are without pension income compared to 0% of men). While both countries provide benefits for an equivalent to a maximum of 1 year per child, women with these life courses are usually engaged in unpaid care work for much longer periods—on average 38 years and 42 years in West Germany and Italy, respectively.

Figure 3.2 Decomposition results – explained shares by work-family life courses and country



Notes: Percent of shares only displayed if minimum significance level of 10%. Results from KOB decompositions from Table 3.5. Based on normalized linear regression models with absolute independent pension income as dependent variable. Adjusted for birth cohort and wave. Confidence intervals depicted in Figure A3.6 and Figure A3.7 in the Supplementary Materials. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The gender segregation on cluster 4, which is characterized by volatile working life courses, is more strongly associated with the GPG in West Germany than in Italy. The overrepresentation of women in this cluster was responsible for 3.7% of the gap in West Germany and 1.4% of the gap in Italy (Table 3.5). Compared to continuous full-time care work (cluster 5), combining care responsibilities with paid part-time employment, periods of unemployment, or early retirement reduces the pension-income penalty for mothers in both countries (returns for cluster 4 and 5, Table 3.6). But, at the same time, it cannot make up for the lack of full-time employment of women (cluster 1). Thus, the overrepresentation of women in these two life-course clusters (4 and 5), which are characterized by strong interdependence between work and family lives, would not lead to the large GPG in both countries if their female-typical biographies were rewarded equally to male-typical ones.

Interestingly, the male majority life courses which drive the Gender Pension Gap differ most between both countries. While the lack of women in the standard life course of men (cluster 1) is only significantly associated with a higher GPG in Italy (9.3%, €621 per year), a large share of the West German gap is due to gender differences in having a civil servant career (10.7%, €1,401). On one hand, this highlights the particularly beneficial pension rights for German civil servants. On the other hand, the contrasting results across countries for cluster 1 results from the return for men to this life course being much lower in West Germany compared to Italy (Table 3.6). This leads to a much lower explained share due to gender differences in cluster 1. Second, this life-course cluster contains a relatively high share of divorced men in Germany (Table A3.8), which likely decreases the average return to this life course for men due to the splitting of pension rights after divorce (Kreyenfeld et al., 2023).

Table 3.5 Results from SA-KOB decompositions

	Italy				West Germany			
Mean pension inc. Men	12,660**				21,119**			
	-57.31				-53.62			
Mean pension inc. Women	6,009**				8,030**			
	-37.13				-28.52			
Absolute GPG	6,651**				13,089**			
	-24.29				-27.03			
Explained, total	2,054**				3,514**			
	-10.07				-8.725			
Unexplained, total	4,597**				9,575**			
	-14.52				-17.65			
	explained	shares	returns	shares	explained	shares	returns	shares
Sum life courses	2106.281	31.67%	386.66	5.81%	3476.08	26.56%	555.203	4.24%
1 FT employed, 2+ chi.	621.2**	9.34%	92.16	1.39%	214.9	1.64%	99.19	0.76%
	-4.88		-1.178		-1.04		-0.741	
2 Civil servant, 2+ chi.	79.72*	1.20%	-84.40*	-1.27%	1,401**	10.70%	2.913	0.02%
	-2.49		(-2.258)		-6.738		-0.0499	
3 Self-employed, 2+ chi.	-245.2**	-3.69%	-69.06+	-1.04%	-181.2*	-1.38%	-33.92	-0.26%
	(-3.929)		(-1.655)		(-2.499)		(-0.598)	
4 PT employed, 2+ chi.	92.08**	1.38%	-15.25	-0.23%	479.3**	3.66%	-58.72	-0.45%
	-3.069		(-0.247)		-3.784		(-0.791)	
5 FT care/other, 2+ chi.	1,568**	23.58%	58.67+	0.88%	1,646**	12.58%	88.07	0.67%
	-13.68		-1.778		-9.03		-1.625	
6 FT empl./care, 1 child	-19.3	-0.29%	477.6**	7.18%	-19.99	-0.15%	518.0*	3.96%
	(-0.633)		-4.48		(-0.735)		-2.394	
7 FT employed, married	-4.059	-0.06%	87.04	1.31%	41.67	0.32%	-2.24	-0.02%
	(-0.463)		-1.275		-1.134		(-0.0251)	
8 FT employed, single	13.84	0.21%	-160.1**	-2.41%	-105.6*	-0.81%	-58.09	-0.44%
	-0.936		(-2.677)		(-2.020)		(-1.574)	
n men	1357		1357		804		804	
n women	1509		1509		815		815	

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. Z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Shares do not add up to 100% since the controls and the intercept are not depicted. Full Table is shown in the Table A3.15. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table 3.6 Baseline coefficients from linear regression analysis used for SA-KOB, normalized

	Men	Italy Women	Pooled	Men	West Germany Women	Pooled
1 FT employed, 2+ children	2061.3	1392.7	1885.9	724.6	-10.36	682.0
2 Civil servant, 2+ children	2778.3	5377.5	3829.7	12454.1	12379.9	12597.4
3 Self-employed, 2+ children	-2473.1	-1083.6	-2112.1	-5532.7	-4426.9	-4968.2
4 PT employed, 2+ children	-1642.4	-1435.4	-1230.6	-3643.4	-2332.1	-1925.9
5 FT care and other, 2+ children	-2299.9	-4574.7	-3997.7	-2668.5	-6395.3	-5542.1
6 FT employed or care, 1 child	2289.7	-927.5	730.9	1160.0	-1495.2	34.79
7 Ft employed, married, no children	819.8	-353.0	396.9	2459.4	2499.1	2686.3
8 Ft employed, single	-1533.6	1604.1	496.9	-4953.5	-219.1	-3564.3

Notes: Linear regression models with absolute independent pension income as the dependent variable. All categorical variables normalized. Controlled for birth cohort and wave. Z-statistics in parentheses. ‘Pooled’ refers to a model using all men and women in the sample. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Gender pension gaps are largely driven by the under- or overrepresentation of women in life-course clusters that include parenthood (clusters 1–5). In line with recent research on the US (Fasang & Aisenbrey, 2022), this highlights the crucial role of the gendered interplay of work and family over the life course. Other scholars have highlighted the importance of parenthood for women’s financial security in old age (e.g., Crespi et al., 2015). But the multichannel SA-KOB decomposition uniquely detects and quantifies the extent to which the interdependence of work and family life courses evident for women, but not for men, drives the Gender Pension Gaps in different welfare state contexts.

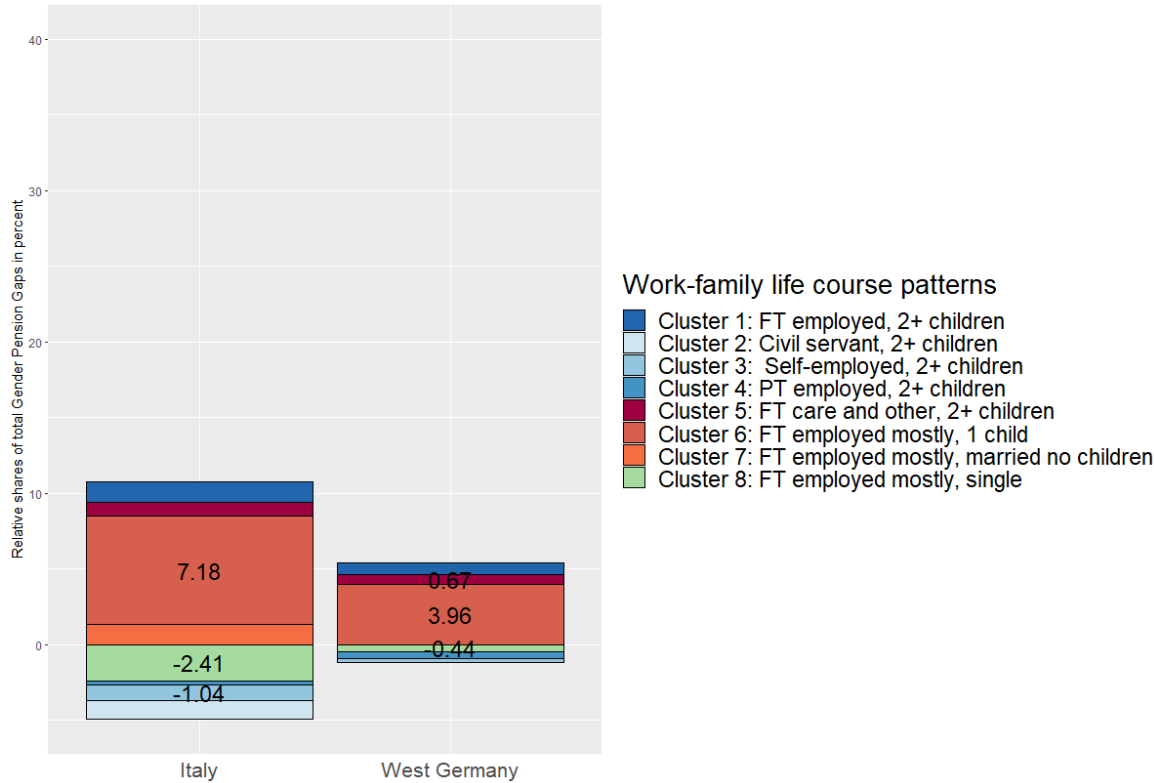
In sum, equalizing gendered work-family life courses, or the pension returns to the gender-specific life courses, would lead to an increase in pensions of €2,106 per year for women in Italy (€175 per month) and of €3,476 per year (€289 per month) in West Germany. To put these numbers in perspective for the Italian case, one could consider, for example, that eliminating 30% of the GPG (keeping men’s pensions constant) would lead to an increase in women’s pension of an amount corresponding roughly to one third of the absolute

poverty threshold (set at €700 per month for a single person aged 60-74, data from ISTAT 2021). This is substantively relevant from a social policy perspective, also given that women's old-age poverty risk exceeds men's by far in most European countries, including Germany and Italy (Haitz, 2015).

Turning to the returns component of the unexplained shares of the SA-KOB decomposition, Figure 3.3 displays the shares arising from differences in pension income returns for membership of the same life-course cluster. As expected, GPGs are much less related to differences in returns to the same work-family life-course clusters than compositional differences of men and women experiencing different life-course clusters.³⁵ The sum of different rewards for the same life-course clusters only accounts for 5.8% of the GPG in Italy and 4.2% in West Germany. In both countries, the main driver is the life-course cluster characterized by having one child in marriage coupled with full-time employment in the private sector (Italy: 7.2%, €478; West Germany: 4%, €518). Alternatively, such differences in returns might be a result of gender-sensitive redistribution in the pension system favoring men, a high degree of gender-specific within-cluster heterogeneity (suggested by our sensitivity analysis), or gender-specific labor-market outcomes that affect pension income and are not controlled for (e.g., earnings, occupational segregation).

³⁵ This is the case when focusing on the returns component of the unexplained shares due to gender differences in returns for the same clusters. However, the largest share of the GPG is due to gender differences in the intercept (see Table A3.15), the second component of the total unexplained share. As a result, the *total* unexplained share (i.e., the sum of both the intercept and returns component) is larger than the *total* explained share. However, given that the contribution of SA-KOB arises from quantifying how the different life-courses clusters are related to group inequalities, we do not discuss the total unexplained shares.

Figure 3.3 Decomposition results – shares from returns component of the unexplained part by work-family life courses and country



Notes: Percent of shares only displayed if minimum significance level of 10%. Results from KOB decompositions from Table 3.5. Based on normalized linear regression models with absolute independent pension income as dependent variable. Adjusted for birth cohort and wave. Confidence intervals depicted in Figure A3.6 and Figure A3.7 in the Supplementary Materials. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Given the higher compositional compared to returns components, our findings thus suggest that modifying pension regulations—without fundamentally changing the entire pension system—is much less effective than comprehensive gender equality and work-family reconciliation policies.

3.4.4 Comparison of SA-KOB with standard KOB using life course summary measures

Empirically comparing the results of the SA-KOB with the standard decomposition approach using one-dimensional summary indicators as proxies for family and work life courses (the standard model) allows us to identify the benefits and limitations of both modelling approaches. We compare the results and discuss the benefits and limitations in detail in Appendix A3.5.2 and Table 3.7 and summarize the main conclusions here.

Table 3.7 Comparing benefits and limitations of the standard and the SA-KOB decomposition

	Standard decomposition approach	SA-KOB decomposition
Benefits	- allows for considering more heterogeneity in the data when it uses continuous variables - reveals associations for specific life-course aspects that are concealed in holistic approach (e.g., differences in returns to educational years and divorce)	- prevents multicollinearity and can consider more life-course states simultaneously - captures association of life-course complexities with gap in outcome (e.g., synergy of care work and part-time) - can model interactions between different life courses domains (e.g., family and work life courses)
Limitations	- implies high risk of multicollinearity - cannot consider all relevant life-course stages and therefore risks concealing important associations (e.g., unpaid care work) - risks concealing interdependences between different life-course states (e.g., synergy of short full-time employment spells and long unpaid care work) - hardly captures interdependences between life-course domains (e.g., work and family) - uses selective sets of life course variables (straightforward and easily extractable, mostly duration in states) - focus on duration in states risks concealing associations with other life-course complexities (e.g., timing and ordering of life-course states)	- uses a holistic categorization that reduces variation to explore in the data - uses clusters that might conceal the role of specific and less dominant life-course aspects (within-cluster heterogeneity)

Both modeling approaches have their strengths and weaknesses. The standard approach reveals the association of a limited number of specific and theoretically preselected life-

course aspects with the GPG. The main limitation of SA-KOB is within-cluster heterogeneity, which might conceal the association of the GPG with specific life-course aspects within the clusters. This might particularly be the case for less prevalent life-course states that could either be concealed by more dominant life-course states within single clusters or could be distributed across different clusters (e.g., in our case, unemployment in cluster 4, duration of education in cluster 2, or duration of divorce in clusters 1, 2 and 6 in the SA-KOB).

However, the SA-KOB decomposition has important theoretical and empirical implications especially when the aim is to show how typical life-course patterns are rewarded (here: in pension systems) and consequently related to gaps in outcomes. First, SA-KOB allows us to include a multitude of different (categorical) life-course states and thus efficiently circumvents multicollinearity issues, whereas standard KOB is prone to multicollinearity and therefore has to use a limited number of life-course covariates. In our case, for example, we had to exclude the duration of unpaid care work due to multicollinearity. As a consequence, the conclusion of the standard approach would be, in line with previous literature, that the highest shares of the GPG can be explained by the lack of women in employment. However, the SA-KOB decomposition points out that this is most likely rather driven by mothers with very few years of full-time employment before taking up unpaid care work for a long period, if not for the rest of their lives. SA-KOB more accurately highlights the dynamics behind women's fewer years in full-time employment. The continuous unpaid care-work provided by these mothers upon family formation prevents them from having as many years in full-time employment as fathers and means they do not have higher pensions in old age due to low returns to care-dominated life courses.

Second, if, in the standard model, one cannot include all life course summary measures due to multicollinearity, the selection of specific variables to be included in the model

depends largely on the researcher's decision. This also poses the question of whether the life-course variables selected are indeed the most important predictors of the outcome of interest (here: mostly the duration in employment for the GPG, Table 3.1, see for this question Chapter 5).

Third, summary measures are likely to conceal how different life-course clusters unfold their associations with inequalities through the order in which only certain life-course states appear along individual trajectories allocated to a given cluster. For example, only SA-KOB shows that part-time employment typically follows family-related employment interruptions as displayed by cluster 4. Therefore, SA-KOB demonstrates that, when considered in this combination with family-related employment interruptions, part-time employment is counterproductive for women's pension income whereas the standard approach suggests that more years in part-time employment benefit women's pension incomes.

Finally, the standard approach does not show the role of the relationship over time between different life-course domains. Only the SA-KOB decomposition (using MSA) reveals that the GPG is almost exclusively driven by the life courses of parents of at least two children as well as the importance of the intersections between work and family life courses that only occur for women.

We conclude that SA-KOB is particularly beneficial if researchers want to understand inequality-generating mechanisms in greater complexity than can be offered by simplified and unidimensional proxies for life courses due to SA-KOB's capacity to consider more life-course states simultaneously and capture (multidimensional) life-course complexities. This becomes more important the more complex and interdependent the life-course domains of interest are and the more this varies between the groups of interest (in our illustration women and men).

3.4.5 Extension: Single-channel or multichannel sequence analysis

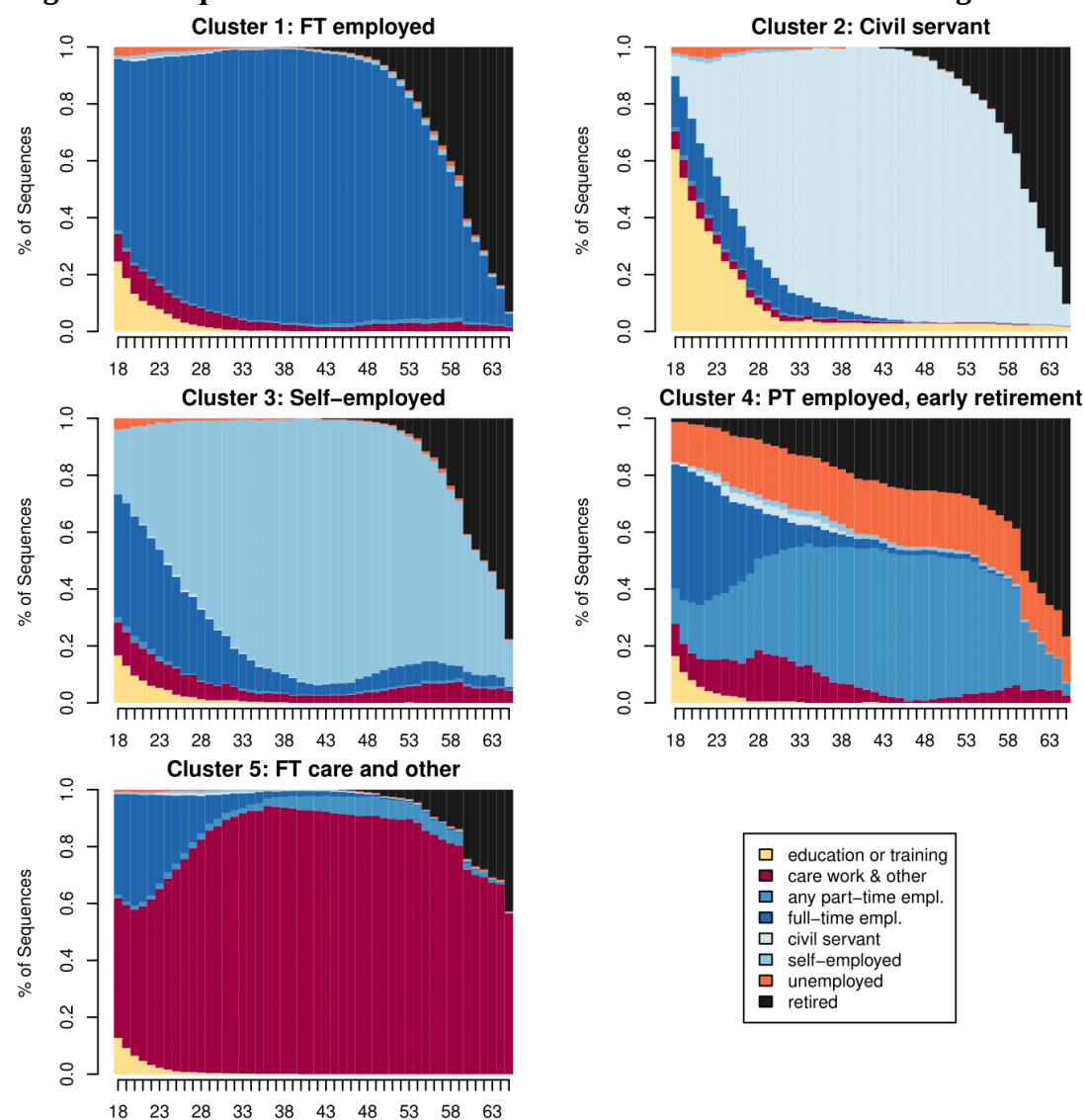
We further estimate single-channel sequence analyses based on the work and family domains separately. First, this is to empirically test the substantive added value and robustness of multichannel compared to a single-channel SA for the understanding of GPGs. Second, we demonstrate the possibility of combining single-channel SA and KOB by applying the three options introduced in section 3.3.1.2 and elaborate on how to decide between single-channel or multichannel SA. Third, using single-channel SA provides additional substantial insight into how work and family life courses are separately related to GPGs.

Given the previous literature and the reliance of most pension systems on employment outcomes (see above), it might be most straightforward to consider the employment life courses only (i.e., just one life course dimension, option 1). However, a proper comparison of single-channel and multichannel SA would involve including the clusters of both single-channel SAs for the work *and* the family domain in the KOB (option 2) and adding interactions between them (option 3). Thus, we apply a stepwise modelling approach (Table 3.8), starting by first only decomposing the GPG based on typical work life-course clusters and typical family life-course clusters in two separate models. As second step we then include the life course clusters of both domains jointly in the KOB decomposition. We estimate a single-channel SA for work and family life courses separately using the same main parameters as for the MSA in the main model (OM, substitution costs based on transition rates, indel costs as $\max(sm)/2$, and Ward clustering).

Work life courses only (option 1a). For the work domain, we selected the five-cluster solution, which is, among others, suggested by the cluster cut-off criteria and reveals similar typical work life-course clusters as the first five clusters in our main analysis (see Figure 3.4). Differences compared to the cluster solution using MSA (see Figure A3.3) are that the

shares of unpaid care work are slightly larger in cluster 1 (characterized by full-time employment) and in cluster 3 (characterized by self-employment) while the share of care work is smaller in cluster 4 (characterized by atypical work life courses such as part-time employment, unemployment, and early retirement). Given the focus on work life courses only, the intertemporal interdependences between work and family life courses, particularly for clusters 4 and 5, are not visible.

Figure 3.4 Sequence distribution for work life-course clusters from single-channel SA

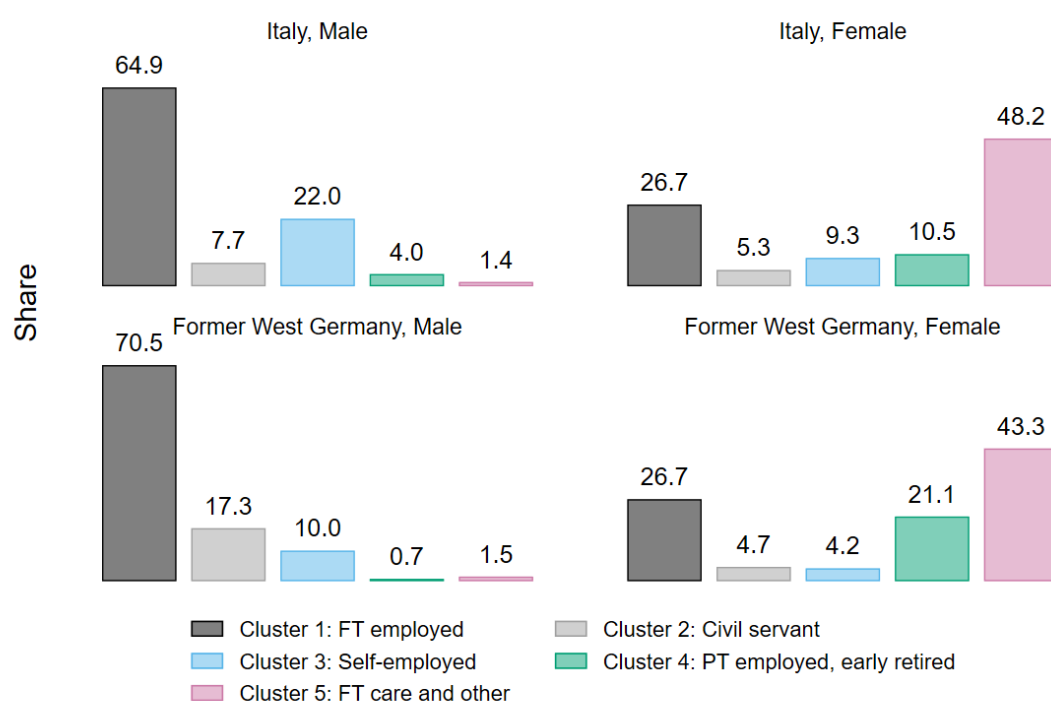


Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The gender-specific distribution over the different clusters (Figure 3.5) is more pronounced for all clusters except for cluster 4 for women in both countries and men in West Germany

compared to the MSA cluster outcome used for the main model (Figure 3.1, Table A3.9). Because family life courses do not play a role in sorting individuals in this cluster solution, individuals with similar work life courses are grouped together independently of their family lives, which increases the shares in most of the clusters. This reveals, that in total 48% of Italian and 43% of West German women in these cohorts have performed unpaid care work almost continuously over their life course.

Figure 3.5 Gender-specific distribution over work life-course clusters (single-channel SA) by country



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Focusing on the work life courses increases the explained shares of life-course cluster 5 (characterized by mostly continuous care work) in both countries compared to the main model (Table 3.8, Table 3.5). This is because the gender difference in occurrence of this life-course increases by including women who spent most of their lives doing unpaid care work while having two children or less³⁶ (by around 10 percentage points in Italy and 15 in

³⁶ These were women allocated to life-course clusters 6–8 in the main MSA cluster solution.

West Germany) and the poor returns for these life courses remain similarly low (not shown, available from the authors). The explained share for cluster 1 (continuous full-time employment) increases by about one quarter in Italy but diminishes from a substantial but statistically nonsignificant level to almost zero in West Germany. While in Italy, the more pronounced gender difference in experiencing these life courses is weighted by a stronger return than in the main model, in West Germany, it seems to be countered and weighted by a much smaller return that is close to zero (not shown, available from the authors). This might be driven by much lower pension income returns of single men for continuous full-time employment as revealed in the main model (cluster 8 of the MSA, Table 3.6) who are assigned to cluster 1 in the single-channel SA. By failing to account for this selectivity and thus increasing the heterogeneity of family life courses by ignoring these when clustering the work life courses, the analysis thus obscures the (insignificant) tendency that parts of the GPG in West Germany are due to mothers' limited access to continuous full-time employment.

Solely considering working life-course trajectories increases some explained shares by reinforcing the gender differences in some life-course clusters; it diminishes or even obscures other associations while not really adding to our understanding of how GPGs are related to life-course complexities compared to the main model. Only the main model shows, among other things, that the highly gendered life courses arise mostly for married couples and particularly parents. Additionally, focusing exclusively on work life courses conceals the policy-relevant interdependence of work and family life courses only occurring for women-majority life-course clusters. This emphasizes that much of the GPG is shaped by the long-lasting effect of past welfare policies in interaction with the current pension systems.

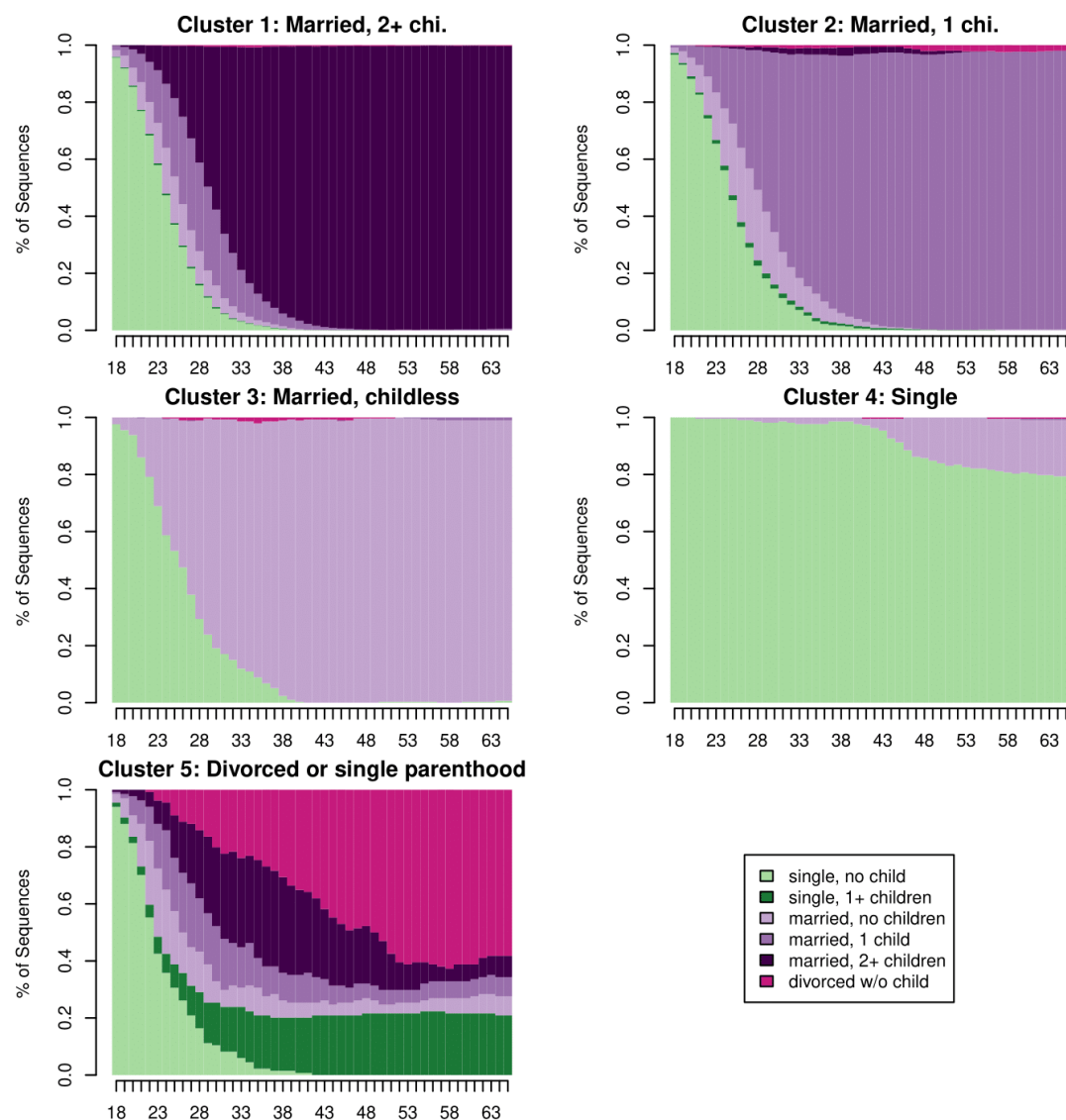
Table 3.8 Single-channel SA.

	Italy						West Germany					
	Work SA		Family SA		Work & family SA		Work SA		Family SA		Work & family SA	
Mean pension inc. men	12,660**		12,660**		12,660**		21,119**		21,119**		21,119**	
	(61.33)		(55.59)		(50.89)		(50.44)		(60.20)		(62.36)	
Mean pension inc. women	6,009**		6,009**		6,009**		8,030**		8,030**		8,030**	
	(43.93)		(37.63)		(36.48)		(28.62)		(31.78)		(27.48)	
Absolute GPG	6,651**		6,651**		6,651**		13,089**		13,089**		13,089**	
	(24.40)		(23.72)		(21.79)		(26.15)		(28.60)		(40.69)	
Explained, total	4,035**		-26.26		4,152**		9,294**		-89.62		9,900**	
	(8.434)		(-0.464)		(10.30)		(11.40)		(-0.528)		(11.36)	
Unexplained, total	2,616**		6,678**		2,500**		3,795**		13,179**		3,189**	
	(7.562)		(23.19)		(8.688)		(5.283)		(26.19)		(3.991)	
	explained	returns			explained	returns	explained	returns			explained	returns
Work: Cluster 1: FT employed	895.0**	45.46			843.8**	38.80	19.97	44.00			-153.2	105.9
	(3.925)	(0.251)			(4.069)	(0.231)	(0.0406)	(0.140)			(-0.331)	(0.330)
Work: Cluster 2: Civil service	74.72+	-98.47			70.35+	-102.1+	1,394**	16.76			1,317**	-22.11
	(1.871)	(-1.364)			(1.823)	(-1.784)	(6.832)	(0.185)			(5.462)	(-0.233)
Work: Cluster 3: Self-employed	-207.3*	-36.77			-220.8**	-41.06	-358.2**	-131.2+			-364.4**	-127.4+
	(-2.250)	(-0.561)			(-3.060)	(-0.599)	(-3.287)	(-1.724)			(-3.472)	(-1.848)
Work: Cluster 4: PT employed, early retired	143.3**	-4.124			132.8**	-6.384	398.0**	-4.365			362.6**	-4.541
	(5.344)	(-0.0803)			(4.162)	(-0.153)	(3.045)	(-0.177)			(2.742)	(-0.212)
Work: Cluster 5: FT care or other	1,759**	30.62			1,701**	33.39	2,322**	47.86			2,144**	55.85+
	(14.98)	(0.931)			(13.24)	(1.385)	(11.19)	(1.207)			(9.602)	(1.670)
Family: Cluster 1: Married, 2+ chi.			1.026	894.7*	0.912	55.45			-121.3	4,048**	-79.99	2,693**
			(0.0865)	(2.447)	(0.104)	(0.158)			(-1.318)	(5.426)	(-1.417)	(4.002)
Family: Cluster 2: Married, 1 chi.			5.075	209.8*	4.623	90.25			26.04	931.8**	22.66	682.8**
			(0.327)	(1.980)	(0.348)	(1.065)			(0.383)	(4.469)	(0.452)	(2.941)
Family: Cluster 3: Married, childless			-1.993	11.83	-2.574	0.217			27.57	55.15	20.96	51.29
			(-0.270)	(0.180)	(-0.219)	(0.00339)			(0.733)	(0.664)	(0.685)	(0.539)
Family: Cluster 4: Single			17.86	-279.2**	15.60	-189.6**			-115.7*	-74.88	-98.22*	-38.73
			(0.989)	(-5.086)	(0.905)	(-3.202)			(-2.025)	(-1.525)	(-2.155)	(-1.118)
Family: Cluster 5: Divorced or single parent			1.144	32.43	1.082	43.28			29.13	-384.5**	20.50	-326.6**
			(0.0830)	(1.439)	(0.0822)	(1.514)			(0.561)	(-3.569)	(0.701)	(-3.969)
Constant		4,205**		6,144**		4,381**		9,195**		8,202**		6,703**
		(6.154)		(12.94)		(6.574)		(7.339)		(7.518)		(4.720)
n male	1357	1357	1357	1357	1357	1357	804	804	804	804	804	804
n female	1509	1509	1509	1509	1509	1509	815	815	815	815	815	815

Notes: All models based on normalized linear regression models with absolute independent pension income as the dependent variable and controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Family life courses only (option 1b). Second, we run a single-channel SA for family life and use the retrieved clusters in a KOB model. The cluster-cutoff criteria clearly suggested a five-cluster solution. As in the MSA clusters, we extracted a family life course characterized by being married and having two or more children, being married and having one child, being married and childless, and being single (or marrying very late in the life course, Figure 3.6). Additionally, the single-channel SA identified a more “atypical” family life-course cluster characterized by high shares of divorce (up to 60%) and a substantial share of single parenthood (up to 20%). We did not find such a group of family life courses in our MSA cluster solution.

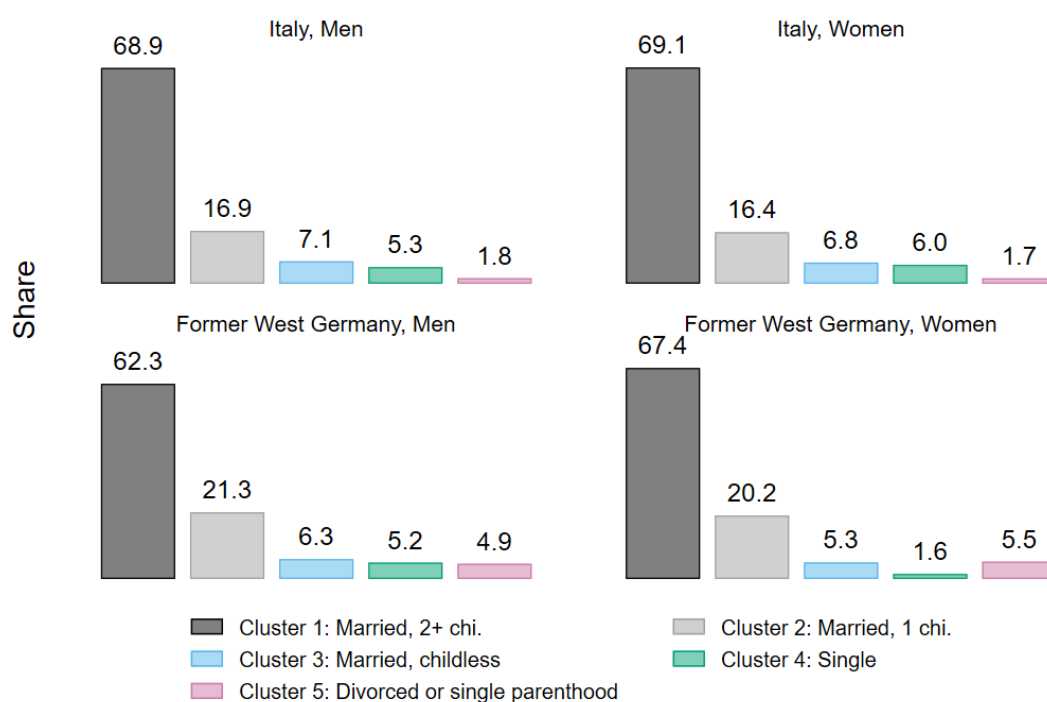
Figure 3.6 Sequence distribution for family life-course clusters from single-channel SA



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Using the family life course clusters as covariates in the KOB reveals that no part of the GPG is attributable to compositional differences in family life course clusters across genders (see Table 3.8, second column per country). This is because there are almost no gender-specific compositional differences in family life courses (see Figure 3.7). However, in both countries, substantial shares of the GPG arise due to gender differences in returns for family life courses characterized by marriage and having children (clusters 1 and 2). Mothers with such life courses receive much lower pensions compared to fathers. The contribution of this mechanism to the GPGs is larger for life courses with two or more children compared to those with only one child.

Figure 3.7 Gender-specific distribution over family life-course clusters (single-channel SA) by country



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

In Italy there is also a considerable share of the gap decreased due to better returns of women compared to men in the singlehood cluster (cluster 4) and in Germany due to the divorce cluster (cluster 5). The divorce benefit in Germany is probably driven by pension claim splitting among couples after divorce, which benefits women (Kreyenfeld et al., 2023). These results add to our findings from the main model using MSA.

Work and family life courses (option 2). Option 2 considers the single-channel cluster solutions from both life-course dimensions simultaneously in the KOB, i.e., with both cluster solutions controlling for each other (Table 3.8, third column).

The main association of the work life courses with the GPG remains very stable in both countries when controlling for the family life-course clusters. Additionally, we find for Germany that even when controlling for typical work life courses, mothers of two or more children receive €2693 lower pensions than fathers (returns component, family cluster 1; €683 pension income difference for parents with one child, cluster 2). The returns components are, however, lower compared to the model, where we do not control for work life courses. Still, jointly the first two family life course clusters make up more of the GPG than the compositional components due to the two women-majority work life course clusters (4 and 5). Moreover, women in West Germany with more atypical life courses (particularly being divorced, family cluster 5) receive a higher pension income than their male peers with similar life courses (the GPG decreases by €327) – even when controlling for work life courses. In Italy, the pension benefit of single women compared to men remains substantial as well, which points to a gendered selection into singlehood with regard to characteristics not included in this analysis, e.g., earnings (Jalovaara & Fasang, 2020).

The results thus highlight the role of family life course characteristics beyond work life courses that previous literature has pointed out (Fasang et al., 2013) and confirms that it unfolds its association with the GPG primarily through differences in returns between men and women (Bardasi & Jenkins, 2010; Ezeyi & Vujic, 2017). However, large parts are mediated by work life course patterns, suggesting that the association largely works through differences in work life courses of men and women with similar family life courses. Overall the results thus indicate different life course opportunities that systematically arise for men and women with similar family lives. Consequently, family life courses, or aspects thereof, should be considered when

analysing gendered pension income inequalities. Future research could disentangle the interrelation of different work and family life course aspects for the GPG in more detail.

Interactions between work and family life courses (option 3). Option 2 does not reveal potential interdependences between work and family life courses. Thus, we aimed to include interactions between all work and family life-course clusters to approximate a replication of the MSA model with single-channel clusters (option 3). However, this was not feasible with our data, as the inclusion of many interactions leads to very small cells that are not substantively or statistically meaningful.

Single-channel or multichannel SA? Whether single-channel or multichannel SA provide a better fit depends on the theoretical foundations, the substantive focus, and the data structure of each study. For our illustrative case, this extension confirms the importance of using MSA to identify the gendered interdependences of work and family lives for understanding the complex mechanisms underlying the formation of GPGs that theory has put forward (see Section 2). Only MSA reveals the clear interdependences between work and family life courses that appear almost exclusively for women and are related to GPG. Including both family and work life-course clusters in the model (option 2) is meaningful if the research aims to reveal the association between the life-course cluster in one domain (e.g., work) with a gap in outcomes while controlling for the life-course cluster of another domain (e.g., family). For example, we find that the associations between the work life-course clusters and the GPG are stable for controlling for family life courses.³⁷

Including interactions between clusters of both domains might easily overstretch the model and make the interpretation very difficult (option 3). The extent to which this is the case depends on the data structure, sample, and number of clusters. If interactions between all life-course clusters of different domains are not feasible due to data restrictions and there are clear

³⁷ We do not observe multicollinearity in this model.

theoretical expectations, theoretically driven interactions between a selection of single-channel clusters might be a feasible alternative. In our case, we cannot reconstruct the work-family interdependences we find in the MSA with interactions between single-channel SA clusters for work and family, most likely because it strains our model due to the low observations in some of the clusters. While a combination of single-channel SA with KOB is technically feasible, in our illustration on GPGs, MSA is clearly more pertinent from the substantive point of view.

3.5 Discussion and conclusions

In this study, we propose a novel combination of SA and KOB to offer a life-course sensitive analysis of group inequalities in outcomes that for the first time allows researchers to distinguish composition and returns effects. To illustrate the potential of this approach and derive a set of practical guidelines, we have presented an illustrative application on work-family life courses to Gender Pension Gaps in Europe, a field that requires further social-policy research in view of population aging, changing family demographics and persistent gender inequalities in old age (Kuitto et al., 2021).

Our contribution is twofold: methodological and substantive. The methodological contribution consists of the introduction of the stepwise SA-KOB decomposition. This is accompanied by detailed guidelines for making core decisions when applying the two analytical steps. These practical guidelines should facilitate applications of SA-KOB decomposition in other fields of research and inform some issues that apply to either method separately as well as to SA-KOB specifically. For example, we provide a systematic comparison of the most popular options for choosing the reference coefficients for KOB regarding the fit with SA-KOB and propose an option that is particularly well-suited to SA-KOB decomposition, that is, choosing the reference coefficients for clusters separately. We outline practical guidelines with regard to cluster-specific reference coefficients that are partly relevant for standard KOB applications as well. An additional benefit of SA-KOB decomposition is that researchers can use a set of established tools in SA for quantifying and visualizing within-cluster

heterogeneity—an issue that applies to covariates in standard SA-KOB applications but has not received much attention so far. We suggest that these tools should guide the interpretation of the KOB estimates. For example, our application suggests that the explained shares tended to be underestimated, because the pension rewards of life-course clusters were obscured by life courses that were poorly assigned to specific clusters. Carefully considering within-cluster heterogeneity in SA applications is important more generally—poorly classified trajectories risk biasing the effect sizes of associations with sequence clusters, or, in severe cases, concealing associations entirely. Further, SA-KOB decomposition provides research on inequalities in life-course sensitive outcomes with a tool to bypass multicollinearity while still considering multiple life courses’ characteristics in decomposition analyses.

The explained part of the decomposition quantifies the role of the group-specific life courses for group inequalities in the outcome of interest. The returns component calls attention to group-specific inequalities in outcomes despite having the same trajectory pattern. We argue that SA-KOB decomposition adds to recent methodological advances in the field of sequence analysis that combine SA with other methods to improve the capacity of SA to address a wider range of research questions (see Liao et al. 2022 and Raab and Struffolino 2022 for a review).

Our second contribution is a substantive one, as we provide new insights on how gendered work-family life courses are associated with GPGs across welfare states. As we show by systematically comparing SA-KOB and the standard KOB approach (using life-course summary indicators), these substantive new insights emerge from the ability of SA-KOB to overcome the methodological flaws of previous studies that adopt the standard KOB approach. SA-KOB combines a multiverse of work and family life courses and at the same time bypasses the multicollinearity issues that arise when including several summary variables in the model to account for different life-course characteristics. By relying on such simplifying summary measures, the standard approach risks concealing some important inequality-generating dynamics that unfold across time.

SA-KOB enabled us to distinguish the part of the GPG that is due to gender differences in typical life-course clusters (the explained or compositional part) and the part that arises due to gender differences in pension returns for the same life-course clusters (the returns component of the unexplained part). Findings from the first step, the multichannel sequence analysis, consistently show strongly gendered life-course clusters in line with the conservative welfare states of Germany and Italy. While fathers can reconcile parenthood with continuous employment, mothers either experience volatile employment life courses dominated by part-time jobs or fully dedicate their lives to lifelong unpaid care work in our study cohorts. Strong relationships between work life and parenthood only exist for life courses that are experienced mostly by women, and the gender differences in these life courses drive the GPG. The SA-KOB decomposition shows that GPGs in both countries are much more strongly associated with gender segregation in life-course clusters that are differently rewarded by the pension system (compositional component) than with different pension rewards between men and women for the same work-family life-course clusters (returns component). In both countries, the lack of full engagement in lifelong care work by fathers is the main driver of the gap. Thus, the high GPGs in Italy and West Germany are produced by the interrelation of different welfare policies that first shape the high level of gender differences in life courses and then reward these life courses unequally, to the detriment of women. Particularly the compositional component of the SA-KOB is highly linked with the theoretical life course literature since it somewhat quantifies the cumulative disadvantage over the life course of one group over another as well as “[w]hat is valued, protected, or rewarded over the life course” (O’Rand 1996: 230).

Thus, from a social policy perspective, SA-KOB decomposition informs intertemporal linkages between policy fields that first shape gendered life courses, and then decades later, pension rewards related to these life courses. The difference between compositional and returns effects cannot be interpreted in causal terms, but it can inform social policy, which generally has to rely on descriptive evidence concerning questions that do not lend themselves to experimental

designs or methods for causal inference. One of these questions concerns how the timing, sequencing, and occurrence of events that unfold over several decades are associated with group inequalities in later life. The elements of the pension systems put in place to address gender inequalities (e.g., care benefits) are apparently too insignificant to lead to a reduction of the GPGs by offering better rewards for life courses characterized by care work periods (returns component). We follow previous findings in pointing out the insufficiency of care benefits in current pension systems (Abatemarco & Russolillo, 2022; Frommert & Strauß, 2013; Lis & Bonthuis, 2019)—at least for the highly gendered contexts of Italy and West Germany for our study cohorts. Given that compositional life-course differences are the main driver of the GPG in Italy and Germany (explained part), policies that equalize access to different life-course types and reduce the strong interrelation between work and family lives for women are the most promising way to reduce the GPG. Such policies generally include defamilizing policies, such as public childcare, well-paid parental leave of about 1 year, and the abolition of joint couple's taxation (e.g., Adema et al., 2020; Ferragina, 2019; Ferragina & Seeleib-Kaiser, 2015; Leitner, 2001; Zagel & Lohmann, 2020). Thus, policymakers in charge of retirement policies need to also consider policies affecting prime-age adults to ensure that women and men enjoy equal and fair financial security in old age.

Our illustrative application is subject to data constraints as, for example, it lacks annual information on other categories that are likely to be relevant for pension income, such as earnings, industries, and occupations. Additionally, in certain clusters, the number of gender-specific observations is notably low. This can compromise the robustness of the coefficients which poses challenges for interpreting the returns components. Some possible extensions of SA-KOB exist, especially with regard to exploring different decomposition techniques in combination with SA. First, our analysis focuses on group differences at the mean. However, mechanisms shaping inequalities are likely to differ along the distribution. This is also the case for GPGs (e.g., Hänisch & Klos, 2014; Nolan et al., 2019). A combination of SA with

decompositions along the distribution (for an overview: Fortin et al., 2011) was recently proposed by Gruijters, Van Winkle, and Fasang (2023). Second, applying Ñopo's decomposition to address the common support problem by using matching procedures might reveal further inequality dimensions like social closure for group differences (Ñopo, 2008; Sprengholz & Hamjediers, 2022). Finally, SA-KOB decomposition does not uncover causal effects as the KOB by design is not causal (Yu & Elwert, 2023). However, recent methodological advances in both the decomposition (Jackson & VanderWeele, 2018; Lundberg, 2022; Yu & Elwert, 2023) and the SA literature (Barban et al., 2020; Lee et al., 2017; for an overview: Raab & Struffolino, 2022, Chapter 7.4) have advanced both approaches in a causal manner. Future research should explore whether these approaches can be combined in a more causal SA decomposition. Given the caveats on causal interpretation, skepticism on whether any policy conclusion can be drawn based on standard KOB decompositions is merited (Yu & Elwert, 2023). However, descriptive evidence on the complex mechanisms generating inequalities is key information that should not be discarded when designing policies, especially concerning research questions that are not amenable to research designs for causal inference from observational data for practical or ethical reasons. SA-KOB provides such policy-relevant findings by showing, for example, how much redistribution current pension regulations would have to undertake to compensate for pension inequalities resulting from prior gendered life courses.

To conclude, the proposed SA-KOB decomposition provides new insights into the extent of inequality in outcomes between groups associated with previous trajectories. As such, it holds potential for applications beyond later-life inequalities, for example, considering outcomes at early life stages and their association with prior educational trajectories or sequences of family situations in which children grow up. Decomposing the gender, race, or educational wage, wealth, or health gap based on previous (family, work, or health) trajectories are other natural applications.

Author's note

The code for running and replicating the analysis is available online: <https://osf.io/ke3mg/>. The SHARE data is easily available upon registration. The SHARE-RV data (only used to conduct robustness checks and confirm the imputation strategy for the retrospective data) is available upon registration as well but is not necessary to replicate the results.

Acknowledgments

We want to thank Anna Hammerschmid for her valuable contributions to previous versions of this Chapter. We are very grateful for the collaboration, discussions, and her engagement in this project. A special thanks to Maik Hamjediers and Maximilian Sprengholz for their valuable feedback, of which the manuscript benefitted greatly. We are further grateful for helpful comments from Alicia Garcia Sierra and Richard Breen as well as the social demography reading group participants at the Department of Sociology and Nuffield College, two anonymous reviewers and the editors of SMR. All remaining errors are ours.

4 Differences in Gender Pension Gaps in public and private pensions in West Germany: What role do work-family life courses play?

Abstract

Even though Gender Pension Gaps (GPG) reach much higher levels than gender inequalities in wages in most European countries, we know much less about how gendered pension income inequalities emerge. This paper contributes to the literature by taking a holistic perspective and comparing the gender gaps in public and private pensions. First, decomposing GPGs based on work-family life courses including the relative earning positions of individuals over the life course in West Germany is an important contribution due to the constantly high Gender Wage Gap. Secondly, differentiating between types of pension income sheds light on the impact of pension privatization on the GPG and provides more targeted suggestions for policymakers. Applying the combination of Multichannel sequence analysis with the Kitagawa-Blinder-Oaxaca decomposition to rich linked survey-register data for Germany improves previous decomposition analyses by considering the interdependences of work, including earnings, and family life course complexities.

The results reveal that large shares of the GPGs are attributable to a) many more fathers than mothers having access to privileged life courses that obtain high pension rewards while b) life courses that are mostly realised by women and dominated by unpaid care work lead to very low pension income. For both, public and private pensions, earnings play a role beyond full-time employment, but to a much higher extent for the gap in private compared

to the gap in public pensions. Given the high levels of the Gender Wage Gap for current cohorts in the labour force as well as the pension privatization in Germany, the results highlight the risk of high GPGs for future retirement cohorts.

Key words: Gender Pension Gap, life course, pension privatisation, Gender Wage Gap, Sequence Analysis, Kitagawa-Oaxaca-Blinder decomposition

4.1 Introduction

Across Europe, women receive on average substantially less pension income than men (e.g., Hammerschmid & Rowold, 2019a; Tinios et al., 2015). In light of the ageing population, these Gender Pension Gaps (GPG) are a pressing matter not just for current cohorts but will increasingly affect a growing number of women in the future. While it has been argued that GPGs are of little concern because women share financial resources on the household level, it needs to be acknowledged that household contexts have become more uncertain, and a growing number of women live without a spouse during retirement. The divorce rates among the old age population increase (gray divorce revolution, Brown & Wright, 2017), and women are much less likely to repartner but face a much larger decline in the standard of living after divorce in old age (Lin & Brown, 2021; Schimmele & Wu, 2016). Thus, a sufficient own pension is increasingly important to prevent women's currently higher vulnerability to old-age poverty compared to men in most European countries in the future (Haitz, 2015). Consequently, understanding the causes of gender inequalities in pension income is of tremendous political relevance (Kuitto et al., 2021).

Nevertheless, thorough empirical evidence on the mechanisms shaping Gender Pension Gaps (GPG) is scarce (Folbre et al., 2005) and empirical studies predominantly applying decomposition methods focus on women's lack of (full-time) employment as a driver of the GPG (Bonnet et al., 2020; Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014; Johnson et al., 1999; Levine et al., 1999; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018 see last column of Table 3.1 for an overview). However, the accumulation of pension entitlements is more complex, goes beyond the sole focus on full-time employment and the employment sphere more generally and is likely to differ by pension type. The present study adds to this and examines underlying heterogeneities

regarding, the differences between public and private pensions, the role of earnings over the life course, and work-family independences. Its focus on West Germany provides a relevant context for high gender inequalities in the labour market and a comparatively high GPG with women receiving only about half of the average pension income of men (Hammerschmid & Rowold, 2019a).

This study contributes to several limitations in the GPG literature. First, private (occupational and personal) pensions are identified as being particularly disadvantageous for women and gender gaps are particularly high in private pensions (e.g., Frommert, 2020; Ginn, 2004; Jefferson, 2009). However, most studies do not empirically address the implications of the primary pension policy shift in the last decades “toward the privatization and marketization of pensions” (Ebbinghaus, 2015, S. 56), i.e., the increasing relevance of private pensions, for gendered pension income inequality. The only studies assessing gendered *private* pension income focus on the Beveridgean pension context of the UK (Bardasi & Jenkins, 2010; Ezeyi & Vujic, 2017; Gardiner et al., 2016). The current literature misses a clear empirical assessment of the consequences of pension privatisation for GPGs in other pension systems and how gender inequalities in public and private pensions are differently related to gender inequalities in pre-retirement life courses. This study fills this gap by decomposing the gender gap in public and private pension income separately. For understanding the future of GPGs it is vital to assess whether distinct associations are in play for shaping gender inequality in private in contrast to public pensions to assess the implications of the increasing share of private pensions on gendered pension income inequality. This addresses the need for research on pension privatisation in Europe (Kuitto et al., 2021).

Secondly, the previous literature was unable to consider earnings over the life course for decomposing GPGs. Thus, we do not know how gendered earnings inequalities, as they move over the life course, translate into gendered pension income inequality in old age. This, however, is pressing given that the gender pay gap is at high levels in most European countries, especially in those with high female labour participation (Schmieder & Wrohlich, 2021), making it likely that female full-time employment alone might not succeed in preventing high GPGs. This study is the first to decompose GPGs based on individuals' relative earnings positions over the life course, beyond the annual employment status. Including differences in income positions is specifically relevant for assessing how gender inequalities in the labour market are potentially differently associated with the gender gaps in public compared to private pensions. Earning differences over the life course might play a particularly important role in the gender gap in private pensions since they are likely to determine access and endowment with occupational and personal pensions (Geyer et al., 2021; Ginn & MacIntyre, 2013). This is highly relevant in Germany, since the Gender Wage Gap remained at a high level for decades (Schmieder & Wrohlich, 2021; Schmitt & Auspurg, 2022), making the study not only crucial for current women in old age but also for women in the labour force nowadays.

Lastly, the scientific discourses are criticised for focusing on employment mostly when it comes to old age income security (Frericks, 2020, see also Chapters 1 and 3) and neglect the theoretically important interdependences of employment and family life course complexities for analysing Gender Pensions Gaps (Allmendinger et al., 1992; Ginn et al., 2001). Recent decomposition studies on GPGs do not consider family lives in detail and life course interdependences or complexities more generally. They i) simplify life courses by only considering one-dimensional summary measures (most commonly the years spent

in employment), ii) focus on a very limited number of (employment) life course elements due to multicollinearity concerns, and iii) do not consider the interaction between family and work lives (see Chapters 1 and 3). Overall, the literature misses a holistic perspective that disentangles how common work-family life courses of individuals are related to gender gaps in different pension types.

This paper takes a life course perspective and analyses how the interaction of gendered work-family life course patterns with different pension pillars in later life relates differently to the gender gaps in public and private pensions. Specifically, I focus on two main questions.

1. To what extent are (gendered) work-family life courses, including earnings over the life course, associated with the Gender Pension Gap?
2. Do these associations vary between the gender gap in public compared to private pensions?

To address these questions, I use rich linked survey-register data. The survey data provides access to private pension income and family life courses which are most often not available in official records, while the administrative record provides rich reliable retrospective working data, including relative earnings over the life course. Such unique data is only available for Germany, and the German pension system represents the historical prototype of a Bismarckian pension logic with a strong status maintenance function and subject to pension privatization since the 90s (Ebbinghaus, 2021; Ebbinghaus et al., 2011). West Germany is a ideal context for contributing to the literature illustrating how high levels of gender inequalities in pre-retirement ages are reproduced by strongly earnings-related pension institutions across pension income types.

Methodologically, I apply the recently developed SA-KOB decomposition which captures the association of multidimensional life course complexities with the GPG (Chapter 3). In the first step, Multichannel sequence analysis identifies and visualizes common work-family life course patterns, that is, it groups life courses in which work and family states similarly unfold. In a second step, I decompose the GPG based on group membership in these life course patterns revealing i) which part of the GPG is due to the gendered segregation on different work-family life courses (explained component) and, ii) which part is due to unequal rewards for the same life course pattern (returns component). This provides a holistic perspective, depicting real life courses and the complexity therein, such as work-family interdependences, and how these are related to the GPG.

The largest shares of the Gender Pension Gaps arise from highly gendered life course patterns (explained component): privileged life courses (stable civil servant careers for public and high-income employment for private pensions) obtaining high pension rewards are dominated by fathers. Furthermore, the gender differences in access to high-income careers for parents drive the GPG in private pensions to a much higher extent than the gap in public pensions.

For public pensions, the returns component of the SA-KOB can specifically reveal whether the gender gap in pension income is decreased because women compared to men with similar life courses benefit from the rewarding structure with higher pension income (e.g., through childcare benefits). I do not find that this is the case, even for life course patterns with large shares of unpaid care and part-time work parallel to parenthood. However, divorced women benefit from higher public pension income compared to divorced men which decreases the GPG and confirms the equalizing power of pension right splitting at couple level (which is in Germany mostly realized upon divorce).

4.2 Theoretical and institutional background

4.2.1 Empirical evidence

The most common stream of literature analyses the gender gap in pension income between men and women directly by decomposing it into a part that is due to differences in endowments with characteristics (mostly, the gender difference in years spent in specific employment types, e.g., full-time employment) and a part that is due to different returns for the same characteristic (e.g., Bardasi & Jenkins, 2010; Bonnet et al., 2020; Even & Macpherson, 2004; Zhao & Zhao, 2018). In line with the international literature, all decomposition studies in Germany found that differences in years men and women have been full-time employed is the main reason for the GPG, explaining up to 92% of the gap (Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014). Some studies also highlight the strong impact of differences in income at different points in the life course (e.g., mid-career earnings, salary in last job) (Bonnet et al., 2020; Even & Macpherson, 2004; Levine et al., 1999; König et al., 2019) or occupational segregation (Levine et al., 1999; Veremchuk, 2020). Whereas the association of earnings with the GPG or women's pension income has not yet been studied for Germany, Cordova et al. (2022) also find evidence that parts of the gender pension wealth gap in Germany are due to the high level of gender segregation on industries and company size. Furthermore, for Germany specifically, a large share of the gap is reduced due to the more years women have spent in part-time employment (Cordova et al., 2022; Frommert & Strauß, 2013).

Studies on the gender gaps in occupational and personal pensions are particularly rare. Most decomposition studies on the GPG focussed on public pension income only or the sum of pension income from public, occupational, and personal pensions. However, studies on the UK are an exception but focussing largely on the association with previous employment

careers. Bardasi and Jenkins (2010) found for Britain that only a small part of the GPG in private pensions is due to gender differences in time spent in various employment categories and that the gap is mainly due to differences in returns for these characteristics. They conclude that the gender gap in private pension might be predominantly driven by further employment characteristics they could not include, most importantly earnings. However, Ezeyi and Vujic (2017) find that the wage in the last job before retirement plays only a role for gender differences in access to private pensions, but not in private pension income in the UK.

For Germany, higher gender gaps in private compared to public pension income are reported (Frommert, 2020). The gender differences in coverage and contributions or assets in private pensions are moderated by career length, hours worked and partly also by the lower income of women (OECD, 2021).

Studies that do not decompose the GPG but analyse determinants of women's or gender inequality in pension income highlight the role of family lives. Going beyond the employment lives only, Fasang et al. (2013) found a negative impact of the duration of being married and a positive association between being never married, divorced or widowed with West German and British women's retirement income. This holds while controlling for the employment history. Kreyenfeld et al. (2023) find similar associations of partnership status with women's pension income and highlight the impact of splitting pension rights resulting in divorce being a 'gender equaliser' for public pension income inequality in West Germany. Apart from the partnership status, fertility plays a role as well. Comparing the pension income of mothers with childless women, Möhring (Möhring, 2018) finds motherhood pension penalties across European countries. Regarding West German mothers, the accumulation of public pension claims lags behind that of childless

women over the years, and this disparity persists throughout their entire life course (Stegmann, 2007). Mika and Czaplicki (2017) find a more modest motherhood penalty when considering childcare benefits for West Germany, suggesting how the childcare benefits compensate for parts of maternal employment interruptions in the public pension system.

However, decomposition studies do not consider family lives in detail and family-work interdependences or life course complexities more generally (see section 1.4 and chapter 3). As a result, they largely not discuss mechanisms beyond the impact of the duration worked full-time. More generally, for Germany possibly different mechanisms in play between public and private pensions and remain uncovered as well.

Madero-Cabib and Fasang (2016) were the first to introduce a holistic perspective of the life courses impacting men's and women's pension income to the field of studies. In their analysis of West Germany and Switzerland, they apply Multichannel sequence analysis to categorise typical work-family life courses and reveal the highest pension penalties for typical female work-family life courses compared to the male standard life course of stable married parenthood and full-time employment. Focussing on the association of employment trajectories on earning points, Tophoven & Tisch (2016) find similar levels of earnings point accumulation at young ages for female full-time and part-time employment trajectories but diverging until age 42 with a lower accumulation for women with part-time careers. Considering employment trajectories only as well, Möhring and Weiland (2022) take a couple's perspective and show the highest pension incomes of women in dual-earner and lowest in male-breadwinner couples.

These studies did not focus directly on the GPG and could not differentiate between which part of the gap is due to compositional differences between men and women and which part

is due to differences in returns. To address these limitations, in Chapter 3 we suggest combining sequence analysis with the Kitagawa-Oaxaca-Blinder decomposition and find in our case study that the gender gaps in total pension income in West Germany and Italy are driven by parents and the interdependences of work and family occurring only in female life courses.

4.2.2 Gender, life courses, and the pension system

Literature theoretically argues and empirically proves that pension systems reward life courses (Ginn et al., 2001; Madero-Cabib & Fasang, 2016; Chapter 3). As a result, the life course perspective has been recommended to approach the GPG (Foster & Ginn, 2018). The life course paradigm highlights how the development of individuals over the whole life course depends on the individual opportunities and constraints that are shaped by the socio-historic context, the interdependences with other individuals (linked lives) as well as the timing of events along an individual's life courses (Elder et al., 2003). Individual trajectories are formed by the influence of norms and culture as well as in the context of economic institutions, labour markets, and welfare policies (Bernardi et al., 2019). Consequently, it is important to consider different socio-historic developments as their cultural and institutional legacies are likely to leave their marks on individuals' life opportunities. In line with this, both, welfare states and cultural context are relevant for structuring individuals' everyday life, pushing or discouraging certain gender arrangements and thus shaping gendered life courses and outcomes (Krüger & Levy, 2001; Pfau-Effinger, 1998; Rosenfeld et al., 2004).

The normative influence unfolds especially through the gendered division of labour. The most traditional form assigns men the status of sole "bread-winner" (Becker-Schmidt & Krüger, 2009, S. 25) and women the role of homemaker and caretaker (Meyer & Pfau-

Effinger, 2006; Rosenfeld et al., 2004). As a result, in male-breadwinner contexts, males' continuous full-time employment and their higher independence of family events are enabled through the unpaid care work done by women (Saraceno 1994; Becker-Schmidt & Krüger 2009). Institutions impact gendered work-family life courses particularly through family policies as these determine the opportunities for parents to reconcile care with employment responsibilities (Adema, 2012).

In line with the notion of linked lives, women's employment life courses are strongly correlated with their family life courses (Allmendinger et al., 1992; Muller et al., 2020; Zagel & Van Winkle, 2022). Specifically, childbirth and having children impact women's employment behaviour and varying outcomes (e.g., Arntz et al., 2017; Boeckmann et al., 2015; Gangl & Ziefle, 2009). The level of this association varies across country contexts (e.g., Aisenbrey et al., 2009; Budig et al., 2016). The implications of such dynamics become visible in various gender inequalities in the working sphere such as the overall labour market participation and working hours, gender-specific occupational segregation as well as the gender gap in wage and lifetime earnings (e.g., Anker, 1998; Blau et al., 2006; Budig et al., 2016; Gangl & Ziefle, 2009; Joshi & Davies, 2002; Kunze, 2018). Based on the life course paradigm and, specifically, the concept of cumulative advantage or disadvantage (CAD) (e.g., Dannefer, 2003), these differences between men and women are likely to accumulate over the life course exceeding retirement age (O'Rand, 1996). In the tradition of CAD "[t]he gender gap in pensions can be understood as the sum of gender inequalities over a lifetime, including differences in the life course (motherhood penalty), segregated labour market and gendered social norms and stereotypes more generally." (European Institute for Gender Equality, 2015, S. 4).

The gendered life courses and outcomes alone do not necessarily lead to gender inequalities in pension income. However, pension institutions can produce GPGs by rewarding the gendered life courses differently. “Accordingly, aged inequality is significantly affected by women's cumulative disadvantage” (O’Rand, 1996, S. 233). This is particularly the case for pension income since pension systems are often structured “based on the norm of continuous full-time employment” (Leitner, 2001, S. 103) or “around an idealized male worker“ (Grady, 2015, S. 454) and do not sufficiently take into account the unequal life-course outcomes of men and women (e.g., Grady, 2015; Quadagno, 1988).

Given all these considerations, the relevance of including the role of family lives in addition to employment biographies for the understanding of gender inequalities in pensions has been highlighted (Allmendinger et al., 1992; Fasang, 2010; Frericks, 2020; Ginn et al., 2001; Meyer & Pfau-Effinger, 2006).

4.2.3 The German context

The life course of the cohorts studied is affected by welfare policies (mostly family policies) and norms in place from 1939 to 2016 (see Figure 1.1). During most of these years, Germany was separated and different gender ideologies and gendered divisions of labour (Rosenfeld et al., 2004) were implemented in both former states. Given that the West German welfare state was long seen “as a prototype of a conservative welfare regime” (Kreyenfeld, 2010, S. 354), providing the perfect setting for a context with high gender inequalities in life courses, this study focuses on West Germany only. Important characteristics were (and are partly in place today) long maternal leaves, a late public regulated and underdeveloped childcare infrastructure with reduced opening hours as well as a joint taxation system. Together, these encouraged women to work part-time if at all (Rosenfeld et al., 2004).

The gender arrangement enforced by the context becomes visible when considering care work or the maternal labour force of the studied cohorts. Women were timewise much more involved than men in unpaid care work (Rosenfeld et al., 2004). For the cohorts born between 1930 and 1960, the maternal employment rates were below 40% in West Germany (Trappe et al., 2015). Overall, the institutional and normative context led to a realisation of the male breadwinner/female homemaker or female part-time model and thus a traditionally gendered division of labour was dominant (Pfau-Effinger, 1998). Table A4.1 and Table A4.2 give an overview of norms, family policies and the pension system in West Germany.

The pension system and gender

For the analysis of gender inequalities in pensions, the consideration of both, access to pension claims and the amount of pensions are important (Ginn et al., 2001). Regarding access to pension claims, eligibility rules and the consideration of time spent for unpaid care work are relevant (Ginn et al., 2001). In Germany, the minimum length of contribution is only 5 years, which enables access for individuals with long employment interruptions. Only 7% of West German women (1% of men) do not receive any independent pension income (see Table A3.2).³⁸

Concerning the amount of pension, redistributive elements in pension systems play an important role (e.g., Ginn et al., 2001), as these diminish the reproduction of (gendered) labour market inequalities in old age (Vlachantoni, 2012). Thus, elements that strengthen the link between contributions and pension income are disadvantageous for women's pension income (Grech, 2013; Horstmann et al., 2009; Samek Lodovici et al., 2011).

³⁸ The discussion of the German pension system refers to the pension policies in place from 2006-2015 (which is when pension income was surveyed in the data used, see Figure 4.1).

The German Public Pension system

Individuals' contributions are recorded as earning points (EP) in the German pay-as-you-go system. The reference for the value of the EP is the average income from all insured individuals: 1 EP equals the average income each year, lower values represent below-average incomes etc. The sum of all EPs determines the retiree's public pension income. Self-employees are usually not covered by the public pension system. Civil servants are insured in a separate, very generous system.

The German pension system is thus overall characterized by a strong link between stable employment outcomes and pension income. Studies focussing on women's pension income indeed find a strong positive association between West German women's length of both, full-time and part-time employment careers, and their income in old age (Fasang et al., 2013; Sefton, Evandrou, & Falkingham, 2011).

However, pension systems often include redistributive tendencies that address gender-specific outcomes of the division of labour more or less directly aiming to offset gender inequalities and thus being particularly important for women's pension income (Horstmann et al., 2009; Leitner, 2001). This includes benefits for care work and the treatment of part-time work.

In the German pension system, EPs can be topped up for low-income periods and employment interruptions. For children born before 1992, as is the case for 99.4% of childbirths in my sample³⁹, earnings points for one year equivalent to an average employment income per child are credited until 2014 (child care benefit). Low EPs acquired before 1992 are topped up until a maximum equivalent of 75% of the average income if the

³⁹ 20 children were born after 1991 and 19 of the parents were males.

individuals reach a minimum contribution length of 35 years.⁴⁰ This benefits particularly part-time employed mothers (Czepek, 2020). However, research found that the risk of a low public pension is higher, the lower the working hours, even when controlling for earnings (Brussig et al., 2017). Lastly, upon divorce, the joint EPs of the couples which were acquired during marriage are split equally between both partners. This tends to benefit women (Kreyenfeld et al., 2023). Overall, the redistribution in the public pension system works slightly in favour of women as the positive gender gap in the replacement suggests (Table A4.2).

Lastly, *occupational and personal pensions* are argued to be particularly disadvantageous for women due to their positions in the labour market (e.g., Ginn, 2004; Jefferson, 2009; Möhring, 2018). Occupational pensions refer to pension schemes on the firm level sponsored by the employer and personal pensions to personal saving plans (Ebbinghaus, 2011a). In this study, I refer to private pensions as the sum of both, occupational and personal pensions (see Table A4.5 for an overview of income types considered).

First, women have lower access to occupations that guarantee generous occupational pension schemes, as these are typical in male-dominated jobs and in larger companies, in which women are underrepresented (Fasang, 2010; Ginn, 2004; Ginn & Arber, 1996). Reduced working hours, lower occupational status, shorter tenure, and lack of financial resources for the high contributions needed are further obstacles (Ginn 2004). Second, private pensions are overall highly linked to stable employment and lifetime earnings as well as final salaries (Ginn & Arber, 1996), have in contrast to most public pillars most

⁴⁰ For contributions since 1992, earning points acquired parallel to child rearing periods of children aged below 10 are upgraded to maximal 0.9996 earning points per year (Frommert & Lizon, 2017) and only a minimum contribution length of 25 years is needed. But only a minority of my sample benefit from this policy.

often no redistributive elements and are therefore likely to reinforce (gender) inequalities that emerged in the labour market (Möhring, 2015).

Occupational and personal pensions are not mandatory in Germany, but the overall coverage increased in the past decades (Bridgen & Meyer, 2014; Ebbinghaus et al., 2011). Occupational pensions differ largely by industry and company size (Ebbinghaus, 2018; Wiß, 2020) and lead to inequalities between men and women due to gender-specific occupational segregation (Schmitz-Kießler, 2020). 39.8% of men but only 21.7% of women in my sample have an occupational pension (Table A4.2).

4.2.4 Expectations

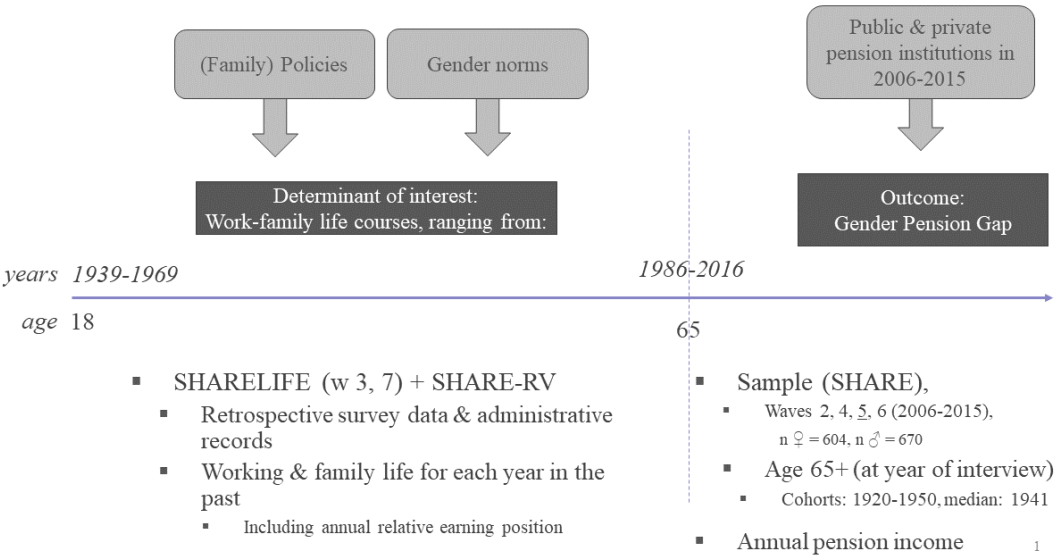
Since in all pillars childcare and part-time employment are not covered equivalently to employment in the public and private pension system, life courses being dominated by care work, employment interruptions and part-time employment are likely to only achieve low pension returns. At the same time, women, and especially (early) mothers, are more likely to have such work-family life courses. As a result, a share of the GPG is likely due to more women than men having life courses being characterised by care work, employment interruptions and part-time employment parallel to parenthood. Furthermore, for the studied cohorts it is more likely that fathers realize stable careers without employment interruptions which are likely to lead to reasonable public and private pension incomes. It follows that a share of the GPG is likely due to life courses characterized by stable careers and dominated by men. Going beyond employment types alone, I expect that particularly careers in high-income jobs are more accessible to men and result in high pension income. Thus, parts of the GPG are due to the lack of women in stable high-income careers.

Due to the stronger link to former labour market characteristics, I assume a higher reproduction of labour market inequalities in private pensions. Gender gaps in access to

and income from private pensions should thus be more associated with gender differences in life courses characterised by unstable working or part-time careers and labour market dropouts compared to the gender gap in public pensions. Particularly, I assume the influence of income inequality over the working career to be more reproduced in private pensions and thus gendered life courses characterised by high- or low-income jobs are more associated with the gap in private pensions.

For public pension income, the reproduction of life course inequalities should be at least partly buffered by the gender-specific redistribution tendencies implemented. I expect the GPG in public pension income to be reduced due to higher rewards for mothers compared to fathers in life courses with high shares of care-related employment interruptions and part-time employment. Focussing specifically on the splitting of public pension claims after divorce, the gender gap in public pensions should be reduced due to better rewards to the life courses of divorced women compared to divorced men independently of the working life course.

Figure 4.1 Observed outcomes along the life course, institutional impacts on the GPG and data.



4.3 Data and Methods

4.3.1 Data

I use cross-sectional data for Germany from the SHARE, retrospective survey data from SHARELIFE (waves 3 and 7 of SHARE) and link it to administrative records of the German Public Pension Insurance (SUF SHARE-RV - source: FDZ-RV; relation of data sources depicted in Figure 4.1).⁴¹ This linked survey-register data combines the strength of both data types. Most importantly for this study, the survey data covers rich retrospective information on family life courses such as partnerships and childbirth as well as detailed information on occupational and personal pension claims that are not available in the administrative records. The administrative data provides rich monthly data on pension-relevant social activities, including the relative earnings, recorded as earning points. Additionally, the administrative records do not risk being affected by the recall bias retrospective survey data often suffer from (Kreyenfeld & Bastin, 2016; Muller et al., 2020). This is especially important given the gender focus of this analysis since women tend to have more complex life courses and these are shown to be oversimplified in retrospective survey data (Bühler et al., 2022). The linked survey-register data thus provides the unique opportunity to study public *and* private pension income with detailed and robust work-family life courses.

The sample is based on the SHARE survey data. I selected all respondents aged 65 or older as my analytical sample instead of applying popular sample criteria such as self-identification with retirement or pension receipt. This avoids systematically excluding (i) individuals who would not consider themselves retirees and (ii) who are not eligible for pension income. As Chapter 2.2 shows, these are predominantly women with discontinuous

⁴¹ See Chapter 8 for details on the data.

work biographies, who often do not experience a clear-cut transition to official retirement (see Chapters 2.1 and 2.2 for a more detailed discussion). Since these women are likely to have no or low pensions, excluding them would lead to an underestimation of the Gender Pension Gap.

I exclude individuals (i) with missing information on any variables used; (ii) who were part of the labour force and received a salary or unemployment benefits, (iii) with entirely missing retrospective data as well as (iv) without linkage to the administrative data (see Table A4.3 for the sample selection steps and number of observations). To maximise the number of observations I pool waves, using wave 5 as a baseline (Table A4.4). I then link this sample to the administrative record. The final sample consists of 604 women and 670 men in Western Germany.

Dependant variable. My outcome measure is *individual annual pension income*, which is surveyed in SHARE.⁴² I differentiate between public pension income and private pension income (see Table A4.5 for all pension types considered). To consider gender differences in pension endowments as well as access to pension claims simultaneously, values for respondents without any of the different pension income types are set to zero. This ensures that individuals without any pension income are not excluded from the analysis, considers the gender coverage gap in the different pension income types, and reveals the gender inequality in pension income of the entire old-age population of 65+-year-olds (see Chapter 2.1.3). The dependent variable is thus a proxy for overall pension income inequality, covering access to and amount of pension income. I focus on independent pension income

⁴² Since the administrative records do not provide any information on private pension income and do not allow me to choose which public pension types to include, I use the survey data on pension income for all pension types.

based on individuals' achievements and therefore, I do not consider pension income types from derived claims, such as survivor pensions.

I use absolute pension income, only consider regular payments (i.e., no lump payments) and top-code the highest 1% annual pension income with the 99th percentile.

Independent variable. My main explanatory variable is a *typology of work-family life courses*.

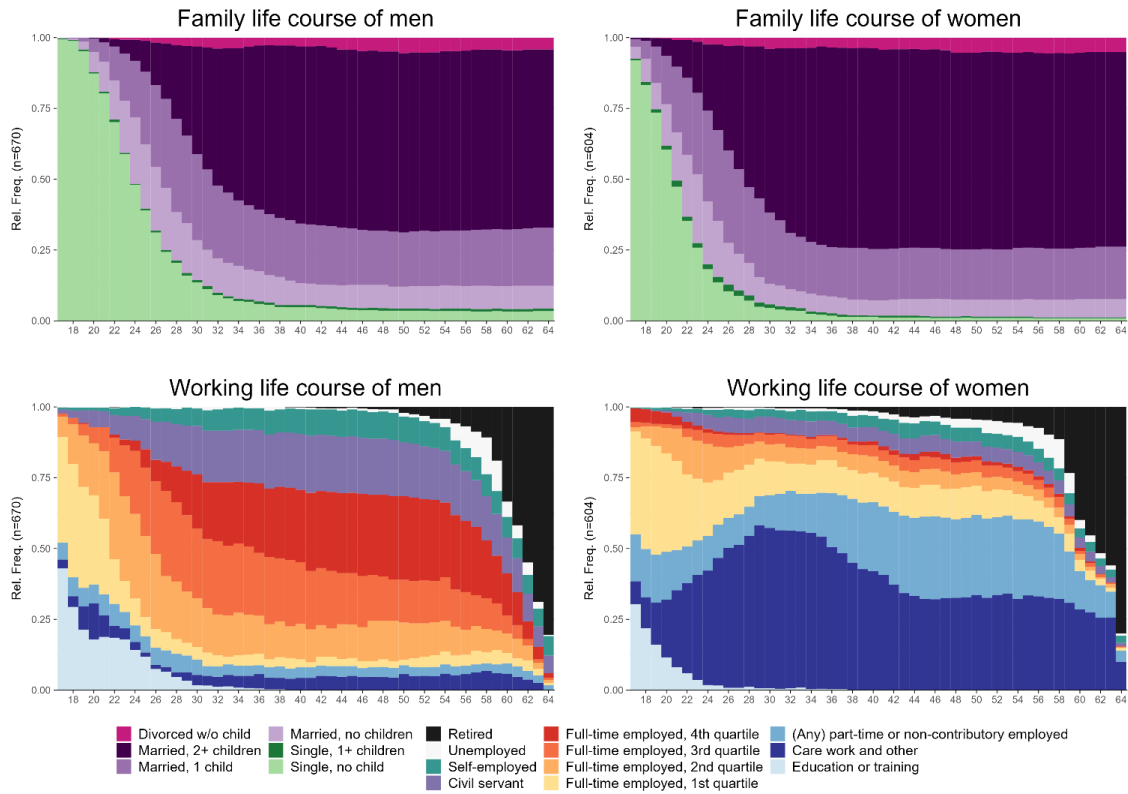
The *family life* categories are based on the retrospective survey data from SHARELIFE. I differentiate between 'single, no child', 'single, 1+ children', 'married, no children', 'married, 1 child', 'married, 2+ children' and 'divorced, w/o child' (see Figure A4.2). 'Single' identifies individuals not cohabiting, married or divorced. 'Married' also includes not cohabiting couples and non-married cohabiting couples. 'Divorce' overwrites all other family states occurring in the same year, except for marriage. Note that the number of children refers to any children alive, including adopted children, in the given year (Brugiavini et al., 2019). No information is available on whether the children lived with the respondent (see section 2.3.2 for a discussion of this issue).

For the *working life* course, I combine administrative records on pension-relevant activities from SHARE-RV with annual retrospective survey data on work and family life from SHARELIFE (using the Job Episodes Panel by Brugiavini et al., 2019). Priority is given to the administrative records to prevent potential impacts of recall bias (Kreyenfeld & Bastin, 2016). However, non-contributory periods to the public pension systems were imputed with the linked survey data as these periods might include various social activity states or employment types not covered by public pensions.

I created 11 working life categories: ‘Education or training’, ‘Care work/other’, ‘(Any) part-time or non-contributory employed’, ‘Full-time employed, 1st EP quartile’, ‘Full-time employed, 2nd EP quartile’, ‘Full-time employed, 3rd EP quartile’, ‘Full-time employed, 4th EP quartile’, ‘Civil servants’, ‘Self-employment’, ‘Unemployed’, and ‘Retired’ (see Figure A4.3). ‘Care work/other’ includes mostly care work (‘Looking after home or family’) and covers all marginal ‘other’ states, most importantly being ‘sick or disabled’, and spells with missing information. ‘(Any) part-time or non-contributory employed’ covers years with non-full-time short employments, any part-time employment periods and any private employment or short-term employment for which individuals did not gain earnings points (this is most likely employment not covered by the public pension system). The income quartiles for full-time employment in the private sector states are based on the distribution of the earning points and thus cover only employment that is recorded in the public pension records (only for such employment earning points are available).⁴³ Figure 4.2 shows the distribution of work and family characteristics for each age by gender. Whereas the family life courses are quite similar across gender apart from women entering parenthood at slightly earlier years, the work life courses differ strongly by gender, as expected for the conservative context of West Germany.

⁴³ Please consult Appendix A4.4 for more information on linking the survey with the register data as well as the procedure to fill recorded non-contributory periods with the survey data.

Figure 4.2 Family and work life courses by gender.



Notes: Sequence distribution plots based on analysis sample showing the share in each life course state per age, ranging from age 18-65.

Controls. Following the argumentation in Chapter 3, I control for cohorts only to prevent adjusting for elements inherent in the work-family life course states (e.g., education).

4.3.2 Methods

I apply a two-step combination of Multichannel sequence analysis with the Kitagawa-Oaxaca-Blinder decomposition (SA-KOB decomposition, see Chapter 3). The outcome of the first step is a typology of typical work-family life course clusters which are then used as independent variables in the second step to decompose the GPG.

Step I: Multichannel Sequence Analysis (MSA)

I first ran a Multichannel sequence analysis and cluster analysis (Pollock, 2007). Sequence analysis is an explorative tool to categorise trajectories (sequences) into groups using

cluster analysis and has become particularly established in the field of life course research (Abbott, 1995; MacIndoe & Abbott, 2004). I make use of its strength to operationalise life-course patterns (Aisenbrey & Fasang, 2010) and focus on two dimensions, family and work. Among the several approaches suggested to account for more than one temporal process at a time (Piccarreta, 2017), I use multichannel sequence analysis (Gauthier et al., 2010; Pollock, 2007). As a result, respondents with a similar working and a similar family life course are clustered in the same work-family life course groups.

To identify the best cluster solution for the aim of my analysis, I refer to statistical cluster quality measurements (Studer & Ritschard, 2016) as well as construct validity (Aisenbrey & Fasang, 2010). For the main results, the clusters are obtained from Optimal Matching and substitution costs based on a common future for comparing the sequences and hierarchical clustering based on Ward's Linkage.⁴⁴ This 10-cluster solution reaches the third-best Cluster Quality Indices values among a set of 18 different cluster solutions covering different combinations of parameters (Average Silhouette Width: 0.266). The two better-performing alternatives suggest 11 clusters. Apart from the additional cluster, the clusters are very similar to the chosen cluster solution. However, since the additional cluster does not seem to provide added theoretical value important enough to justify the inclusion of an additional cluster, I favour the 10-cluster solution. The chosen cluster solution is overall relatively robust against the cost specification applied. For a detailed documentation of the cluster selection, please consult Appendix A4.5.

⁴⁴ For a detailed discussion of the different parameters used in Sequence Analysis please refer to Studer & Ritschard (2016).

Step II: Kitagawa-Oaxaca-Blinder decomposition (KOB)

In the second step, I use the established work-family life course typology of the MSA to decompose the average absolute Gender Pension Gaps by applying the Kitagawa-Oaxaca-Blinder decomposition (Blinder, 1973; Kitagawa, 1955; R. Oaxaca, 1973).

Based on separate linear regression models of men and women I apply the 'twofold' decomposition' (Jann, 2008, S. 455) with the methodological choices suggested for combining it with sequence analysis (see Chapter 3)⁴⁵:

$$\bar{y}_M - \bar{y}_W = \underbrace{\alpha_M - \alpha_W}_{\text{Intercept}} + \underbrace{\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W(\hat{\beta}^* - \hat{\beta}_W)}_{\text{Returns}} + \underbrace{(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*}_{\text{Explained}}$$

$$\text{with } \hat{\beta}^* = \begin{matrix} b_{1M} \\ b_{2M} \\ b_{3M} \\ b_{4M} \\ b_{5F} \\ b_{6F} \\ b_{7M} \\ b_{8M} \\ b_{9M} \\ b_{10M} \end{matrix}, \hat{\beta}_M = \begin{matrix} b_{1M} \\ b_{2M} \\ b_{3M} \\ b_{4M} \\ b_{5M} \\ b_{6M} \\ b_{7M} \\ b_{8M} \\ b_{9M} \\ b_{10M} \end{matrix}, \hat{\beta}_F = \begin{matrix} b_{1F} \\ b_{2F} \\ b_{3F} \\ b_{4F} \\ b_{5F} \\ b_{6F} \\ b_{7F} \\ b_{8F} \\ b_{9F} \\ b_{10F} \end{matrix}$$

where $\bar{y}_M - \bar{y}_W$ stands for the average GPG. $\bar{X}_M$ and $\bar{X}_W$ denote vectors of mean values of the independent variables and $\hat{\beta}_M$ and $\hat{\beta}_W$ denote coefficient vectors of men and women and α_M and α_W the intercepts, respectively. $\hat{\beta}^*$ is the reference coefficient vector that is

⁴⁵ This includes choosing the reference coefficient based on which gender is the clear majority group in the life course cluster (i.e., coefficient from female regression models as reference for the women-majority life course cluster and the coefficients from the regression models for men for the rest), normalization of categorical variables and the reassignment of single trajectories after the MSA to bypass extreme cases of the common support problem (for the latter, see description in A4.6). For a more detailed discussion of combining MSA with the KOB decomposition, its specifics and limitations please refer to Chapter 3.3.2.

based on the coefficients from the linear regression models of females for the women-majority work-family life courses and based on male coefficients otherwise.

Explained or compositional component: $(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*$. The explained share quantifies the share of the gap that emerges from the gender-specific distribution on the different work-family life course clusters. It thus reveals which part of the GPG is due to the under- or overrepresentation of women compared to men in life-courses, that are rewarded with particularly high or low pension income.

Returns component of the unexplained share: $\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W(\hat{\beta}^* - \hat{\beta}_W)$. The returns component shows which part of the GPG is due to differences in pension income of men and women who have the same work-family life-course pattern. The returns parts of the SA-KOB decomposition have thus the potential to reveal how the GPG is related to differences in pension rewards for the same life course pattern.

For more details on the explained and returns component, please refer to Chapter 3.3.2.

4.4 Results

4.4.1 Step I: Work-family life course patterns

Figure 4.3 shows the results of the first step of the analysis, the Multichannel sequence analysis. Each typical life course pattern consists of a working (left sequence plot) and a family life course (right sequence plot). One line in these two sequence plots represents a life course trajectory as observed in the data that has been identified to be representative of the respective cluster (Fasang & Liao, 2014). On the outer side of each cluster, the gender-specific distribution on each life course pattern is plotted.

I find ten distinct clusters, with the first six being characterised by stable marriages and having 2 or more children for most of the individuals in the life course patterns (see Table 4.1 for the cluster characteristics by gender).

Cluster 1, *High-income full-time employed & 2+ children*, is characterised by stable full-time employment with earnings in the highest quartile or upward mobility leading to the highest earnings quartile. Family formation starts in the mid-30s, whereby the age at birth of the first child varies. This life course is dominated by fathers (17.6% of all men, see Figure A4.5), while almost no mothers in West Germany experience such high-income careers.

Cluster 2, *Low-medium income full-time employed & 2+ children*, reveals a more volatile employment life pattern. Even though most individuals experience nearly continuous full-time employment in the private sector, there are many shifts in the relative earnings over individuals' life course, including both, upward and downward earnings mobility. Short unemployment periods, particularly before retirement, are more common compared to cluster 1 (see Table 4.1), and family formation as well as retirement starts earlier. Occasionally, individuals have short unpaid care interruptions in their careers. While only a low share of women (4.3%) can combine volatile full-time employment with parenthood, 20.7% of men experience this life course. The gender difference is similar to cluster 1.

Cluster 3, *Civil servant & 2+ children*, includes the most stable employment pattern characterized by relatively long educational periods followed by continuous civil servant careers. Most individuals are married and have 2 or more children, but several individuals have only one child or remain childless (see also higher shares of these family life course

states, Table 4.1). This life course pattern is most common among men (20%) while only 5.5% of women have careers in the civil service.

Cluster 4, *Self-employed & 2+ children*, is dominated by continuous self-employment after mostly volatile periods in other forms of employment or care work. These first volatile periods are of different lengths leading to differences in age when self-employment starts ranging from the early 20s to the early 50s. This typical work-family life course is, again, more prevalent among men compared to women but the gender difference is less pronounced (6.1% of men, 3.5% of women).

In cluster 5, *Part-time or non-contributory employed & 2+ children*, most life courses start with a very short period of full-time employment, mostly in lower-earning jobs. Parallel to the start of family formation, often at a young age, most individuals shift from full-time employment to full-time care. After these employment interruptions of about 5 to 20 years, the individuals return to labour force in part-time or non-contributing overwhelmingly, or low-income jobs. Once individuals enter part-time employment, they mostly stay in part-time employment until retirement which stands in contrast to the notion of part-time employment as a 'bridge' returning to full-time employment (e.g., Kelle et al., 2017; Stegmann, 2007).

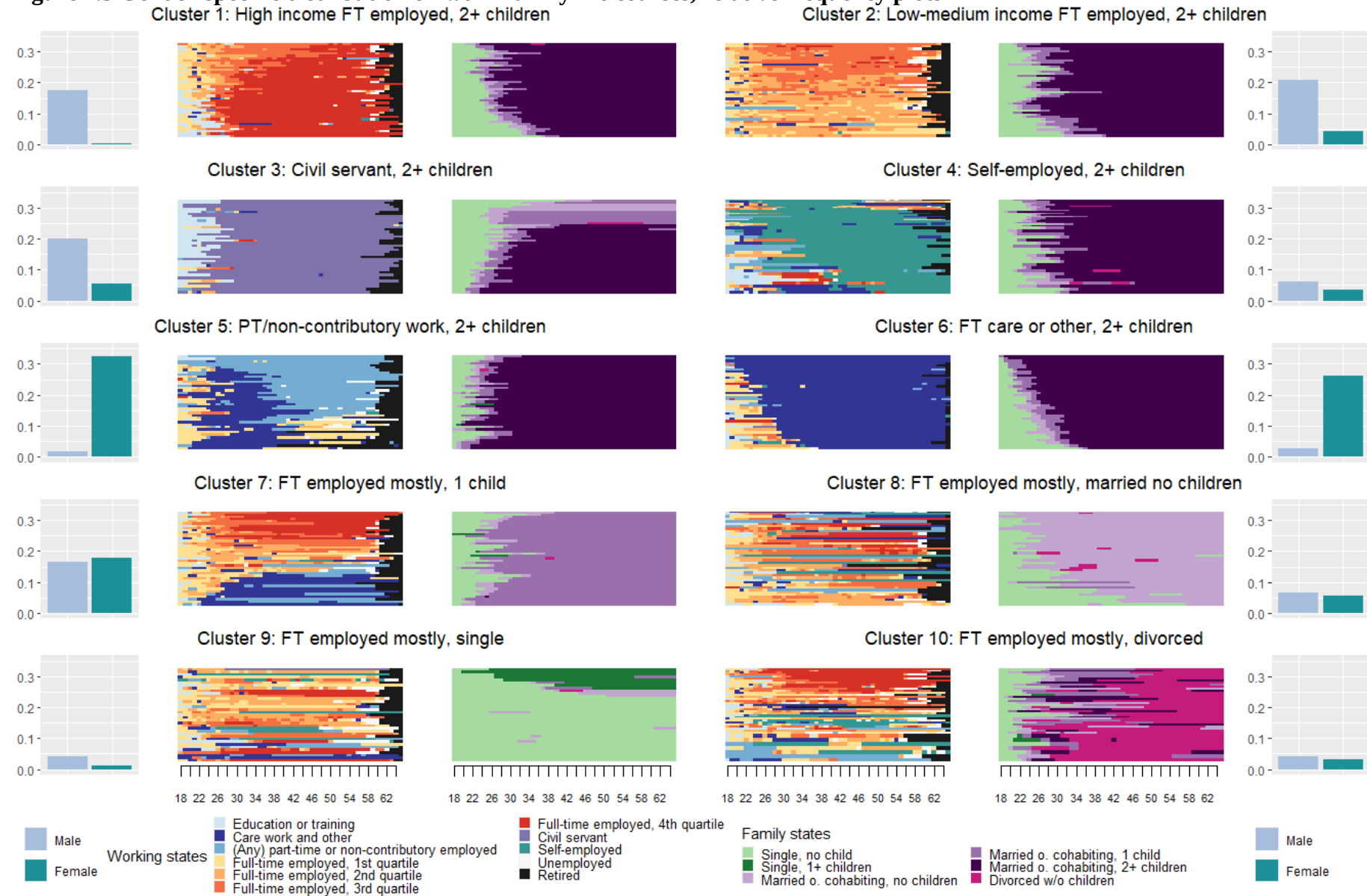
One-third of all women have this life course, whereas almost no men reconcile parenthood with care and part-time work.

Cluster 6, *Full-time care or other & 2+ children*, starts with short education or employment trajectories. Like cluster 5, individuals transition to care work parallel to marriage and childbirth, but in this case, they start lifelong trajectories of care work upon family formation. This is the second most prevalent life course pattern among women (26% of all

women), but only 2.7% of all men have similar life courses. Clusters 5 and 6 show thus a clear temporal interaction between family and work lives, highlighting the concept of linked lives for these female life course patterns while men-typical life courses do not show such interactions at all.

Clusters 7-10 have very heterogeneous working life courses that are mostly dominated by different types of full-time employment but differ by family life. Individuals in cluster 7 “*Full-time employment mostly & 1 child*” have 1 child, and respondents in cluster 8 “*Full-time employment mostly & married, no children*” get married partly at relatively old ages and have no children. Respondents in cluster 9 “*Full-time employment mostly & single*” remain singles and a higher share of all men experience this life course. Individuals in Cluster 10, *Full-time employed mostly & divorced*, experience a divorce and, in most cases, remain divorced afterwards. However, the timing of divorce differs largely with some individuals experiencing divorce already in their early 20s and others in their early 40s. Clusters 7, 8, and 10 are gender-neutral, i.e. the share of men and women is similar.

Figure 4.3 Gender-specific distribution on work-family life courses, relative frequency plots



Notes: Relative frequency sequence plots (displaying representative sequences for each cluster) with the working domain on the left-hand side and the family domain on the right-hand side in the centre of the graph. Gender-specific distribution on the life course patterns on the outer side, showing the share of all men (blue bar) and all women (green bar). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table 4.1 Summary characteristics of life course clusters by gender

	Cluster 1: High Income Ft Empl., 2+ Chi.		Cluster 2: Low-Medium Income Ft Empl., 2+ Chi.		Cluster 3: Civil Servant, 2+ Chi.		Cluster 4: Self-Empl., 2+ Chi.		Cluster 5: Pt/Non-Contributory Work, 2+ Chi.		Cluster 6: Ft Care, 2+ Chi.		Cluster 7: Ft Empl. Mostly, 1 Child		Cluster 8: Ft Empl. Mostly, Married No Children		Cluster 9: Ft Empl. Mostly, Single		Cluster 10: Ft Empl. Mostly, Divorced	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Work life course (share of spells)																				
FT Employed, 1st Quartile	5.72	3.13	11.78	33.73	1.68	0.63	4.47	3.67	12.15	14.69	7.64	6.20	7.78	17.00	8.96	19.18	12.58	29.95	12.35	19.85
FT Employed, 2nd Quartile	6.78	5.21	32.61	21.15	1.73	0.44	7.11	1.39	9.72	3.20	2.89	1.62	13.88	10.28	22.14	12.68	22.61	9.38	10.64	17.87
FT Employed, 3rd Quartile	14.55	21.88	30.22	10.50	0.61	0.13	4.78	0.79	3.47	0.73	5.44	0.91	24.79	5.28	21.80	16.18	15.90	4.69	16.59	9.87
FT Employed, 4th Quartile	55.54	43.75	5.41	1.92	0.68	0.00	2.18	0.00	2.26	0.72	5.32	1.32	24.55	0.88	15.89	4.84	13.19	3.91	23.36	3.51
PT Or Non-Contributory Emp.	1.31	8.33	3.24	5.61	2.86	1.64	6.40	7.84	41.67	36.43	5.67	7.45	3.90	20.05	5.67	12.13	3.78	13.28	6.25	10.75
Self-Employed	0.48	0.00	0.91	0.72	0.02	0.19	57.27	62.60	2.60	0.47	7.18	0.33	7.25	3.15	8.72	5.33	9.10	10.68	16.29	1.75
Civil Servant	0.16	0.00	0.04	0.96	73.99	76.14	0.20	0.79	0.00	0.14	0.23	0.34	0.38	0.16	0.00	0.49	2.70	0.00	0.00	0.00
Care Work And Other	1.45	10.42	2.70	9.13	0.95	2.59	9.30	13.10	8.85	26.50	58.56	77.40	4.87	29.09	1.55	15.93	4.40	17.45	2.90	15.68
Retired	6.81	6.25	9.22	11.22	7.88	7.58	3.25	5.95	14.24	12.31	4.05	2.25	6.88	9.11	7.61	9.50	6.94	6.51	4.61	10.86
Education Or Training	5.54	1.04	1.21	2.00	9.50	10.61	4.12	3.27	3.82	1.26	2.78	1.77	3.75	1.60	4.65	2.45	3.09	4.17	2.38	4.93
Unemployed	1.66	0.00	2.65	3.04	0.11	0.06	0.91	0.60	1.22	3.55	0.23	0.41	1.97	3.41	3.00	1.29	5.71	0.00	4.61	4.93
Family life course (share of spells)																				
Single, No Child	16.35	0.00	15.00	11.86	19.48	16.16	14.63	10.71	16.67	7.75	15.51	10.73	15.55	10.63	26.74	15.38	81.10	67.71	12.05	10.86
Single, 1+ Children	0.44	0.00	0.07	0.00	0.19	0.00	0.05	0.00	0.00	0.30	0.00	0.78	0.55	1.13	0.00	0.00	12.89	24.22	0.00	1.32
Married O. Cohabiting, No Children	4.63	0.00	3.21	2.80	10.26	9.72	3.96	2.08	2.08	3.22	1.74	3.10	7.82	5.41	70.16	78.92	4.86	4.17	9.30	11.51
Married O. Cohabiting, 1 Child	7.03	2.08	8.08	7.69	17.75	12.69	6.15	9.72	6.60	7.41	6.13	6.05	71.99	78.84	1.84	2.70	0.77	3.91	16.44	5.26
Married O. Cohabiting, 2+ Chi.	70.30	94.79	72.54	76.20	50.72	58.71	73.88	76.88	73.26	79.01	75.58	78.77	2.14	1.13	0.00	0.12	0.00	0.00	16.22	9.21
Divorced W/O Children	1.25	3.13	1.09	1.44	1.60	2.71	1.32	0.60	1.39	2.32	1.04	0.57	1.95	2.86	1.26	2.88	0.39	0.00	45.98	61.84
Education																				
Max. lower secondary (Isced 0-2)	0.85	0.00	9.42	23.08	4.48	0.00	0.00	33.33	8.33	27.55	5.56	39.87	10.00	28.30	11.63	14.71	11.11	37.50	3.57	10.53
Max. post-secondary (Isced 3, 4)	41.53	50.00	79.71	61.54	27.61	12.12	60.98	52.38	75.00	63.78	55.56	50.63	64.55	60.38	60.47	61.76	59.26	25.00	57.14	63.16
Tertiary education (Isced 5,6)	57.63	50.00	10.87	15.38	67.91	87.88	39.02	14.29	16.67	8.67	38.89	9.49	25.45	11.32	27.91	23.53	29.63	37.50	39.29	26.32
Birth cohort																				
1920-1938	28.81	0.00	47.48	38.46	35.07	6.06	58.54	52.38	41.67	31.63	44.44	46.84	31.82	33.64	37.21	29.41	29.63	62.50	21.43	26.32
1939-1944	46.61	0.00	32.37	30.77	40.30	36.36	24.39	23.81	33.33	38.27	33.33	37.34	36.36	27.10	44.19	38.24	40.74	37.50	35.71	42.11
1945-1950	24.58	100.00	20.14	30.77	24.63	57.58	17.07	23.81	25.00	30.10	22.22	15.82	31.82	39.25	18.60	32.35	29.63	0.00	42.86	31.58

Table 4.1. Summary characteristics of life course clusters by gender (continuation)

	Cluster 1: High Income Ft Empl., 2+ Chi.		Cluster 2: Low- Medium Income Ft Empl., 2+ Chi.		Cluster 3: Civil Servant, 2+ Chi.		Cluster 4: Self- Empl., 2+ Chi.		Cluster 5: Pt/Non- Contributory Work, 2+ Chi.		Cluster 6: Ft Care, 2+ Chi.		Cluster 7: Ft Empl. Mostly, 1 Child		Cluster 8: Ft Empl. Mostly, Married No Children		Cluster 9: Ft Empl. Mostly, Single		Cluster 10: Ft Empl. Mostly, Divorced	
Coverage public pensions																				
No public pension receipt	1.69	0.00	0.72	0.00	2.99	6.06	2.44	9.52	8.33	5.10	0.00	8.23	0.91	10.28	2.33	5.88	0.00	12.50	3.57	0.00
Any public pension	98.31	100.00	99.28	100.00	97.01	93.94	97.56	90.48	91.67	94.90	100.00	91.77	99.09	89.72	97.67	94.12	100.00	87.50	96.43	100.00
Coverage private pensions																				
No private pension receipt	27.12	0.00	47.48	53.85	96.27	93.94	68.29	90.48	75.00	71.94	88.89	93.04	48.18	66.36	39.53	61.76	55.56	75.00	50.00	42.11
Any private pension	72.88	100.00	52.52	46.15	3.73	6.06	31.71	9.52	25.00	28.06	11.11	6.96	51.82	33.64	60.47	38.24	44.44	25.00	50.00	57.89
Sum of EPs From FT Emp	57.80	53.04	40.07	22.65	2.14	0.84	11.76	1.55	11.45	5.87	10.36	4.16	43.51	13.87	38.30	25.84	32.22	13.66	38.21	21.31
Public pension income	18,820	18,360	13,996	9,931	31,044	23,201	11,309	4,265	14,720	7,087	12,899	3,654	17,198	8,658	15,820	10,796	12,450	5,995	12,388	12,014
Private pension income	6,868	14,400	2,346	1,536	359	179	4,168	1,076	3,400	951	918	152	4,071	824	4,778	1,162	3,193	923	2,485	1,388
n	118	2	139	26	134	33	41	21	12	196	18	158	110	107	43	34	27	8	28	19

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The gender-specific distribution on the typical work-family life courses mirrors the traditional division of labour that has been facilitated by gender norms and the welfare context in the second half of the 20th century in West Germany. In the second step, I assess how this high level of gender segregation on life courses is associated with the average gender gap in pension income.⁴⁶

4.4.2 Step II: Decomposition of the Gender Pension Gaps

The GPG in public pension amounts to 58.3% (€10,829). In line with previous literature (Frommert, 2020), the GPG in private pensions is higher in relative terms (76.6%, €2,539). Table 4.2 depicts the decomposition results in absolute as well as relative terms respectively of the gender gap in public (M1) and private pensions (M2). Figure 4.4 and Figure 4.5 illustrate the shares of the GPGs that are due to the explained and returns components of the SA-KOB decomposition for the different life course patterns by pension type.

Explained part.

Positive coefficients indicate that the gender gap would have been smaller if there had been no gender difference in the prevalence of the respective life course. Negative ones suggest that the gap would have been even larger (as the average pension income of women would have been lower) if as many women as men had had the respective biography pattern. Overall, I find that the high level of gender segregation on the life course patterns 1, 2, 3, 5, and 6 (visible in Figure 4.3) is related to the gender gap in both, public and private pension income. The association of life courses with the GPG in the sense of the explained

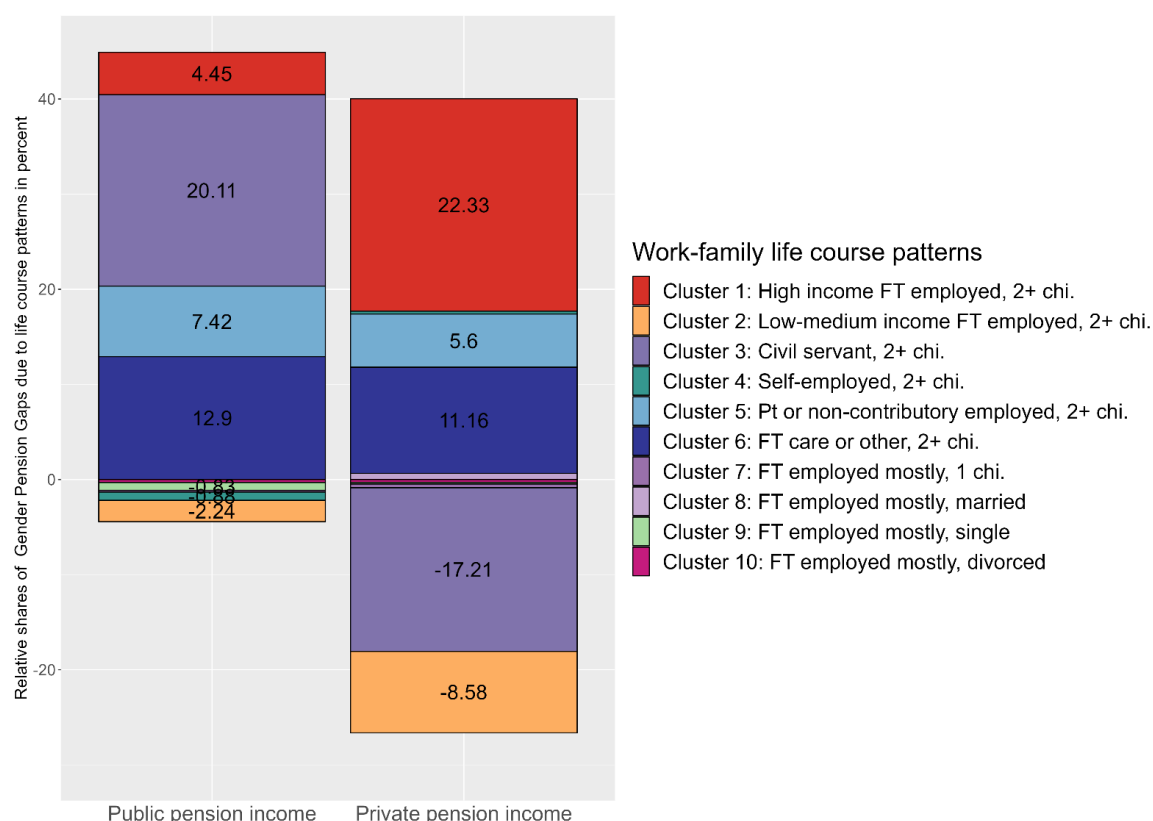
⁴⁶ For the decomposition in the second step of the analysis however, the highly gendered life courses result in a methodological challenge (common support problem). For example, for the association of the high-income life course, I can compare a substantial group of men with two women only. To improve the robustness of the decomposition, I systematically select and assign single trajectories to those low-n cluster for each group as suggested in Chapter 3.3.2. The procedure and the result are documented in A4.6.

part of the decomposition is not solely a result of the high level of gender segregation, but also of the gender-biased evaluation through the pension system: men-typical life courses are generally rewarded with more pension income than female-dominated ones (Table A4.6).⁴⁷ In total, the high level of gender segregation on typical life courses explains 40% of the gender gap in public pensions, but only 13% of the gap in private pensions.

On one hand and as expected, the overrepresentation of women in the two care-dominated and poorly rewarded clusters (5, 6) are among the main reasons for the high GPGs in both pension types. West German women would receive an average of €1,396 more public pension income if as few women as men had engaged in continuous unpaid care work (12.9% of the GPG, see Table 4.2, M1). Another €804 of the gender gap in public pensions is due to the German part-time legacy encouraging mothers to combine care duties with paid part-time employment (7.4%). The relative shares of the GPGs explained by these two women-dominated life course patterns are slightly higher for public compared to private pensions. This highlights the poor recognition of care work in the German public pension system and goes against the assumption that life courses dominated by labour-market dropouts and part-time employment are particularly penalized in private pensions.

⁴⁷ Except for cluster 4, self-employed and parenthood for both pension income types and civil servant careers for private pensions.

Figure 4.4 Shares of the GPG due to the gender-specific distribution on work-family life courses (explained part)



Notes: Value depicted = minimum significance level of 10%. Results from Table 4.2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

On the other hand, and in line with the expectation, the lack of mothers in stable careers independent of employment type plays a dominant role for women's lower pension income compared to men. However, the direction of the association depends strongly on the career type and earnings level. For the dominance of fathers in high-income and low-income full-time employment careers in the private sector I find contrary relationships with the GPG. While the lack of mothers in high-income careers drives the GPG, the gap would be even higher if as many mothers as fathers experienced full-time, but low-earnings careers (2 percentage points for public and 9 percentage points for private pensions). This highlights the relevance of earnings beyond full-time employment for gendered pension inequality. Additionally, as expected, the associations differ largely between pension pillars. The lack

of mothers in the high-earnings (1) and the low-earnings cluster (2) is much more strongly associated with the gap in private compared to public pensions. 4% of the gap in public compared to 22% of the gap in private pensions is due to more men being able to combine parenthood of two or more children with stable careers in high-income jobs in the private sector (see Table 4.2, M2).

While this is the main reason for the gender gap in private pensions, most of the gap in public pensions is due to the much higher prevalence of men having stable civil servant careers that are highly rewarded in the public pension system (cluster 3, 20% of the gap). Reversely, the lack of civil servant careers for women cushions the gap in private pensions (-17% of the gap). This is most likely because the high public pension income for civil servants disincentivises them to invest additionally in voluntary pension schemes (see different returns by pension type for cluster 3, Table A4.6). While the gender gap of parents being self-employed does not play a role for private pensions, the public pension gap would be slightly larger if as many mothers as fathers had careers in self-employment. The latter is straightforward given that self-employment is not well covered in the public pension system. However, this is why one could have expected that self-employees would invest more in personal pensions and thus the gender inequality in cluster 4 to be positively associated with the gender gap in private pensions.

Overall, the explained shares demonstrate that large portions of the GPGs arise from “the valuing of unpaid work and the sharing of it” (Lewis, 1992, S. 170). This is evident in the comparatively lower public and private pension rewards for the female-dominated life course, which are characterised by unpaid care work.

Table 4.2 Decomposition results of the Gender Pension Gap in public and private pension

VARIABLES	Public pension income (M1)				Private pension income (M2)			
	Absolute	Shares	Absolute	Shares	Absolute	Shares	Absolute	Shares
Average pension income men	18,588** (47.46)				3,316** (11.98)			
Average pension income women	7,759** (24.52)				776.7** (7.222)			
Absolute Gender Pension Gap	10,829** (21.51)	100.00%			2,539** (8.549)	100.00%		
Total part explained	4,452** (9.624)	41.10%			379.9+ (1.676)	15.00%		
Total part unexplained	6,377** (12.29)	58.90%			2,160** (6.135)	85.00%		
Sum by life course patterns	explained		returns		explained		returns	
		40.46%		3.93%		13.39%		9.50%
Cluster 1: High-income FT employed, 2+ chi.	482.3** (3.746)	4.45%	-6.657 (-0.228)	-0.06%	567.0** (4.731)	22.33%	-27.21 (-1.617)	-1.07%
Cluster 2: Low-medium income FT employed, 2+ chi.	-243.0* (-2.174)	-2.24%	-54.78 (-0.943)	-0.51%	-217.8* (-2.188)	-8.58%	-44.41 (-1.340)	-1.75%
Cluster 3: Civil servant, 2+ chi.	2,178** (7.496)	20.11%	94.34 (1.289)	0.87%	-436.9** (-4.142)	-17.21%	-94.20* (-2.112)	-3.71%
Cluster 4: Self-employed, 2+ chi.	-94.98+ (-1.675)	-0.88%	46.82 (0.749)	0.43%	7.728 (0.329)	0.30%	25.79 (0.626)	1.02%
Cluster 5: Pt or non-contributory employed, 2+ chi.	803.5** (4.379)	7.42%	-4.621 (-0.086)	-0.04%	142.1* (2.045)	5.60%	-8.550 (-0.190)	-0.34%
Cluster 6: FT care or other, 2+ chi.	1,396** (7.465)	12.90%	66.82 (1.120)	0.62%	283.4** (4.617)	11.16%	50.20 (1.021)	1.98%
Cluster 7: FT employed mostly, 1 chi.	-19.52 (-0.582)	-0.18%	488.2** (2.607)	4.51%	-9.971 (-0.538)	-0.39%	254.8+ (1.956)	10.03%
Cluster 8: FT employed mostly, married	1.693 (0.131)	0.02%	-43.07 (-0.527)	-0.40%	16.11 (0.744)	0.63%	102.5+ (1.739)	4.03%
Cluster 9: FT employed mostly, single	-90.27+ (-1.896)	-0.83%	7.748 (0.235)	0.07%	-4.165 (-0.126)	-0.16%	4.119 (0.211)	0.16%
Cluster 10: FT employed mostly, divorced	-32.78	-0.30%	-168.9* (-1.56%)		-7.573	-0.30%	-21.70	-0.85%

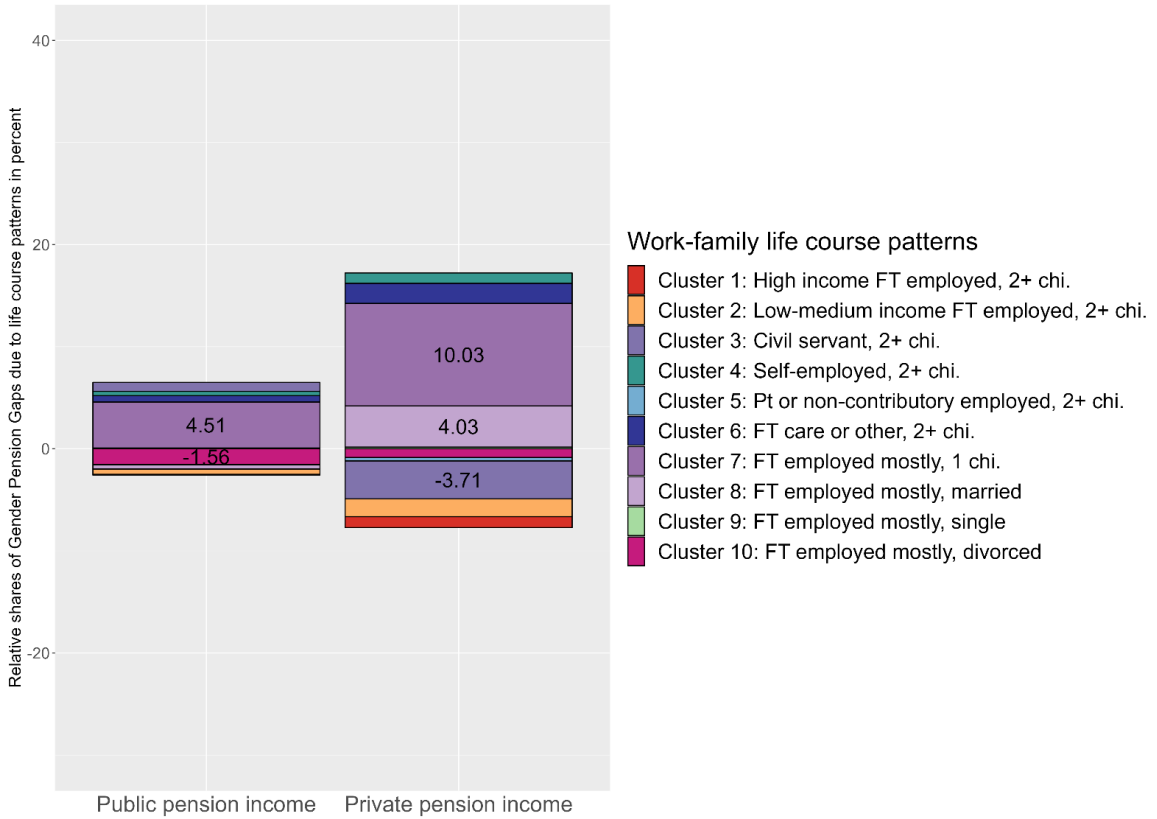
Notes: Results from regression-based Kitagawa-Oaxaca-Blinder decompositions. Based on normalized linear regression models with absolute independent pension income as dependent variable. Reference coefficients from the coefficients of regression models of the gender by which the respective life course is empirically dominated, see Table A4.6 for coefficients. Controlled for birth cohort. z-statistics in parentheses. Confidence intervals in Figure A4.7 and Figure A4.8** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Returns component. The only significant returns component in line with my expectations relates to cluster 10. I find that women's better pension returns compared to men's in life courses with long divorce trajectories reduce the gender gap in public pensions by 1.6% even though men in these life course groups spent longer years in highly rewarded employment types such as high-income careers and civil service (see Table 4.1). This is most likely due to the splitting of couples' pension claims of which women profit more and thus in line with Kreyenfeld et al. (2018).

However, other than that I do not find that the gender-sensitive redistribution implemented in the German public pension system reduces the GPG through gender differences in pension returns. Even for the clusters with high shares of unpaid care among mothers, I do not find that mothers benefit from higher pension income compared to fathers. This underlines the insufficiency of the pension system to combat gender inequality in old age due to the gendered division of care and stands in contrast to previous studies focussing on the motherhood penalty instead of comparing women's with men's outcomes (Mika & Czaplicki, 2017).

This finding must undergo validation regarding gender-specific unobserved heterogeneities as well as within-cluster heterogeneity because both can drive gender differences in returns. While I can rule out that gender-specific within-cluster heterogeneity is the reason for the returns components, I cannot control for additional characteristics that could drive gender differences in returns for the same life course (e.g., occupational segregation).

Figure 4.5 Shares of the GPG due to gender-specific returns for the same work-family life course pattern (returns component)



Notes: Value depicted = minimum significance level of 10%. Results from Table 4.2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

More generally, I find that a higher share of the gender gap in private compared to public pension is due to differences in returns for the same life course patterns (especially for clusters 7 and 8, see Figure 4.5). Given the high volatility in the working life courses of clusters 7 and 8, the gender-specific heterogeneity in characteristics such as earnings, part-time employment, and care work in these life courses most likely (see Table 4.1) drives the result. Private pensions thus reproduce such gender inequalities within life course patterns more strongly through more gender-unequal pension returns than public pensions (see Table A4.6).

Sensitivity analysis

The main results are robust for controlling for survey waves, nativity, and educational level (see Appendix A4.7), using different reference coefficients or weights in the KOB and excluding poorly assigned individuals from the clusters (available upon request). Using the original cluster solution instead of slightly equalizing the gendered distribution on the clusters to make the decomposition more robust and address the common support problem, as expected, the magnitude of the explained shares becomes larger. Excluding poorly assigned individuals from the cluster to make the life courses within clusters and between genders more similar suggests that the returns components were not driven by gender-specific within-cluster heterogeneity. In line with Chapter 3, effect sizes of explained shares tend to be slightly larger.

4.4.2.1 Analytical extensions: decomposition along the distribution and the Gender Coverage Gap

I include further analyses to address two major conceptual choices of the main analysis. First, the main analysis decomposes the GPG on the mean, though the relationship between life course patterns and the Gender Pension Gap is likely to differ along the distribution. Second, I include additional analyses to assess the sensitivity of the conceptual decision to include zero pensions as well as to analyse the Gender Coverage Gap, i.e., the gender differences in public and private pension receipt.

4.4.2.1.1 Decomposition along the distribution of public pension income

The literature on the Gender Pension Gap so far mostly focuses on the inequalities at the mean. Bonnet (2020, for France), Nolan et al. (2019, for Ireland), Ezeyi & Vujic (2017, for England) and Hänisch & Klos (2014, for total pension income in Germany) are the only ones decomposing the GPG along the distribution. Considering the distribution beyond the

mean is highly relevant from a policy perspective. Policymakers might want to target women at the bottom of the pension distribution since these women are most vulnerable to poverty if they have to rely on their own pensions.

Decomposing the GPG along the distribution exacerbates the problem of the low number of observations per gender for the clusters. I thus use quartiles instead of more fine-grained percentiles to mitigate this issue (see Appendix A4.3.1 for the results using deciles). Still, the results must be interpreted with caution as the limited number of observations restricts the robustness, particularly for smaller clusters.

I apply the RIF (recentered influence function) decomposition which reveals the explained and returns shares, similarly to the KOB, for each characteristic separately (Firpo et al., 2018). However, given that the available implementations of the RIF decomposition do not allow specifying the reference coefficient as flexible as the KOB⁴⁸, I present the results using the male coefficients as reference for all variables (see A4.3.1 for details). Given the large share of no private pension income, I only decompose the GPG in public pension along the distribution.

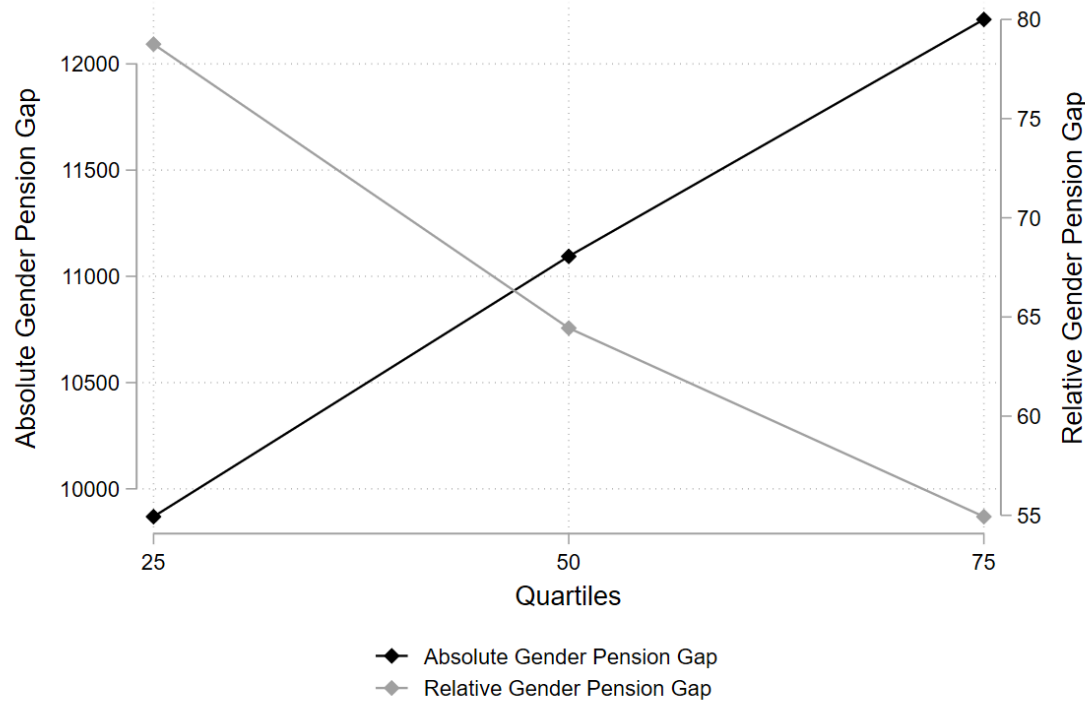
The absolute gap between men's and women's pension income increases from just below €10,000 at the lowest quarter of pension income, to above €12,000 at the 75th percentile (see Figure 4.6). Conversely, the relative GPG decreases from just below 80% at the 25th percentile to about 55% at the 75th percentile.⁴⁹ This corresponds to the notion of 'sticky floors' mostly applied to the phenomena of the gender wage gap being the largest at the bottom of the distribution (Arulampalam et al., 2007). The trend is similar for total pension income (Hänisch & Klos, 2014). This also highlights the need to disentangle the driver of

⁴⁸ This is at least the case for the `oaxaca_rif` command (Rios-Avila, 2020) used in stata.

⁴⁹ Similar patterns are documented for deciles in the Appendix (see Figure A4.10).

the gap at the bottom of the distribution since women with the lowest pensions not only face the largest inequalities but are also most likely to face old-age poverty (Bonnet et al., 2020).

Figure 4.6 Gender gap in public pensions over the distribution

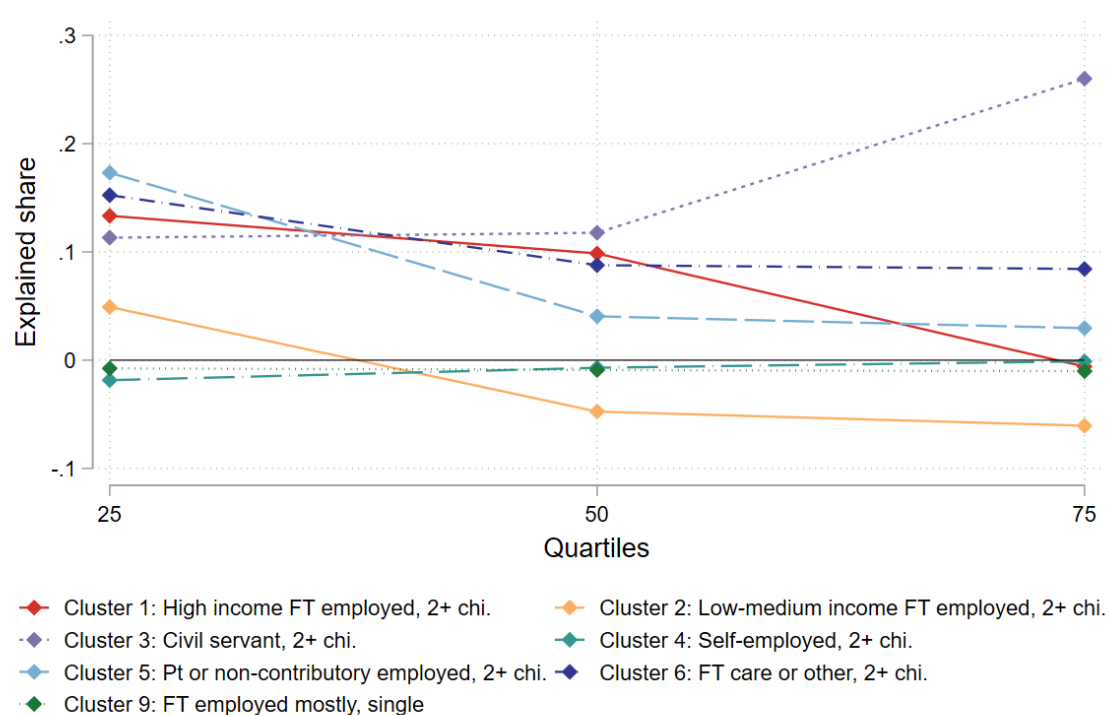


Notes: For the 25th, 50th and 75th percentiles. See Figure A4.10 for the deciles. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Decomposing these varying gaps across the distribution reveals that some associations vary strongly along the distribution (Figure 4.7 and Figure 4.8). For the explained shares, the decomposition at the mean revealed that the gap is reduced because more fathers compared to mothers have stable but low-income careers (cluster 2). However, this relationship is driven by the top half of the distribution. For the bottom of the distribution (1st quartile, i.e., the quarter with the lowest pensions), the association is reversed: about 5% of the gap is due to the lack of mothers with low-income careers. At the median, the 50th percentile, the relationship is already negative. This highlights that any full-time employment in the private sector, even if in low-income jobs, is important to increase women’s compared to

men's public pension – but only at the lower end of the pension distribution.⁵⁰ The association of the GPG with cluster 1 (parenthood combined with high-income careers) peaks at the bottom of the distribution and decreases along the distribution. For the life courses with careers in the private sector (clusters 1 and 2), my results align with the pattern identified for total pension income and employment years previously (Hänisch & Klos, 2014). However, my results confirm this pattern for both, high and low-income careers.

Figure 4.7 Decomposition over the distribution: selected explained shares



Notes: Results from RIF decomposition. Based on normalized models with absolute independent pension income as dependent variable. Only life course clusters with significant explained shares for at least one decile are depicted. Explained shares are shown in relative terms. Coefficients from male baseline models were used as reference. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The association of civil servant careers with the GPG peaks for the highest quartile: the lack of mothers with stable civil servant careers makes up for about one-fourth of the gap

⁵⁰ The analysis over deciles confirms this trend: a career in less stable and low-income jobs would be particularly relevant for decreasing the GPG at the very bottom of the distribution (1st decile, see Figure A4.11).

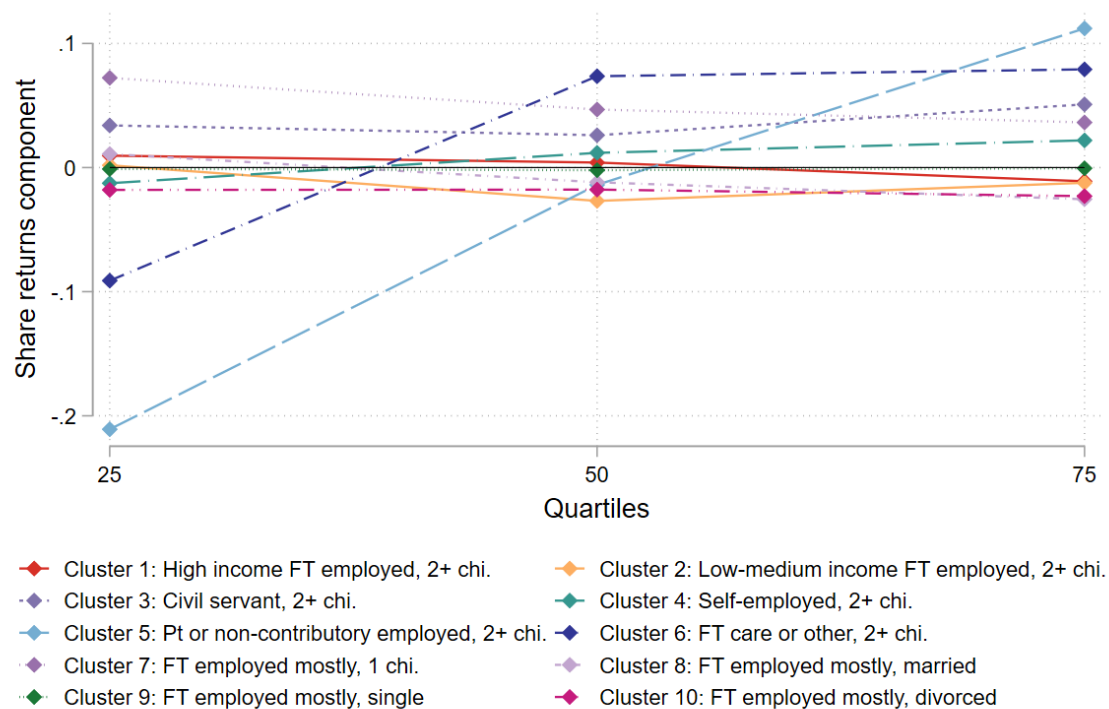
at the top quarter of the pension distribution alone. Thus, for equalising the gender gap in public pensions at the top of the distribution, equal access of women to civil service careers is by far the most important component.

The share of the gap that is attributable to the dominance of mothers in life courses characterized by unpaid care work or care-related employment interruptions and part-time employment is most pronounced at the lower end of the pension distribution and decreases along the distribution. While it remains significant for life courses of mothers who were consistently engaged in care work (cluster 6), it becomes insignificant for mothers who balance their caregiving responsibilities with part-time employment (cluster 5) from the median pension level onward. Simultaneously, however, the relationship between the returns component and the GPG along the distribution shows a reverse pattern (see Figure 4.8). At the bottom of the distribution mothers with part-time careers fare better than fathers in the same life course pattern, which more than compensates for the portion of the gap attributed to the much higher representation of mothers in this life course pattern (explained component). However, at the upper end of the distribution, the returns component turns positive. Consequently, women in cluster 5 and at the top of the pension distribution are worse off compared to men with similar life courses due to worse pension rewards, despite the reduced compositional component.

Considering both the compositional and the returns components jointly the results imply that mothers with part-time careers are only at a pension disadvantage compared to men at the top of the income distribution. Elements implemented in the pension system for women with such life courses (e.g., childcare benefits and upgrading of part-time employment reflected in the returns component) apparently help to equalize the overall disadvantage experienced by women at the lower end of the income distribution due to more mothers

reconciling their parenthood with part-time employment compared to men (explained part). A similar pattern is evident for mothers consistently engaged in unpaid care work, albeit to a lesser extent. The pension penalty of mothers doing unpaid care work is mitigated due to better returns compared to fathers with similar life courses, but not entirely cancelled out. However, from the median pension level onward, men receive higher pensions compared to women in cluster 6, thereby reinforcing the part of the GPG associated with cluster 6. These are important dynamics that become not evident in the analysis at the mean, partly because they cancel each other out due to opposing associations along the distributions (returns component).

Figure 4.8 Decomposition over the distribution: returns component



Notes: Results from RIF decomposition. Based on normalized models with absolute independent pension income as dependent variable. Only life course clusters with significant explained shares for at least one decile are depicted. Shares are shown in relative terms. Coefficients from male baseline models were used as reference. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

In summary, these results suggest that the key for reducing the GPG among individuals with the lowest public pension income is enabling mothers access to any full-time

employment. Jointly, 60% of the GPG at the bottom of the distribution are due to the majority of women reconciling their motherhood with continuous unpaid care work, or part-time employment after care-work-related interruptions, instead of having any stable full-time employment or civil service career (clusters 1-3, 5-6). For the center and top parts of the distribution it then becomes more relevant to create access for mothers to particularly well-rewarded employment types such as high-income jobs in the private sector or the civil service to decrease the GPG. This finding is overall in line with results for France (Bonnet et al., 2020) and links the results of the main analyses to previous evidence in Germany suggesting that the lack of female employment is the main driver of women's low pension income (Fasang et al., 2013).

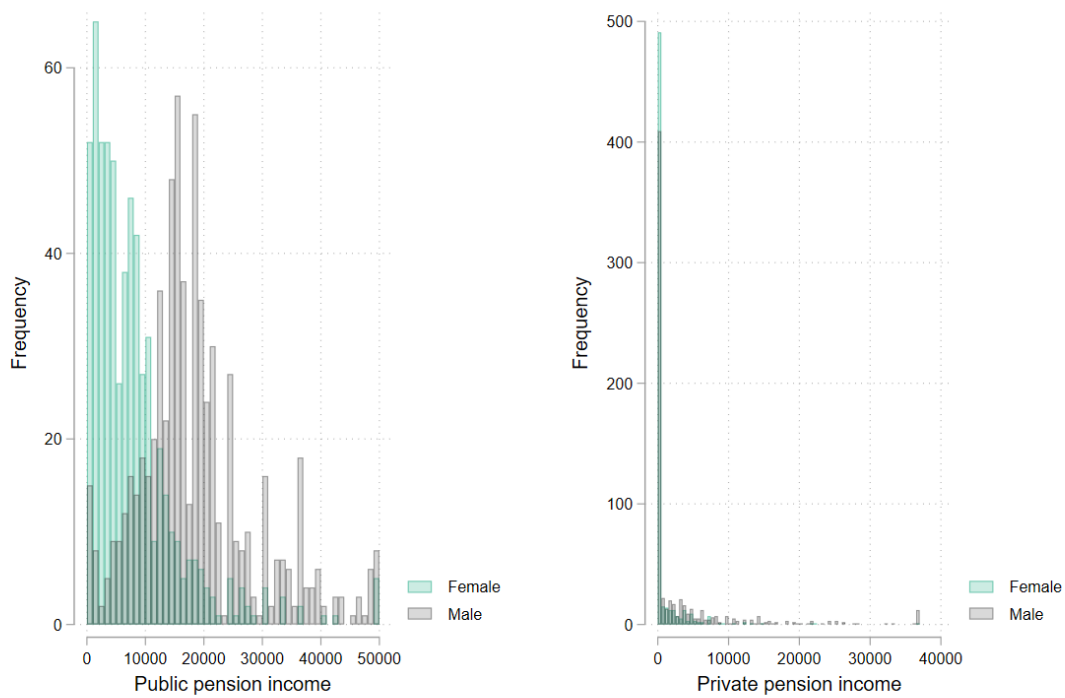
Moreover, the results beyond the mean suggest that women with women-majority life courses dominated by unpaid care work do profit from redistribution in the pension systems when it comes to pension income compared to men, but only at the bottom of the distribution. While mothers with part-time careers after care-related employment interruptions do even have small pension advantages compared to fathers with similar life courses due to better pension rewards, the disadvantage of mothers with continuous unpaid care work is not fully buffered through more favourable pension rewards. However, it has to be kept in mind that the reference at the bottom of the distribution is at a very low level with an annual pension of €12,531 (men's pension at the 25th percentile, not shown), which is only slightly above the threshold for the poverty risk (€11,749 for singles in 2013, Statistisches Bundesamt (Destatis), 2023). In fact, women's pension income is below the poverty risk even at the median and top of the distribution (€6,122 and €10,016, not shown), which highlights the relevance for policymakers to increase women's pension income not only at the bottom of the distribution.

4.4.2.1.2 The role of the Gender Coverage Gap

By including individuals with no pension income with a pension income of zero, the main analysis assesses the gender gap in pension receipt (whether individuals receive any pension income) and the gender gap in pension entitlements (the pension amount individuals receive) jointly. However, previous research shows that determinants of the gender coverage gap and the gender gap in pension entitlements among recipients differ (Even & Macpherson, 2004; Ezeyi & Vujic, 2017). Additionally, the conceptual decision to include zero pensions leads to zero-inflated distributions, particularly for private pensions (see Figure A4.9). Almost 96% receive an independent public pension, but only 34% receive their own private pension (Figure A4.13). OLS regressions, which are used as baseline models for the KOB, are not best equipped for modeling such zero-inflated distributions.

There are several ways to deal with zero-inflated distributions (see Chapter 2 for a general overview and Appendix A4.3.2 for a discussion applied to this analysis). Following the most popular approach in the literature if zero pension income was included and methodologically addressed (Bardasi & Jenkins, 2010; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Gardiner et al., 2016; Johnson et al., 1999; Zhao & Zhao, 2018, see Table 2.1), I decompose the gender gap in pension receipt (i.e., the Gender Coverage Gap) and the gender gap in pensions among pension recipients separately. Disentangling the mechanism for the gender coverage gap compared to the gender gap in pension entitlements among pension receivers helps to refine the main results, especially regarding private pensions.

Figure 4.9 Public and private pension distribution by gender



Notes: €1000 bins for public and €500 bins for private pensions. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

In the following, I first decompose the Gender Coverage Gap (for private pensions only). As the second step, I analyse the Gender Pension Gap among pension recipients only by excluding individuals without pension receipt.⁵¹

Private pensions

The Gender Coverage Gap in private pensions amounts to 44.5%. More than half of this gap is due to gender differences in work-family life courses (see Table 4.3, M3). The gender gap in private pensions among private pension recipients is smaller in relative terms but larger in absolute terms (58%, €4,422, see Table 4.3, M4) compared to the main model, when considering gender differences in access to and entitlement of private pensions jointly (76.6%, €2,539, see Table 4.2, M2). However, the subsample of private pension recipients

⁵¹ I do not consider the selection into pension receipt in the modelling approach for the second step. See Appendix A4.3.2 for challenges of the exclusion restriction of selection models for this analysis.

is very small and highly selective (291 men and 146 women) which is why the decomposition in M4 must be read with caution.

55% of the Gender Coverage Gap alone is attributable to the much higher share of women with life courses dominated by continuous unpaid care work (cluster 6): 38% of women without any private pension receipt have these life courses, but only 8% of women who receive own private pensions (see Figure A4.9). For the gender gap in private pensions among private pension recipients, however, continuous care work does not play a significant role (Table 4.3, M4). Surprisingly, the other female-typical life course pattern characterised by mothers reconciling their care work responsibilities with employment interruptions and part-time employment does not play a role for access to private pensions. Gender differences in the realisation of such life courses tend to be mostly responsible for differences in private pension income once access is achieved, even though the share is not significant.⁵²

⁵² The insignificance is most likely driven by the lack of statistical power due to the small subsample of private pension recipients. The effect size of the cluster 5 is in M4 even slightly larger compared to the main model. The interpretability is limited due to the low number of observations per cluster. Merging small clusters did not help but rather skewed the meaning of the clusters due to an increase of within-cluster heterogeneity.

Table 4.3 Decomposition results of the gender gap in private pension receipt and private pension among recipients

VARIABLES	Private pension income, main model (M2)				Gender Coverage Gap (M3)				Private pensions, recipients only (M4)			
	Absolute	Shares	Absolute	Shares	Absolute	Shares	Absolute	Shares	Absolute	Shares	Absolute	Shares
Average pension income/receipt men	3,316** (11.98)				0.434** (21.80)				7,635** -14	7,635		
Average pension income women	776.7** (7.222)				0.242** (13.85)				3,213** -8.246	3,213		
Absolute gender pension/ coverage gap	2,539** (8.549)	100.00%			0.193** (7.272)				4,422** -6.596			
Total part explained	379.9+ (1.676)	15.00%			0.101** (3.893)	52.55%			-18.29 (-0.0304)	-0.004137		
Total part unexplained (including constant)	2,160** (6.135)	85.00%			0.0914** (2.813)	47.45%			4,440** -5.131	1.0041368		
Sum by life course patterns		explained 13.39%		returns 9.50%		explained 52.30%		returns 9.08%		explained -8.62%		returns -7.92%
Cluster 1: High income FT employed, 2+ chi.	567.0** (4.731)	22.33%	-27.21 (-1.617)	-1.07%	0.0734** (3.628)	38.11%	-0.00304 (-1.113)	-1.58%	35.09 -0.116	0.79%	-179 (-1.521)	-4.05%
Cluster 2: Low-med. income FT empl., 2+ chi.	-217.8* (-2.188)	-8.58%	-44.41 (-1.340)	-1.75%	0.0293* (2.524)	15.22%	0.000716 -0.208	0.37%	-947.1** (-3.325)	-21.42%	-264.5+ (-1.715)	-5.98%
Cluster 3: Civil servant, 2+ chi.	-436.9** (-4.142)	-17.21%	-94.20* (-2.112)	-3.71%	-0.115* (-2.520)	-59.74%	-0.00655 (-0.756)	-3.40%	2.23 -0.149	0.05%	19.62 -0.302	0.44%
Cluster 4: Self-employed, 2+ chi.	7.728 (0.329)	0.30%	25.79 (0.626)	1.02%	-0.00197 (-0.812)	-1.02%	0.00641 -1.261	3.33%	72.99 -0.924	1.65%	-55.93 (-0.858)	-1.26%
Cluster 5: Pt or non-cont. employed, 2+ chi.	142.1* (2.045)	5.60%	-8.550 (-0.190)	-0.34%	0.00421 (0.209)	2.19%	-0.00592 (-1.620)	-3.07%	261.1 -0.863	5.91%	85.54 -1.201	1.93%
Cluster 6: FT care or other, 2+ chi.	283.4** (4.617)	11.16%	50.20 (1.021)	1.98%	0.106** (5.350)	54.86%	0.00663 -1.637	3.44%	99.94 -1.165	2.26%	153.9 -1.586	3.48%
Cluster 7: FT employed mostly, 1 chi.	-9.971 (-0.538)	-0.39%	254.8+ (1.956)	10.03%	-0.00235 (-0.582)	-1.22%	0.0156 -1.302	8.10%	58.04 -0.71	1.31%	96.85 -0.246	2.19%
Cluster 8: FT employed mostly, married	16.11 (0.744)	0.63%	102.5+ (1.739)	4.03%	0.00338 (0.837)	1.75%	0.00687 -1.341	3.57%	-22.2 (-0.441)	-0.50%	0.356 -0.00212	0.01%
Cluster 9: FT employed mostly, single	-4.165 (-0.126)	-0.16%	4.119 (0.211)	0.16%	0.00249 (0.796)	1.30%	0.000667 -0.32	0.35%	-53.18 (-0.724)	-1.20%	-18.44 (-0.346)	-0.42%
Cluster 10: FT employed mostly, divorced	-7.573 (-0.516)	-0.30%	-21.70 (-0.517)	-0.85%	0.00167 (0.807)	0.87%	-0.00389 (-1.135)	-2.02%	112.2 -0.922	2.54%	-188.8 (-0.901)	-4.27%
n male	670		670		670		670		291		291	
n female	604		604		604		604		146		146	

Notes: Results from regression-based Kitagawa-Oaxaca-Blinder decompositions. Based on normalized linear regression models with absolute independent pension income as dependent variable. Reference coefficients from the coefficients of regression models of the gender by which the respective life course is empirically dominated. Controlled for birth cohort. z-statistics in parentheses.

** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Another 38% of the Gender Coverage Gap is due to the lack of mothers with stable high-income careers (cluster 1) and 15% due to fewer mothers with unstable low-income careers (cluster 2, Table 4.3, M3). Given that the gender differences in clusters 1 and 2 are almost identical, equalising the share of mothers with high-income careers would reduce to Gender Coverage Gap in private pensions to a higher degree (twice as much). However, while the gender differences in stable high-income careers do not play a role for the gender gap in private pensions among pension receivers, the gender gap would be even larger, if as many mothers as fathers had low-income careers (Table 4.3, M4). Thus, while having low-income careers helps to create access to private pensions for women, it does not help to equalise the private pension entitlements between men and women but would even worsen the GPG among private pension receivers. The negative explained component in the main model for stable low-income careers of fathers (Table 4.2) is thus driven by the much lower pension entitlements of men with such life courses, once they have access. Furthermore, among the private pension recipients, women with low-income careers have higher private pension returns compared to men, which cushions the gender gap in private pensions (Table 4.3, M4, returns component).

Lastly, the results confirm the assumption that the negative composition part due to women's lack of civil service in the main model is driven by the lack of civil servants to invest in private pensions in the first place. The association of the civil service cluster with the Gender Coverage Gap is strongest among all life course patterns (-60%, cluster 3, Table 4.3, M3) while there is no association with the gender gap in private pensions among pension recipients (Table 4.3, M4).

Previous studies applying two-part models have found mixed results. My results somewhat confirm previous results (Johnson et al., 1999) that wage matters for both, gender

differences in access to as well as entitlement with private pensions and stand in contrast to the finding that wage plays a role for the gender gap in private pension receipt only (Ezeyi & Vujic, 2017). Other findings suggested that the gender differences in the years in (full-time) employment are more important for explaining the gender coverage gap, while wage differences are more important for explaining the Gender Pension Gap among pension recipients (Even & Macpherson, 2004). My results suggest that having a stable full-time employment career and not being engaged in continuous care work is important for gaining access to private pensions in the first place. But earnings differences matter at least as much since raising the share of women with high-income compared to low-income careers would decrease the Gender Coverage Gap twice as much. Wage differences, the employment type and continuous unpaid care do also matter for gender differences in private pension receipt in West Germany. Further, taking a holistic life course perspective the results of this analysis do not confirm the negative role of women's more years of part-time work that previous work has found for women's compared to men's private pensions receipt in England (Ezeyi & Vujic, 2017).

Public pensions

For public pensions, I do not decompose the Gender Coverage Gap because of the almost universal coverage of public pensions.

When excluding individuals without any public pensions, the gender gap in public pensions decreases slightly from €10,829 (58,26%, Table 4.2, M1) to €10,603 (56.02%, Table 4.4, M5) because mostly women are affected by zero pension (41 women compared to 12 men). Overall, the results from the main analysis are confirmed. Thus, the results for public pensions are not sensitive to pension receipt and are comparable to studies that exclude individuals without any pension claims (see Table 2.1).

Table 4.4 Decomposition results of the gender gap in public pension among public pension recipients

VARIABLES	Public pension income - main (M1)				Public pensions, recipients only (M5)			
	Absolute	Shares	Absolute	Shares	Absolute	Shares	Absolute	Shares
Average pension income men	18,588** (47.46)				18,927** (49.02)			
Average pension income women	7,759** (24.52)				8,324** (25.48)			
Absolute Gender Pension Gap	10,829** (21.51)				10,603** (20.96)			
Total part explained	4,452** (9.624)	41.10%			4,592** (9.697)	43.31%		
Total part unexplained	6,377** (12.29)	58.90%			6,011** (11.90)	56.69%		
Sum by life course patterns		explained 40.46%		returns 3.93%		explained 43.00%		returns 3.50%
Cluster 1: High income FT employed, 2+ chi.	482.3** (3.746)	4.45%	-6.657 (-0.228)	-0.06%	490.1** (3.981)	4.62%	-20.65 (-0.715)	-0.19%
Cluster 2: Low-medium income FT employed, 2+ chi.	-243.0* (-2.174)	-2.24%	-54.78 (-0.943)	-0.51%	-269.0* (-2.544)	-2.54%	-37.15 (-0.626)	-0.35%
Cluster 3: Civil servant, 2+ chi.	2,178** (7.496)	20.11%	94.34 (1.289)	0.87%	2,225** (7.350)	20.98%	90.57 -1.249	0.85%
Cluster 4: Self-employed, 2+ chi.	-94.98+ (-1.675)	-0.88%	46.82 (0.749)	0.43%	-97.70+ (-1.713)	-0.92%	52.55 -0.874	0.50%
Cluster 5: Pt or non-contributory employed, 2+ chi.	803.5** (4.379)	7.42%	-4.621 (-0.0864)	-0.04%	910.1** (4.785)	8.58%	11.32 -0.227	0.11%
Cluster 6: FT care or other, 2+ chi.	1,396** (7.465)	12.90%	66.82 (1.120)	0.62%	1,432** (7.465)	13.51%	68.74 -1.192	0.65%
Cluster 7: FT employed mostly, 1 chi.	-19.52 (-0.582)	-0.18%	488.2** (2.607)	4.51%	-6.011 (-0.214)	-0.06%	383.2* -2.195	3.61%
Cluster 8: FT employed mostly, married	1.693 (0.131)	0.02%	-43.07 (-0.527)	-0.40%	2.827 (0.270)	0.03%	-19.13 (-0.235)	-0.18%
Cluster 9: FT employed mostly, single	-90.27+ (-1.896)	-0.83%	7.748 (0.235)	0.07%	-105.3* (-2.127)	-0.99%	-0.844 0.0273	-0.01%
Cluster 10: FT employed mostly, divorced	-32.78	-0.30%	-168.9*	-1.56%	-23.05	-0.22%	-157.9*	-1.49%
n male	670		670		658		658	
n female	604		604		563		563	

Notes: Results from regression-based Kitagawa-Oaxaca-Blinder decompositions. Based on normalized linear regression models with absolute independent pension income as dependent variable. Reference coefficients from the coefficients of regression models of the gender by which the respective life course is empirically dominated. Controlled for birth cohort. z-statistics in parentheses.

** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Differences in the results compared previous literature, most importantly regarding care-work and part-time employment-dominated life courses, are not due to conceptual decisions on how to deal with pension receipt. On the contrary, the explained shares tend to increase even slightly, especially for the two female-dominated clusters (5 and 6).

4.5 Discussion

This study applied an innovative two-step approach and rich linked survey-register data to assess how work-family life course patterns relate to the gender gap in public and private pensions in West Germany. In summary, I contribute to the recent literature on the GPG, by decomposing the GPG based on work-family life course patterns including annual relative earnings (I.), and differentiating between mechanisms shaping the GPG in public and private pensions (II.). The research design enables discussing in detail the implications of pension privatisation, concerning work-family interdependences. The first step of the analysis reveals high levels of gender segregation on work-family life courses. The segregation on the differently well-rewarded life course patterns makes up for 40% of the gender gap in public and 13% of the gap in private pensions. The SA-KOB decomposition shows that only life courses of parents with two or more children drive the gender gap in pension income since only these life courses are highly gendered. Four main findings stand out.

First, 20% of the gap in public pensions alone is due to almost fully female life courses in which mothers drop out of the labour force for the rest of their lives after childbirth or a shorter period before reconciling care work with part-time employment (similar associations for private pensions). Previous studies highlighted the lack of women's full-time employment as the main reason for the GPG in Germany (Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014). Considering unpaid care work linked to parenthood from a holistic life course perspective on top of formal employment shifts

the focus to the relevance of dividing labour between genders for the GPG. Given the empirical evidence suggesting that childbirth drives gender inequalities along the pre-retirement life course (Aisenbrey et al., 2009; Boeckmann et al., 2015; Budig et al., 2016; Gangl & Ziefle, 2009), the results confirm this inequality to be reproduced in old age by the German pension system. Equalising the division of unpaid care work is particularly important to mitigate the gender disparities in public pensions at the bottom of the distribution, where the GPG is most pronounced and the poverty risk is highest.

The compositional gender difference in engaging continuously in care work is also the main reason for the gender gap in private pension coverage, while the female-typical part-time life course type tends to mostly play a role for gender differences in private pension incomes among private pension recipients. These different mechanisms through which both women-majority life course types relate to gender inequalities in private pensions will have to be considered and further assessed in future studies of gender inequalities in private pensions.

Second, gender inequalities in earnings, once women are active in the labour force, are relevant for the GPG as well. This complements previous research in West Germany highlighting the role of (full-time) employment for the GPG without considering earning positions (e.g., Fasang et al., 2013; Frommert & Strauß, 2013). If mothers have access to full-time employment, it is to lower-income jobs which leads to lower pension income. While the lack of mothers with high-income careers drives the GPG, the gap would be even larger if the share of mothers with unstable low-income careers were as high as that of fathers (particularly for private pensions). Thus, only access to high-income employment ensures mothers' financial independence in old age. Even though previous data has suggested the opposite (Fasang et al., 2013), full-time employment itself is overall not sufficient.

Low-income employment improves, however, mothers' public pension income at the lower end of the pension distribution. But to increase women's pensions compared to men's along the whole distribution, mothers need access to better-rewarded employment life courses. Additionally, some studies found that parts of the GPG are predominantly due to differences in pension returns (returns component of decomposition) for years worked full-time (Cordova et al., 2022; Hänisch & Klos, 2014). My results suggest that at least part of these findings is due to the gender differences in earnings over the life course. Not differentiating between earnings positions, in Chapter 3 we did not find the lack of mothers with full-time employment careers being associated with the GPG in West Germany. The results of this Chapter suggest that the contrary associations for the high- and the low-income life course patterns (positive compositional component of cluster 1 and the negative one of cluster 2) offset each other when merged as overall full-time employment as in Chapter 3. Clearly, not considering earning differences risks concealing the important role of the gender wage gap for gender inequalities in pensions.

Third, these earnings-related associations are much stronger pronounced for the gender gap in *private* pensions, confirming the stronger reproduction of earning inequalities in the second and third pension pillars even compared to the heavily earnings-related public pension system in Germany. The association of mothers' lack of high-income careers with the gender gap is five times higher in private compared to public pensions and explains 22% of the GPG alone. This finding confirms the assumption of Bardasi & Jenkins (2010) that for private pensions the driving mechanisms for the high gender inequalities in access and entitlements are characteristics beyond labour market participation. The most straightforward reason is that individuals with stable high-income careers are much more likely to accumulate financial resources that could be invested in private pensions compared to care-dominated lives or life courses with low-income careers, whereas public pensions

rely less on monetary investments. Indeed, the similar level of gender difference in high-income compared to low-income careers of parents attributes to more than twice of the gender difference in access to private pensions. However, equalising the access of mothers to low-income careers would still help close the Gender Coverage Gap. At the same time, it would increase the gender gap in private pension benefits among recipients since low-income careers are detrimental to the entitlement with private pensions.

Fourth, the returns component of the SA-KOB suggests that the only efficient pension policy tool for reducing gender inequality in pension income is the splitting of pension claims after divorce (confirming Kreyenfeld et al., 2023). Otherwise, I do not find that the GPG at the mean is reduced by better rewards of mothers compared to fathers with the same life course pattern due to gender-sensitive adjustments in the pension rewarding structure. However, at the bottom of the distribution, women benefit from better pension returns compared to men for the two women-majority life course patterns characterized by continuous care work or part-time employment. Redistributive elements such as childcare benefits and topping-up low EPs thus lead to female public pension advantages which mitigate the compositional disadvantages of women for these life courses and thus the Gender Pension Gap among individuals with low pensions. Still, they cannot make up for the pension loss of women with female-typical life courses due to the lack of careers in private employment or civil service – the GPG is highest at the bottom of the distribution. Consequently, to reduce the overall GPG, redistributive elements would have to be enhanced.

This study is subject to several limitations due to data constraints. First, given that the results highlight the relevance of more detailed labour characteristics over the life course, such as earnings, other pronounced gendered labour market inequalities, like gender-specific occupational segregation, might also contribute to the GPG. Consequently, future

research should assess the impact of gender segregation on occupations or firm size, especially amidst the privatisation of pensions, given that both are related to gendered private pension income in the UK (Ezeyi & Vujic, 2017; Gardiner et al., 2016) and a much smaller share of the gender gap in private pension is explainable by the available data in this study. Secondly, some of the results are subject to a very low number of observations in the linked-survey-admin data. This leads to limited statistical power and robustness, especially for small clusters (e.g., cluster 10 characterised by divorce) and for the extension analyses along the distribution and among pension recipients only. Hence, future research might want to replicate the analysis for contexts providing richer data (e.g., more recent cohorts or countries with higher private pension coverage).

A limitation of the holistic approach is the within-cluster heterogeneity that risks concealing associations of specific life course aspects with the GPG. For example, it might be crucial to consider the age at divorce or the first care work period as well as the duration of care work to assess the role of these heterogeneities within clusters 5, 6 and 10 for gendered pension inequalities. Thus, isolating and systematically assessing different aspects of life courses might be beneficial to reveal such associations and contribute to our understanding of links between life course inequalities and the GPG (see Chapter 5).

This Chapter could also be extended by (i.) applying the design to countries with different pension systems (e.g., Netherlands, UK, Denmark, or Sweden), (ii.) including cohort and period comparisons to better understand the dynamic of the GPG over time, and (iii) simulating pension income for future cohorts in old-age. In particular, taking advantage of the unique context in Germany and comparing the dynamics across East and West German cohorts whose life courses were shaped by very different welfare contexts during the division of Germany, but are rewarded by the same pension systems after the reunification allows disentangling the interaction of the same pension system with different levels of

gendered life courses (Möhring & Weiland, 2022; Rowold, 2022). However, this was not feasible for this study using SHARE-RV due to the small number of observations for East Germany. Other country comparisons were not feasible due to the unique benefit of SHARE-RV which provides earnings over the entire life course.

The results of this study highlight the relevance of the traditional gender division of labour for the reproduction of gender inequalities in old age. Specifically, the GPG largely resulted from German policies that initially incentivised a traditional division of labour and subsequently provided inadequate pension rewards for life courses with many years of unpaid care work. The strongly gendered life courses alone could not drive the gender gap in pensions if the German public and private pension systems would reward these different life courses equally. This shows how the pension system in later life interacts with the welfare regimes at early life stages in shaping the high level of GPGs. Ignoring this inter-temporal link of welfare policies, the current pension system particularly punishes women who acted in line with the gender norms and welfare state contexts and devoted most of their lives to fulfilling socially unaddressed care needs.

In old age, these women must rely on the redistribution in the pension system to achieve financial independence. However, my results confirm previous studies highlighting the inability of public pensions to buffer the reproduction of gender inequalities over the life course in old age (e.g., Frommert & Strauß, 2013; Grabka et al., 2017; Horstmann et al., 2009). Only women at the bottom of the pension distribution benefit from redistributive elements given their overall low pension level. Still, the childcare benefits equivalent to one or two years of employment are nothing else than a mirage for the high shares of women who have provided on average 35 years of unpaid care work until age 65. The only equalizing influence of pension policies on the GPG, both at the mean and along the entire distribution, is observed for the intra-couple pension claim splitting after divorce, which

makes it the most attractive policy tool for equalizing pension rewards for gender-typical life course patterns. Such ex-ante policies are important to ensure independent pension income for women whose life courses were shaped in conservative contexts that incentivised a traditional division of labour between men and women.

However, for future cohorts, a more sustainable approach to prevent high GPGs is to implement welfare policies aiming to equalize the gendered division of labour and gender inequalities in the labour market. Firstly, facilitating an equal division of unpaid and paid work among men and women is important to guarantee access for women, especially mothers, to the labour force in the first place. However, particularly in the light of pension privatisation, this is most likely not sufficient to equalize pension income across genders: earning inequalities are especially relevant for gender inequality in private pensions. Given the current privatisation of pensions as well as the constant level of the gender wage gap in Germany, gender inequalities in the labour market such as earning gaps, are likely to be more strongly reproduced for future retirement cohorts. Thus, secondly, it is of high importance to combat inequalities in the labour market beyond labour market participation once women are employed. Additionally, redistribution measures could be implemented in private pension schemes to buffer the reproduction of remaining gender inequalities. Such an attempt has been made in Germany with the private but state-subsidised Riester program. However, even though more women than men are Riester savers, they receive only about half the pension income as men from Riester (Geyer et al., 2021).

Whether the overall increasing labour market participation of women in Germany can still lead to a decrease of the gendered pension income inequality in the future will largely depend on whether the German welfare state succeeds in equalizing gendered labour market inequalities and how low-income and part-time employment, as well as unpaid care work, will be rewarded in the pension system.

Acknowledgements

I want to thank Nicole Kapelle for very helpful feedback on previous versions of this Chapter. I am also grateful for helpful comments from the social demography reading group participants at the Department of Sociology and Nuffield College, particularly Alicia Garcia Sierra, Antonino Polizzi, Christiaan Monden and Zachary Van Winkle. All remaining errors are mine.

5 Full-time employment is all that matters? Decomposing Gender Pension Gaps based on relevant life course features in Germany & the Netherlands

Abstract

Gender Pension Gaps (GPG) are so far either analysed based on selective sets of life course summary measures, mostly the duration in full-time employment, or based on life course typologies. Whereas the first approach neglects other dimensions such as the timing or ordering of events over the life course of other work and family categories, the latter risks concealing specific attributes of life courses particularly relevant for pension income inequality. This study contributes by, first, applying the Life Course Feature Selection to identify which work and family life course aspects are the most important pension income predictors. The second step decomposes the GPG based on these relevant life course aspects, revealing how they are related to gendered pension income inequality. Comparing the results across the different pension regimes of Germany and the Netherlands provides insights on whether similarly gendered life course aspects are differently related to GPG across different pension designs.

Results show that aspects of all life course dimensions (duration, timing, ordering, and complexity) play a role for pension inequality. Life course aspects related to unpaid care work, particularly the years engaged in care work, are among the most relevant pension predictors in both countries. Contrary to what previous literature suggested, the duration of full-time employment plays a much less important role.

The non-parametric Ñopo decomposition reveals that the extremely gender-specific combination of the observed life course experiences, particularly regarding care work, is

the main contributor to the Gender Pension Gap in both countries. Highlighting the relevance of specific life course aspects allows policy makers to employ more targeted policies to prevent the reproduction of gender inequalities over the life course in old age. Furthermore, the results advice to consider life course proxies beyond the years spent in (full-time) employment as well as the extremely gender-specific combination of life course aspects.

Key words: Gender Pension Gap, life course, feature selection, pension inequality, pension system, decomposition

5.1 Introduction

Research recently highlights that pension income and more specifically, gender gaps therein, are highly determined by life courses (Foster & Ginn, 2018; Madero-Cabib & Fasang, 2016; Rowold et al., 2023). Key aspects of life courses are the (i) duration, (ii) timing, and (iii) order of events (Halpern-Manners et al., 2015; Liao et al., 2022; Studer & Ritschard, 2016). However, the literature on Gender Pension Gaps (GPG) is divided into two main approaches. The first, and most common empirical stream in the literature, focusses on one life course aspect, namely the duration in specific life course states, mostly the years spent working full-time (Bardasi & Jenkins, 2010; Bonnet et al., 2020; Cordova et al., 2022; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Hänisch & Klos, 2014; König et al., 2019; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018). Generally speaking, these studies find that the lower years women have spent being (full-time) employment compared to men is the main reason for the high GPGs. However, they usually do not consider other life course complexities beyond the duration in (full-time) employment that might play a role (see Table 3.1). The evolving trend towards more complex careers (Barbieri & Filippo, 2022; Van Winkle & Fasang, 2017) puts added pressure on pension systems to adapt to increasingly diverse life courses. Additionally, pension privatisation challenges the predictability of pension income and the notion that continuous full-time employment is the main predictor of pensions. Considering life course complexities, such as the timing or ordering of family-related employment interruptions, becomes increasingly vital in both conceptualizing pensions and conducting analyses.

The second empirical stream in the literature takes a holistic perspective and applies sequence analysis to assess the association between typical life course patterns and pension income (e.g. Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022) or the GPG directly (Rowold et al., 2023). Hereby, life course trajectories are clustered into groups of similar life courses and then these groups are used as covariates in regression or

decomposition analyses. However, since the life course clusters often entail some degree of within-cluster heterogeneity concerning the duration, timing, or ordering of life course events, they risk concealing which specific life course aspects matter most for gendered pension income inequality.

This study is located between both approaches and aims to address their limitation of concealing important associations of the GPG with single life course aspects by introducing an innovative 2-step approach to the field of gendered pension income inequalities. As the first step, I use Machine Learning tools to identify which specific aspects of life courses are important determinants of pension income. As the second step, the GPG is decomposed using matching based on these relevant aspects.

I assess two main questions:

- 1) Which work and family life course features are relevant predictors of pension income inequality?
- 2) How are these life course features associated with the gender gap in pension income?

Combining the Life Course Feature Selection with decomposition techniques for the first time in the field of old age inequalities, this study contributes to the literature in at least two ways.

First, applying the Life Course Feature Selection (Bolano & Studer, 2020) allows the identification of the most important variables (i.e., ‘features’) for pension income covering all working and family life course aspects (duration, timing, ordering, complexity). This contributes to the first stream of literature in revealing whether the duration in different working states, primarily full-time employment, is indeed the best predictor of pension income when considering other life course aspects as well. Furthermore, the feature selection might uncover life course aspects the first literature stream has neglected so far

or that have been obscured in heterogeneous life course clusters by the more holistic life course approach of the second stream of literature. For example, studies using sequence analysis have found that female-typical life courses with long periods of care work are an important reason for women's lower pension income (Madero-Cabib & Fasang, 2016; Chapters 3 and 4). While the first literature stream has not considered care work life course aspects yet, it is still unclear whether the duration, the timing (e.g., at what age did the first care work period start) or the ordering of care work matters most. Furthermore, as the systematic and empirical identification of key pension predictors has not been undertaken previously, this step may unveil critical factors influencing pension income that have not been explored in research previously. Empirically uncovering pension predictors has gained greater relevance as the pension income calculation becomes more opaque, particularly due to the growing prevalence of private pensions.

Secondly, I decompose the Gender Pension Gaps based on the life course variables that have been identified as the most important pension predictors in the first step. The Kitagawa-Oaxaca-Blinder (KOB) decomposition is commonly employed to analyse mechanisms shaping GPGs (e.g. Bardasi & Jenkins, 2010; Bonnet et al., 2020; Even & Macpherson, 2004; Frommert & Strauß, 2013; König et al., 2019; Zhao & Zhao, 2018). However, it has been criticised for comparing men and women who are not comparable based on their characteristics, i.e., for whom there is no counterfactual (common support problem, Ñopo, 2008). Comparing men and women without a counterfactual introduces the risk of biasing the decomposition components (Hamjediers & Sprengholz, 2023).

I argue that this concern is particularly pertinent for outcomes where life courses are the covariates of interest, such as the GPG since life course aspects are likely to differ largely in their prevalence across genders given highly gendered work-family life courses (e.g., Madero-Cabib & Fasang, 2016; Uccheddu et al., 2022, Chapters 3 and 4). For example,

many women might transition from stable full-time careers to unpaid care work during their 30s, but almost no men. As a result, for these women, there is no male counterfactual to compare their pensions with. To overcome this shortcoming and reveal additional mechanisms related to gender-exclusive life course experiences, I apply the Ñopo (2008) decomposition which uses matching techniques for a non-parametric decomposition of group differences in outcomes. It decomposes the GPG into the two parts of the KOB (compositional and returns effects) but over the common support (i.e., matched men and women only) and contributes additionally by quantifying which parts of the GPG stem from highly gender-specific (life course) segregation (e.g., almost no men transiting from full-time employment to unpaid care work). The latter reveals which parts of the GPG are attributable to the gender-specific combination of life-course experiences which has not been assessed previously.

Together, this two-step approach disentangles whether the association of the years being full-time employed with the GPG has been overestimated previously and which part of GPGs is due to the duration, timing, or ordering of other relevant work and family life course aspects instead, and gender-exclusive experiences thereof.

I apply this innovative methodological combination to West Germany and the Netherlands. The country comparison design allows analysing the interrelation of similar welfare state contexts in place along the life course and different pension systems at the end of one's life course for shaping the gender inequality in pension income. Assessing whether the results differ by the pension systems (e.g., Bismarckian versus Beveridgean and low versus high coverage of occupational pensions) is important due to the current shift "toward the privatization and marketization of pensions" (Ebbinghaus, 2015, S. 56) in most European countries. While the research design does not allow for a detailed discussion of the impact of single pension policy elements, it reveals whether the most important pension predictors

as well as in their association with the Gender Pension Gap vary across pension systems, holding the level of gender inequalities in life courses stable. Whereas a detailed decomposition of GPGs for the Netherlands is still missing, for Germany, this study enhances previous decomposition studies of the GPG (Frommert & Strauß, 2013; Hänisch & Klos, 2014) as well as more indirect approaches discussing gender inequality in pension income from a holistic life course perspective (Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022).

Discussing which pension design rewards or penalizes which life course aspects allows policymakers to better match pension policies to (work-family) life course inequalities that were shaped in the former or current welfare regimes and thus prevent the reproduction of inequalities in old age.

5.2 Background

5.2.1 Life courses and pension outcomes from a gender perspective

The life course approach highlights that individuals' biographies unfold within the context of opportunities and constraints that are shaped by the socio-historic context, the interdependences with other individuals (linked lives) as well as the timing and ordering of events along life courses (Elder et al., 2003). For this study, three elements of the life course paradigm are pertinent for the overall understanding of gendered life courses. First, the socio-historic context is relevant for understanding how policies and norms shaped the gendered division of labour and, consequently, gendered life courses. Second, the notion of linked lives is important as it highlights the relevance of considering not only employment, but other types of unpaid work linked to other individuals, such as child or elderly care, as well as family lives more general for the emergence of gendered inequalities over the life course. Recent studies show that work-family interdependences are highly

gendered – interdependences are low or do not exist for men but high for women (Fasang & Aisenbrey, 2022; Madero-Cabib & Fasang, 2016; Rowold et al., 2023).

Third, the Cumulative Advantage/Disadvantage (CAD) theory helps understand how gender inequalities over the life course are reproduced. It refers to the dynamic that (dis)advantages of individuals or groups, and thus inequality, accumulate over time, for example over life courses (Dannefer, 2003; DiPrete & Eirich, 2006; O’Rand, 1996). Pension systems reward or penalize different life course patterns differently (Ginn et al., 2001; Madero-Cabib & Fasang, 2016; Möhring, 2021). Given its effective conceptualization of life courses and their complexities, it is thus not surprising that the life course approach has been highlighted as a leading theoretical lens via which to examine gender inequalities in pension income (Foster & Ginn, 2018).

Applied to the GPG, the strength of the life course approach is to reveal how unequal pension regulations reward gender-specific life courses or aspects which were formed by other public policies decades before pension receipt (Madero-Cabib & Fasang, 2016; Chapter 3). It highlights how different welfare state policies across several decades interact with each other to reproduce or buffer inequalities: “Patterns of inequality within and among cohorts emerge over time as products of the interplay between institutional arrangements and individual life trajectories.” (O’Rand, 1996, S. 230). Applying the path-dependency of CAD to economic outcomes in old age, individuals' previous work and family characteristics, as well as their interrelation, impact their economic situation at old age a) indirectly by influencing access to specific career and family characteristics that are relevant for economic outcomes and b) directly (Halpern-Manners et al., 2015), for example, if they are differently rewarded in the pension system.

While life courses can be seen and operationalized as holistic entities, focussing on key aspects of life courses can disentangle which specific elements of life courses reproduce

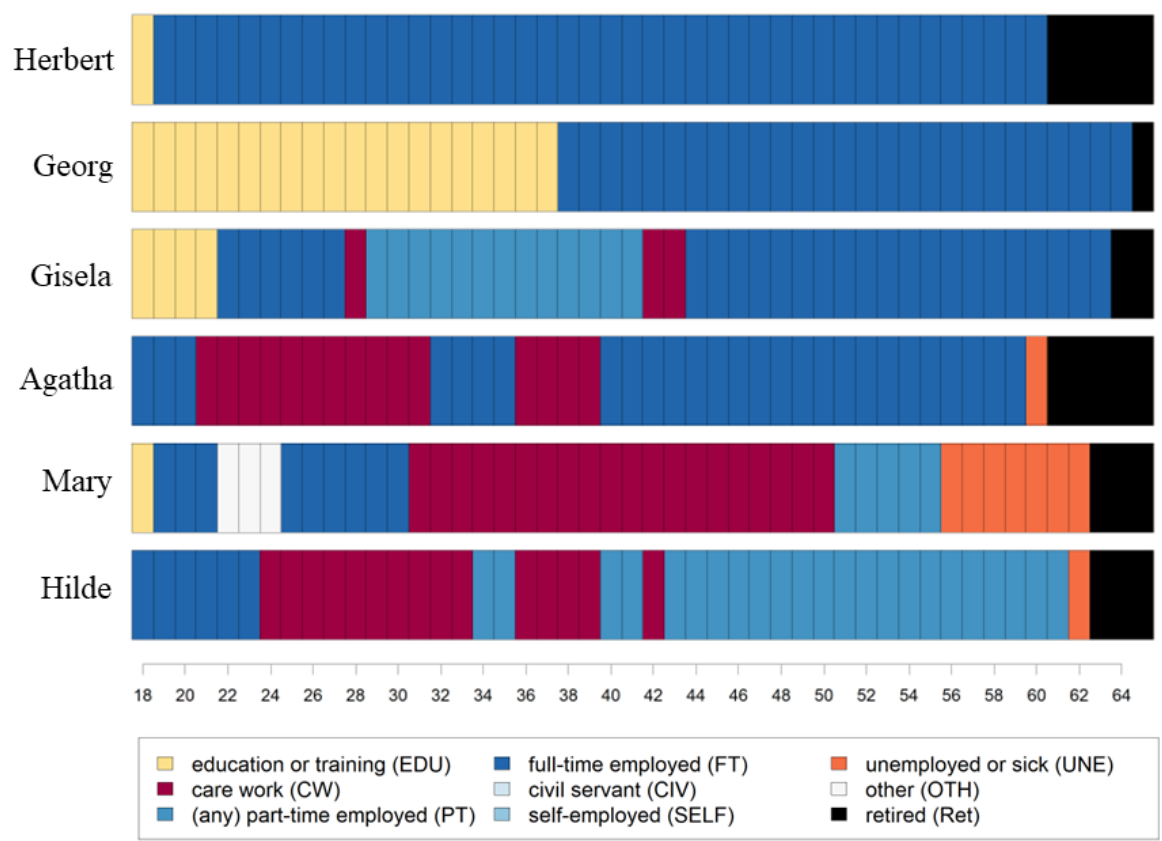
(dis)advantages at later ages which is important to understand the accumulation of (dis)advantages. The life course aspects considered in this study cover the (i) duration, (ii) timing, and (iii) order of life course states or transitions (Halpern-Manners et al., 2015; Liao et al., 2022; Studer & Ritschard, 2016), as well as its overall complexity (Elzinga, 2010). Even though there are strong theoretical reasons for why different life course complexities play a role in pension accumulation, it is often not clear which specific life course aspects and elements these are. For example, is the significance of events over the life course determined more by their chronological order or their timing? If the latter holds, then which specific timing is crucial, and how should it be operationalized?

Given the multitude of aspects related to different life course states, I cannot discuss all of them in detail. Instead, I describe the different life course aspects and discuss potential example mechanisms related to (gendered) pension income. Figure 5.1 depicts example work trajectories from ages 18-65 selected from the analysis sample to visually introduce the different life course aspects and their theoretical relevance.

(i) *duration*. The life course aspect ‘*duration*’ captures the time spent in the life course states. For example, Herbert was full-time employed for 42 years and George for 27 (Figure 5.1). Applied to pension income inequality, focussing on the duration of (full-time) employment appears straightforward because most pension systems in Europe are built “based on the norm of continuous full-time employment” (Leitner, 2001, S. 103). In line with this, the duration of full-time employment is arguably a good proxy for the continuous employment life courses of Herbert and George and has been used predominantly in the empirical literature (see section 5.2.2). Additionally, given the much lower (full-time) employment rate of women (e.g., Kreyenfeld, 2015), women’s duration in full-time employment will be lower on average, for example, Mary and Hilde were less than 10 years full-time employed (Figure 5.1). Consequently, and based on the previous empirical

literature, it appears straightforward to assume that the duration of full-time employment is the most important predictor of pension income inequalities and the gender gap therein (*expectation 1a*).

Figure 5.1 Example work life courses



Note: Example working life courses selected from the analysis sample (see section 5.3.1). The colour code shows which working life states the individuals experienced for each year, spanning ages 18-65.

According to the duration of full-time employment, the life courses of George, Gisela, and Agatha are very similar since they spent almost the same years being full-time employed. However, Gisela and Agatha did not experience continuous full-time employment careers comparable to George, but several care-related employment interruptions and different work states. Thus, relying on the duration of employment as the majority of the literature has done simplifies the more complex life courses of Agatha, Gisela, Hilde, and Mary when it comes to the frequency, order, duration, and timing of employment interruptions and types. This goes back to the “Steady Worker Bias [which] conceptualizes the career as a

relatively sustained sequence of full-time jobs of longer as opposed to shorter durations over the adult life span” (O’Rand, 1996, S. 235). However, the duration of full-time employment is systematically less applicable to “women’s more interrupted, mobile, and truncated work career patterns” (O’Rand, 1996, S. 235). For example, Kuivalainen et al. (2020) found in Finland that the length of working life is particularly related to pension income for men, but less so for women. Especially in conservative, liberal, or Southern European welfare states, there is a clear pattern of gender-specific life courses: most men have continuous full-time employment careers, while women’s careers are characterized by unpaid care work, other employment interruptions or part-time employment (e.g., Madero-Cabib & Fasang, 2016; Uccheddu et al., 2022; Chapter 3). Gender-sensitive research on old age highlights the relevance of family life courses as well as the interdependences between work and family lives for women’s pension income specifically (e.g., Fasang, 2010; Frericks, 2020; Ginn et al., 2001). Additionally, the pension privatisation challenges increasingly that continuous full-time employment is the main predictor of pensions, and the predictability of pension income more generally. Thus, in line with the life course paradigm, other life course aspects beyond the time spent being (full-time) employed (timing, ordering, complexity, duration in other work and family states) are likely to play a role for pension income, directly or indirectly, and the gender gap therein (*expectation 1b*).

(ii) *timing*. One of these other aspects, ‘*timing*’, considers at what time or age individuals were in a specific state or experienced a transition. For example, Agatha had her first care period (i.e., the first occurrence of doing care work in her life course) at age 21, Hilde at 24, and Mary at 31 (Figure 5.1). For pension income, the *timing* of retirement matters directly since pension claims are deducted (in the case of early retirement) or subsidized (in the case of postponing retirement) in most pension systems (e.g. for Germany, OECD,

2013). More indirectly and in line with the concept of path dependency, the *timing* of specific states that can be identified as turning points is relevant, as such states are likely to alter “opportunities and experiences immediately thereafter” (Bernardi et al., 2019, S. 4) and thus impact the remaining trajectory. For example, labour market interruptions (e.g., due to childcare) as well as their length, the likelihood and *timing* of resuming employment, and the determination of subsequent working hours (e.g., Arntz et al., 2017; de la Vega, 2022; Ziefle & Gangl, 2014) matter for the subsequent career (e.g., Aisenbrey et al., 2009) and accumulation of lifetime earnings (de Castro Galvao, 2023). Early unemployment or care work might lead to negative signaling or potential discrimination of employers, while later employment interruption might have a negative impact due to a potential higher human capital devaluation (G. S. Becker, 1962; Blau et al., 2006; Kunze, 2018; Mincer & Polachek, 1974). Women’s lower labour market attachment likely further negatively affects the bargaining position when it comes to household-level specialization and facilitates a traditionally gendered division of labour with negative consequences for women’s independent pension income accumulation (Möhring & Weiland, 2022). As such, the age at (first) childbirth is likely a relevant turning point constraining mother’s careers given that parenthood often sets the starting of a more traditionally gendered division of labour (Anxo et al., 2011; Baxter et al., 2008). It follows, that the age at which such turning points are experienced matters for the pension claim accumulation of mothers, particularly in conservative contexts. Thus, the *timing* of employment interruptions, as well as marriages and childbirths are likely to play a role for (gendered) pension claims.

(iii) *Ordering*. In line with the concept of path dependency (Bernardi et al., 2019) the *ordering* of specific life course events might be important for pension accumulation. Order refers to the sequencing of multiple events along the life course. For example, Gisela was full-time employed, before she had her first care work period, and then worked part-time

before she experienced her second care work period (Figure 5.1). For individuals taking up care duties, whether they have been (full-time) employed before the unpaid care period likely sets the opportunity structure to enter the labour market at later life course stages. Additionally, whether individuals have direct transitions from any employment to retirement or transit to retirement from less privileged states like unemployment (e.g., Agatha, Mary, and Hilde) or care work might be a proxy for retirement pathways that are often gendered (Fasang, 2010; Madero-Cabib et al., 2021; Riekhoff & Järnefelt, 2017). Given that these different pathways are related to different types of income (Fasang, 2010; Riekhoff & Järnefelt, 2017, 2018), having a direct transition from privileged employment to retirement might be related to pension income and gender inequalities therein. In this study, *ordering* also covers which states are experienced within the life course at all as well as the number of periods in a specific (order of) state over the life course (e.g., George has no care work period, Mary has one, Gisela two and Hilde three). On the one hand, multiple periods of employment interruptions (due to care work, unemployment, or other) might be a proxy for weak labour market attachment and thus low opportunities to build up high pension claims. On the other hand, this might imply the capability and strong motivation to return to the labour force and thus lead to higher pension claims compared to single, but much longer employment interruptions.

(iv) *complexity*. The overall *complexity* of life courses aims to capture the complexity, unpredictability, or stability of the observed life course period. It can be estimated in several ways but is most often defined based on i) the number of distinct life course states, ii) the number of transitions between such states, or the unpredictability (entropy) (see for an overview: Pelletier et al., 2020). Similarly to the number of specific states or subsequences (*ordering*), *complexity* can be both positively or negatively related to pension income

accumulation. For example, a certain degree of complexity can prevent an accumulation of pension claims if it restricts access to employment.

5.2.2 Empirical evidence on Gender Pension Gaps

There are two main approaches in the literature to assess gender inequality in pension income directly. The first and most common approach assesses Gender Pension Gaps or women's pension income using different variables that aim to summarize life courses. Most of the studies focus on the life course concepts of *duration*. Commonly, the studies apply decomposition techniques to differentiate between a part of the gap that is due to compositional gender differences and due to differences in returns. These studies usually use the sum of years spent in relevant states until old age, most often, the years in (full-time) employment (Bardasi & Jenkins, 2010; Bonnet et al., 2020; Cordova et al., 2022; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Frommert & Strauß, 2013; Hänisch & Klos, 2014; Johnson et al., 1999; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018, see Table 3.1 for an overview). Consequently, they predominantly highlight that large parts of gendered pension income inequality are due to women's fewer years spent in (full-time) employment and partly also women's lower earnings and lower education for the observed cohorts (compositional gender differences).

Fewer studies show that family life course aspects such as the number of children, the partnership status, or the duration of being married, never married, widowed, and divorced matter for women's pension or old age income in different welfare regimes (Evandrou et al., 2009; Fasang et al., 2013; Kreyenfeld et al., 2023). Some studies aim to capture aspects of life courses beyond the duration by operationalizing for example career types or volatility or being married at specific age ranges (Möhring, 2015; Sefton, Evandrou, & Falkingham, 2011).

Still, the first approach in the literature on GPGs clearly focuses on the *duration* of specific states, mostly full-time employment, and highlights the importance of years in full-time employment (or the lack thereof) for women's low pension income (see Table 3.1).

For Germany, studies decomposing the gender gap in pension income or wealth identify the gender difference in full-time employment as the main reason (Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014). Additionally, women spending more years in part-time jobs helps shrink the GPG in Germany (Cordova et al., 2022; Frommert & Strauß, 2013). For the Netherlands, the empirical evidence on the GPG is scarce. The only decomposition analyses of the GPG in the Netherlands are covered in multi-country studies that do not discuss the country-specific results in detail. The gender gap in total pension income is mostly due to gender differences in paid work and tertiary education (Bettio et al., 2013; Veremchuk, 2020). Dotti Sani & Luppi (2021) focus on mothers and assess the length of maternal absence from the labour market after the first childbirth on a) the mother's pension income and b) the within-couple Gender Pension Gap. They found overall that the longer the mothers abstain, the lower their pension income and the higher the within-couple gap. However, while this is the case in Germany, in the Netherlands they do not find these associations – even though Dutch mothers spent on average more years out of the labour force.

The *second*, more recent approach to analyse gendered pension income inequality uses sequence analysis to first find and describe common patterns of life courses and then assess how these are related to women's pension income (Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022) or the Gender Pension Gap directly (Chapters 3 and 4). These studies find that women-typical life courses characterised by care-related employment interruptions or part-time employment lead to lower pension incomes in Italy, West Germany, and Switzerland (Madero-Cabib & Fasang, 2016; Chapter 3) and, for West

Germany, that women in male-breadwinner couple life courses receive the lowest pension income (Möhring & Weiland, 2022). Furthermore, in Chapter 3 we reveal that only gender-specific work-family life courses of married parents are associated with the GPG in West Germany and Italy because life courses are only highly gender-segregated if individuals have children. For parents, the dominance of women in the life courses characterised by employment interruptions, unpaid care work and part-time employment drives the GPG because these life courses result in low pension incomes. On the other hand, and mostly in Italy, a lower share of the gap is due to the lack of mothers reconciling their motherhood with full-time employment.

However, there is not yet a systematic understanding of which elements of life courses are most important for pension income inequality broadly and thus most likely also for the gender inequality dimension therein. The first approach in the literature rather simplifies life courses based on summary measures, mostly focusing on the life course concept of duration, and thus neglecting other life course concepts. In fewer cases, single other life course aspects are considered but without a systematic selection. Using selective sets of life course summary measures has been criticised in the literature previously (Halpern-Manners et al., 2015). Particularly, when the life courses of interest span several decades, multiple life course domains (such as work and family), and many different life course categories (e.g., full-time and part-time employment, care work, unemployment, etc.) it is hard to theoretically disentangle which specific ones of these elements should be considered. The result is that the literature mostly relies on the most straightforward and easily extractable concept: *duration*.

While the second approach reveals the complexities and potential multidimensional linkage of different life course domains, it does usually not isolate specific life course aspects since the life course groups will always be subject to some degree of within-group heterogeneity

concerning one or several life course aspects. Life courses with the same *ordering* of specific life course stages but different *timing* of the transition might be clustered in one group (e.g., individuals with transition from full-time to part-time employment at different ages). The holistic approach thus risks concealing which specific aspects within common life course types are specifically important for (gendered) pension income inequality. This, however, is important to understand for developing targeted public policies aiming at reducing the accumulation of inequalities over the life course.

5.2.3 Welfare state context & country comparison design

The country comparison aims to assess whether similar gendered life course aspects are differently related to the Gender Pension Gaps across different pension systems. Furthermore, the first analytical step explores whether similarly gender-stratified work and family life course aspects differ in their predictive importance for pensions across the two different pension systems. The country comparisons design is in line with the life course approach, which argues to consider the sum of all institutions within a country as ‘packages’ as well as the interactions of different policy domains (Fasang, 2012; Fasang et al., 2013, S. 2). I thus do not discuss single elements of the welfare and pension regimes but describe and assess them as well as their interaction with each other as whole ‘packages’.

West Germany and the Netherlands are ideal cases for the research design as they share similar gender inequalities before old age (for individuals aged 65+ at the year of the interview) but have very different pension systems implemented. In particular, both countries are examples of contexts that shape highly gendered life courses (see below), which are associated with Gender Pension Gaps from a holistic perspective (see Chapters 3 and 4). Previous research suggests that low levels of gender segregation on different life course types are not related to the GPG in more egalitarian country contexts (Rowold,

2022). Therefore, in order to examine how the relationship between gendered life course aspects and the gender pension gap is channeled through different pension systems, it makes sense to first focus on countries with high levels of gendered life courses. In addition, for the introduction of the combination of the Life Course Feature Selection with the Ñopo decomposition, a certain degree of gender segregation on different life courses is advisable to illustrate the contribution of the methods. Both countries are particularly useful showcases for the common support problem related to life courses and thus for introducing the novelty of quantifying parts of the GPG due to gender-exclusive life course experiences (DA and DB components of the Ñopo decomposition, see section 5.3.2.2).

In the following, I give an overview of the welfare state contexts and pension policy that serves this bigger-picture perspective. The welfare state contexts are relevant to understanding how the work and the family lives of the individuals, and gender inequality therein, were shaped. The description of the pension regimes is to understand how the pension regulations reward these life courses with pension income.

Gender inequalities in work and family life courses before old age: the welfare state regimes

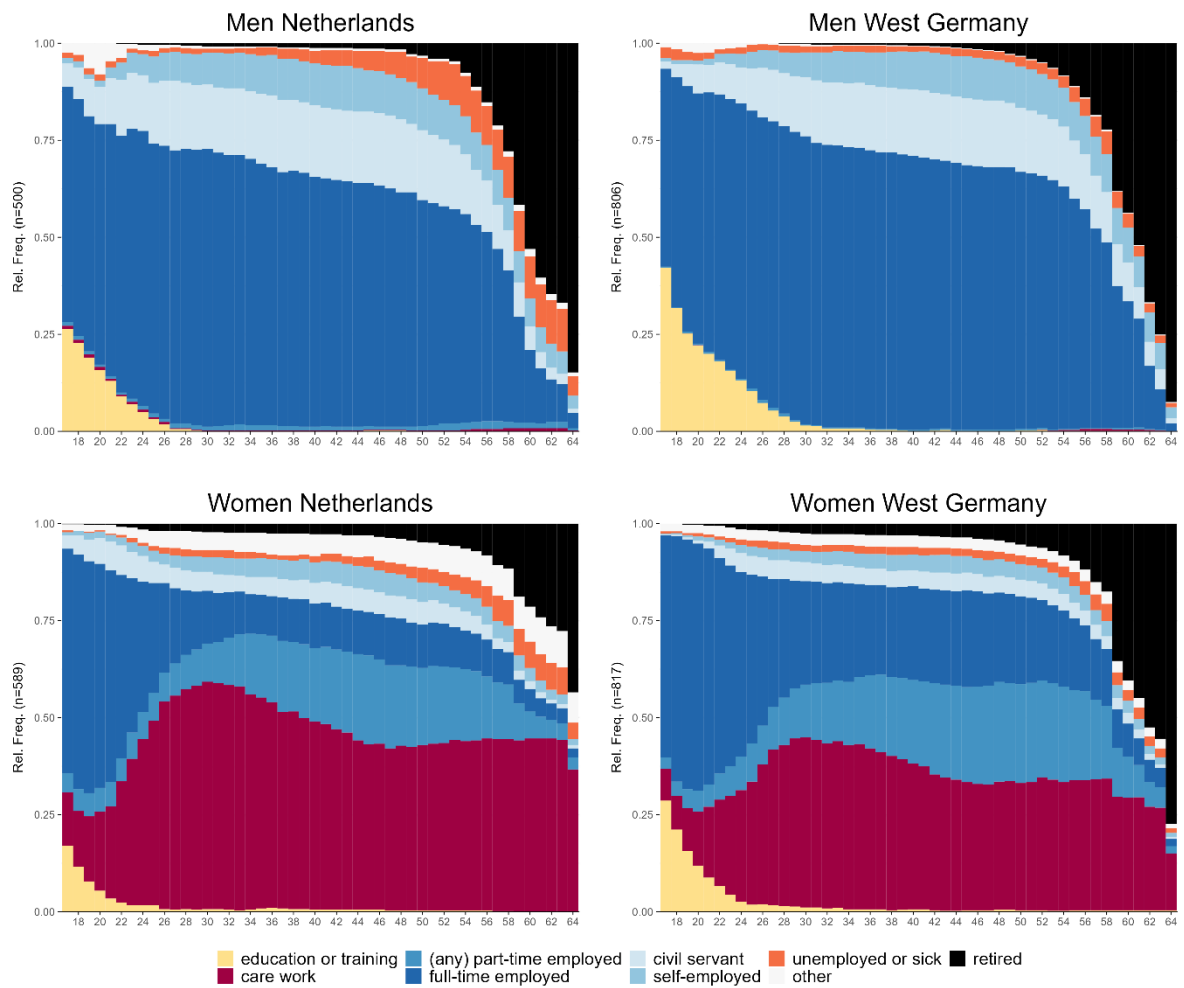
The welfare state regimes of the Netherlands and Germany are both considered conservative (e.g., Komp-Leukkunen, 2019; Uccheddu et al., 2022) and incentivised a traditional gendered division of labour, that is men are the main breadwinner and women the main caretaker.⁵³ Key policies are poor public childcare, generous maternal leave policies (e.g., for up to 3 years in Germany) as well as couple-level joint taxation (e.g., Adema et al., 2020; Bach et al., 2011; Daly & Ferragina, 2018; Neyer, 2003). Incentivising

⁵³ The studied cohorts in Germany lived most of their preretirement lives in two different welfare regimes of the separated Germany: the conservative FRG and the socialist GDR. Since these welfare regimes are strongly correlated with the level of gender segregation of work-family life courses (Rowold, 2022) as well as other prominent indicators for gender inequality in the labour market as well as gender role attitudes (e.g., Barth et al., 2020; Sprengholz et al., 2022; Trappe et al., 2015), I focus on the regions of Germany that previously belonged to the FRG (i.e., West Germany).

a traditional division of labour, such policies set the ground for an accumulation of gender inequalities over the life course, which exacerbates potentially initially minor labour market inequalities between men and women at early ages. Within such a policy context, the *timing* of union formation as well as transitions to marriage and parenthood is highly relevant for the rest of an individual's career and, consequently, pension claim accumulation. This is empirically confirmed for women only, whose family and work life courses are highly intervened in conservative contexts, whereas this is not the case for men (e.g., Madero-Cabib & Fasang, 2016; Chapter 3).

Mirroring the similar welfare policies linked to the gender division of labour, prominent gender inequalities have been documented at similar levels for the past century in the Netherlands and West Germany. For example, both countries have rather low female and maternal employment in the second half of the 20th century (Möhring, 2018). Additionally, in the 1990s the gender wage, employment, and part-time employment gaps as well as the female and maternal labour market participation were similarly high (Anttonen & Sipilä, 1996; Hammerschmid & Rowold, 2019a). Thus, the gender inequalities in life course aspects are likely to be similar across both countries.

Figure 5.2 Work life courses across genders in West Germany and the Netherlands



Notes: State distribution plots of working lives from ages 18-65 based on analysis sample (see section 5.3.1). Depicts the distribution over the observed work life course categories for each year of age between 18-65. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Indeed, Figure 5.2 descriptively confirms a very similarly gendered distribution of work life course states from age 18 to 65 for the study sample. Dutch and German men's working life courses were dominated by full-time employment in the private sector. In the Netherlands, the share of unemployment or being sick or disabled is higher than in Germany from the mid-thirties, and marginal shares of male part-time employment exist. While most women started their careers as full-time employees as well, the most dominant life course state from about age 25 on is unpaid care work in both countries (about 40-50% at age 26). Part-time employment becomes the second or third important working state from mid-age on. While for Dutch women care work and 'other' states were more prevalent at

all ages compared to their German peers, West German women were to a higher share part-time or full-time employed. Despite these marginal differences, the working life course of women in the Netherlands and West Germany looked very similar for the current old-age population.

Gender inequalities in pension income: the pension systems

Contrary to similar levels of gender inequalities, the Dutch and the German pension systems are classified as two opposing ideal-typical pension systems, the Beveridgean and the Bismarckian system (Ebbinghaus, 2021). The aim of the Beveridgean system is historically to provide a minimum income and to reduce poverty, while the Bismarckian was implemented to reproduce the pre-retirement living standard in old age, including its inequalities (Ebbinghaus, 2021). Both pension systems consist of three pillars: the public, occupational, and personal pensions.

When it comes to gender inequality in pensions it is important to consider both the *amount* of claimed pensions, as well as the *access* to pensions in the first place (Ginn et al., 2001). In the following, I give a brief overview of both dimensions of pension claims for all three pension pillars.

Public pensions. In contrast to the Bismarckian system in Germany where a minimum contribution record of five years is required for accessing public pensions, most residents from age 65 are eligible for the basic public pension income in the Beveridgean system in the Netherlands (Ebbinghaus, 2021; OECD, 2015). Thus, in contrast to the universal access to public pension in the Netherlands, in Germany, individuals without stable employment careers, often women, are systematically excluded (Leitner, 2001).

The Dutch public pension system provides a flat-rate basic pension (OECD, 2015). The amount of flat-rate public pensions depends on the household situation (living alone or not),

and the years of residence in the Netherlands and is set based on the overall minimum wage (Anderson, 2011; OECD, 2013). In Germany, individuals collect pension claims throughout their employment career and based on their relative earnings. For each year they were employed, they receive earning points that mirror the individual's relative position in the earnings distribution of all eligible employees. Benefits for childcare periods are considered in the German pension system but only make up for about one year of a care-related employment interruption per child for the studies' birth cohorts. Some redistributive elements are included in the public pensions in Germany, such as topping up low earning points and imposing a maximum value of earning points. But overall, the highly employment-linked public pension system in Germany likely reproduces pre-retirement inequalities to a much higher degree compared to the progressive public pension in the Netherlands.

Private pensions (occupational and personal). Occupational pensions are pension schemes on the company level that are sponsored by the employer whereas personal pensions refer to entirely personal saving plans (Ebbinghaus, 2011a).⁵⁴ The access to and amount of such private pensions is restricted by labour market characteristics since they are particularly linked to employment outcomes (Ginn & Arber, 1996; Rutten et al., 2023). For example, access to and claims from occupational pensions are highly stratified by (lifetime) earnings, occupation, and industry as well as company size for example – labour market characteristics that are also highly gender stratified leading to lower access for women (Fasang, 2010; Ginn, 2004). However, when occupational pensions are mandatory, and thus cover the whole workforce, this might benefit women (Ebbinghaus & Neugschwender, 2011). For example, Möhring (2018) shows that the motherhood pension penalty is lower in countries with (quasi-)mandatory occupational pension schemes, such as the

⁵⁴ With 'private pensions' I refer to both, occupational and personal pensions jointly.

Netherlands. On the other hand, it is argued, that private pensions are generally more likely than public pensions to reproduce (gendered) labour market inequalities due to the stronger link of claims to stable employment and earnings as well as the lack of redistributive elements (Ginn & Arber, 1996; Möhring, 2015).

For example, in the Netherlands, the overwhelming majority of occupational pensions are based on lifetime earnings (OECD, 2013). The Dutch pension system reaches almost full coverage of the (quasi)mandatory occupational pension scheme (Ebbinghaus, 2015): more than 90% of workers are covered in an occupational pension scheme (Tyros et al., 2023). Still, full coverage is not fully achieved and men had a higher coverage in occupational schemes compared to women across all employment types in 2001 (Frericks et al., 2006). In line with this, 85% of men and 40% of women in the analysis sample of this study had any occupational pension in the Netherlands. While the coverage of occupational pensions in Germany has increased over the years (Bridgen & Meyer, 2014; Ebbinghaus et al., 2011), the level is still far away from the one in the Netherlands, reaching 42% of men and 29% of women in the studied sample.

Overall, the design of the pension systems leads to low-medium levels of income inequality and poverty risks for the Dutch population aged 65+ in a European comparison, while in Germany the inequality and poverty risks are at a medium-high level (Ebbinghaus, 2021). The (gendered) pension income inequality in Germany comprises all types of pension income, whereas it is driven by private pensions in the Netherlands. In summary, the German pension system is an illustrative case for a Bismarckian pension system with earnings-linked public pensions, and a still low but increasing importance of private pensions, whereas the Dutch Beverigean pension system illustrates a context with generous flat rate public pensions, but very high levels of private pensions. The Dutch context gives thus an idea of whether gendered life course aspects are reproduced in a pension system

with high levels of private pension, which is what many European pension systems head for (Ebbinghaus, 2011b), even if the public pensions are highly progressive.

5.2.4 Expectations

The research design is highly explorative and is not suitable for classical hypothesis testing. However, I summarize some overarching expectations based on the theory, previous literature, and policy context that aim to guide through the results.

Based on the previous empirical literature as well as the design of both pension systems, it appears straightforward to assume that *duration* in full-time employment is the most important predictor of pension income inequalities (*expectation 1a*). However, based on the life course paradigm I think it is more plausible that life course aspects covering the *ordering* and *timing* of life courses are at least of similar importance for pension income (*expectation 1b*). The *timing*, *ordering*, and *duration* of care-related employment interruptions are among the most important life course pension predictors and also important for the GPG (*expectation 2*). This is because they are almost not rewarded in the pension systems and are additionally likely to limit career prospects over the life course and thus pension contributions. Concerning family life courses, I expect that particularly the *timing* matters (*expectation 3*). The *timing* of marriage and childbirth is expected to be relevant as these events are shown to support a gendered division of labour and matter, particularly for women's pension income. For the Ñopo decomposition, I expect that large parts of the GPG are attributed to highly gender-exclusive combinations of life course aspects that generate different pension income levels (components DA and DB, see section 5.3.2.2) due to the strong evidence on gendered work-family life courses (*expectation 4*).

5.3 Data & Methods

5.3.1 Data

This paper applies data from the Survey of Health, Ageing and Retirement in Europe (SHARE) on cross-sectional information such as pension income and retrospective data from SHARELIFE providing annual information on family and work life courses. The sample consists of individuals aged 65 or older, who were not employed or received payments from employment and had no missing information for the variables used (see Table A5.2 for the sample selection).⁵⁵ It contains the birth cohorts 1911-1950 (median 1940). Merging waves 2, 4, 5, and 6 resulted in 1623 observations in West Germany and 1089 in the Netherlands (wave 5 is used as the base). The measure used for pension income sums up regular payments from public, occupational, and personal pensions (see Table A5.1) and considers individuals without any independent pension income with an income of 0. The dependent variable thus covers gender differences in i) access to and ii) amount of pension income. This pension operationalisation puts the focus on the financial independence and security of women, whereas derived claims are associated with a loss of autonomy (Ginn et al., 2001; Muller et al., 2020). This perspective is relevant for reducing the risk of female poverty in old age, upon union dissolution for example, as well as for social security reasons and is largely applied in the field (e.g. Fasang et al., 2013; Kuivalainen et al., 2020; Möhring, 2015).

The family and working life categories are based on the retrospective survey data from SHARELIFE. For family states, I differentiate between ‘single, no child’, ‘single, 1+ children’, ‘married⁵⁶, no children’, ‘married, 1 child’, ‘married, 2 children’, ‘married, 3+ children’, and ‘divorced, w/o child’ (see Figure A5.1 and Figure A5.3). ‘Single’ identifies individuals not cohabiting or married. Divorce overwrites all other family states occurring

⁵⁵ Non-employed respondents who report spare time work are kept in the analysis sample.

⁵⁶ ‘Married’ also includes not cohabiting couples and non-married cohabiting couples.

in the same year, except remarriage. The data do not distinguish between children living with the respondent or not (see section 2.3.2 for a discussion of this issue). Thus, the number of children refers to any children alive, including adopted children, in the given year (Brugiavini et al., 2019).

The work trajectory accounts for nine mutually exclusive sequence states: ‘full-time employed’ and ‘part-time employed’ (both private sector only), ‘civil servants’, ‘self-employment’, ‘education or training’, unpaid ‘care work’, ‘unemployed’, ‘other’ and ‘retired’ (see Figure A5.2 and Figure 5.1). ‘Part-time employment’ includes any part-time spells as an employee or as short-term employment, including years for which respondents report changes in their working status to or from part-time at least once. ‘Unemployment’ includes unemployment and inactivity. ‘Care work’ is based on the response option ‘*Looking after home or family*’ and thus related to family lives. This measure exclusively captures *full-time* care work that was done for at least 6 months a year and therefore underestimates the actual degree of unpaid care work given per year. ‘Other’ mostly covers ‘*Leisure, travelling or doing nothing*’ and ‘*Military services, war prisoner or equivalent*’.

5.3.2 Analytical strategy

The first step of the analysis identifies and ranks the most important pension predictors among a set of hundreds of life course variables. The second step then reveals how these life course aspects are related to the GPG. Both analysis steps are run separately by country to reveal potential country-specific patterns.

5.3.2.1 Step I: Life Course Feature Selection

For the Life Course Feature Selection, I first extract a multitude of life course indicators covering the life course dimensions of duration (I), timing (III, IV, V), ordering (II, V) and overall complexity (VI) of work and family life courses. Sequence analysis tools allow fragmenting trajectories in their elements covering a range of different properties and

extract them automatically (Studer, 2018). I focus on (I) the duration (i.e., the total time spent in a life course state), (II) the number of frequent sub-sequences of states, (III) the age at which each *n*th spell per state occurred along the trajectory, (IV) the age period when transitions between states occurred⁵⁷, (V) the age at the first transition of a subsequence, and (VI) complexity measures of life courses. The different properties are described based on example trajectories in detail in Appendix A5.6.1.

261 variables of family life courses in Germany and 215 in the Netherlands are automatically extracted as well as 524 and 391 variables of working life courses respectively. I remove some variables based on empirical and theoretical considerations and multicollinearity concerns aiming for the “smallest but exhaustive possible set of theoretically relevant indicators, which are interpretable on their own” (Bolano & Studer, 2020, S. 12). This results in a set of 221 variables in West Germany and 212 variables in the Netherlands. Before applying the feature selection, I operationalise the duration and complexity as continuous variables (in line with literature stream 1). The rest of the variables are categorised because of the data structure (see Appendix A5.5.3 for more details).

To identify the most important pension predictors I use the Boruta algorithm. Boruta is preferred for this study because it considers potential interactions between life course variables for assessing their importance (Bolano & Studer, 2020). This is relevant to consider potential interactions between variables of the family and work dimension which has shown to be particularly important for the gender dimension of pension inequality (Rowold et al., 2023). Additionally, Boruta has recently been shown to perform best among several feature selection approaches, especially for applications with many predictors (for

⁵⁷ The code for automatically extracting this life course aspect is kindly provided by Matthias Studer.

classification tasks, Speiser et al., 2019) and for low-dimensional data sets (Degenhardt, Seifert, & Szymczak, 2019) like it is the case for this study.

The Boruta algorithm is based on a random forest (RF) approach and, other than normal RFs, it provides a cut-off point for distinguishing between important and unimportant features (i.e., variables) by generating so-called shadow features (randomized versions of the original features). RF is run over the original data as well as over the shadow features. The original features then compete against the shadow features and are only considered important for the outcome if they perform better in the RF than the best-performing shadow feature.⁵⁸

The outcome of Boruta is a set of variables deemed as important for pensions (i.e., those performing better than the shadow features) ranked by their importance for predicting pension income. Mean importance is calculated based on the loss of accuracy due to permutation averaged over all random trees (Kursa & Rudnicki, 2010). This estimation of importance is shown to be less biased regarding the variable format (Probst et al., 2019). The mean importance does not imply a meaningful interpretation in absolute terms but is to be interpreted in relative terms.

Boruta is a so-called ‘all-relevant’ approach aiming to reveal all relevant features which is the preferred type of feature given that this study aims to identify all important life course predictors for pension income. However, this might lead to Boruta identifying too many features that cannot all be included in regression models due to overfitting or multicollinearity issues (Bolano & Studer, 2020). I complement the Boruta outcome with a detailed multicollinearity analysis for the decomposition analysis as well as alternative feature selection algorithms as robustness checks (see Appendices A5.6.1 and A5.5.4).

⁵⁸ For a more detailed description of the procedure see A5.6.2.

However, multicollinearity issues are more problematic for the Kitagawa-Oaxaca-Blinder compared to the Ñopo decomposition since the former is based on OLS regressions while the latter is non-parametric. Even though RF approaches perform reasonably well with the default parameters, I follow recent advice for tuning the most common parameters of RF models to identify the optimal parameters to be used in the final model (Boehmke & Greenwell, 2019; Probst et al., 2019, see Appendix A5.6.3 for details).

5.3.2.2 Step II: Decomposition analysis

The most common decomposition in the field is the Kitagawa-Oaxaca-Blinder decomposition (e.g. Bardasi & Jenkins, 2010; Bonnet et al., 2020; Even & Macpherson, 2004; Frommert & Strauß, 2013; König et al., 2019; Zhao & Zhao, 2018). It decomposes the average GPG ($\bar{y}_M - \bar{y}_F$) based on linear regressions in an explained and unexplained part:

$$\bar{y}_M - \bar{y}_W = \underbrace{\alpha_M - \alpha_W}_{\text{Intercept}} + \underbrace{\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W(\hat{\beta}^* - \hat{\beta}_W)}_{\text{Returns}} + \underbrace{(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*}_{\text{Explained}} \quad (1)$$

$\bar{X}_M$ and $\bar{X}_F$ denote vectors of mean values of the independent variables, the selected life course variables from step I, and α_M and α_W denote the intercepts and $\hat{\beta}_M$ and $\hat{\beta}_F$ the coefficient vectors of the separate OLS regressions for men and women, respectively. $\hat{\beta}^*$ is the reference coefficient vector that is based on the coefficients from the linear regression models of females or males. The *explained* share, $(\bar{X}_M - \bar{X}_F)' \hat{\beta}^*$, quantifies the share of the gap that emerges from the gender-specific distribution over the different life course aspects. It reveals the part of the gap that is due to compositional differences between men and women regarding the observed life course aspects, for example, because women have worked on average fewer years full-time compared to men (and full-time employment is highly rewarded in the pension system). The *returns component* of the unexplained share, $\bar{X}'_M(\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_F(\hat{\beta}^* - \hat{\beta}_F)$, shows which part of the GPG is due to differences in

pension income of men and women for the same life course aspect (differences in returns). For example, a share of the gap emerges, because men receive a higher pension income for the same years worked full-time compared to women.

However, the KOB approach has been criticised for comparing wage equations of men and women who lack comparability – technically speaking, common support is absent in the distribution of observed characteristics (common support problem, Ñopo, 2008). I expect that the lack of common support particularly occurs for life course variables, since men and women are shown to have very different work-family life courses, especially in conservative or liberal welfare states (Madero-Cabib & Fasang, 2016; Uccheddu et al., 2022; Chapter 3). For example, in a traditional gender context, there might be many women with several care work periods after an initial full-time employment career, but no men. Thus, we lack a counterfactual (the pension income of comparable men) to compare these women's pension income with. In cases of lacking common support, the unexplained and explained shares might be biased (Hamjediers & Sprengholz, 2023).

To address the common support problem Ñopo (2008) suggests applying a nonparametric decomposition technique by using one-to-many exact matching and decomposing gender gaps into four additive parts. Similar to the KOB, the reference group has to be specified for the Ñopo decomposition which affects Dx and $D0$ but not DA and DB . I define women as group A and men as group B (see Appendix A5.5.5, for sensitivity analyses with swapped groups).⁵⁹

⁵⁹ Ideally, I would prefer to choose the reference category for the covariates separately based on the majority group for each category. However, due to the non-parametric nature of the Ñopo decomposition, the choice of reference cannot vary across covariates. I chose women as group A because they are more likely than men to receive care benefits in pension systems. Thereby care work benefits are captured in the pension returns for characteristics, which is relevant for care work related life courses aspect that I expect to play an important role for the GPG (for a detailed discussion of reference coefficient options and when using females as reference is meaningful for pension income, see Chapter 3.3.2.1).

First, all women in the sample (group A) are each matched to all men in the sample (group B) who have the same combination of observed characteristics as the selected women. Each unique combination of characteristics is termed a ‘stratum’. The matching procedure creates four new samples: matched men and matched women, who have a comparable counterpart and are thus in common support, as well as unmatched men, and unmatched women who are out of support. Over the sample of matched women and men, the women’s pension income is then reweighted based on the ratio of men to women concerning the distribution of characteristics. This creates a counterfactual group A^B whose pension income is the average pension income of matched women if they had the exact same characteristics as matched men (Hamjediers & Sprengholz, 2023). Based on the average pension income of the four different samples resulting from the matching as well as the counterfactual A^B resulting from the reweighting, the Gender Pension Gap can be decomposed in four parts:

$$\bar{Y}_M - \bar{Y}_W = Dx + D0 + DA + DB \quad (2)$$

Dx and $D0$ are analogues to the explained part and unexplained parts of the KOB but calculated only over the common support (i.e., matched men and women only).

$Dx = \bar{Y}_{A^B,m} - \bar{Y}_{A,m}$. The explained part is the difference in the average pension income between matched (m) and reweighted women A^B and the matched women with their original average characteristics. It reveals which part of the gap emerges because men and women are differently distributed across the combinations of observed characteristics (strata). With women as group A (and thus as the reference in terms of returns to characteristics) Dx is the part due to differences in characteristics when the returns of women to these characteristics are applied.

$D0 = \bar{Y}_{B,m} - \bar{Y}_{A^B,m}$. The unexplained part is the difference in average pension income between the matched men as well as the reweighted group of matched women. Given that

the matched men and women have the same observed characteristics, any differences in pension income cannot be explained by these. With men as group B (and in my specification thus as the reference regarding the characteristics) $D0$ is due to differences in returns if men and women had the characteristics of men.

The other two components of the GPG are due to the gender differences in support.

$DA = \bar{Y}_{A,m} - \bar{Y}_{A,u} * n_{A,u}/n_A$. DA is the part of the gap that occurs because the matched sample of women has on average higher or lower pension income compared to the unmatched sample of women (weighted by the ratio of unmatched (u) women).

$DB = \bar{Y}_{B,m} - \bar{Y}_{B,u} * n_{B,u}/n_B$. DB is attributable to differences in average pension income between the matched and the unmatched sample of men.

The components DA and DB provide additional insights compared to the KOB revealing how much of the gap stems from extreme group-specific segregation (Sprengholz & Hamjediers, 2022), which highlights the impact of extremely unequal life course experiences of men and women, i.e., combinations of life course aspects that only occur for one group, in combination with the Life Course Feature Selection.

The additional decomposition components DA and DB are the main benefit of the $\tilde{N}opo$ compared to the KOB decomposition for the aims of this study, besides its robustness against common support problems and functional-form model specification (Hamjediers & Sprengholz, 2023). Additionally, through matching based on several life course aspects, the $\tilde{N}opo$ decomposition considers interactions between these life course aspects and potentially different life course domains (e.g., work and family).

Still, the $\tilde{N}opo$ decomposition has some limitations. Particularly in small samples the curse of dimensionality might result in the lack of common support being solely due to the limited sample size which may lead to biased components of the $\tilde{N}opo$ decomposition (Hamjediers

& Sprengholz, 2023). Additionally, unlike the KOB, the explained and unexplained shares cannot be decomposed per variable (detailed decomposition), and continuous variables must be categorised for matching. At the same time, the KOB is robust to the sample size (Hamjediers & Sprengholz, 2023).

I follow the recent suggestion in the literature to start any decomposition of group differences with a Ñopo decomposition adding sets of variables stepwise to reveal how much of the gap is due to a lack of common support (Hamjediers & Sprengholz, 2023; Strittmatter & Wunsch, 2021). If there is no substantial lack of common support, I can proceed with the KOB since there is sufficient overlap between men and women.

For this study, I find a substantial lack of common support (see section 5.4.2). Consequently, for the analysis, I use Ñopo decompositions with stepwise models containing different sets of the selected variables from the Life Course Feature Selection. Variables are chosen in descending order by their variable importance as estimated by the Boruta algorithm and cut-offs between variable sets are defined based on relatively large differences in mean importance (see Appendix A5.5.4 for details).

As a supplementary analysis, I estimate a KOB over the common support of men and women to quantify which variables are the main contributors to DX of the Ñopo decomposition (i.e., detailed decomposition) for the set of variables with sufficient common support and over the matched men and women (i.e., over common support). I thus make use of the complementary benefits of both decomposition strategies.

5.4 Results

5.4.1 Step I: Life Course Feature Selection

Figure 5.3 shows the mean importance of the selected pension life course predictors (in total 35 in West Germany and 19 in the Netherlands) in descending order and differentiated by the different life course aspects. It reveals that life course variables covering all life course aspects considered—*timing*, *duration*, *ordering*, and *complexity*—play a role for pension income. This is important to highlight since previous research has overwhelmingly used the *duration* in different states as the main predictor of pensions. Even though in both countries the two most important pension predictors indeed relate to the *duration* in states, in West Germany the top-10 pension predictors are dominated by the *ordering* and *timing* of life course experiences, and in the Netherlands, the overall *complexity* of work life courses stands out as main pension predictor.

Focussing on the life course states, contrary to *expectation 1a*, the *duration* of full-time employment is not among the top pension predictors in both countries, even though identified as relevant. On the contrary, the years spent in full-time employment play a rather subordinate role in pension income inequality in West Germany (ranked 7th in mean importance) and the Netherlands (ranked 17th in mean importance, Figure 5.3). Instead, in both countries, the *duration*, *timing*, and *ordering* of unpaid care work emerge as key factors contributing to pension income inequality (*expectation 2* confirmed, Figure 5.4). In West Germany, the sum of years engaged in unpaid care work (*duration*) is the second most important pension predictor, followed by the number of care work periods across the life course (*ordering*) and the age at the start of the first care work period in individuals' lives (*timing*). In the Netherlands, the total years spent in care work is the most important life course aspect for predicting pension income. The *timing* and *ordering* of care work are not among the 10 most important pension predictors but are still more important than the

duration of full-time employment. In West Germany, the age at the transition from care work to part-time employment if later in the life course a transition from part-time employment to retirement followed is among the top-10 pension predictors as well (*timing* and *ordering*). Slightly less important (11th rank) is the age at the transition from care work to full-time employment if later in the life course individuals returned to care work (*timing* & *ordering*).

Figure 5.3 Mean importance of selected variables by country



Notes: Results for all features selected in the Netherlands or West Germany based on the Boruta algorithm, sorted by mean importance. Mean importance as z scores. See Table A5.3 for minimum, maximum and median importance. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Alphabet: Family life courses: SNC = single, without child(ren); S1C = single parenthood; MNC = married without children; M1C = married with one child; M2C = married with two children; M3C = married with three or more children; DIV = divorced w/o child.

Work life courses: EDU= education or training; CW = unpaid care work; PT = (any) part-time employed; FT = full-time employed; CIV = civil servant; SELF = self-employed; UNE = unemployed or sick/disabled; OTH = other; Ret = retired.

Figure 5.4 Mean importance of selected variables by country and life course state



Notes: Results for all features selected in the Netherlands or West Germany based on the Boruta algorithm, sorted by mean importance. Mean importance as z scores. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Apart from care work, other life course states stand out. The years spent in education (after age 17, *duration*) is the second most important pension predictor in the Netherlands, whereas it ranks only 13th in West Germany. Thus, educational inequalities are more strongly reproduced in the Dutch pension system. Moreover, the overall stability of working life courses (measured as entropy, *complexity*)⁶⁰ is one of the main predictors of pension income in the Netherlands (ranked 4th important) while it is less important in Germany (ranked 19th, see also Table 5.2). In the German pension system, the most important pension predictor is the years spent as a civil servant (*duration*). The *timing* of civil service as well as whether and how often civil service spells occurred (*ordering*), are among the top-10 pension predictors as well. This is due to the highly generous pension benefits for German civil servants and in line with previous evidence revealing that the lack

⁶⁰ This indicator particularly emphasizes the (un)predictability of working life courses.

of women with civil service careers is strongly associated with the gendered dimension of pension inequality (see Chapters 3 and 4).

Lastly, the age at retiring directly from full-time employment or civil service is an important pension predictor in West Germany (*timing & ordering*). However, this might be driven by the gendered notion of retirement that leads to a gendered reporting bias of retirement (see Chapter 2.2). Consequently, these life course variables might capture the gender difference in identification with retirement in old age. This is further suggested since the direct retirement from full-time employment as well as the duration of retirement are identified to be particularly important predictors of gendered pension inequality and not for gender-specific pension incomes (see A5.3 for a more detailed discussion). Similarly, the direct retirement transitions from full-time employment or civil service might be proxies for gendered retirement pathways that are associated with old age income (Fasang, 2010; Riekhoff & Järnefelt, 2017, 2018).

Overall, *expectation 1b* is thus confirmed: life course aspects covering the *timing* or *ordering* of various other life course categories are important pension predictors beyond the *duration* of full-time employment. Many of these aspects have not been identified or assessed empirically previously, at least when it comes to Gender Pension Gaps.

However, it stands out that Boruta identifies more variables related to the *duration* in life course states as the most important pension predictors in the Netherlands, whereas more variables related to the *timing* and *ordering* are selected in West Germany. For example, more aspects related to the *timing* and *ordering* of care work are important predictors of pensions in West Germany compared to the Netherlands. There could be several explanations for this.

First, different life course aspects related to the *timing* and *ordering* of events may be more correlated in the Netherlands than in West Germany, which can reduce their estimated importance by the Boruta algorithm. For example, the relative importance of the *timing* and *ordering* of care work experiences across the life course is higher in the Netherlands when using alternative feature selection algorithms that are more robust to correlated variables (see Table A5.6). As a result, higher levels of feature correlations are likely to account for the lower importance of other care work proxies than the *duration* of care work in the Netherlands.

Second, the different pension systems could be an explanation.⁶¹ In West Germany, pension income differences are more pronounced regarding the number of care work spells over the life course (*order*) as well as the age at the first care work experience (*timing*). In the Netherlands, on the other hand, the pension penalty increases only with the duration of care work, while the care work pension penalty does not vary strongly with the number of care work interruptions (*order*) or the *timing* (not shown). Given that the pension inequality in the Netherlands is mainly driven by private pensions, this may suggest that for private pensions it is indeed primarily the *duration* of the different states that matters.

Third, as Boruta takes into account interactions between variables when calculating variable relevance, stronger interrelations of care work life course aspects with other life course variables which are important for pensions could lead to their higher importance in West Germany. Conversely, the importance of variables other than care work could be enhanced in the Netherlands for the same reason (partly suggested in section 5.4.2.1). However, all interactions between variables which are relevant for predicting pensions are difficult to disentangle due to the large number of variables considered.

⁶¹ An alternative explanation would be that the gender distribution on the affected life course proxies differs systematically by countries. However, in line with the research design, the gender distribution is similar across countries (not shown, tested for the different care work variables).

The explanation cannot be conclusively clarified in this study. Most likely, a combination of all three explanations applies.

Previous literature on the GPG in Germany had highlighted the role of part-time employment (Frommert & Strauß, 2013, Chapters 3 and 4). Surprisingly, apart from the age at the transition from care work to part-time employment in life courses with a subsequent transition from part-time employment to retirement in West Germany (*timing & ordering*), life course aspects related to part-time employment play only a subordinate role in overall pension inequality in West Germany (age at the first part-time spell, the number of part-time spell occurrences and the duration of part-time employment). No part-time life course aspects are identified as important pension predictors in the Netherlands. This is astonishing, since in both countries, part-time employment is relatively common, and highly gender segregated.

In both countries, only one variable related to family life courses is selected as an important pension predictor. In the Netherlands, the *duration* of being single and childless is relevant (ranked 6th). The age at the transition from being married without having children to being married and having a child is only moderately relevant in West Germany (i.e., *ordering & timing* of first childbirth, ranked 17th). Potentially the association of family lives with pension income is channeled through unpaid care work linked to family transitions (see Chapters 3 and 4) and this is why almost no family life course aspects are among the most important pension predictors. Additionally, family life course experiences are likely to exhibit opposing associations with pensions for men and women (e.g., Kreyenfeld et al., 2023; Möhring, 2015) which might cancel each other out in the feature selection for pension income pooled over gender. Indeed, this seems to be the case for the occurrence of divorce in West Germany (see sensitivity analyses below). However, I cannot find clear evidence

for the relevance of the *timing* of family events, especially not for women (*expectation 3* not confirmed).

5.4.1.1 Sensitivity: Life course predictors for *gendered* pension inequality

As a sensitivity analysis, I examine which life course pension predictors are particularly important for *gendered* pensions. This is relevant as this study analyses the Gender Pension Gap, and the life course predictors may differ across *overall* and *gender* pension inequality. Since, to my knowledge, there are no feature selection algorithms for gaps in outcomes, I use two different approaches to identify life course proxies particularly important for gender inequalities in pensions.

5.4.1.1.1 Including gender as a predictive feature

As a first approach, I include gender as a variable in the life course feature selection. The aim is to identify variables whose importance for predicting pensions is channeled through gender. Variables whose relative importance decreases compared to the main model, which does not include gender as a variable, are likely to be important predictors of GPG.

The first striking finding is that gender is by far the most important pension predictor in West Germany (almost twice as important as the second most important predictor, not shown) and the second most important pension predictor in the Netherlands (Table 5.1). This underlines the extent of gender inequalities in pensions compared to any pension inequalities arising from life course proxies, independently of gender.

Table 5.1 Selected features for overall pension inequality, including gender as predictor and not, ranked.

West Germany			Netherlands		
feature_labels	Rank Main	Rank gender predictor	feature_labels	Rank Main	Rank gender predictor
Duration civil service (CIV)	1	2	Duration care work (CW)	1	6
Duration care work (CW)	2	3	Duration education (EDU)	2	1
Number of CW periods	3	15	Age at 1st OTH->FT transition	3	5
Age at 1st CW spell	4	9	Complexity work life (entropy)	4	3
Age at 1st CIV->RET transition	5	4	Duration retirement (Ret)	5	11
Age at 1st FT->Ret transition	6	16	Duration self-employ.(Self)	6	8
Duration full-time employ. (FT)	7	8	Number of CW-OTH subsequences	7	4
Age at 1st CW->PT_PT->Ret transition	8	14	Duration single, no child (SNC)	8	10
Age at 1st CIV spell	9	6	Age at 1st EDU spell	9	9
Number of CIV periods	10	12	Age 23-27: end of EDU	10	13
Age at 1st CW->FT_FT->CW transition	11	21	Duration civil service (CIV)	11	14
Number of EDU-CIV subsequences	12	5	Age at 1st CW spell	12	22
Duration education (EDU)	13	11	Number of OTH-FT subsequ.	13	17
Duration retirement (Ret)	14	35	Number of CW periods	14	23
Number of FT-CW subsequences	15	24	Age at 1st EDU->CIV transition	15	16
Age at 1st PT spell	16	29	Number of UNE-FT subsequ.	16	18
Age at 1st MNC->M1C trans.	17	22	Duration full-time employ. (FT)	17	19
Number of PT periods	18	27	Number of FT-CW subsequences	18	21
Complexity work life (entropy)	19	20	Age at 1st UNE->FT trans.	19	15
Age at 1st EDU->FT transition	20	31	Number of EDU periods	20	20
Age 23-27: end of EDU	21	13	Duration married, 3+ chi. (M3C)	21	12
Duration self-employment (Self)	22	7	Duration other (OTH)	22	7
Complexity work life (turbulence)	23	17	gender		2
Age 58-62: start of Ret	24	23			
Age at 1st EDU->FT_FT_Ret transition	25	10			
Duration part-time employ. (PT)	26	32			
Complexity work life (C)	27	18			
Duration single, no child (SNC)	28	19			
Age 58-62: end of CIV	29	34			
Age at 1st FT->CW transition	30	37			
Age at 1st CW->RET transition	31	40			
Age at 1st EDU->CIV transition	32	26			
Number of EDU periods	33	36			
Number of SELF periods	34	25			
Complexity work life (number transitions)	35	30			
Age at 1st EDU->CIV_CIV->Ret transition	36	33			
Age 38-42: end of M1C	39	28			
Age 63-65: end of CIV	40	38			
gender		1			

Notes: Results for all features selected in the Netherlands or West Germany ranked by mean importance based on the Boruta algorithm. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Second, gender channels the predictive power of some life course proxies (lower relative importance compared to the main model; see grey highlighted rows in Table 5.1). In both countries, the *duration* of retirement as well as the *ordering* and *timing* of care work experiences are among the life course predictors with the most striking decline in relative importance. The ranking of the *duration* of care work is affected to a lesser extent in the Netherlands and not at all in West Germany. In West Germany, further, the *timing* and *ordering* of care work in combination with full-time employment as well as with part-time employment seem to be particularly channeled through gender. This underlines the fact that it is not only the *duration* of care work but also its *timing* and *ordering*, as well as its combination with other forms of work, that are important for gendered pensions.

In West Germany largest differences emerge further for the age at the first part-time experience (*timing*), the age at the first transition from education to full-time employment (*ordering & timing*), the age at direct retirement from full-time employment (*ordering & timing*), and the age at the first transition from care work to full-time employment if a transition back to full-time employment follows later in life (*ordering & timing*).

Overall, gender interacts with certain life course experiences, and these appear to be particularly related to the *timing* and *ordering* of life course experiences, making these important life course proxies to consider when analysing GPGs.

5.4.1.1.2 Gender-specific pension income predictors

As a second approach, I run gender-specific feature selection models. They reveal the most important features for pension income inequality among men and among women respectively, and thus not for gender inequality. This is important for two reasons.

First, the main difference between the gender-specific and the overall inequality dimension is that only the overall inequality (as used in the main model) incorporates gendered pension

inequality. Consequently, comparing the result of gender-specific pension predictors to predictors of overall inequality is informative for getting a sense of which predictors are particularly important for gender inequality in pensions and which might be driven by one of the genders. Life course aspects are likely to be particularly relevant for the GPG if they are much more important for overall pension inequality, as the gender dimension of pension inequality is by definition not captured by gender-specific pension disparities, e.g., pension inequality among women only. This is of interest because the feature selection formally reveals the most important predictors of *overall* pension inequality, but not specifically of the Gender Pension Gap.

Second, I expected contrary associations of family life course aspects with pensions by gender, especially for the *timing* of family transitions. Moreover, the literature documents contrary associations of family characteristics such as divorce or never being married with pension income between men and women (e.g., Kreyenfeld et al., 2023; Möhring, 2015). The feature selection based on overall pension income inequality risks concealing such contrasting associations by gender since they might cancel each other out when pooled together. Gender-specific Life Course Feature Selection likely unveils such variables with different pension returns by gender.

Table 5.2 Selected features for overall, female- and male-specific pension inequality, ranked.

West Germany				Netherlands			
features	Rank Main	Rank Men	Rank Women	features	Rank Main	Rank Men	Rank Women
Duration civil service (CIV)	1	3	1	Duration care work (CW)	1	31	17
Duration care work (CW)	2	36	2	Duration education (EDU)	2	1	8
Number of CW periods	3	33	23	Age at 1st OTH->FT transition	3	30	21
Age at 1st CW spell	4	40	25	Complexity work life (entropy)	4	14	1
Age at 1st CIV->RET transition	5	5	3	Duration retirement (Ret)	5	9	15
Age at 1st FT->Ret transition	6	21	38	Duration self-employment (Self)	6	3	32
Duration full-time employment (FT)	7	6	7	Number of CW-OTH subsequences	7	25	22
Age at 1st CW->PT_PT->Ret trans.	8	36	33	Duration single, no children (SNC)	8	19	12
Age at 1st CIV spell	9	14	6	Age at 1st EDU spell	9	12	2
Number of CIV periods	10	19	21	Age 23-27: end of EDU	10	11	5
Age at 1st CW->FT_FT->CW trans.	11	36	37	Duration civil service (CIV)	11	17	10
Number of EDU-CIV subsequences	12	4	16	Age at 1st CW spell	12	27	19
Duration education (EDU)	13	1	11	Number of OTH-FT subsequences	13	23	20
Duration retirement (Ret)	14	30	48	Number of CW periods	14	32	28
Number of FT-CW subsequences	15	38	15	Age at 1st EDU->CIV transition	15	4	11
Age at 1st PT spell	16	42	30	Number of UNE-FT subsequences	16	26	30
Age at 1st MNC->M1C transition	17	8	43	Duration full-time employed (FT)	17	16	27
Number of PT periods	18	41	46	Number of FT-CW subsequences	18	24	18
Complexity work life (entropy)	19	29	19	Number of EDU periods	19	13	16
Age at 1st EDU->FT transition	20	11	45	Complexity work life (C)	20	18	6
Age 23-27: end of EDU	21	17	5	Number of EDU-CIV subsequences	21	15	7
Duration self-employment (Self)	22	2	40	Age at 1st SELF->Ret transition	22	2	25
Complexity work life (turbulence)	23	32	20	Age 58-62: start of Ret	23	7	26
Age 58-62: start of Ret	24	35	17	Number of MNC-M1C-M2C-DIV subsequences	24	20	4
Age at 1st EDU->FT_FT->Ret trans.	25	15	4	Age at 1st OTH spell	25	22	14
Duration part-time employment (PT)	26	34	9	Number of MNC-M1C-DIV subsequences	26	28	13
Age at 1st EDU->CIV transition	32	7	28	Age at 1st CIV->RET transition	27	10	29
Number of FT periods	39	12	29	Number of MNC-M2C-DIV subsequences	28	21	3
Number of FT-FT subsequences	40	9	35	Duration married, no child. (MNC)	29	6	24
Age 1st SNC->M1C_M1C->M2C transition	45	10	47	Duration part-time employment (PT)	30	29	9
Duration single, children (S1C)	46	43	10	Duration married, 2+ child. (M2C)	31	5	23
Age 63-65: end of CIV	47	25	22	Duration married, 1 child (M1C)	32	8	31
Number of DIV_M1C subsequences	48	37	8				

Notes: Results for features selected in the Netherlands or West Germany ranked by mean importance based on the Boruta algorithm. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

I. Differentiation between pension predictors driven by gendered pension inequality and by one gender.

Comparing the relative importance across the pooled and the gender-specific model (via the ranks) indeed reveals several pension predictors that are only important in the pooled model (see Table 5.2, grey highlighted rows). This implies that these life course aspects hold particular significance in the context of gendered pension inequality, as gender-specific pension disparities do not entail the gender dimension of pension inequality. For example, in both countries, the *timing* and *ordering* of care work are ranked much more important for predicting the overall inequality compared to the pension inequality among men, and women both. In the Netherlands this is also the case for the *duration* of care work and in Germany for several life course aspects related to the *timing* of transitions from care work (for a more detailed discussion, see Appendix A5.3). The results are thus very much in line with the first approach.

Both approaches reinforce the impression that the results are not only valid for overall pension inequality but also for the Gender Pension Gap and prove the robustness of the relevance of *timing* and *ordering* of life course events, particularly of unpaid care work. Furthermore, binary associations of single variables with the GPG confirm that care work-related life course aspects are more strongly associated with the GPG than the duration of full-time employment (for more details, see Appendix A5.3).

Moreover, there are life course aspects which are much more or exclusively important for gender-typical pension inequality. This is for example the case for the duration of part-time employment in both countries (only among top-10 pension predictors for women, see Table 5.2) and the number of full-time employment periods in West Germany (only important for

pension inequality among men, see Table 5.2). As expected, this also applies to several family life course characteristics in both countries (see below).

Other life course aspects seem to be driven by one gender (i.e., they are important for the overall inequality and inequality among one gender). For example, education-related life course aspects are particularly important for pension inequality among men (see Table 5.2). However, this does not rule out that these aspects are also important for gendered pension inequality.

II. Family life course aspects concealed in pooled income inequality model

As expected, some family life course aspects are important predictors for male- or female-specific pension inequality in both countries whereas they were not selected as relevant predictors in the pooled model. Several family life course aspects make it to the top ten predictors for gender-specific pension inequality in both countries. For example, life course aspects related to divorce are very important for pension inequality among women in both countries (see Table 5.2, columns 2 and 3 per country, see Appendix A5.3 for more details). This result suggests that feature selection approaches applied to group inequalities in outcomes, such as gender pension or wage gaps (e.g., Bonaccolto-Töpfer & Briel, 2022), risk concealing variables that might be associated with the gender gaps through gender differences in returns.

5.4.1.2 Sensitivity of the Life Course Feature Selection

The sensitivity and other additional analyses are described in detail in Appendix A5.5.

For social policy-relevant sensitivity analyses, I include survivor pensions in the total pensions, exclude individuals with zero pensions (i.e., focusing on pension recipients only) and, for the Netherlands only, control for household size. The first two considerations are

critical conceptual parameters for the level of gendered pension inequalities (see Chapter 2) and differ across studies in the field (Table 2.1).

The overall findings remain the same in the Netherlands and West Germany (see Appendix A5.5.2). The most considerable change occurs in the relative importance of the duration of full-time employment which increases when excluding individuals without pension receipt in West Germany. However, the duration of care work is still a much more important predictor of pension inequalities.

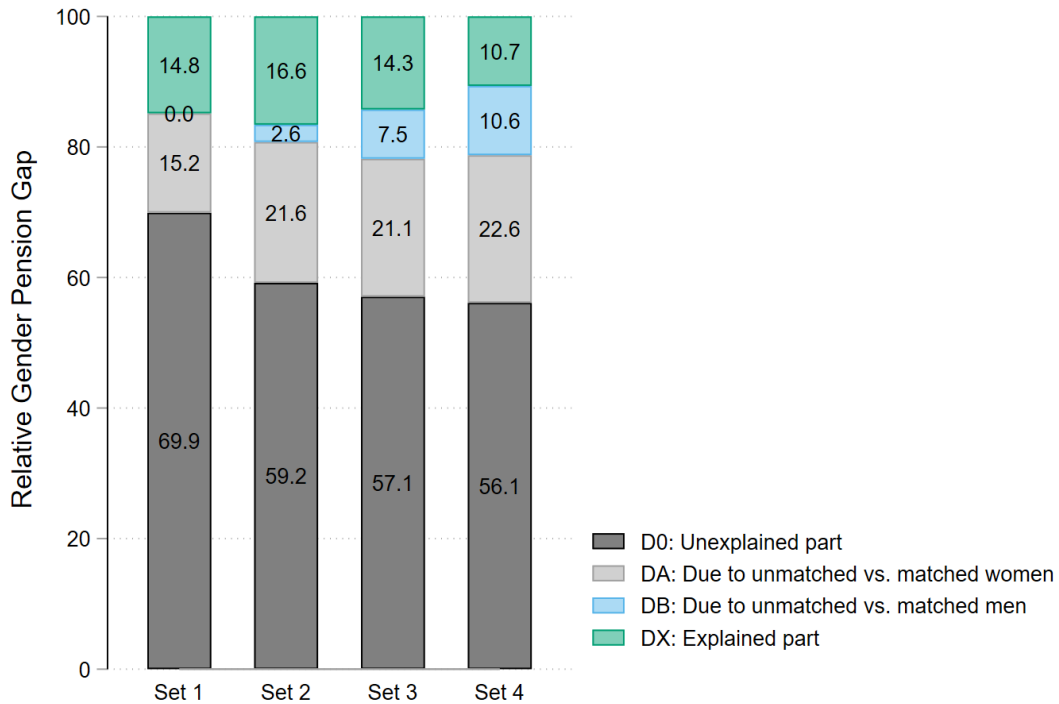
I further test three alternative feature selection algorithms (recursive feature elimination, Lasso, Vita). While clear differences in the ranking of relevant pension predictors occur, the overall pattern is the same (see Appendix A5.6.1, Table A5.6). All algorithms identify life course aspects related to care work among the top pension predictors, and, as more important predictors compared to the duration of full-time employment. The results are further robust to some changes in the functional form of the life course variables and the set of variables considered (see A5.5.3).

Consequently, the main finding is robust: care work life course aspects are more important life course pension predictors than full-time employment across several sensitivity analyses. Future research should therefore consider care work characteristics when analysing pension inequalities, independent of the conceptual choices taken.

5.4.2 Step II: Decomposition of Gender Pension Gaps

The Gender Pension Gap amounts to €13,449 (60.81%) in West Germany and €12,489 (46.37%) in the Netherlands. Table 5.3 and Table 5.4 show the variables used for the stepwise decomposition, the matched sample by gender, and the decomposition results for the four components of the Ñopo decomposition. They document a decreasing size of the matched samples of men and women with the increasing number of matching variables used in the model. The “curse of dimensionality” (Hamjediers & Sprengholz, 2023, S. 12) becomes evident at a certain point, where only a very small proportion of men and women are matchable, and these highly selective groups result in limited interpretability (see Appendix A5.5.5 for details). The risk of selective samples is particularly high for this analysis because of the overall small samples used. Consequently, I discuss the outcomes of the models with a reasonable share of matched men and women: models 1-3 for the Netherlands (20% of women matched in model 3) and models 1-4 for West Germany (30% of women matched in model 4). Figure 5.5 and Figure 5.6 depict the relative share of the GPG attributable to the decomposition components. Overall, for West Germany, I can attribute up to 44% of the GPG to gender differences in common support or compositional differences in the most important pension predictors (model 4) and up to 69% in the Netherlands (model 3).

Figure 5.5 Ñopo decomposition of Gender Pension Gap in West Germany



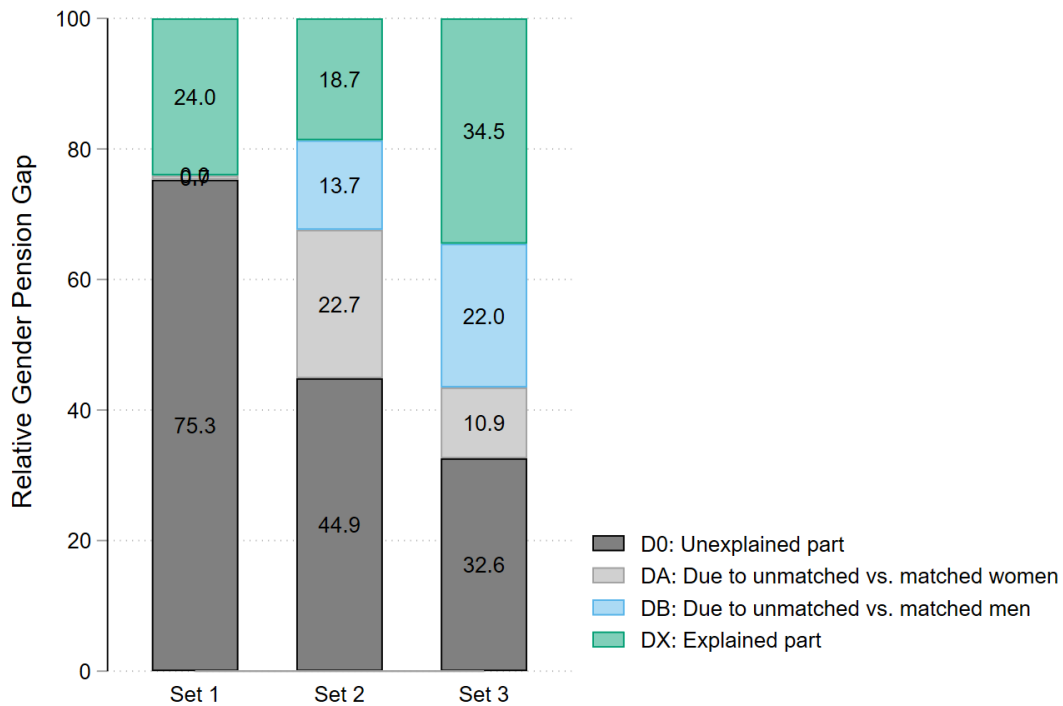
Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.3). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table 5.3 Ñopo decomposition of Gender Pension Gap in West Germany

Models	Variable set 1	Variable set 2	Variable set 3	Variable set 4
	Duration care work, duration civil services	set 1+ age at first CW spell, age at first transition of CIV->Ret & FT->Ret subsequence	set 2 + duration FT, age at first transition of CW->PT_PT->RET and CW->FT_FT->CW subsequences, age at 1st civil service period	set 3 + duration EDU, number of subsequence FT->CW
Variables				
% matched women	64	41	35	30
% matched men	100	95	86	72
N of strata	13	79	280	475
N of strata matched	6	19	46	62
D0	9405	7959	7674	7547
DA	2048	2902	2842	3043
DB	0	356	1015	1426
DX	1997	2232	1918	1433

Note: Ñopo decomposition of the Gender Pension Gap into the four components presented in equation 2. % matched refers to the share of women and men respectively matched for the respective characteristics controlled for. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure 5.6 Ñopo decomposition of Gender Pension Gap in the Netherlands.



Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.4). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table 5.4 Ñopo decomposition of Gender Pension Gap in the Netherlands

Models	Variable set 1	Variable set 2	Variable set 3
Variables	Duration care work, duration education	Set 1 + Complexity work life courses (entropy)	Set 2 + duration self-employment
% matched women	72	25	20
% matched men	100	89	75
N of strata	24	116	188
N of strata matched	10	36	37
D0	9403	5603	4071
DA	82	2841	1357
DB	0	1712	2749
DX	3003	2332	4311

Notes: Ñopo decomposition of the Gender Pension Gap into the four components presented in equation 2. % matched refers to the share of women and men respectively matched for the respective characteristics controlled for. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Compositional differences over common support (DX): The share of the GPG that is ‘explained’ by differences in the distribution of characteristics for the matched samples of men and women (i.e., over common support) ranges between 10.7-16.6% for models 1-4 in West Germany (Figure 5.5, DX). In the Netherlands, the share of the GPG due to compositional differences over common support reaches higher levels with 18.7-36.5% (Figure 5.5, DX). These are overall lower compositional components than previous literature has found using the KOB decomposition and assuming common support, especially for Germany (up to 56%, Frommert & Strauß, 2013; Hänisch & Klos, 2014) but also internationally (see Table 3.1 for an overview).⁶² Thus, previous literature likely overestimated the compositional component, whereas there are significant shares of the GPG due to the lack of common support (DA, DB; see below).

In West Germany, 14.8% of the GPG alone is due to differences in the *duration* of civil service and care work between men and women with comparable characteristics (64% of women and 100% of men matched based on the two most important pension predictors, model 1). I find the largest explained part for model 2, which adds *timing* of care work as well as the age at the transition to retirement directly from civil service or full-time employment (*timing & ordering*) based on which only 41% of all women and 95% of men can be matched. The GPG would be reduced by 16.6% if the matched men and women had the same distribution of the observed characteristics (Figure 5.5). In the Netherlands, almost one-quarter of the GPG alone is due to differences in the *duration* of care work and education from age 18 over the life course between matched women and men (72% of women and 100% of men matched, model 1, Figure 5.6). I find the largest compositional component of the GPG in the Netherlands in model 3 which includes additionally the

⁶² This still the case when swapping the reference groups and making the parameters more similar to the parameters used in most of the KOB decomposition applications even though the explained shares become larger in Germany (see A5.5.5). However, previous studies have used different sets of variables.

unpredictability of working life courses as well as the years spent in self-employment and ‘explains’ more than one-third of the GPG (34.5%). However, the group of matched women is likely selective (only 20% of women are matched).

To discern the key drivers contributing to the explained share, I run Kitagawa-Oaxaca-Blinder decompositions over the common support for the models with sufficient common support (1-3 in West Germany and model 1 in the Netherlands, see Appendix A5.4). In both countries, large shares of the remaining GPGs over the matched samples of men and women of model 1 are due to the more years women have spent on average doing care work compared to men (21% of the remaining GPG in the Netherlands and 9% in West Germany). In the Netherlands, another 19% are because of women’s fewer years in education and in West Germany, 11% are due to fewer years in civil service. For the smaller sample over common support for model 2 in West Germany, the explained share due to differences in the duration of care work becomes insignificant (because women with many years of care work are not matched, see discussion of DA and DB component below). The share due to differences in years worked in civil service drops to 8% but remains significant. 14% are due to gender differences in the age at which individuals transition from full-time employment to retirement (driven by women not experiencing such a pattern).

However, the DX component highly depends on the matching outcome, i.e., how the common support is defined. The DA and DB components of the decomposition provide insight into how differences between the matched and unmatched samples attribute to the GPG.

Gender differences in support (DA & DB). The Ñopo decomposition contributes to previous literature by uncovering which shares of the GPG are due to pension income

differences of matchable and unmatched men and women respectively. This provides insights into how much of the GPG is attributable to gender-exclusive life course experiences (i.e., combinations of life course experiences ('strata') that appear only for one gender).

In both countries, the DA and DB components of the GPG play a significant role, but more consistently across models in West Germany. In West Germany (model 1), a larger part of the GPG (15.2%) is attributable to lower pension income among the unmatched compared to the matched women (DA, see Table 5.3) compared to the compositional component DX. The share rises to 22.6% in model 4, suggesting that the extremely gender-specific combination of the observed life course aspects plays the main role for the Gender Pension Gap in West Germany. In other words, the pension income of the selective group of women who are matchable on their life course experiences with men (30-41% of women in models 2-4) differs strongly from the pension income of women who are not matchable to men, so that the resulting pension income difference attributes to more than one-fifth of the GPG (models 2-4). In the Netherlands, the lower pension income of unmatched women contributes to 22.7% of the GPG in model 2 and 10.9% in model 3, while in model 1 gender differences in support are not relevant yet.

In both countries, higher pension income among the unmatched compared to the matched men (DB) only becomes relevant for the GPG from model 2 onwards because all men are matched in model 1. In West Germany, the share of the GPG due to DB increases from 2.6% to 10.6% in model 4. In the Netherlands, it reaches higher levels and increases from 13.7% in model 2 to 22% in model 3 (for a potential reason, see A5.6.4).

The results strongly suggest that life course experiences of men and women differ severely (low shares of men and women matched) and gender-specific combinations of such

experiences are related to particularly high or low pensions and thus to large shares of the GPG (DA and DB). Consequently, the combinations of life course aspects experienced by only one of the genders and different pension returns for these, are an important reason for the Gender Pension Gap (*expectation 4* confirmed).

5.4.2.1 Exploring the mechanisms behind DA and DB

1. Summary characteristics of the different samples

To further understand the mechanisms behind the DA and DB components, Hamjediers and Sprengholz (2023) advise comparing the characteristics across groups and matching status. Table 5.5 and Table 5.6 show the distribution over the selected variables for the samples of matched and unmatched men and women for model 3. In both countries, the main difference in characteristics between the unmatched and matched women is that the matched women engaged only for relatively few years in unpaid care work or not at all. In other words, matched women most likely receive higher pensions because they did not spend substantial time in care work (and, for West Germany, started their care work engagement relatively late in their life courses). For example, in West Germany, 92.4% of matched women did not spend any year in unpaid care work and among those who did, almost all had their first care period late in their life course (Table 5.5). In the Netherlands, 89% of matched women did not report unpaid care work and 10% spent a relatively low number of years doing care work (Table 5.6).

Table 5.5 Observed characteristics by gender and matching status for West Germany (model 3)

	Men unmatched	Men matched	Women matched	Women unmatched
N	109 (6.7%)	697 (42.9%)	290 (17.9%)	527 (32.5%)
Pension income, mean	28,605.58	21,099.72	11,507.32	7,101.99
Duration civil service				
No	0 (0.0%)	642 (92.1%)	259 (89.3%)	504 (95.6%)
1st tercile	37 (33.9%)	10 (1.4%)	7 (2.4%)	19 (3.6%)
2nd tercile	51 (46.8%)	11 (1.6%)	8 (2.8%)	3 (0.6%)
3rd tercile	21 (19.3%)	34 (4.9%)	16 (5.5%)	1 (0.2%)
Duration care work				
No	109 (100.0%)	691 (99.1%)	268 (92.4%)	2 (0.4%)
1st quartile	0 (0.0%)	4 (0.6%)	12 (4.1%)	123 (23.3%)
2nd quartile	0 (0.0%)	2 (0.3%)	10 (3.4%)	126 (23.9%)
3rd quartile	0 (0.0%)	0 (0.0%)	0 (0.0%)	138 (26.2%)
4th quartile	0 (0.0%)	0 (0.0%)	0 (0.0%)	138 (26.2%)
Age 1st care work				
Not such pattern	109 (100.0%)	691 (99.1%)	268 (92.4%)	2 (0.4%)
very early	0 (0.0%)	1 (0.1%)	2 (0.7%)	136 (25.8%)
rather early	0 (0.0%)	0 (0.0%)	0 (0.0%)	138 (26.2%)
rather late	0 (0.0%)	0 (0.0%)	0 (0.0%)	138 (26.2%)
late	0 (0.0%)	5 (0.7%)	20 (6.9%)	113 (21.4%)
Age transition CIV->Ret				
Not such pattern	28 (25.7%)	648 (93.0%)	264 (91.0%)	522 (99.1%)
very early	28 (25.7%)	7 (1.0%)	6 (2.1%)	0 (0.0%)
rather early	16 (14.7%)	16 (2.3%)	6 (2.1%)	2 (0.4%)
rather late	13 (11.9%)	14 (2.0%)	11 (3.8%)	2 (0.4%)
late	24 (22.0%)	12 (1.7%)	3 (1.0%)	1 (0.2%)
Age transition FT->Ret				
Not such pattern	92 (84.4%)	233 (33.4%)	190 (65.5%)	465 (88.2%)
very early	1 (0.9%)	104 (14.9%)	38 (13.1%)	18 (3.4%)
rather early	6 (5.5%)	110 (15.8%)	30 (10.3%)	15 (2.8%)
rather late	5 (4.6%)	129 (18.5%)	18 (6.2%)	9 (1.7%)
late	5 (4.6%)	121 (17.4%)	14 (4.8%)	20 (3.8%)
Duration FT employ.				
No	38 (34.9%)	55 (7.9%)	63 (21.7%)	85 (16.1%)
1st quintile	27 (24.8%)	31 (4.4%)	49 (16.9%)	170 (32.3%)
2nd quintile	24 (22.0%)	43 (6.2%)	32 (11.0%)	178 (33.8%)
3rd quintile	10 (9.2%)	126 (18.1%)	62 (21.4%)	78 (14.8%)
4th quintile	5 (4.6%)	211 (30.3%)	49 (16.9%)	11 (2.1%)
5th quintile	5 (4.6%)	231 (33.1%)	35 (12.1%)	5 (0.9%)
Age start of CW->PT_PT->RET				
Not such pattern	109 (100.0%)	697 (100.0%)	290 (100.0%)	425 (80.6%)
very early	0 (0.0%)	0 (0.0%)	0 (0.0%)	26 (4.9%)
rather early	0 (0.0%)	0 (0.0%)	0 (0.0%)	26 (4.9%)
rather late	0 (0.0%)	0 (0.0%)	0 (0.0%)	25 (4.7%)
late	0 (0.0%)	0 (0.0%)	0 (0.0%)	25 (4.7%)
Age 1st civil service				
Not such pattern	0 (0.0%)	642 (92.1%)	259 (89.3%)	504 (95.6%)
very early	41 (37.6%)	10 (1.4%)	2 (0.7%)	2 (0.4%)
rather early	17 (15.6%)	15 (2.2%)	12 (4.1%)	11 (2.1%)
rather late	14 (12.8%)	22 (3.2%)	12 (4.1%)	6 (1.1%)
late	37 (33.9%)	8 (1.1%)	5 (1.7%)	4 (0.8%)

Notes: Based on Nopo decomposition of the Gender Pension Gap presented in equation 2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

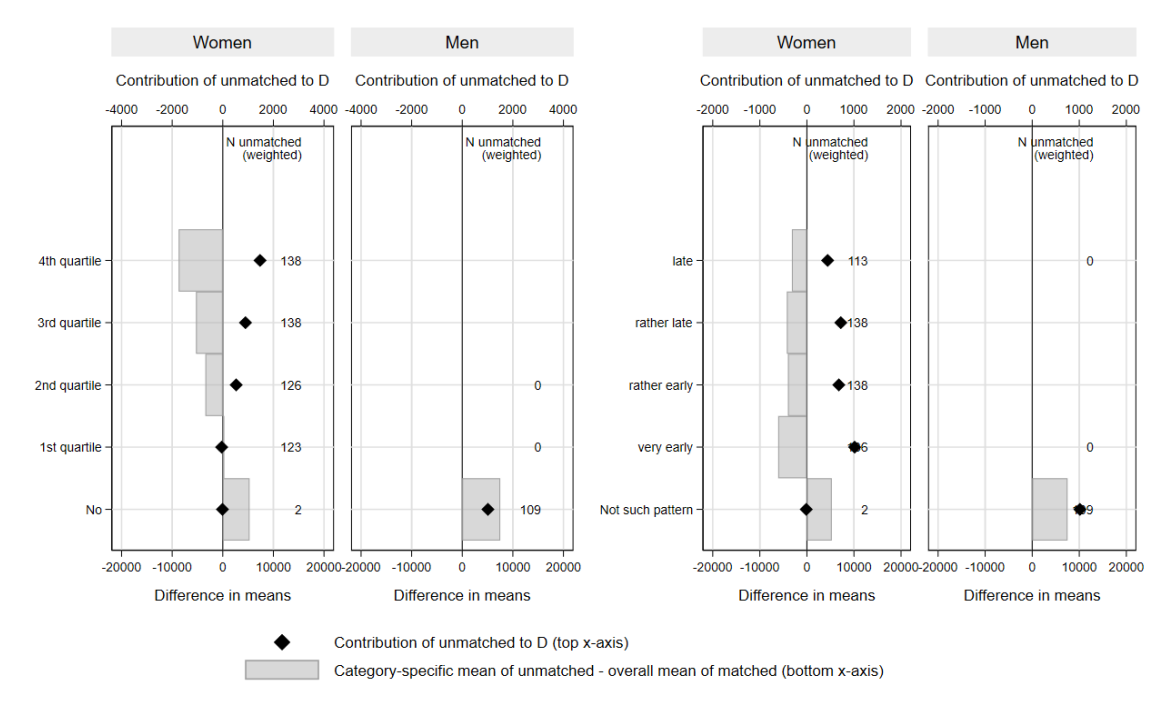
Table 5.6 Observed characteristics by gender and matching status for the Netherlands (model 3)

	Men unmatched	Men matched	Women matched	Women unmatched
N	126 (11.6%)	374 (34.3%)	119 (10.9%)	470 (43.2%)
Pension income, mean	35,091.61	24,182.95	15,800.02	14,099.95
Duration care work				
No	121 (96.0%)	370 (98.9%)	106 (89.1%)	13 (2.8%)
1st quartile	4 (3.2%)	3 (0.8%)	12 (10.1%)	101 (21.5%)
2nd quartile	0 (0.0%)	1 (0.3%)	1 (0.8%)	118 (25.1%)
3rd quartile	0 (0.0%)	0 (0.0%)	0 (0.0%)	120 (25.5%)
4th quartile	1 (0.8%)	0 (0.0%)	0 (0.0%)	118 (25.1%)
Duration education				
No	44 (34.9%)	319 (85.3%)	95 (79.8%)	386 (82.1%)
1st quartile	11 (8.7%)	13 (3.5%)	7 (5.9%)	31 (6.6%)
2nd quartile	21 (16.7%)	7 (1.9%)	7 (5.9%)	26 (5.5%)
3rd quartile	18 (14.3%)	24 (6.4%)	7 (5.9%)	12 (2.6%)
4th quartile	32 (25.4%)	11 (2.9%)	3 (2.5%)	15 (3.2%)
Complexity work life course				
1st decile	2 (1.6%)	38 (10.2%)	13 (10.9%)	56 (11.9%)
2nd decile	3 (2.4%)	55 (14.7%)	12 (10.1%)	39 (8.3%)
3rd decile	6 (4.8%)	47 (12.6%)	10 (8.4%)	46 (9.8%)
4th decile	6 (4.8%)	55 (14.7%)	10 (8.4%)	38 (8.1%)
5th decile	17 (13.5%)	42 (11.2%)	15 (12.6%)	35 (7.4%)
6th decile	37 (29.4%)	37 (9.9%)	7 (5.9%)	28 (6.0%)
7th decile	28 (22.2%)	32 (8.6%)	15 (12.6%)	34 (7.2%)
8th decile	3 (2.4%)	43 (11.5%)	21 (17.6%)	42 (8.9%)
9th decile	12 (9.5%)	22 (5.9%)	13 (10.9%)	62 (13.2%)
10th decile	12 (9.5%)	3 (0.8%)	3 (2.5%)	90 (19.1%)
Duration self-employment				
No	62 (49.2%)	348 (93.0%)	103 (86.6%)	430 (91.5%)
1st tercile	25 (19.8%)	3 (0.8%)	4 (3.4%)	17 (3.6%)
2nd tercile	15 (11.9%)	15 (4.0%)	6 (5.0%)	13 (2.8%)
3rd tercile	24 (19.0%)	8 (2.1%)	6 (5.0%)	10 (2.1%)

Notes: Based on Ñopo decomposition of the Gender Pension Gap presented in equation 2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure 5.7 confirms these patterns by showing the difference between the mean pension income of the unmatched women per variable category and the overall pension income of matched women, as well as the contribution to the GPG for each category. They thus indicate which categories per variable are most important for the overall DA component of the decomposition (Sprengholz & Hamjediers, 2022). For the *duration* of care work, a linear association is depicted: the more years unmatched women have spent doing care work, the higher the difference in pension income compared to the matched women overall and the larger the contribution to the GPG (D, see Figure 5.7, Figure A5.4).

Figure 5.7 Contributions of years spent in care work (left panel) and age at first care work spell (right panel) to GPG (DA/DB), West Germany (model 3)



Notes: Bars depict the difference between the mean pension income of the unmatched individuals per variable category and the overall pension income of matched individuals by decomposition groups. The squares depict the contribution of the category-specific pension difference to the overall GPG (D) and thus to DA/DB by weighting it by the number of unmatched individuals in each category relative to all individuals for each group. Based on Sprengholz & Hamjediers (2022) and the *nopo* package in *stata* (Sprengholz & Hamjediers, 2023). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

In both countries, DA is driven by particularly large pension penalties for women who spent many years in care work. In West Germany, the pension penalties are additionally largest for unmatched women who had their first care work spells already at young ages (*timing*; these contribute most to the GPG, see Figure 5.7).

In West Germany, the large DA component is further driven by the lack of any years in civil service of unmatched women (*duration*; e.g., Figure A5.5) as well as no direct transitions from full-time employment or civil service to the retirement of unmatched women (*timing* and *ordering*; e.g., Figure A5.6). Fewer years in full-time employment and no educational years after age 17 likely play a role as well (*duration*; see Table 5.5, not shown for education).

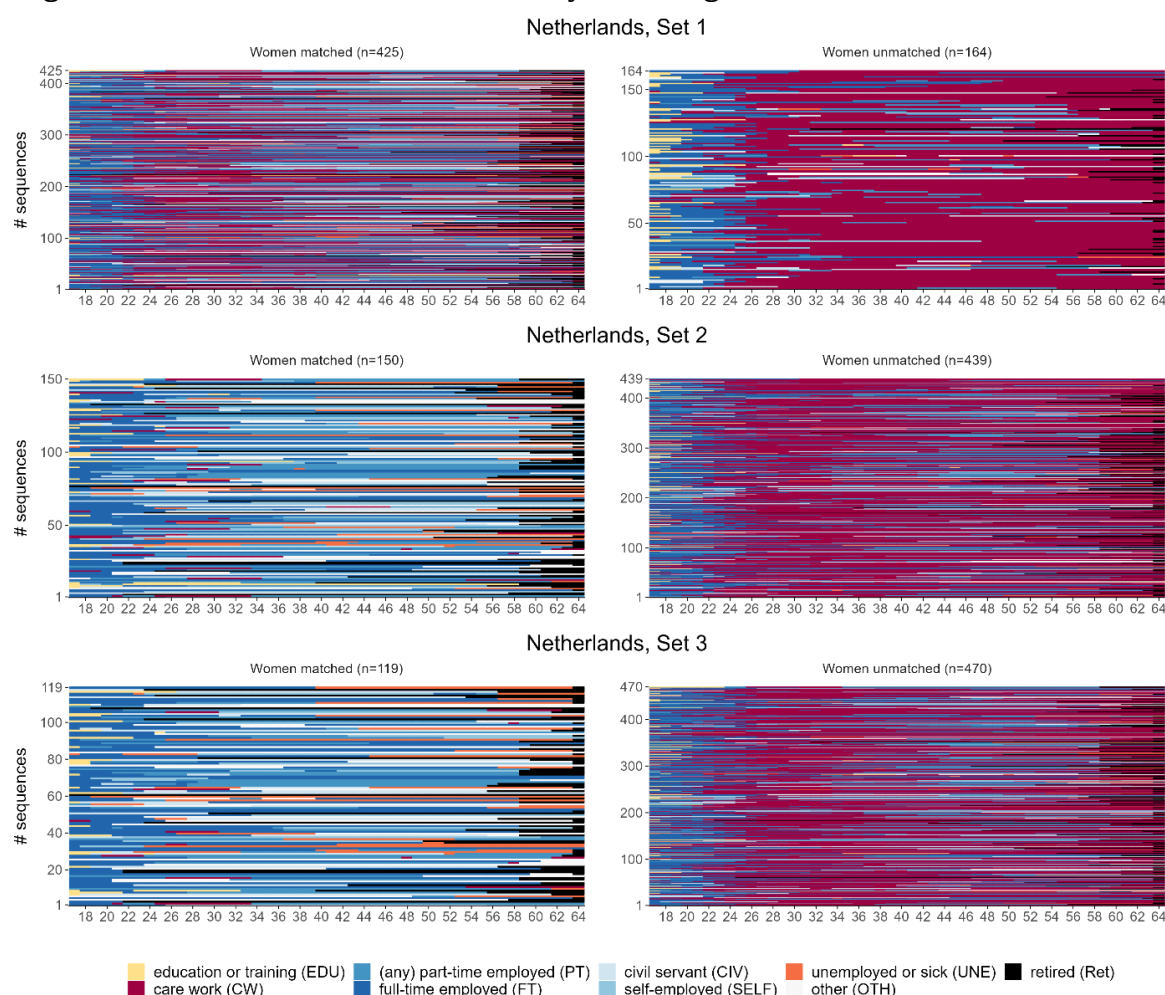
For the Netherlands, the substantial part of the gap resulting from higher pensions for the unmatched compared to matched men (DB) is due to more years of education of the unmatched men (e.g., no educational years after age 17 for 85% of the matched compared to 35% of the unmatched men, Table 5.6). In West Germany similarly, the DB component is due to more years of education for the unmatched men (model 4, not shown), but also more years spent as civil servants instead of private full-time employment (Table 5.5).

2. Life course visualisation of the different samples by gender and matching status

When using life course aspects as characteristics in the Ñopo decomposition, it is beneficial to visualize the life courses by groups and sample status. This additional step helps in interpreting the DA and DB components because the distribution over different life course aspects alone does not fully uncover how the combination of events unfolds throughout the life course. Figure 5.8 compares the life courses of matched and unmatched women for models 1-3 in the Netherlands. Matching on the combination of the overall complexity of working life courses with the years spent in care work and education (model 2) effectively

distinguishes between women with multiple or long care work-related employment interruptions (unmatchable with men and lower pension income, see Figure A5.4) and women with no or only short care work periods but otherwise continuous employment (matchable with men). In contrast, matching based on the duration of care work and education only (model 1) does not succeed in clearly differentiating patterns of life courses with different pension incomes and therefore does not lead to a significant DA component. For example, in both samples, there are life courses with long care-related employment interruptions.

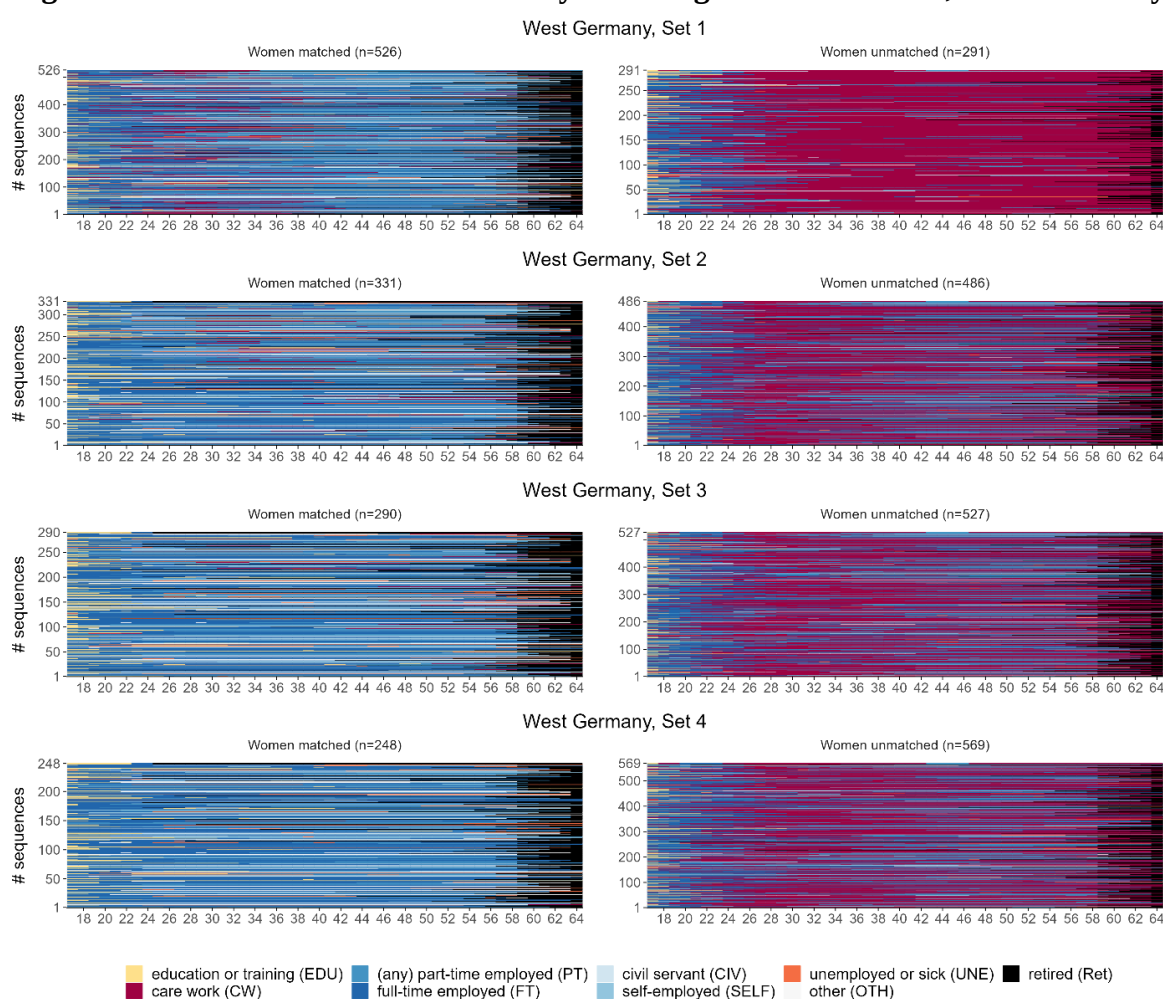
Figure 5.8 Work life courses of women by matching status and model, Netherlands



Notes: Samples based on $\tilde{N}$ opo decomposition of the Gender Pension Gap presented in equation 2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Introducing more dimensions for matching (duration of self-employment in model 3) makes women with short care work interruptions unmatchable and increases the share of matched women with long periods of sickness or ‘other’ states. Due to lower pensions for sickness or ‘other’ periods, this essentially leads to a decrease in the DA component because the pension incomes of matched and unmatched women converge. This visualises the curse of dimensionality underlying exact matching applied in the $\tilde{\text{Nopo}}$ decomposition.

Figure 5.9 Work life courses of women by matching status and model, West Germany



Notes: Samples based on $\tilde{\text{Nopo}}$ decomposition of the Gender Pension Gap presented in equation 2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure 5.9 visualizes that DA in model 1 for West Germany compares the pension income of women with no or relative early care-related employment interruptions (matched) with women who barely return to any employment after their first care work period

(unmatched). Adding the *timing* of the first care work episode in model 2, women with early care-work-related employment interruptions are not matchable anymore, and among matched women, there are fewer and later care work interruptions. This increases the DA component (see Figure 5.5).⁶³

Thus, matching on the combination of the duration and the *timing* of care work (among others) in West Germany leads to a similar differentiation of work life courses for women compared to matching on the duration of care work and the overall complexity of work life courses in the Netherlands. Consequently, the conclusion is the same in both countries: the DA component of the GPGs is due to lower pensions of women with long, continuous care work trajectories. Additionally, the overall *complexity* of working life courses might substitute parts of the importance of the *timing* and *ordering* of care work in the Netherlands, since complexity in women's life courses seems mostly related to the *timing* and *ordering* of care work experiences. This reinforces the explanation that the overall *complexity* of working life courses scores higher when it comes to variable importance than the *timing* and *ordering* of care work for predicting pensions in the Netherlands (see section 5.4.1). In interaction with other variables, the overall complexity seems to be a better predictor of pensions.

5.4.2.2 Sensitivity of the decomposition

Hamjediers and Sprengholz (2023) stress the risk that the lack of common support, and substantial shares of the gap due to it (DA and DB) in Ñopo decompositions might be an artifact of low sample size. As a sensitivity check, I follow their suggestion to apply less strict matching techniques. Using propensity score matching overall confirms the lack of common support: large shares of women are not matchable (DA). However, almost all men

⁶³ Adding more matching variables fine grains the selectivity of the matched sample but does not change the main pattern (Figure 5.9, models 3 and 4).

are matchable when relaxing the matching approach. As a result, the DB component cannot be confirmed robust whereas substantial shares of the GPG due to DA remain for all relevant models of the main analysis, even though at a decreased level (particularly for the Netherlands). Overall, this confirms the main finding of the differently rewarded gender-exclusive combinations of life course aspects related to care work (DA) playing a major role for GPGs (for more details see Appendix A5.5.5).

Moreover, as usual in decomposition analyses, the results depend to some extent on the reference chosen. When the references are swapped, the explained shares tend to be larger in West Germany for all models and in the Netherlands for model 1, while they decrease in the Netherlands for the other models (see A5.5.5).

5.5 Discussion

This paper applies an innovative combination of machine learning, sequence analysis, and decomposition approaches to study gender inequalities in pension income in West Germany and the Netherlands. In the first step, the Life Course Feature Selection identifies the most important predictors for pensions out of hundreds of variables, that cover the duration, timing, ordering, and complexity of work and family life courses. The second step decomposes the gender gap in pension income based on the life course variables selected in the first step to reveal how the most important predictors of overall pension inequality are associated with the gender gap therein. Combining the Life Course Feature Selection with decomposition techniques contributes to the previous literature that either focussed on a selective set of life course summary measures, primarily the duration of employment, or on holistic life course ideal types for assessing gendered pension inequalities. The Ñopo decomposition goes beyond standard decomposition approaches by emphasizing and quantifying the role of gender-exclusive life course experiences while overcoming the common support problem (Ñopo, 2008).

The Life Course Feature Selection detects that life course complexities beyond the *duration* of full-time employment, as commonly used as the main predictor of GPGs so far, matter for pension income inequalities. This is generally in line with studies on gendered pension inequality that have applied a more holistic perspective and related whole life course patterns to old-age outcomes (Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022; Rowold et al., 2023). However, the feature selection manifests that specifically the *timing* and *order* of life course events matter for overall pension inequality in West Germany, and the *complexity* of the overall work life courses in the Netherlands. Sensitivity tests further highlight that the *timing* and *ordering* of life course experiences, especially for unpaid care work, make a difference in *gender* inequalities in pensions in particular. Future studies on GPGs should therefore consider such life course proxies in their analyses.

In both countries, life course aspects related to unpaid care work are among the most important predictors of pension income. The *duration* and also the *timing* and *ordering* of care work matter more for pension income than the years spent in full-time employment over the life course. This finding holds in both contrary pension logics, the Bismarckian and the Beveridgean system. Even though unpaid care work can be seen as merely the different side of the coin of (full-time) employment, this finding is highly relevant as pension systems have the measures to compensate for care work engagement—but these are just too inefficient to counteract the disadvantages of care work in pensions. Applied to gendered pension inequality, one component of GPGs is how men and women divide paid and unpaid work over the life course, but the other component is how (un)equal both sides of the same coin are rewarded in the pension system (see Chapters 3 and 4).

For the overall pension inequality, family life course aspects such as marriage and parenthood status as well as work life course aspects related to part-time employment only play a marginal role. Additional gender-specific analyses show that such life course aspects

are more relevant for the male- and female-specific pension inequalities. The country comparison further reveals that the overall instability of working life courses as well as educational inequalities are more reproduced in the Dutch Beveridgean system with the (quasi) mandatory occupational pension scheme. For the German system, the highly generous pensions for civil servants unfold their association with pension inequality through the duration, timing, and ordering over the life course – the years spent in civil service are the most important pension predictor in West Germany.

In the second analytical step, the introduction of the Ñopo decomposition to the study of gender inequalities in old age confirms the significance of the life course dimensions of *order*, *timing*, and *complexity* as well as life course aspects related to care work in understanding the Gender Pension Gap, as identified by the feature selection within the broader context of pension inequality. This becomes visible through two components of the decomposition: differences in observed characteristics among comparable men and women (explained share, DX) as well as pension income differences between matchable and unmatched women (DA), though the magnitude depends on the number of covariates considered. The results suggest that life course aspects related to care work are the main characteristic of unmatched women, i.e., almost exclusively occurred for women. The more years unmatched women engaged in care work (*duration*) and (for West Germany only) the earlier they had their first care work period (*timing*), the higher the pension penalty compared to matched women. Substantial parts of the GPGs are ascribed to these pension differences between matched and unmatched women (DA). Extending the decomposition analyses with the KOB over the matched sample reveals that differences in care work *duration* (and *timing*) are not only among the main reasons for the DA component but also contribute to large parts of the explained component DX – if the matched sample is sufficient. This enhances the existing body of literature, which has often

relied on a limited range of life course summary measures primarily centered around the *duration* of full-time employment, thereby focussing on an unsuitable set of predictors. On the other hand, the finding is in line with studies taking a holistic perspective using sequence analyses to reveal associations of typical life course patterns with i) (women's) pension income (Madero-Cabib & Fasang, 2016; Tophoven & Tisch, 2016) or ii) the Gender Pension Gap directly (Rowold et al., 2023). These have found that typical female life courses dominated by care work are associated with the largest pension penalties or largest share of the GPG.

The Ñopo decomposition further quantifies substantial shares of the GPG due to higher pensions of men that are not matchable to women compared to those who are (DB). This component of the decomposition reaches higher levels in the Netherlands. Unmatchable men have more beneficial characteristics such as more educational years (*duration*). Given that data constraints do not allow for the consideration of wage over the life course, differences in educational years might capture parts of the association of earnings differences with gendered pension inequalities (see Chapter 4).

Overall, about half of the GPG in West Germany and up to two-thirds in the Netherlands are attributable to the most important life course aspects as identified by the Life Course Feature Selection. In the Netherlands, higher shares are attributable to differences in the endowment with life course characteristics among matchable men and women compared to West Germany (compositional component DX). Still, compared to previous decomposition studies that assumed gender comparability over the distribution of the covariates the Ñopo decomposition suggests that the compositional component of the GPG is smaller. Instead, in both the Netherlands and West Germany the largest component of the GPG is due to gender-exclusive combinations of life course aspects (DA, DB). The low shares of men and women that can be matched reveal that the life course experiences

severely differ across genders and gender-specific combinations of such experiences are related to particularly high or low pensions. Jointly, the gendered-exclusive life course experiences and different pension returns for these, are the main reasons for the Gender Pension Gap.

This stresses the pertinent impact of highly gendered work-family life courses, mostly driven by the traditionally gendered division of labour, on gender inequalities in old age. The traditional division of labour unfolds its “group-segregating fashion” (Hamjediers & Sprengholz, 2023, S. 2) by ascribing care work almost exclusively to women and resulting in highly gendered life course aspects related to care work for which no male counterparts exist. Gendered pension inequalities are then produced by a second social process of pension systems ascribing a much lower (pension) value to unpaid care work compared to paid work. Together, the welfare state policies in place over the life course incentivising a gendered division of labour, and the unequal evaluation of the gendered life course experiences through pension systems produce high GPGs. Even the Beveridgean pension system with progressive public pensions in the Netherlands cannot prevent such reproduction of gendered inequalities in old age as it is flanked with a high coverage of regressive occupational pensions.

Overall, the results thus highlight the relevance for considering life course aspects beyond the *duration* in states, namely *ordering*, *timing*, and *complexity*, to uncover the complex mechanisms related to the GPG. This is the case for the compositional component of the Gender Pension Gap (DX), but especially due to the components arising from gender-exclusive life course experiences (DA and DB). As pension systems are built on the assumption of *continuous* full-time careers (Allmendinger et al., 1993; Fasang, 2010; Foster & Ginn, 2018; Grady, 2015; Kuitto et al., 2021; Leitner, 2001), the *timing* and

ordering of life course events that deviate from this ideal type are important for pensions. For the Gender Pension Gap, this is particularly true for employment interruptions that are highly gendered and poorly rewarded in pension systems, such as the *timing* and *ordering* of care work, and its combination with other forms of work. Uncovering how the relative importance of life course proxies differs across pension systems sheds light on how they reproduce different life course inequalities to different degrees. However, despite these differences across countries, the conclusion is the same: large shares of the GPG emerge because no men are taking up the unpaid care need and unpaid care work is not valued equivalently to other forms of work in the pension system.

Empirically, the combination of the Life Course Feature Selection with the Ñopo decomposition proves very suitable for the application to GPG and can easily be extended to other group-specific and life-course-sensitive inequalities.

First, the results highlight that future research on gendered pension income inequality should consider life course aspects beyond the *duration* in (full-time) employment states. Many of the most important life course pension predictors in the Netherlands and Germany that the Life Course Feature Selection uncovered relate to the *timing* or *ordering* of life course events or the life course's overall *complexity* and have not been identified or assessed empirically previously, at least when it comes to Gender Pension Gaps. Consequently, the Life Course Feature Selection approach can solve the current “steady worker bias” (O’Rand, 1996, S. 235) tendencies in the literature of considering the *duration* of (full-time) employment as the main predictor of GPGs by default. More generally, Life Course Feature Selection provides a systematic and data-driven selection of variables from various life course aspects and prevents concealing less obvious but still relevant predictors. Otherwise, variable selection depends fully on the researcher, and it is particularly difficult to select the relevant predictors out of a pool of several hundred life

course variables. As the recent literature on GPG shows, the variable selection will be biased towards rather easily extractable and straightforward life course aspects, mostly covering the aspect of *duration*. This tendency however is contrasted by the increasingly complex and destandardised life courses (e.g., Van Winkle, 2018; Van Winkle & Fasang, 2017, 2021), which makes it increasingly important to systematically consider more diverse and less easily extractable life course elements, such as the *timing*, *ordering* and overall *complexity*.

Second, ranking the predictors by the relevance for the outcome helps to identify different sets of variables for more impartial matching. This is important since matching and thus the Ñopo decomposition is highly sensitive to the variable sets, and consequently, to the researchers' choices. Additionally, in empirical cases where there are large latent and theoretically relevant constructs such as life courses, it is difficult to disentangle which proxies out of hundreds of options are theoretically and empirically superior. We often assume that life courses matter, but we usually are less certain about which specific proxies matter most. For example, does timing matter? If it matters, how to operationalise it, and which age periods are particularly relevant? As the results show, the most important predictors are not always the most straightforward ones, and they differ across inequality dimensions (overall pensions, gender inequality in pensions, and gender-specific inequalities).

Third, one of the main contributions to the literature on GPGs is the finding that the largest parts of the GPG can be attributed to the lack of male counterparts for unpaid care work for which the largest pension penalties exist. Future research will benefit from the application of Life Course Feature Selection with the Ñopo decomposition to other group inequalities in other outcomes by quantifying the parts of the gaps that are driven by extremely unequal (i.e., not matchable) experiences over the life course. As such,

combining feature selection approaches with decomposition techniques might be a promising approach for other inequality dimensions in old age that are based on life courses, such as health or wealth inequalities.

Other extensions would be to include information on occupations and earnings over the life course to quantify the role of, for example, overall job and earnings security or the timing of specific job or earnings changes for pension income inequality and the GPG. This would be particularly important for private pensions. More generally, future research might want to focus solely on private pensions, for which we know less about the inequality-generating mechanisms compared to public pensions.

Finally, it is important to consider some limitations. While the Life Course Feature Selection approach contributes to the holistic approach by identifying which specific aspects of life courses are relevant, it does not capture typical life course patterns as entities. For example, full life courses with all their complexities and interrelations between life course domains (here: family and work) do not become visible even though we know that the strong linkage between work and family life for women only is an important predictor of the GPG (see Chapters 3 and 4). Researchers have to keep this in mind when choosing the life course feature selection approach and fragmenting life courses in its elements. It remains unclear whether some of the most important features might just be proxies for specific life courses that become not visible. For example, one could question whether it makes a big difference whether and when individuals have direct transitions from full-time employment to retirement or whether it is not more likely that this might be a good proxy for specific life courses (e.g., care work-related activities also at older ages, lack of continuous full-time employment, and, consequently, lack of identifying with the notion of retirement, see Chapter 2.2).

Furthermore, the feature selection approach applied does not assess which life course aspects are the most important for *gendered* pension income inequality. These, however, might differ from the predictors of overall pension income inequality. Robustness checks indicate that using feature selection for the overall inequality in pensions risks concealing life course variables that unfold their association with GPGs through gender differences in returns for the same characteristics. Additionally, the key predictors of overall pension inequality do not completely align with the ranking of the predictors' (bivariate) associations with the GPG. Still, since gender inequality is inherent in overall pension income, selected predictors for overall pensions prove well for reflecting important predictors of GPGs, as sensitivity analyses show.

Overall there is a need for better data beyond employment participation over the life course as highlighted previously (Bardasi & Jenkins, 2010; Rowold, 2023a), particularly for country comparisons. In both countries, most likely a part of (gendered) pension income inequality is related to earnings inequalities over the life course (see Chapter 4), and other labour market inequalities might be reproduced in pension inequalities as well, such as occupational gender segregation. In this analysis, such inequalities might be partly captured by the educational years.

Additionally, for the Ñopo decomposition, a larger sample size would be important to rule out that the large lack of common support (DA and DB components) is simply due to the low sample size (Hamjediers & Sprengholz, 2023). However, at least for the studied cohorts in West Germany and the Netherlands, it is highly unlikely that the share of men who spent many years doing unpaid care work would increase in larger samples and thus decrease DA largely.

Finally, the study can only draw conclusions for the highly gender-segregated contexts of West Germany and the Netherlands. From a social policy perspective, it would be useful

to apply the novel approach to contexts with more gender-equal life courses, such as East Germany and Sweden. This would provide evidence on whether less gendered life courses still lead to significant shares of GPG due to gender-exclusive life course experiences, or whether this finding is unique to highly gender-segregated societies. Furthermore, it would be an important addition to this study to assess how different levels of gender-segregated life courses interact with different pension designs. This study has revealed that previous literature has been biased due to the focus on the duration of full-time employment, and due to ignoring the lack of common support of life course experiences of men and women. Gender differences in the duration (and timing) of care work over the life courses are the most important reason for gendered pension income inequality. Highlighting the relevance of specific life course aspects for the GPG allows policymakers to employ more targeted policies to prevent the reproduction of gender inequalities over the life course in old age. In both pension systems, this would be primarily to equalize the bulk of unpaid care work between men and women over the life course, or to reward care work more equally in pension systems.

Acknowledgements

I am grateful to Matthias Studer for introducing me to the Life Course Feature Selection and the R code. I am appreciative of ongoing discussions with him and Danilo Bolano regarding the method. Further, I am grateful to Maik Hamjediers and Maximilian Sprengholz for a vivid and beneficial exchange about the KOB and the Ñopo decomposition. I also want to thank Christiaan Monden, Hanna van Solinge, Kène Henkens, and Edith Darin for helpful feedback on previous versions of this Chapter. Finally, I am grateful for helpful comments from the social demography reading group participants at the Department of Sociology and Nuffield College. All remaining errors are mine.

6 Discussion & Conclusion

6.1 Aims and Chapters overview

This thesis aimed to assess how work and family life course complexities are related to gender inequalities in pensions. To study this question and contribute to previous literature, I propose novel combinations of advanced methods. These innovative approaches allow, as I argue, a more gender- and life-course-sensitive analysis of gendered pension inequalities, and other inequalities more broadly. The application of these methods offers new substantive insights into how Gender Pension Gaps are linked to complex inequalities arising over the life course.

As a starting point, I map out the most relevant conceptual choices for studying Gender Pension Gaps as well as their consequences in Chapter 2 and give an overview of which of those the previous literature has implemented. In a descriptive investigation, I further assess the implications of two popular sample criteria for studies on the elderly. Both, the self-identified retirement state, and pension receipt result in partially dramatic gendered sample biases with levels varying across countries. The implications derived from the reflections and results in Chapter 2 are implemented in all three empirical studies (Chapters 3-5). In short, this thesis puts emphasis on the financial independence of women and defines the outcome, pension income, accordingly.

In Chapter 3, together with Emanuela Struffolino and Anette Fasang, I introduce the combination of the Kitagawa-Oaxaca-Blinder (KOB) decomposition with sequence analysis (SA-KOB) to realize more life-course-sensitive analyses of group inequalities in life-course-sensitive outcomes. Applying the SA-KOB to the Gender Pension Gaps in West Germany and Italy and replicating the standard KOB which uses the sum of years spent in

the different life course states highlights the methodological and substantive advantages of SA-KOB (see below).

Chapter 4 applies the SA-KOB to richer data for West Germany that includes earnings over the life course and differentiates between the gender gap in public and private pensions. Its results contribute to the discussion of the consequences of pension privatisation, the role of redistribution in public pensions, and the Gender Wage Gap for gendered pension inequality.

Chapter 5 combines the Life Course Feature Selection with the Ñopo decomposition and changes the perspective from the holistic approach of Chapters 3 and 4 to that of single aspects of whole life courses. I first disentangle which aspects of life courses (the duration, timing, or ordering of states, or the overall complexity of life courses) are the most important predictors of pension income. In the second step, I decompose the GPGs based on the most important life course pension predictors using matching.

6.2 Discussion of findings and contributions

6.2.1 Conceptual contribution

Beyond new substantive findings and methodological innovations, I identify components of the research design that are important for ensuring gender- and life-course-sensitive analyses of Gender Pension Gaps and other old-age outcomes. These parameters entail the conceptualisation of the analysis unit, the definition of pension income as well as the sample. Such decisions are often not discussed in detail in published studies, partly even fully untransparent. Systematically discussing their implication and summarising the status of the literature on gender inequalities (see Table 2.1) equips the field with an up-to-date overview of which conceptual decisions predominate and how comparable the studies, including the empirical results of this thesis, are.

Additionally, I reveal gender biases in the sample definition and data procession that apply to any studies on the elderly more broadly (Chapter 2.3.1). Examining and empirically exposing the gendered sample bias when using self-reported retirement and pension receipt in Chapter 2.2 contributes to the broader discourse on an “appropriate conceptualisation of retirement” (Fasang, 2009, S. 290). It shows descriptively and across country contexts how the two popular sample criteria are highly (gender) selective as has been argued previously (Allmendinger et al., 1993; Fasang, 2009). Primarily women with continuous or very long periods of care work were excluded when any of both sample criteria were applied. I argue that this is because women’s life courses are more heterogeneous and less often follow the male standard of continuous full-time employment that is connected to the overall understanding of retirement and institutionalised as a condition for pension receipt in many pension systems. Applying these measures, and thus omitting individuals without personal pension claims or self-identification with retirement, largely influences the gendered composition of the sample across most countries. Still, various studies on gender inequalities in pensions adopt one or both sample criteria (see Table 2.1) and thus introduce the “steady worker bias” (O’Rand, 1996, S. 235) through their conceptual choices. Studies with broader frames that apply self-assessed retirement or pension receipt as sample criteria (e.g., Belloni et al., 2020; Bridgen et al., 2022; Dingemans & Henkens, 2019; Madero-Cabib & Fasang, 2016; Palomäki, 2017; Ponomarenko, 2016; Tambellini, 2023) are likely to underestimate inequalities among the elderly likewise.

The analysis equips future research with an indication of whether the gendered sample selectivity applies to the studied context. It further calls for caution and transparency on the selectivity when applying the self-assessed retirement status or pension receipt as sample criteria or any measure in the analysis. Otherwise, research risks underestimating (gender) inequalities and skewing country comparisons. Consequently, the use of self-

assessed retirement status or pension receipt requires particular attention and discussions of the implications for the results, at the very least.

For the sake of comparability, future research on gender inequalities in pensions should be transparent about the main conceptual decisions that I have identified as the main parameters for the level of gendered pension inequalities (e.g., the conceptualisation of the sample definitions, gender differences in pension receipt, and pension income).

These conceptual considerations complete the methodological and substantive contributions of this thesis to provide more gender- and life-course-sensitive tools for research to study GPGs, and other life-course-sensitive outcomes in and beyond old age.

6.2.2 Methodological contribution

The main methodological contribution is to introduce and systematically apply methodological combinations for more life-course- and gender-sensitive analyses of Gender Pension Gaps. These methodological innovations mark a significant departure from traditional studies that often rely on summary measures, unveiling the complex intricate labor and family dynamics that contribute to gendered inequalities.

This contributes to the empirical literature on GPGs which has analysed the GPG based on one-dimensional summary measures (e.g., number of years worked full-time) in regression or decomposition approaches and focuses on the employment sphere (literature streams I and II, see Table 1.1) or could not differentiate between compositional and returns effects for shaping the Gender Pension Gap (literature stream III).

This dissertation suggests combining 1) sequence analysis with the Kitagawa-Oaxaca-Blinder decomposition (termed SA-KOB) as well as 2) Life Course Feature Selection with the Ñopo decomposition. While both combinations of methods are particularly well

equipped to analyse Gender Pension Gaps and other group-based inequalities in old age, they can be applied to any life-course-sensitive outcomes at any point in time. Applied to Gender Pension Gaps and work and family life courses, both combinations of methods implement the “need [for] fuller information on the work career — and on its discontinuities as well as its continuities” (O’Rand, 1996, S. 236) and methodologically help to prevent the “steady worker bias” (O’Rand, 1996, S. 235).

1) *SA-KOB decomposition*. The SA-KOB decomposition applied in Chapters 3 and 4 consists of two steps. As the first step, sequence analysis and cluster analysis categorise similar life course trajectories into the same clusters pooled over the groups of interest and visualize them effectively to analyse the group differences in such, including their complexities. As a second step, we use this outcome, i.e., the common life course patterns, as independent variables in a regression-based decomposition analysis. The SA-KOB decomposition enables the differentiation between shares of group-specific gaps in outcomes, such as the Gender Pension Gap, a) due to group-specific life courses (explained or compositional component) and parts b) due to differences in returns for similar life courses (returns component).

The *compositional component* contributes to previous methods by quantifying how much of a group gap in an outcome would change if the access to typical life course patterns were equalised across the groups of interest. It shows, which of the group-specific life course patterns contribute most to the group inequalities in the assessed outcome. Applied to pension income as the outcome, it reveals how much of the Gender Pension Gap is attributable to the gender-specific distribution on life course groups that are unequally rewarded by the pension system. It thus captures the interaction of work-family life courses

and the pension system for shaping GPGs, since gender-specific life course patterns would not play a role if the pension system would reward all life courses equally.

The *returns component* of the SA-KOB quantifies how much of the group-specific gap is attributable to better returns for the same life course pattern of one of the groups. It can capture the share of the gap that is mitigated by policies that aim to produce group differences in returns, for example, by resulting in better returns for the disadvantaged group. On the other hand, it can reveal how group-specific inequalities are accelerated through worse returns of the disadvantaged group for the same life course pattern. As such, for pension income, the returns component assesses whether redistributive elements in the pension system succeed in reducing inequalities by enforcing better returns for the disadvantaged group, here women.

By combining the advantages of both methods, sequence analysis and the KOB, the SA-KOB decomposition addresses various challenges of the previous literature (see Table 1.1 for an overview). The ability of SA to create life course profiles circumvents issues of standard KOB applications. This includes the risk of concealing life course complexities, the risk of multicollinearity, and therefore, overlooking important life course aspects that could not be included in models, as well as interdependences between several life course domains. The strength of KOB to estimate how inequalities in observed outcomes are characterised by differences in these life course profiles and to differentiate between compositional and returns effects contributes to the lack of counterfactual analysis in standard sequence analysis applications. By directly comparing the SA-KOB to the standard KOB (which uses the duration in the different life course states as covariates), we can systematically derive the methodological and substantive contributions of SA-KOB

and show how new substantive findings arise compared to previous literature (see Chapters 3.4.4 and 6.2.3).

Apart from the overall methodological contribution, we offer some technical contributions to the Kitagawa-Oaxaca-Blinder or sequence analysis more generally, extending the utility to separate applications of the methods.

First, we discuss the choice of the reference coefficient used in the KOB and suggest a new option: the category-specific reference coefficient based on the majority group of each category (i.e., life course clusters for SA-KOB). We argue that the category-specific coefficient works particularly well for the SA-KOB, and especially for highly group-specific life course clusters, e.g., in our case female majority clusters. However, the approach is beneficial in any KOB decomposition applications, where single categories of covariates are highly group-specific, and the dominant group varies across categories.

Second, we aim to assess the influence of gender-specific heterogeneity within clusters for the results – a main challenge of SA-KOB. However, we argue that standard covariates used in KOB might be subject to within-category heterogeneity as well albeit it might remain concealed. In contrast, SA brings such heterogeneity to the front through the visualisation of life course and provides tools to address within-cluster heterogeneity. We exclude incoherent trajectories based on within-cluster coherence measures (Silhouette Width) and suggest applying cluster-specific, relative thresholds (e.g., 5th or 10th percentile) of the cluster-specific coherence (average silhouette width, ASW) to account for differences in coherence between clusters.

Third, we discuss the choice of applying a single-channel (i.e., focussing on only one life course domain, e.g., work) or multichannel sequence analysis (MSA, i.e., considering

several domains, e.g., work and family). We argue that the MSA has one benefit which is of advantage for research questions that fit the SA-KOB. Group differences in outcomes may arise from group-specific variations in the connections between several life course domains. These interdependences between life course domains are most effectively categorised and visualised using MSA. Their association with the group-based inequality can then be quantified by the SA-KOB decomposition. For example, in our case, the important role of the gendered level of work and family life course interdependencies for the GPG only becomes apparent when applying MSA. Only the SA-KOB empirically captures the for the GPG pivotal work-family life course interaction (Allmendinger et al., 1992; Ginn et al., 2001).

2) *Life Course Feature Selection & Ñopo decomposition*. Chapter 5 applies the novel two-step combination of two innovative methods: the Life Course Feature Selection & Ñopo decomposition. The first step, Life Course Feature Selection (Bolano & Studer, 2020), improves the standard approach in the literature that relies mostly on the duration in specific states, by assessing which are the most important life course predictors for the outcome (feature selection) for a better-informed variable selection. It contributes to the previous literature by showing that across different pension systems life course aspects of timing, ordering and the overall complexity matter beyond the duration in states. Specifically, contrary to the predominant focus in previous literature on GPGs, the years spent in full-time employment are clearly not the most important life course predictor of pensions. Instead of relying on easily extractable and straightforward life course aspects such as the duration in different states, literature is well advised in considering other life course aspects, e.g., covering the timing and ordering of events. Furthermore, the approach contributes to the holistic perspective of Chapters 3 and 4, and sequence analysis approaches more broadly, by showing which of the life course aspects grouped into clusters

are most relevant. Life Course Feature Selection thus somewhat addresses challenges in the interpretation of clusters due to the within-cluster heterogeneity.

A better-informed variable selection is particularly helpful for matching approaches, where the number of variables to select is limited. Applying a Ñopo decomposition as the second step (Ñopo, 2008), Chapter 5 introduces matching to the GPG literature. It contributes to the KOB by considering only comparable men and women for estimating the components of the KOB. Calculating the compositional and returns components only based on the matched sample of men and women, the Ñopo decomposition presents estimates that are not biased by comparing incomparable men and women based on the main life course predictors of pensions. Moreover, it adds to the SA-KOB as well as standard KOB by quantifying the part of the gap attributable to group-exclusive combinations of life course experiences, that is combinations of life course experiences that are not matchable across groups. I argue and empirically show that the latter is particularly important for Gender Pension Gaps in contexts with highly gender-segregated life courses, such as West Germany and the Netherlands.

Jointly, both novel combinations of methods contribute to the research on Gender Pension Gaps by shedding light on how life courses or their facets are gendered and are differently rewarded in pension systems, consequently resulting in gender disparities in pension income. They address some of the major challenges of standard methods in the field of gendered pension inequalities. Most importantly, they fill a substantive gap by offering more life-course- and gender-sensitive analysis of group inequalities and promise to reveal new findings in other fields. Such life-course-sensitive methods will likely become more relevant for future work since life-course complexities tend to increase over cohorts (e.g., Van Winkle, 2018; Van Winkle & Fasang, 2017, 2021) and thus make it increasingly

relevant to consider them in empirical analyses. Otherwise, important but complex mechanisms in play for shaping inequalities over the life course will remain concealed.

6.2.3 Substantive contribution and discussion

The substantive contributions to the literature on Gender Pension Gaps are manifold and facilitated through the application of novel combinations of methods and research designs. Thus, often the methodological and substantive contributions are closely linked. Below, I elaborate on the most crucial findings, organized based on their substantive contributions. Table 6.1 gives an overview of the substantive findings as well as the main implications derived for policy and future research.

Exploring the foundations: Gendered work-family life courses as the starting point

Highly gendered work-family life courses from age 18-65 following the traditional gendered division of labour emerge as the overarching pattern for the studied cohorts. The sequence analysis also reveals very similarly gendered work-family life course patterns for Italy and West Germany (Chapter 3). Married fathers can reconcile their parenthood with continuous (full-time) employment (termed ‘male-majority’ life course clusters), while mothers either engage fully in consistent unpaid care work upon childbirth or reconcile their family care needs with part-time employment after a care-related employment interruption (termed ‘female-majority’ life course clusters). Consequently, considering work and family life course patterns jointly, also reveals that fathers’ work life courses are independent of their family lives, while mothers’ family lives are strongly intertwined with their family lives, thus much more ‘*linked*’ with others. This is largely in line with other studies that found extremely gendered life courses for conservative, liberal, and Southern European welfare states (Komp-Leukkunen, 2019; Uccheddu et al., 2022).

Chapter 5 adds to the holistic exploration of typical life course clusters, by assessing which specific aspects of life courses do actually matter for pensions. Contrasting the conceptual choice of previous literature, the duration of full-time employment is by far not the most relevant pension predictor in the Netherlands and West Germany. Instead, life course aspects related to care work stand out as the most important predictors of overall pension inequality, and the timing and ordering of transitions over the life course matter as well.

These thick-descriptive results of the first steps of the applied methodological combinations in Chapters 3-5 already suggest that previous research has focussed on simplifying proxies for these more complex realities individuals experience over their life courses. In combination with the decomposition techniques, the exploration and operationalising of such life course complexities have unveiled important substantive implications for the gendered pension income inequality.

Unmasking the relevance of uncompensated care work.

Relying on life course summary measures and mostly the duration in single life course states previous literature has suggested that women's lack of (full-time) employment, and, to a lower extent, wage differences, were the main reasons for the Gender Pension Gap (e.g., Cordova et al., 2022; Frommert & Strauß, 2013; Hänisch & Klos, 2014; Johnson et al., 1999; Nolan et al., 2019; Veremchuk, 2020; Zhao & Zhao, 2018, see Table 3.1). In line with this evidence, in Italy, a substantial part of the GPG is attributable to the lack of mothers in life courses characterised by the well-rewarded continuous full-time employment (9%). On the other hand, we do not find this association with the GPG in West Germany (Chapter 3).

However, all empirical Chapters of this thesis strongly highlight the important role of unpaid care work for gendered inequalities in pensions. Continuous unpaid care work over

the life course is almost exclusively performed by mothers in Italy and West Germany and roughly one-third of women in the studied cohorts have such life courses (Chapter 3). In both countries, the largest compositional component of the GPG stems from the large gender difference in having such life courses and the very poor pension returns for them: 24% of the GPG in Italy and 13% in West Germany.⁶⁴ This finding stands independent of the different pension pillars: the compositional share due to continuous care work life courses is almost identical for public compared to private pensions in West Germany (Chapter 4). Lastly, extending the decomposition over the distribution for public pensions (in West Germany only), reveals that the dominance of women with care-dominated life courses is particularly important for the Gender Pension Gap at the bottom of the distribution, thus particularly for women facing the highest poverty risk.

The Life Course Feature Selection confirms the relevance of unpaid care work, even if it is operationalised as life course summary measures and compared to full-time employment aspects as they were applied in previous research (i.e., diverging from the holistic perspective in Chapters 3 and 4). Chapter 5 reveals that the duration of care work for West Germany and the Netherlands, as well as its timing (for West Germany only), is even one of the main predictors of *overall* pension income inequality – and far more important compared to the years spent in full-time employment. The Ñopo decomposition then strengthens the results of SA-KOB: matching based on the most important life course predictors of pensions leads to matched and unmatched samples of women with similar characteristics compared to the life course clusters. Bringing in an additional inequality dimension, it suggests that large shares of the GPGs in both countries are due to gender-exclusive combinations of life course experiences, which evolve primarily due to the lack

⁶⁴ These numbers only refer to mothers with 2 or more children. When not considering differences in family life courses, almost 50% of women have these continuous care work life courses, and the compositional component due to the gender differences in these life courses rises by another 10%.

of comparable care work engagement of men. Since these gender-exclusive life course experiences are highly unequal rewarded in the pension systems, they are attributable to larger shares of the GPG compared to the compositional differences among the matched, i.e., comparable men and women. The pension penalties are larger for unmatchable women the longer and the earlier their care work periods are.

The results of this thesis thus show how multicollinearity or the ‘steady worker bias’ has prevented previous KOB studies from revealing the relevance of care work: only Frommert & Strauß (2013) have included care work as a covariate previously (see Table 3.1). Still, some might argue that unpaid care work and (full-time) employment are different sides of the same coin, i.e., individuals do either one or the other. However, drawing attention to the relevance of care work is crucial from a social policy perspective, at the very least. While one pathway to higher pension income among women, especially mothers, is taking up full-time employment, policymakers must recognize that increasing the female labour force must be accompanied by substituting the caregiving responsibilities that these women would otherwise have fulfilled. The other pathway to achieving equal pension income for women, in fact, the only solution for women who have reached old age already, is to acknowledge and reward their care work more substantially, which underscores the significance of the reward system embedded in pension designs. Thus, emphasizing the role of care work for Gender Pension Gaps, highlights that the substantial gender disparities in pension income also evolve around the question of which types of life course patterns – or which side of the coin – are rewarded by the pension system – and which are not.

Consequently, Chapters 3-5 reveal, that other than what previous literature has suggested, the endurance of the GPG is not only a story of who is doing ‘the work’, but also, which work “is valued, protected, or rewarded over the life course” (O’Rand 1996: 230), and

subsequently, in the pension system (see for a discussion of social policy implication, Chapter 3.3). This finding is not sensitive to the analytical approaches applied in this thesis: it remains consistent when relying on holistic life course patterns or assessing the most important pension predictors among a multitude of life course summary measures.

Furthermore, Chapter 2.2 suggests that continuous care work also plays a major role in the selection into self-identification with retirement as well as for access to any pension claims. Chapter 4 confirms this link for private pension receipt applying an SA-KOB decomposition: more than half of the gender gap in private pension receipt is attributable to the dominance of mothers with continuous care work compared to fathers. This is not trivial given that studies focussing on gender inequalities in pensions tend to exclude individuals without pension receipt or retirement identification from their studies (e.g., Bonnet et al., 2020; Evandrou et al., 2009; Frommert & Strauß, 2013; König et al., 2019; Madero-Cabib & Fasang, 2016; Möhring & Weiland, 2022). Removing primarily women with continuous care work from the analysis this way (Chapter 2.2) might lead to an underestimation of the role that unpaid care work plays in forming Gender Pension Gaps.

Part-time employment: A catalyst or barrier for gender equality?

A related contribution of this thesis is to disentangle the role part-time employment plays in creating gender inequalities in pension income. This is important since parallel to women's overall labour market participation, female part-time employment has increased to high levels in many countries (e.g., Barbieri et al., 2019; Kreyenfeld, 2015; Schmitt & Auspurg, 2022), and has been considered as a 'bridge' for mothers returning to full-time employment after employment interruptions (e.g., Kelle et al., 2017; Stegmann, 2007). The literature on gendered pension disparities so far has mostly operationalised it as the duration in part-time employment if it was considered at all (see for example Table 3.1 for KOB

studies) and found mixed results dependent on the country context. While in the UK a part of the GPG is indeed due to women's more years in part-time employment, it mitigates the GPG in Germany (Cordova et al., 2022; Ezeyi & Vujic, 2017; Frommert & Strauß, 2013).

However, the sequence analysis results reveal that part-time is largely realised in synergy with care-work-related employment interruptions of mothers (Chapters 3 and 4). This empirically confirms the reputation of part-time employment to enable mothers access to the labour force after care-work-related employment interruptions. However, considering part-time work together with care work periods of several years shows that it rather drives than mitigates the gender gap in total pensions in Italy and West Germany (Chapter 3). Again, this is the case for both, the gender gap in public and private pensions (for West Germany, Chapter 4). However, the part-time life course cluster drives the GPG to a lower extent compared to the other women-majority cluster, characterised by continuous care-work, in both, Italy and West Germany, as well as across public and private pensions. Additionally, in contrast to the care-work cluster, the reconciliation of part-time employment with care duties seems to be not detrimental to women's access to private pensions (Chapter 4).

Chapter 5 further reveals that the duration of part-time employment is not among the most important pension predictors but rather relevant for pension inequalities among women only in West Germany and the Netherlands. In line with the findings of Chapters 3 and 4, for overall pensions, including gendered inequalities, part-time employment is most important in combination with the timing of the transition from care work periods in West Germany (and not at all in the Netherlands). This emphasizes that the accumulation of gender inequalities with regard to part-time work unfolds in more complex synergies that are only captured by sequence analysis or by other life course summary measures beyond

the duration. Future research might thus want to implement such alternative operationalization of part-time employment, and also to incorporate the recently uncovered wage dynamics associated with part-time employment stalling the Gender Wage Gap in West Germany (Schmitt & Auspurg, 2022) which was not feasible in this thesis.

In conclusion, the results of this thesis suggest that women's more prevalent part-time employment compared to men rather drives the GPG – at least if it is realised in combination with the substantial care-work related employment interruption as is the case for the majority of elderly women with part-time employment experiences. This is confirming previous findings in the UK (Ezeyi & Vujic, 2017). While my findings are not in line with evidence on the GPG in Germany (Cordova et al., 2022; Frommert & Strauß, 2013), they confirm the pension risk evolving from part-time employment highlighted previously (Brussig et al., 2017). Whether this will still be the case for younger cohorts who might reduce the duration of the care-work interruptions will have to be assessed in future studies. However, whether part-time helps mitigate the accumulation of gender inequalities over the life course depends on the reference. Compared to continuous care work, which is worst rewarded in the pension system compared to the typical part-time employment life courses, it does. However, part-time employment following care-work-related employment interruptions of mothers still yields much lower pensions compared to continuous (high- and low-income) full-time employment in the private sector as well as civil service.

Gender Pension Disparity: Simply a reflection of the Gender Wage Gap?

By including the relative earnings over the life course, Chapter 4 adds to the question of to what extent the popular Gender Wage Gap (e.g., Kunze, 2018) is reproduced in pension income. Previous literature has included cross-sectional measures of wages at certain points

in the life course to account for wage differences and mostly found that the lower average wage of women contributes to the Gender Pension Gaps.

The explorative sequence analysis in Chapter 4 differentiates between two common life course patterns related to earnings: parents with i) stable and high-income full-time employment and ii) more mobile career paths changing between different, but lower income levels of full-time employment. Using these clusters as covariates, the compositional part of the SA-KOB decomposition reveals a more nuanced relationship of earnings with the GPG compared to previous research: while women would benefit from stable high-income careers, the unstable, low-income careers would be even counterproductive for their pension income compared to that of men. The pattern is the same for public and private pensions, but the role of compositional gender differences is much larger for private pensions. Differentiating full-time employment clusters of parents in less and more prestigious career patterns based on earnings also contributes to Chapter 3, where we could not find a link between the lack of mothers with full-time employment careers independent of earnings levels and the gender gap in total pension income in West Germany. Not distinguishing low- and high-income full-time employment conceals the mechanism due to the contrary associations of underlying career heterogeneities (a positive link for high-income and a negative one for low-income careers of parents, demonstrated in Chapter 4).

The differentiation between earning profiles proves to be particularly important for understanding the mechanisms at different points of the distribution as well as the gender differences in pension receipt. An increase of mothers with low-income careers would indeed mitigate the gender gap in public pensions at the bottom of the distribution but increase the GPG at the centre and top of the distribution. Thus, women at the bottom of

the distribution would actually benefit from any kind of full-time employment. However, to reduce the gender gap in public pensions beyond the bottom of the distribution, women need access to more privileged careers like high-income full-time employment or civil service.

Similarly, for private pensions, the lack of mothers in both, low- and high-income full-time careers explains the lower pension receipt of women compared to men (the Gender Coverage Gap, Chapter 4). However, for the gender gap in private pensions among private pension receivers, a higher share of mothers with low-income careers would be detrimental again. Thus, low-income careers would be helpful for mothers to gain access to private pensions, but not to equalise private pension entitlements between genders.

Differentiating between the different earning profiles over the life course thus contributes to our understanding of the Gender Pension Gaps with nuanced links of different components: different pension pillars, parts of the distribution and pension receipt. The Gender Pension Gap is not merely a reproduction of the Gender Wage Gap, because other characteristics such as working hours and care work play a role as well and mechanisms are more nuanced when it comes to differences along the distribution as well as pension receipt. This is most likely why simple country-level correlations do not suggest a link between the Gender Pension with the Gender Wage Gap (Hammerschmid & Rowold, 2019a), but rather with the Gender Earnings Gap (Lis & Bonthuis, 2019) which also considers differences in working hours and labour market participation. Still, gender differences in earnings do play a role beyond full-time employment, and they are the main reason for the gender gap in private pensions. Consequently, debates on gender differences in wages and earnings should also consider its likely increasing importance for gender inequality in old age due to the privatisation of pensions.

Pension privatisation and its implications for the Gender Pension Gap.

The pension privatisation particularly unfolds its implications for the GPG through the much stronger reproduction of earning differences between men and women (Chapter 4, considering access to and entitlement with private pension jointly). For example, 22% of the gender gap in private pensions is due to the lack of mothers with high-income careers compared to only 5% of the gender gap in public pensions. At the same time, the links between the gender gap in private pensions with the women-majority life course patterns, dominated by unpaid care work and part-time employment parallel to motherhood, are similar to the gender disparities in public pensions.

Moreover, 55% of the gender difference in *access* to private pensions is attributable to the compositional gender difference of parents with continuous care work, as well as to the lack of mothers in high-income (38%) and low-income (15%) careers. However, while women would benefit from any full-time employment careers in the private sector when it comes to private pension *access*, low-income careers would be detrimental to the private pension *income* among private pension recipients. Thus, increasing the share of women with any type of full-time employment in the private sector would help mitigate the gender differences in private pension receipt. However, only equalising the share of women with high-income careers would significantly decrease the gender gap in private pensions, once access to them is reached.

Consequently, policymakers will have to either ensure women's access to generously rewarded careers, such as stable high-income careers in the private sector, or civil service, or implement more redistributive elements in private or public pensions (see 6.3), to reduce the gender gap in in private or in total pensions, respectively. To what extent the links

between gendered work-family life courses and private pension change upon the increasing coverage of private pensions will have to be subject to future research.

Redistribution in pension systems: Does it make a difference for the Gender Pension Gap?

The returns component of the SA-KOB provides a tool to quantify how much of the Gender Pension Gap is reduced due to better returns for women, e.g., due to redistributive elements that pension systems have implemented. The expectation was that women, especially those with life course characteristics that are rewarded in the pension system beyond employment (such as care work), benefitted from better returns compared to men and that these better returns would alleviate the GPG. However, the results largely do not support a significant reduction of the GPG due to better pension returns of women for total pensions in West Germany and Italy (Chapter 3). This finding remains consistent when focussing on public pensions only, where redistribution is more pronounced compared to private pensions (Chapter 4, for West Germany only). Consequently, the results strongly suggest that the redistributive elements in the pension systems are overall too insignificant to effectively reduce the GPG. However, two important exceptions exist.

First, I do find that married women benefit from better public pension returns upon divorce compared to men which mitigates the GPGs slightly in West Germany (Chapter 4). This is in line with recent evidence that the couple-level splitting of pension claims upon divorce is a gender equalizer when it comes to pensions (Kreyenfeld et al., 2023; Schmauk, 2023). My results indicate that pension splitting is indeed an effective tool for reducing the GPG since it results in higher pension income for divorced women compared to divorced men.

Second, exploring the link between work-family and the GPG along the public pension distribution discloses that women with the two women-majority life course patterns characterised by unpaid care work and part-time employment do have better returns

compared to men, but only at the bottom of the distribution (Chapter 4). These differences in returns in favour of women mitigate the gender gap in public pensions at the first pension quartile for both, the continuous care work as well as the part-time work cluster. While the favourable pension returns of women compensate for having the overall poorly rewarded part-time life course pattern (compositional component of SA-KOB), they cannot make up for the pension loss of women compared to men for having engaged in full-time care work continuously.

The methodological approach cannot subsequently conclude or quantify the role of single redistributive elements. However, my results suggest that due to the low levels of redistributive elements in the pension system, they only visibly affect the GPG at the bottom of the pension distribution – where redistribution is particularly important to prevent poverty but also the overall pension level is low. Still, in line with previous research (e.g., Frommert & Strauß, 2013; Grabka et al., 2017; Horstmann et al., 2009; Lis & Bonthuis, 2019; Stegmann, 2007), the results indicate that targeted redistributive elements such as child care benefits insufficiently mitigate the accumulation of gender inequalities over the life course in general, and consequently, the Gender Pension Gap.

Do we need to consider family lives?

My results suggest that family life courses unfold their relationship with the Gender Pension Gap directly exclusively through gender differences in pension returns. For West Germany, the results confirm that the couple-level pension claim splitting upon divorce in the German public pension systems serves as “gender equaliser” (Kreyenfeld et al., 2023, S. 15; see also Schmauk, 2023) in West Germany (Chapter 4). Better public pension returns for divorced women compared to men help reduce the GPG at the mean, but only marginally given the low share of individuals experiencing divorce for the studied cohort.

Focussing on family life course clusters only (derived from single-channel sequence analysis) confirms this association for *total* pension income (see 3.4.5). Furthermore, large parts of the GPG are because married mothers receive much lower pensions compared to married fathers in Italy and West Germany. This link holds in West Germany, when controlling for work life course patterns, highlighting the relevance of more profound gender inequalities among married parents that are not captured by the work characteristics considered in the model.

Still, the direct links between family life courses, or aspects thereof, with GPGs are clearly less pronounced compared to work life courses. Family life course aspects are also barely among the relevant predictors of overall pension inequality in West Germany and the Netherlands (Chapter 5). However, if life course aspects unfold their relationship with the GPG through opposing, gender-specific returns such as for divorce, pooling the analysis over gender risks concealing their relevance since the gender-specific returns might offset each other. Indeed, for gender-specific pension inequality, e.g., pension inequality among women only, some life course aspects related to family lives are among the top-10 pension predictors.

Additionally, life course aspects related to unpaid care work are the most important pension predictors (Chapter 5), and even though care work is subsumed as an element of work life courses, it is strongly linked to family events, such as marriage and childbirth, as becomes visible in the two most common life course for women (Chapters 3 and 4). Additionally, all work-family life courses that are strongest related to GPGs through compositional effects are those of parents (Chapters 3 and 4). The results thus suggest that family lives unfold their relevance for gendered pension income inequality, particularly through their interdependence with work lives. This indicates highly gender-specific work life course

opportunities for men and women with similar family life courses which have consequences for gendered pension income.

Consequently, understanding the complex interconnections between family life courses and gendered inequalities across the lifespan is crucial. First, considering family lives allows us to grasp the interrelations of family life courses with overall gendered inequalities over the life course, and, sheds light on how these are linked to GPGs. Second, it allows us to explore to what extent integrated measures that pension policies have implemented to account for these gendered work-family interactions succeed in equalising pension income across gendered opportunities following the same family lives.

The implications of the gendered division of labour: Gendered interdependences of work and family lives

By considering the family life course dimension, this thesis thus stresses the relevance of the gendered level of interrelation of work and family life courses. Such interdependencies only occur for women-majority life course patterns, in which care-work periods occur parallel to family formation. At the same time, men-majority life courses, characterized by continuous forms of employment, are largely unrelated to family events. The SA-KOB decomposition captures and quantifies how these interdependencies, embedded in the overall life course clusters, are linked to GPGs (Chapters 3 and 4). The decompositions then reveal that the gendered life courses characterized by differences in the level of the intersections between work and family domains are the main drivers of the GPG.

Capturing and modelling such interdependencies between domains is hard to achieve if not taking the holistic approach but focussing on isolated life course summary measures as done in previous literature and Chapter 5. Highlighting the relevance of the gendered work-family interdependences for the Gender Pension Gap is not only an important contribution

to the literature in the field but also points to important interactions between social policies of different domains.

The relevance of intertemporal interdependences of welfare state and pension policies for Gender Pension Gaps: A mismatch for women.

All substantive contributions have one underlying mechanism in common: the interaction between gendered life courses and pension systems. The gendered life courses unfold within opportunities and constraints largely set by welfare state policies, such as work-family policies (e.g., Krüger & Levy, 2001; Pfau-Effinger, 1998; Trappe et al., 2015). All three countries studied have socially and politically facilitated a traditional gender division of labour and the male-breadwinner model, incentivising the highly gendered work-family life courses of the cohorts in old age nowadays.

My results highlight how the highly gendered outcome of such policies is rewarded by pension systems, which independent of their overall system type, favour continuous employment in the private sector or civil service, and thus produce large gender inequalities in pensions. This follows the notion of “the interplay between institutional arrangements and individual life trajectories” (O’Rand, 1996, S. 230) highlighting the relevance of intertemporal interaction of different institutions over a period of more than 50 years for the emergence of GPGs. Men who complied with the incentivised role of the male breadwinner benefit due to the high pension rewards associated with continuous full-time employment. In contrast, women with life courses in line with the role that the welfare state assigned to them (female caretaker or secondary (part-time) earner) are largely penalized in old age. This exposes a systematic, gendered misalignment between welfare state policies from decades ago and the current pension systems – to the detriment of women. Consequently, the unbalanced interaction between welfare state policies implemented

during the second half of the 20th century and the rewarding structure of pension systems nowadays produces substantial Gender Pension Gaps in many countries.

In conclusion, the main takeaway of this thesis is that the gendered division of labour has long-lasting impacts on gender inequalities in old age. The developed combinations of sequence analysis tools and decomposition techniques enable capturing and analysing more in detail the interaction between gendered life course inequalities shaped by welfare institutions and pension institutions later in life. This detects that gender inequalities in the interdependences of work and family lives and performed unpaid care work, especially among parents and related to childbirth (Aisenbrey et al., 2009; Boeckmann et al., 2015; Budig et al., 2016; Gangl & Ziefle, 2009), are strongly reproduced by the pension systems in old age. In order to decrease the high Gender Pension Gaps, pension systems must increase their awareness of these intertemporal interactions with previous welfare state policies and adjust the policies accordingly.

Table 6.1 Overview of substantive findings and derived policy implications

Substantive contribution	Main finding	Policy recommendation	Future research
Work-family life courses	1. Highly gendered work-family life courses following the traditional gendered division of labour -> Men-majority life courses: continuous employment; independent of family formation -> Women-majority life courses: continuous unpaid care work or unstable, interrupted employment incl. part-time employment; highly dependent on family formation	1. Equalising work-family life course opportunities & interdependences across gender	- Assess for different normative & welfare state contexts & cohorts
Care work	1. Typical life courses of women characterised by continuous care work of married mothers' main reason for GPG due to low pension rewards -> In Italy (IT) & West Germany (WG) -> In public & private pensions (WG) -> Particularly important at bottom of public pension distribution (WG) -> Main driver of gender difference in private pension coverage (WG) 2. Life course aspects related to care work most important predictors of pension income in WG & the Netherlands (NL) -> More relevant than duration in full-time employment 3. Gender-exclusive and unequal rewarded combinations of life course experiences, primarily lack of comparable care work engagement of men, as main contributor to GPG	1. Higher rewards for unpaid care work -> Particularly for cohorts of mothers with highly constraint labour market access -> Life-course-sensitive pension policy 2. Equalise gendered division of labour	- Consider care-work proxies in future analysis - Prevent steady worker bias (i.e., primarily relying on the duration of employment)
Part-time employment	1. Typical life courses of women characterised by part-time employment following care-work-related employment interruptions of mothers contribute substantially to GPG due to low pension rewards -> In Italy (IT) & West Germany (WG) -> In public & private pensions (WG) 2. Life course aspects related to part-time employment no relevant pension predictors in WG & NL -> More relevant for pension inequalities among women	1. Higher rewards for part-time employment -> Particularly for cohorts of mothers with high constraints to access to full-time jobs (i.e., limited childcare opening hours, couple-level joint taxation) -> Life-course-sensitive pension policy 2. Equalise gendered division of labour	- Consider part-time employment in synergy with care-work interruptions - Further disentangle the interplay of part-time employment with other gender inequalities, e.g., earnings & occupational segregation

Gender earning disparities	1. High-income careers of parents: equalising access for women would secure higher pensions for women & decrease GPGs 2. Low-income careers of parents: equalising access for women would be counterproductive for women's pensions & increase GPGs -> But would mitigate GPG in public pensions at bottom of distribution -> Would increase women's access to private pensions but decrease their endowment with private pensions once they have access	1. Addressing & preventing gender inequalities in pay & equalising access of mothers to high-income careers 2. Strengthening redistributive elements concerning earnings in pension system -> For cohorts that are affected by large gender inequalities in pay (i.e., including cohorts currently in labour force)	- Consider earnings & differentiate between (more fine-grained) earning profiles over life course - Disentangle role of earnings versus employment - Assess role of earnings in different pension systems - Assess role of other gendered labour market inequalities
Pension privatisation	1. Gender earnings differences main reason for GPG in private pensions -> Much stronger reproduction of earnings differences in private compared to public pensions 2. Gender differences in private pension coverage: due to mothers' continuous care work, & lack of high- & low-income careers in private employment	1. Counter stronger reproduction of inequalities in private stronger through redistribution in public pensions 2. Include redistributive elements in private pensions 3. Equalise access to private pensions	- Changes for future cohorts? -> Monitor how reproduction of gender inequalities in private pensions changes with increasing coverage of private pensions
Redistribution in pension systems	1. Overall too insignificant to buffer the reproduction of inequalities in pensions & prevent high GPGs -> But: help mitigating the GPG among individuals with low public pensions (better pension returns of mothers compared to fathers with women-majority life course patterns at bottom of distribution) 2. Pension claim splitting upon divorce helps diminishing the GPG through higher pensions for women	1. Increase redistributive elements, particularly with regard to care work & part-time employment (i.e., most often part-time care) 2. Application and extension of couple-level pension claim splitting on more couples	- Focus on impact of single policy reforms - Assessing heterogeneities along the pension distribution
Family lives	1. Only few single family life course aspects directly related with GPG -> Through difference in pension returns for the same life course (most pronounced for divorce) 2. Family life courses mostly unfold impact on GPG in interaction with work life courses -> Family transitions such as marriage & childbirth linked to unpaid care work, the most important pension predictor	1. Equalise gendered life course opportunities with regard to family lives -> Promote equal life chances of men and women, irrespective of their family lives	- Include family life course proxies - Systematically disentangle & quantify interactions of family life course with work life course aspects

Gendered interdependences of work & family lives	1. Highly gender-specific interdependences of work & family lives -> Only for women-majority life course patterns: pivotal care-related employment interruptions parallel to family formation -> Men's careers unrelated to family lives -> Gender differences in work-family interactions are key features of gendered life courses strongly linked to GPG	1. Equalise work-family life course opportunities & interdependences	- Consider interactions between work & family life course proxies - Couple-level analyses or focus on parents for further disentangling mechanisms concerning gendered work-family interaction
Intertemporal interactions of welfare state & pension policies	1. GPG result of interaction between i) welfare state policies shaping life course opportunities decades ago & ii) pension policies nowadays reward the resulting gendered life courses differently 2. Misalignment for women -> Low pensions for women-majority life courses fulfilling traditional gender norms incentivised by policies -> High pensions for men-majority life courses fulfilling traditional gender norms incentivised by policies	1. Awareness about intertemporal policy interactions over decades 2. Sensitivity towards these interactions & contexts in which life courses were shaped	
Overall conclusion	1. Gender inequalities over the life course strongly reproduced by pension system -> Gender-typical life courses in line with welfare state context & norms unequally rewarded by pension system	Harmonising pension policies with work-family policies prior retirement age 1. Retirement policies engaging with work-family policies targeted before retirement age -> Equalise work-family life course opportunities & interdependences and division of labour across gender 2. Life-course-sensitive pension policy, in accordance with welfare state policies that shaped life courses -> E.g., cohort-specific redistributive elements, e.g., birth-cohort-specific care work benefits	- More life-course- & gender-sensitive analyses of GPGs - Replication of methods for younger cohorts & different country contexts

6.3 Social policy implications

The findings point to two pathways for reducing the Gender Pension Gaps. First, to ensure mothers' access to stable high-income, or civil service careers that yield generous pensions, by decreasing the gender segregation on different work-family life courses. Second, by including much more efficient redistributive elements, especially regarding performed unpaid care work, in the pension systems. I conclude that pension policies must be more life-course-sensitive and be aware of their strong interactions with welfare state policies from decades ago since they reward highly gendered life courses unequally. I argue in favour of a better-harmonised match between welfare state and pension policies, consequently, combining the two pathways to more gender-equal pension income in a cohort-sensitive fashion.

1. Pathway to more gender-equal pensions: Equalizing gendered life course opportunities.

Given that this study highlights the strong association between gendered work-family life courses emerging from the gendered division of labour, and particularly women's sole role as caretakers, with the Gender Pension Gap, equalising the division of care among genders will have a positive effect on women's financial independence in old age. Equalizing women's, and particularly mothers', access to well-rewarded life courses, would be the most sustainable pathway to long-lasting, equal pensions for women compared to men since it would ensure women's own, independent pension income. Consequently, if policymakers want to decrease the gender gap in pension without significantly shifting the pension policies, they must aim at reducing (gender) inequalities prior to pension receipt. This would include providing a more gender-equal reconciliation of unpaid and paid work, enabling equal access to employment. Increasing women's, and especially mothers', access to the labour force would not only help decrease gender inequalities over the life course

and in pensions but also provide the pension system with additional economic contributions.

To reach this aim, pension policymakers should engage more with other policies and promote policies that incentivise a more equal division of labour across genders. Several policy measures can be derived from previous research that highlighted the crucial impact of policies on the gendered division of labour, and women's labour force. Increasing the public childcare infrastructure, as well as implementing equality-oriented parental leave policies, which also targets fathers with leave periods reserved for fathers only (e.g., Almqvist & Duvander, 2014; Arntz et al., 2017; Bauernschuster & Schlotter, 2015; Boll & Lagemann, 2019; Geyer et al., 2015; Haan & Wrohlich, 2011; Huebener et al., 2016; K.-U. Müller & Wrohlich, 2020; Ziefle & Gangl, 2014) have shown to impact mother's labour force positively. For Germany specifically, abolishing the joint taxation of married couples which economically highly disincentivises dual full-time earner couples would likely increase the labour force participation of women, and would provide more inflows for the welfare state (Bach et al., 2011, 2020).

Additionally, pension policymakers must consider the increasing risk of reproducing the Gender Wage Gap in pensions, particularly through pension privatisation. Given that the level of the Gender Wage Gap remained fairly stable since the early 2000s in many countries (e.g., England et al., 2020; Kunze, 2018; Schmieder & Wrohlich, 2021; Schmitt & Auspurg, 2022), they cannot hope that the problem of gender differences in access and investments in private pensions will dissolve soon. Thus, to decrease the Gender Pension Gap it is pertinent to prevent gender inequalities in the labour market beyond access to employment, such as the Gender Wage Gap.

Consequently, any life-course-sensitive pension policy must aim at reducing (gender) inequalities prior to pension receipt. The most sustainable approach to prevent high GPGs is to reform the welfare policies that shape the gendered division of labour to prevent gender inequalities in the labour market. The increasing contributions to the pension system resulting from enabling more women access to the labour force could be used for financing more generous childcare benefits for women whose labour force was prevented by insufficient family policies in earlier years.

2. Pathway to more gender-equal pensions: More redistribution and rewards for unpaid care work.

However, women who are close to or already in old age can no longer alter their life courses. If they lack a sufficient own pension income, they must rely on redistributive elements in the pension system or, frankly speaking, on their partner. Given that this dissertation emphasizes gender inequalities in unpaid care work over the life course as the main driver of GPG, it highlights the relevance of care-work-related benefits. While current benefits are way too small to decrease the GPG on the mean (e.g., in Italy and Germany, see Chapter 3), they tend to buffer the GPG through better returns for women at the bottom of the distribution – but without fully equalising the gender pension differences (Chapter 4).

Redistributive elements in the pension system are important tools to prevent female old-age poverty. When creating these, policymakers should factor in that the continuous care work performed by women in contexts without public childcare infrastructures most often not only enabled continuous pension contributions of husbands but was also primarily linked to childbirth, thus the upbringing of new cohorts of pension contributors (Chapters 3 and 4). In times of population ageing and discussions on how to finance the pension system in the future, policymakers might want to reconsider the value of this unpaid care

work provided. Thus, the question becomes pressing whether the current childcare benefits implemented in pension systems are sufficient. Unquestionable is, that for current cohorts in old age, the childcare benefits would have to be much more generous to efficiently prevent high GPGs through equal pension returns for female-typical compared to male-typical life courses. Receiving one to two years of pension entitlements equivalent to an average earning employment contribution for childcare per child (in the case of Germany) cannot make up for the pension disadvantage of the majority of women with either continuous unpaid care work or care-related employment interruptions upon childbirth that naturally exceed two years in welfare contexts without a sufficient public childcare infrastructure.

Another important improvement would be to upgrade the pension rewards for part-time employment in the pension system. This is important for two reasons. First, my results show that part-time employment is often realised in combination with care-work- and family-related interruptions of mothers. And, given the low pension returns, it attributes substantial parts of the GPG. Second, given that female part-time employment increases across cohorts and there is a trend toward decreasing wages for part-time (e.g., Barbieri et al., 2019; Kreyenfeld, 2015; Schmitt & Auspurg, 2022), the Gender Pension Gap in the future will likely be more linked to women-majority life courses characterized by part-time employment – if the pension rewards for it do not change.

My results point further to a strong reproduction of gender earnings inequality in pension systems. Given that this is particularly the case in private pensions and the Gender Wage Gap has remained relatively constant for decades (e.g., England et al., 2020; Kunze, 2018; Schmieder & Wrohlich, 2021; Schmitt & Auspurg, 2022), the reproduction of earnings inequalities will tentatively increase in the future with the pension privatisation. Thus, implementing measures that guarantee a redistribution of earnings will help mitigate an

increasing reproduction of gender earnings inequalities in pensions as well. In Germany, the recently introduced ‘basic pension’ (*Grundrente*) is a first step and tends to support primarily women (Geyer et al., 2020). However, even though childcare periods of children up to age 10 are considered for eligibility, the minimum contribution of 33 years is likely still difficult to reach for women with almost continuous unpaid care work. Any reform that relies on a minimum contribution of many years will have to consider a realistic amount of unpaid care work provided to include the majority of elderly women.

Lastly, an open question remains how to prevent a severe production of gender inequalities in private pensions. One way would be to create private programs that entail some degree of redistribution such as the state-subsidised Riester program in Germany. While Riester successfully increased women's access to the program compared to men, women still only receive approximately half of the pension income that men do from Riester (Geyer et al., 2021). Otherwise, stronger redistributive measures in the public pension pillar will have to counter the more severe reproduction of inequalities in private pensions.

Besides all these commonly incorporated redistributive elements, pension systems could further institutionalise the redistribution at the couple level. The couple-level splitting of pension claims upon retirement in Germany is the only pension policy element that has been confirmed to significantly reduce the GPG and ensures women's fair pension income stake for the unpaid care work done within the traditional gender arrangement. Partners in Germany have the opportunity to equally share the pension claims that they have collected jointly through the gendered division of labour. However, the pension claim splitting is so far mostly realised by couples upon divorce. An extension to cohabiting couples and further incentives to choose pension income splitting has the potential to decrease gendered pension disparities and make women's pension income independent of future partnerships. Other countries might want to consider this tool as well. Additionally, expanding pension

claim splitting to private pensions might circumvent the lack of redistribution in private pensions.

In highlighting the importance but insufficiency of care work benefits, the conclusion is in line with longstanding evidence across different countries (e.g., Frommert & Strauß, 2013; Grabka et al., 2017; Horstmann et al., 2009; Stegmann, 2007). As such, a statement from the early 90s is largely still valid nowadays: “family work is not considered to be a contribution to the economy and is barely rewarded in retirement. Only activities outside the family are taken into account. Consequently, the system is constructed [...] for male full-time workers in gendered labour markets and family systems.” (Allmendinger et al., 1993, S. 193). This emphasizes the need to prevent the current gendered misalignment of welfare policies for the emergence of the GPGs and the suggestion of a life-course-sensitive and cohort-specific pension policy.

Overcoming the gendered mismatch between welfare state policies decades ago and pension policies nowadays: cohort-specific benefits for life-course-sensitive prevention of GPGs

Both pathways to achieving a more gender-equal pension income are not mutually exclusive but could be combined sustainably through cohort-specific and, therefore, life-course-sensitive pension policies. This would address the currently gendered misalignment between the incentive of highly gendered life courses through policies in the second part of the 20th century and unequal rewards through pension systems to those, that this thesis highlighted. From an individual perspective, these intertemporal policy interactions work well for men with typical male life courses only. For women, however, the mismatch is tremendous: women who fulfilled the role the welfare state assigned to them (female caretaker or secondary (part-time) earner), are largely penalized in old age. The policies

might align from the household perspective if these women have a partner with sufficient pension income and an equal sharing of household resources is realised. However, this perpetuates women's lifelong dependence and is not sustainable given the current demographic, social and policy developments. Thus, hoping for household-level redistribution will certainly become less realistic in the future. If pension systems pursue the path of pension privatisation and individualisation, then they need to make it accessible for everyone.

The intertemporal institutional mismatch could be addressed by designing retirement policies in accordance with the welfare state context in which the specific birth cohorts of the elderly spent their pre-retirement lives, particularly during their prime family formation ages (20-40). A cohort-specific redistribution would, for example, entitle higher childcare pension benefits for cohorts during whose family formation phase there were no or very limited childcare facilities. Similarly, pension rewards for part-time work while having children of school age could be enhanced significantly, as long as childcare facilities and schools predominantly offer part-time childcare during the caregiver's children's early years. These benefits would then gradually be cut for younger cohorts following increasing welfare support and defamiliarization of care needs in society. Other measures that could be used for defining the cohort-specific pension benefits and that would take into account women's disadvantage through pension privatisation would be the degree of gender inequalities in the labour force during the cohorts' active labour market age, such as the Gender Wage Gap or gender-specific occupational segregation.

Aligning the benefits based on the welfare support for a reconciliation of unpaid and paid work in which life courses unfolded can also help prevent potential negative anticipation effects for mothers' labour supply (e.g., S. Becker et al., 2021). Furthermore, it might also

incentivise a more rapid investment in policies that will enable a more gender-egalitarian division of labour, and consequently, lower welfare expenditures for care benefits in pensions in the long-term.

Such policies should still be flanked with an extension of couple-level pension splitting, where feasible (i.e., long-term partnerships), and potentially implementing means-testing that ensures decreasing benefits for the top pension receivers to prevent the reproduction of other inequality dimensions.

Life-course-sensitive pension benefits would support current and future elderly women's financial independence in old age while enabling a more sustainable pathway to own, independent pension for younger and future cohorts of women through a more gender-equal access to highly rewarded life course patterns. This way, all individuals, independent of the type of work they have performed over their life courses and the gendered opportunity contexts, can rely on their own, independent pension.

6.4 Limitations & avenues for future research

6.4.1 Limitations

Several limitations are important to stress. Some constraints relate to the data. First, the analyses are subject to a relatively low number of observations. This results in partly very small clusters (Chapters 3 and 4) or variable categories (Chapter 5) per gender. Smaller categories are more prone to be skewed by outliers and therefore to bias the results. Also, some associations might not be statistically significant due to a lack of statistical power. Larger sample sizes would be particularly important to replicate the analyses in Chapter 4 concerning private pension receipt and the decomposition along the distribution with more statistical power.

Second, there is a need for richer data containing more labour market characteristics over the life course that are typically gendered such as proxies for the horizontal and vertical gender-specific segregation (i.e., industries, occupations, positions). As I observe gender-based disparities in earnings throughout the life course to contribute significantly to gendered pension inequality, it is prudent to examine the influence of additional prevailing gender labour market inequalities.

Third, the SHARE data focuses on individuals in old age nowadays. However, we know that the life courses of individuals as well as occupational and personal pension coverage have changed. Thus, the analyses do not capture such recent trends. As a consequence of these data-related limitations, applying the analysis to more recent cohorts or countries with higher private pension coverage would be advisable.

Other limitations are related to the generalizability of the findings. This thesis focussed on countries with high levels of gendered life courses to introduce novel combinations of methods and prove their empirical and substantive contribution to quantifying the relationship of gendered life courses with the GPG. Thus, the results are not necessarily generalizable to other countries and policy contexts. I found that the main results remain independent of the pension system for the countries studied (West Germany, Italy, the Netherlands). I would expect the findings to be similar in countries with similarly gendered work-family life courses, except for contexts with highly progressive pension systems and with low coverage of private pensions. However, future research will have to prove this. Certainly, this thesis cannot derive implications for different contexts, such as more gender-equal ones with more destandardised life courses.

Moreover, the findings for the differences across public and private pensions as well as the role of earnings profiles rely solely on the pension context of West Germany since there

was no comparable linked survey-register data containing earnings over the life course available for other countries in SHARE (Chapter 4). However, the mechanisms might be different for other contexts and particularly, for contexts with higher shares of private pensions. Thus, to what extent the results are a temporal snapshot of a specific period and cohort dynamics still must be assessed by applying the analytical approach to younger cohorts or different countries.

Lastly, the analyses performed do not claim to be causal. Rather the results draw a broad picture of how gender inequalities over the life course are linked with the Gender Pension Gap which can be used as a starting point for more causal research in the future.

6.4.2 Future research

The methodological innovations and substantive contributions of this thesis lay the ground for follow-up research. Applying the two methodological innovations, SA-KOB and Life Course Feature Selection combined with Ñopo, to other life-course-sensitive outcomes promises new substantive findings with regard to how group inequalities in outcomes are related to life course complexities. Furthermore, an overall suggestion derived from this thesis is to avoid the steady work bias at any part of the empirical step, since it likely omits a selective group of women, or their life courses, in highly gendered contexts. First, the conceptualisation of the sample and measurements should be performed with caution, especially concerning the two measures pension receipt and self-assessed retirement. Second, both methodological combinations suggested help mitigate the steady worker bias in empirical models.

Some more specific avenues for future research stand out.

First, it is relevant to assess to what extent the results of this thesis change for younger cohorts, since this dissertation relies on data on the population in old age nowadays. This is important to take into account the role of current dynamics beyond the cohorts studied in this dissertation that are pertinent for the Gender Pension Gap: a slow shift towards more egalitarian models concerning the division of labour and a much stronger increasing female labour force participation (e.g., Nieuwenhuis et al., 2020; Sullivan, 2021; Zoch & Heyne, 2023), increasing complexities of life courses (Rowold & Van Winkle, 2023; Van Winkle, 2018; Van Winkle & Fasang, 2017, 2021) and the pension privatisation. The methodological applications developed in this thesis will most likely become even more rewarding applied to younger cohorts, as they consider these increasing life course complexities. For example, combining the methods with techniques to prognose the pension income and life courses of pre-retirement cohorts (Vecgaile, 2023) can improve our understanding of how the Gender Pension Gap develops in the future. Replicating the analysis for other contexts, such as younger cohorts or countries with a more egalitarian division of labour or different pension systems, is also important to assess whether the results hold beyond the studied cohorts and contexts with highly gendered work-family life courses. Similarly, future research might want to assess the role of earnings in other contexts, such as in pension systems that are less linked to earnings by design than the German pension system.

Second, and similarly, the methodological approaches should be extended to other combinations of varying gender arrangements and pension systems. Given the conceptual focus on a few countries with highly gendered life courses, extensions to other country comparison designs focussing on different levels of gendered life course segregation would be relevant for additional policy-relevant contributions. Such studies would allow to further disentangle the interrelation of different pension systems with these different levels of

gendered life courses. Examples would be Sweden, with a regressive pension system but more gender-equal life courses alongside relatively high gender wage gaps, or France, with somewhat less pronounced gender inequalities compared to the countries studied in this paper, but a similarly regressive pension system to Germany (Hammerschmid & Rowold, 2019a). Additionally, multicountry comparisons would provide a first idea of overarching patterns and identify relevant comparison groups. This could be followed up by country comparisons that focus on a few countries and allow for more detailed analyses.

Third, the developed methods could be applied to the household or couple-level dimension, which is largely understudied when it comes to gendered pension income (a couple perspective is only fully implemented by Möhring & Weiland, 2022, see Table 2.1). For example, decomposing the within-couple gap in pension income with the methods developed in this thesis might bring new evidence on how inequalities within couples develop over the life course, and which periods are most relevant for producing gender inequalities. Related to that, future research might want to focus on parents, given that this thesis suggests that the GPG is mostly driven by life course differences between mothers and fathers. Research focussing on the household level can also derive more direct implications for old-age poverty.

Fourth, this thesis suggests that it is fruitful to consider earnings over the life course and to categorise between different earning profiles to assess nuanced links to pension receipt as well as pension income at the mean and along the distribution. However, earnings over the life course are operationalised in direct linkage to full-time employment. There is still a need to quantify and disentangle which parts of gendered earnings versus employment inequalities over the life course are reproduced in Gender Pension Gaps. Furthermore, there is a need to assess the impact of other gendered labour market inequalities that develop

over the life course, such as occupational segregation which might be particularly relevant to assess for gender differences in private pensions.

Fifth, in line with this, there is an overall need for more evidence on how pension privatisation will affect gender disparities, as well as other forms of inequalities, in pension income. This might give a better understanding of how gender inequality in pension income will develop in the future. Moreover, future studies could assess the impact of new pension reforms on the GPG. For Germany, these would include the Riesterrente (subsidized private pension) and basic pension ('Grundrente'), both of which were largely not implemented yet for the years and cohorts studied in this thesis. Related to this, policy-relevant simulations would be helpful (e.g., Dekkers et al., 2022).

Sixth, future research might want to further disentangle the complex role of part-time employment for gender disparities in pension income. This thesis highlights the relevance of considering complexities beyond the duration of part-time employment, e.g., the strong link of part-time employment to care-work interruptions, for a more nuanced picture. However, future research might want to also consider how other pertinent labour market inequalities that are possibly linked to part-time employment such as earnings (Schmitt & Auspurg, 2022), but also occupational segregation, completes our understanding of part-time employment as a promoter or barrier to gender equality in pensions.

Methodologically, both introduced methods combinations, SA-KOB and Life Course Feature Selection combined with the Ñopo decomposition, can be further sophisticated and extended. First, the field of sequence analysis develops quickly and several methodological advancements are currently in development (Liao et al., 2022). Thus, the SA-KOB might benefit from the implementation of very recent innovations, such as concerning the Multichannel sequence analysis, heterogeneous sequence analysis clusters, and validation

of cluster outcomes (e.g., Helske et al., 2023; Piccarreta & Struffolino, 2023; Ritschard et al., 2023; Studer, 2021). These will likely ease the interpretation and make the sequence analysis, and consequently SA-KOB, more robust. Second, recent developments in the KOB and related decomposition techniques promise more classical causal counterfactual analyses (Jackson & VanderWeele, 2018; Lundberg, 2022; Yu & Elwert, 2023). Even though achieving causal analyses for processes that unfold over several decades presents challenges, exploring a combined approach with the SA-KOB may prove worthwhile. Third, the Life Course Feature Selection identifies the most important life course predictors of overall pension inequality, but not specifically of the Gender Pension Gap. Extensions of feature selection algorithms that disentangle the most important predictors of *gaps* in outcomes directly, will be more straightforwardly linkable to decomposition of group inequalities.

Overall, the proposed combinations of sequence analysis, Machine Learning and decomposition tools establish a foundation for further combinations of life courses and decomposition techniques, fostering the development of more life-course-sensitive decompositions.

6.5 Concluding remarks and outlook

The overarching finding of this thesis is that the unequal engagement in unpaid care work between genders over the life course is one of the main reasons for high Gender Pension Gaps. Unpaid care is only done by women and poorly rewarded in the different pension systems studied. This finding complements previous research which highlighted the role of (full-time) employment and did not consider care work. However, as my results show, such steady worker bias in empirical research risks concealing important inequality-generating dynamics. By providing more gender- and life-course-sensitive tools for

research to study GPGs, the thesis shows how we can overcome such biases and, consequently, reveal the association of inequalities with more complex mechanisms arising over the life course.

Thus, the thesis confirms that the Gender Pension Gaps are, to a large extent, the result of an accumulation of gender inequalities over the life course, that is expressed through the highly gender-specific life course patterns or even gender-exclusive combinations of life course experiences. However, this reproduction is only enabled through highly unequal pension returns for gender-specific life courses, as outcome of the gendered division of paid and unpaid work.

The results thus point to the strongly gendered misalignment of welfare state policies decades ago and pension institutions nowadays. Men who fulfilled the roles the welfare states have assigned to them receive generous pensions, while women who complied with their roles as caretakers are “only a husband away from poverty” (Orloff 1993, p. 319). The misalignment of the interaction of public policies over the life course is thus to the detriment of women and facilitates the long-lasting impacts the gendered division of labour has for gendered financial inequalities at older ages.

However, should policymakers be genuinely concerned about Gender Pension Gaps in the future? Previous research indicated a slight decrease of the Gender Pension Gap, with the prospect of further reduction attributed to the rising participation of women in the labour force (e.g., Fasang, 2009; Hänisch & Klos, 2014; Kreyenfeld et al., 2023; Tophoven & Tisch, 2016). This outlook aligns with recent predictions of a declining GPG across various countries in the future (Dekkers et al., 2022; Vecgaile, 2023), which mostly rely on public pensions only. However, whereas public pensions are indeed well predictable, the

predictability of private pensions – and as a consequence of the pension privatisation, total pension income – is diminishing.

The findings of this thesis cast doubt on such optimistic trends. While on one hand the share of women with continuous unpaid care work likely decreases over cohorts, the share of women reconciling their care duties with shorter care work-related employment interruptions and part-time employment likely increases. Both women-majority life course patterns are attributable to large parts of the GPG. At the same time, gender wage differences are relatively persistent (e.g., England et al., 2020; Kunze, 2018; Schmieder & Wrohlich, 2021; Schmitt & Auspurg, 2022) and my findings indicate that these are particularly reproduced in private pensions, the significance of which will increase. Since women's increasing part-time employment, and lower pensions for these, as well as the consistent gender earnings disparities, counters the trend towards the overall increasing female labour force participation, we might see a 'stalled' Gender Pension Gap, similar to the Gender Wage Gap in Germany (Schmitt & Auspurg, 2022). However, this will have to be assessed in future studies.

Whether or not the Gender Pension Gap can be significantly reduced in the future, and financial independence will be secured regardless of gender, will largely depend on the implementation of social policy reforms and their success. To reduce Gender Pension Gaps swiftly, I suggest a more life-course-sensitive pension policy with cohort-specific care benefits, coupled with investment in progressive work-family policies enabling more gender-equal division of labour, and consequently, life courses.

There is a long way to go to achieve gender equality in pension income. But it can be achieved if welfare state policies before retirement age and the pension design are better intertemporally coordinated – for all genders.

7 Bibliography

- Abatemarco, A., & Russolillo, M. (2022). The Dynamics of the Gender Gap at Retirement in Italy: Evidence from SHARE. *Italian Economic Journal*, 9, 445–473. <https://doi.org/10.1007/s40797-022-00201-7>
- Abbott, A. (1995). Sequence Analysis: New Methods for Old Ideas. *Annual Review of Sociology*, 21(1), 93–113.
- Abbott, A., & Forrest, J. (1986). Optimal Matching Methods for Historical Sequences. *Journal of Interdisciplinary History*, 16(3), 471–494. <https://doi.org/10.2307/204500>
- Abbott, A., & Hrycak, A. (1990). Measuring Resemblance in Sequence Data: An Optimal Matching Analysis of Musicians' Careers. *American Journal of Sociology*, 96(1), 144–185. <https://doi.org/10.1086/229495>
- Adema, W. (2012). Setting the scene: The mix of family policy objectives and packages across the OECD. *Children and Youth Services Review*, 34(3), 487–498. <https://doi.org/10.1016/j.childyouth.2011.10.007>
- Adema, W., Clarke, C., & Thévenon, O. (2020). Family Policies and Family Outcomes in OECD Countries. In R. Nieuwenhuis & W. Van Lancker (Hrsg.), *The Palgrave Handbook of Family Policy* (S. 193–217). Springer International Publishing. https://doi.org/10.1007/978-3-030-54618-2_9
- Aisenbrey, S., Evertsson, M., & Grunow, D. (2009). Is There a Career Penalty for Mothers' Time Out? A Comparison of Germany, Sweden and the United States. *Social Forces*, 88(2), 573–605. <https://doi.org/10.1353/sof.0.0252>
- Aisenbrey, S., & Fasang, A. E. (2010). New Life for Old Ideas: The „Second Wave“ of Sequence Analysis Bringing the „Course“ Back Into the Life Course. *Sociological Methods & Research*, 38(3), 420–462.

- Allmendinger, J., Brückner, H., & Brückner, E. (1992). Ehebande und Altersrente. Oder: Vom Nutzen der Individualanalyse. *Soziale Welt*, 43(1), 90–116.
- Allmendinger, J., Brückner, H., & Brückner, E. (1993). The Production of Gender Disparities over the Life Course and Their Effects in Old Age—Results from the West German Life History Study. In A. B. Atkinson & M. Rein (Hrsg.), *Age, Work and Social Security* (S. 188–223). Palgrave Macmillan UK. https://doi.org/10.1007/978-1-349-22668-9_8
- Almqvist, A.-L., & Duvander, A.-Z. (2014). Changes in gender equality? Swedish fathers' parental leave, division of childcare and housework¹. *Journal of Family Studies*, 20(1), 19–27. <https://doi.org/10.5172/jfs.2014.20.1.19>
- Amarante, V., Colacce, M., & Manzi, P. (2017). The gender gap in pensions in Latin America. *International Social Security Review*, 70(2), 57–85. <https://doi.org/10.1111/issr.12135>
- Anderson, K. M. (2011). 11 The Netherlands: Adapting a Multipillar Pension System to Demographic and Economic Change. In B. Ebbinghaus (Hrsg.), *The Varieties of Pension Governance: Pension Privatization in Europe* (S. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586028.003.0011>
- Anker, R. (1998). *Gender and jobs: Sex segregation of occupations in the world*. International Labour Organization.
- Antonova, L., Bucher-Koenen, T., & Mazzonna, F. (2017). Long-term health consequences of recessions during working years. *Social Science & Medicine*, 187, 134–143. <https://doi.org/10.1016/j.socscimed.2017.06.025>
- Anttonen, A., & Sipilä, J. (1996). European Social Care Services: Is It Possible To Identify Models? *Journal of European Social Policy*, 6(2), 87–100. <https://doi.org/10.1177/095892879600600201>

- Anxo, D., Mencarini, L., Pailhé, A., Solaz, A., Tanturri, M. L., & Flood, L. (2011). Gender Differences in Time Use over the Life Course in France, Italy, Sweden, and the US. *Feminist Economics*, 17(3), 159–195. <https://doi.org/10.1080/13545701.2011.582822>
- Arntz, M., Dlugosz, S., & Wilke, R. A. (2017). The Sorting of Female Careers after First Birth: A Competing Risks Analysis of Maternity Leave Duration. *Oxford Bulletin of Economics and Statistics*, 79(5), 689–716.
- Arulampalam, W., Booth, A. L., & Bryan, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Ilr Review*, 60(2), 163–186.
- Atlas of European Values. (2020). *Atlas of European Values*. <http://www.atlasofeuropeanvalues.eu/new/europa.php?ids=164&year=1990>
- Bach, S., Fischer, B., Haan, P. W., & Wrohlich, K. (2020). Reform des Ehegattensplittings: Realsplitting mit niedrigem Übertragungsbetrag ist ein guter Kompromiss. *DIW Wochenbericht*, 87(41), 785–794. https://doi.org/10.18723/diw_wb:2020-41-1
- Bach, S., Geyer, J., Haan, P., & Wrohlich, K. (2011). Reform of income splitting for married couples: Only individual taxation significantly increases working incentives. *DIW Economic Bulletin*, 1(5), 13–19.
- Barban, N., de Luna, X., Lundholm, E., Svensson, I., & Billari, F. C. (2020). Causal Effects of the Timing of Life-course Events: Age at Retirement and Subsequent Health. *Sociological Methods & Research*, 49(1), 216–249. <https://doi.org/10.1177/0049124117729697>
- Barbieri, P., Cutuli, G., Guetto, R., & Scherer, S. (2019). Part-time employment as a way to increase women's employment: (Where) does it work? *International Journal of*

<https://doi.org/10.1177/0020715219849463>

Barbieri, P., & Filippo, G. (2022). *Work-life instability and life-course income volatility. An age-cohort analysis of the stratified economic penalty of de-standardized work trajectories in Germany and the UK.*

Bardasi, E., & Jenkins, S. P. (2010). The Gender Gap in Private Pensions. *Bulletin of Economic Research*, 62(4), 343–363. <https://doi.org/10.1111/j.1467-8586.2009.00336.x>

Barth, D., Jessen, J., Spieß, C. K., & Wrohlich, K. (2020). Mothers in Eastern and Western Germany: Employment Rates and Attitudes Are Converging, Full-Time Employment is Not. *DIW Weekly Report*. https://doi.org/10.18723/DIW_DWR:2020-38-1

Bauernschuster, S., & Schlotter, M. (2015). Public child care and mothers' labor supply—Evidence from two quasi-experiments. *Journal of Public Economics*, 123, 1–16.

Baxter, J., Hewitt, B., & Haynes, M. (2008). Life Course Transitions and Housework: Marriage, Parenthood, and Time on Housework. *Journal of Marriage and Family*, 70(2), 259–272. <https://doi.org/10.1111/j.1741-3737.2008.00479.x>

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of political economy*, 70(5, Part 2), 9–49.

Becker, S., Buslei, H., Geyer, J., & Haan, P. (2021). Employment Responses to Income Effect: Evidence from Pension Reform. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3834988>

Becker-Schmidt, R., & Krüger, H. (2009). Krisenherde in gegenwärtigen Sozialgefügen: Asymmetrische Arbeits-und Geschlechterverhältnisse–vernachlässigte Sphären

- gesellschaftlicher Reproduktion. In *Arbeit. Perspektiven und Diagnosen der Geschlechterforschung*. (S. 12–41). Verlag Westfälisches Dampfboot.
- Belloni, M., Brugiavini, A., Buia, R. E., Carrino, L., Cavapozzi, D., Orso, C. E., & Pasini, G. (2020). What do we learn about redistribution effects of pension systems from internationally comparable measures of Social Security wealth? *Journal of Pension Economics & Finance*, 19(4), 548–566.
<https://doi.org/10.1017/S1474747219000118>
- Belotti, F., Deb, P., Manning, W. G., & Norton, E. C. (2015). twopm: Two-part models. *The Stata Journal*, 15(1), 3–20.
- Bergmann, M., Kneip, T., De Luca, G., & Scherpenzeel, A. (2019). Survey participation in the Survey of Health, Ageing and Retirement in Europe (SHARE), Wave 1-7. *Munich: Munich Center for the Economics of Aging*.
- Bernardi, L., Huinink, J., & Settersten, R. A. (2019). The life course cube: A tool for studying lives. *Advances in Life Course Research*, 41, 100258.
<https://doi.org/10.1016/j.alcr.2018.11.004>
- Bettio, F., Tinios, P., & Betti, G. (Hrsg.). (2013). *The gender gap in pensions in the EU*. Publications Office of the European Union.
- Blau, F. D., Ferber, M. A., & Winkler, A. E. (2006). *The economics of women, men and work*. Pearson Higher Ed.
- Blau, F. D., & Kahn, L. M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature*, 55(3), 789–865.
<https://doi.org/10.1257/jel.20160995>
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*, 8(4), 436–455.

- Boeckmann, I., Misra, J., & Budig, M. J. (2015). Cultural and Institutional Factors Shaping Mothers' Employment and Working Hours in Postindustrial Countries. *Social Forces*, 93(4), 1301–1333. <https://doi.org/10.1093/sf/sou119>
- Boehmke, B., & Greenwell, B. M. (2019). *Hands-on machine learning with R*. CRC press.
- Bolano, D., & Studer, M. (2020). *The link between previous life trajectories and a later life outcome: A feature selection approach*. <https://doi.org/10.12682/LIVES.2296-1658.2020.82>
- Boll, C., & Lagemann, A. (2019). Public Childcare and Maternal Employment—New Evidence for Germany. *LABOUR*, 33(2), 212–239. <https://doi.org/10.1111/labr.12143>
- Bonaccolto-Töpfer, M., & Briel, S. (2022). The gender pay gap revisited: Does machine learning offer new insights? *Labour Economics*, 78, 102223. <https://doi.org/10.1016/j.labeco.2022.102223>
- Bonnet, C., & Geraci, M. (2009). Correcting gender inequality in pensions. The experience of five European countries. *Population and Societies*, 453(4).
- Bonnet, C., Hourriez, J.-M., & Depledge, R. (2012). The Treatment of Couples by the Pension System: Survivor's Pensions and Pension Splitting. *Population*, Vol. 67(1), 147–162.
- Bonnet, C., Hourriez, J.-M., & Reeve, P. (2012). Gender Equality in Pensions: What Role for Rights Accrued as a Spouse or a Parent? *Population (English Edition, 2002-)*, 67(1), 123–146.
- Bonnet, C., Meurs, D., & Rapoport, B. (2020). Gender pension gaps along the distribution: An application to the French case. *Journal of Pension Economics & Finance*, 21(1), 1–23.

- Bonnet, C., & Rapoport, B. (2020). Is There a Child Penalty in Pensions? The Role of Caregiver Credits in the French Retirement System. *European Journal of Population*, 36(1), 27–52. <https://doi.org/10.1007/s10680-019-09517-0>
- Börsch-Supan, A., Brandt, M., Hunkler, C., Kneip, T., & Korbmacher, J. (2013). Data Resource Profile: The Survey of Health, Ageing and Retirement in Europe (SHARE). *International Journal of Epidemiology*, 42(4), 992–1001.
- Börsch-Supan, A., Czaplicki, C., Friedel, S., Herold, I., Korbmacher, J., & Mika, T. (2018). SHARE-RV: Linked Data to Study Aging in Germany. *Journal of Economics and Statistics*.
- Börsch-Supan, A., Czaplicki, C., Friedel, S., Herold, I., Korbmacher, J., & Mika, T. (2020). SHARE-RV: Linked data to study aging in Germany. *Jahrbücher für Nationalökonomie und Statistik*, 240(1), 121–132.
- Börsch-Supan, A., & Schröder, M. (2011). Retrospective data collection in the Survey of Health, Ageing and Retirement in Europe. *SHARELIFE Methodol*, 5.
- Boulton, A. J., & Williford, A. (2018). Analyzing Skewed Continuous Outcomes With Many Zeros: A Tutorial for Social Work and Youth Prevention Science Researchers. *Journal of the Society for Social Work and Research*, 9(4), 721–740. <https://doi.org/10.1086/701235>
- Bravo, J. M., & Herce, J. A. (2022). Career breaks, broken pensions? Long-run effects of early and late-career unemployment spells on pension entitlements. *Journal of Pension Economics & Finance*, 21(2), 191–217. <https://doi.org/10.1017/S1474747220000189>
- Bridgen, P., & Meyer, T. (2014). The Liberalisation of the German Social Model: Public–Private Pension Reform in Germany since 2001. *Journal of Social Policy*, 43(1), 37–68. <https://doi.org/10.1017/S0047279413000597>

- Bridgen, P., Meyer, T., & Davison, L. (2022). It's not late entry: Human capital, welfare states and the pension penalty experienced by post-war migrants who retired in the European Economic Area. *Ageing & Society*, 1–33. <https://doi.org/10.1017/S0144686X21001811>
- Brown, S. L., & Wright, M. R. (2017). Marriage, Cohabitation, and Divorce in Later Life. *Innovation in Aging*, 1(2), igx015. <https://doi.org/10.1093/geroni/igx015>
- Brugiavini, A., Cavapozzi, D., Pasini, G., & Trevisan, E. (2013). *Working life histories from SHARELIFE: a retrospective panel*. 14.
- Brugiavini, A., Orso, C. E., Genie, M. G., Naci, R., & Pasini, G. (2019). *Combining the retrospective interviews of wave 3 and wave 7: The third release of the SHARE Job Episodes Panel*.
- Brussig, M., Postels, D., & Zink, L. (2017). Niedrige Renten trotz langer Versicherungszeiten. Eine empirische Analyse der Risikofaktoren. *WSI-Mitteilungen*, 70(4), 248–258. <https://doi.org/10.5771/0342-300X-2017-4-248>
- Brzinsky-Fay, C., Kohler, U., & Luniak, M. (2006). Sequence Analysis with Stata. *The Stata Journal*, 6(4), 435–460.
- Budig, M. J., Misra, J., & Boeckmann, I. (2016). Work–Family Policy Trade-Offs for Mothers? Unpacking the Cross-National Variation in Motherhood Earnings Penalties. *Work and Occupations*, 43(2), 119–177. <https://doi.org/10.1177/0730888415615385>
- Bühler, B., Möhring, K., & Weiland, A. P. (2022). Assessing dissimilarity of employment history information from survey and administrative data using sequence analysis techniques. *Quality & Quantity*. <https://doi.org/10.1007/s11135-022-01333-9>
- Butts, L. (2020). A Life-long Culmination of Inequalities: Mechanisms Behind the Gender Pension Gap. *The Public Sphere: Journal of Public Policy*, 8(1).

- Carr, E., Murray, E. T., Zaninotto, P., Cadar, D., Head, J., Stansfeld, S., & Stafford, M. (2016). The Association Between Informal Caregiving and Exit From Employment Among Older Workers: Prospective Findings From the UK Household Longitudinal Study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, gbw156. <https://doi.org/10.1093/geronb/gbw156>
- Cohn-Schwartz, E., & Naegele, L. (2023). Employment over the life course and post-retirement social networks: A gendered perspective. *International Psychogeriatrics*, 1–11. <https://doi.org/10.1017/S1041610223000558>
- Cooke, M. (2006). Policy Changes and the Labour Force Participation of Older Workers: Evidence from Six Countries. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, 25(4), 387–400. <https://doi.org/10.1353/cja.2007.0015>
- Cordova, K., Grabka, M. M., & Sierminska, E. (2022). Pension Wealth and the Gender Wealth Gap. *European Journal of Population*, 38, 755–810. <https://doi.org/10.1007/s10680-022-09631-6>
- Corsi, M., & D’Ippoliti, C. (2009). Poor old grandmas? A note on the gender dimension of pension reforms. *Brussels Economic Review*, 52(1), 35–56.
- Cotton, J. (1988). On the Decomposition of Wage Differentials. *The Review of Economics and Statistics*, 70(2), 236–243. <https://doi.org/10.2307/1928307>
- Crespi, I., Zanier, M. L., Santoni, C., Fermani, A., & D’Ambrosi, L. (2015). Family, work and old women’s situation in Italy and Spain: New gender inequalities. *About Gender*, 4(8), 20–52.
- Czepek, J. A. (2020). *Flexibler Arbeitsmarkt, unflexible Renten? Diskontinuierliche Lebensverläufe und ihre Absicherung im Alter.* 85–100. <https://doi.org/10.5771/9783748906889-85>

- Daly, M., & Ferragina, E. (2018). Family policy in high-income countries: Five decades of development. *Journal of European Social Policy*, 28(3), 255–270. <https://doi.org/10.1177/0958928717735060>
- Damman, M., Henkens, K., & Kalmijn, M. (2014). *Family Histories and Women's Retirement: The Role of Childbearing and Marital Experiences* (SSRN Scholarly Paper ID 2471444). Social Science Research Network. <https://doi.org/10.2139/ssrn.2471444>
- Dannefer, D. (1987). Aging as intracohort differentiation: Accentuation, the Matthew effect, and the life course. *Sociological Forum*, 2(2), 211–236. <https://doi.org/10.1007/BF01124164>
- Dannefer, D. (2003). Cumulative Advantage/Disadvantage and the Life Course: Cross-Fertilizing Age and Social Science Theory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 58(6), 327–S337. <https://doi.org/10.1093/geronb/58.6.S327>
- de Castro Galvao, J. (2023). Gender Inequality in Lifetime Earnings. *Social Forces*, 101(4), 1772–1802. <https://doi.org/10.1093/sf/soac060>
- de la Vega, N. (2022). The differential effect of childbirth on men's and women's careers. *Labour Economics*, 78, 102249. <https://doi.org/10.1016/j.labeco.2022.102249>
- Degenhardt, F., Seifert, S., & Szymczak, S. (2019). Evaluation of variable selection methods for random forests and omics data sets. *Briefings in Bioinformatics*, 20(2), 492–503. <https://doi.org/10.1093/bib/bbx124>
- Dekkers, G., Van den Bosch, K., Barslund, M., Kirn, T., Baumann, N., Kump, N., Liégeois, P., Moreira, A., & Stropnik, N. (2022). How Do Gendered Labour Market Trends and the Pay Gap Translate into the Projected Gender Pension Gap? A Comparative

- Analysis of Five Countries with Low, Middle and High GPGs. *Social Sciences*, 11(7), Article 7. <https://doi.org/10.3390/socsci11070304>
- Del Boca, D. (2002). The effect of child care and part time opportunities on participation and fertility decisions in Italy. *Journal of Population Economics*, 15(3), 549–573. <https://doi.org/10.1007/s001480100089>
- Denton, F. T., & Spencer, B. G. (2009). What Is Retirement?: A Review and Assessment of Alternative Concepts and Measures. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, 28(1), 63–76.
- Dingemans, E., & Henkens, K. (2019). Working After Retirement and Life Satisfaction: Cross-National Comparative Research in Europe. *Research on Aging*, 41(7), 648–669. <https://doi.org/10.1177/0164027519830610>
- DiPrete, T. A., & Eirich, G. M. (2006). Cumulative Advantage as a Mechanism for Inequality: A Review of Theoretical and Empirical Developments. *Annual Review of Sociology*, 32(1), 271–297. <https://doi.org/10.1146/annurev.soc.32.061604.123127>
- Dotti Sani, G. M., & Luppi, M. (2021). Absence from Work after the Birth of the First Child and Mothers' Retirement Incomes: A Comparative Analysis of 10 European Countries. *Work, Employment and Society*, 35(3), 470–489. <https://doi.org/10.1177/0950017020937935>
- Ebbinghaus, B. (Hrsg.). (2011a). 1 Introduction: Studying Pension Privatization in Europe. In *The Varieties of Pension Governance: Pension Privatization in Europe* (S. 0). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586028.003.0001>
- Ebbinghaus, B. (2011b). *The Varieties of Pension Governance: Pension Privatization in Europe*. Oxford University Press.

- Ebbinghaus, B. (2015). The Privatization and Marketization of Pensions in Europe: A Double Transformation Facing the Crisis. *European Policy Analysis*, 1(1). <https://doi.org/10.18278/epa.1.1.5>
- Ebbinghaus, B. (2018). Privatisierung und Vermarktlichung der Altersvorsorge: Eingetrübte Aussichten des deutschen Mehssäulenmodells. *WSI-Mitteilungen*, 71(6), 468–475. <https://doi.org/10.5771/0342-300X-2018-6-468>
- Ebbinghaus, B. (2021). Inequalities and poverty risks in old age across Europe: The double-edged income effect of pension systems. *Social Policy & Administration*, 55(3), 440–455. <https://doi.org/10.1111/spol.12683>
- Ebbinghaus, B., Gronwald, M., & Wiß, T. (2011). Germany: Departing from Bismarckian Public Pensions. In *The Varieties of Pension Governance*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199586028.003.0005>
- Ebbinghaus, B., & Neugschwender, J. (2011). The public-private pension mix and old age income inequality in Europe. In *The Varieties of Pension Governance: Pension Privatization in Europe*. (S. 384–422). Oxford University Press.
- Elder, G. H., Johnson, M. K., & Crosnoe, R. (2003). The Emergence and Development of Life Course Theory. In J. T. Mortimer & M. J. Shanahan (Hrsg.), *Handbook of the Life Course* (S. 3–19). Springer US. https://doi.org/10.1007/978-0-306-48247-2_1
- Elzinga, C. H. (2010). Complexity of Categorical Time Series. *Sociological Methods & Research*, 38(3), 463–481. <https://doi.org/10.1177/0049124109357535>
- England, P., Privalko, I., & Levine, A. (2020). Has the Gender Revolution Stalled? *The Economic and Social Review*, 51(4, Winter), Article 4, Winter.
- European Institute for Gender Equality. (2015). *Gender gap in pensions in the EU: Research note to the Latvian presidency*. European Institute for Gender Equality.

- European Institute for Gender Equality. (2018). *European Institute for Gender Equality*. Gender differences in the aggregate replacement ratio. http://eige.europa.eu/gender-statistics/dgs/indicator/ta_fineco_inc_pension__ilc_pnp11
- Eurostat. (2018). *Aggregate replacement ratio (excluding other social benefits) by sex*. Aggregate replacement ratio (excluding other social benefits) by sex. <https://ec.europa.eu/eurostat/databrowser/view/tespn070/default/table?lang=en>.
- Evandrou, M., Falkingham, J., & Sefton, T. (2009). *Women's Family Histories and Incomes in Later Life in the UK, US and West Germany* (LSE STICERD Research Paper No. CASE138 ID 1547614; S. 1–31). Social Science Research Network. <https://papers.ssrn.com/abstract=1547614>
- Even, W. E., & Macpherson, D. A. (2004). When Will the Gender Gap in Retirement Income Narrow? *Southern Economic Journal*, 71(1), 182–200. <https://doi.org/10.2307/4135320>
- Ezeyi, O., & Vujic, S. (2017). *Twin peaks: An analysis of the gender gap in pension income in England*. Twenty-Second Annual Conference of the Society of Labor Economists.
- Faik, J., & Köhler-Rama, T. (2012). Der Gender Pension Gap—Ungeeigneter Indikator. *Wirtschaftsdienst*, 92(5), 319–325. <https://doi.org/10.1007/s10273-012-1382-3>
- Fasang, A. E. (2009). *Family biographies and retirement processes: A comparative analysis of West Germany and the United Kingdom* [PhD Thesis]. Jacobs University Bremen.
- Fasang, A. E. (2010). Retirement: Institutional Pathways and Individual Trajectories in Britain and Germany. *Sociological Research Online*, 15(2), 1–16. <https://doi.org/10.5153/sro.2110>

- Fasang, A. E. (2012). Retirement Patterns and Income Inequality. *Social Forces*, 90(3), 685–711. <https://doi.org/10.1093/sf/sor015>
- Fasang, A. E., & Aisenbrey, S. (2022). Uncovering Social Stratification: Intersectional Inequalities in Work and Family Life Courses by Gender and Race. *Social Forces*, 101(2), 575–605. <https://doi.org/10.1093/sf/soab151>
- Fasang, A. E., Aisenbrey, S., & Schomann, K. (2013). Women’s Retirement Income in Germany and Britain. *European Sociological Review*, 29(5), 968–980. <https://doi.org/10.1093/esr/jcs075>
- Fasang, A. E., & Liao, T. F. (2014). Visualizing Sequences in the Social Sciences. *Sociological Methods & Research*, 43(4), 643–676. <https://doi.org/10.1177/0049124113506563>
- Fasang, A. E., & Ulrich Mayer, K. (2020). Lifecourse and social inequality. In J. Falkingham, M. Evandrou, & A. Vlachantoni, *Handbook on Demographic Change and the Lifecourse* (S. 22–39). Edward Elgar Publishing. <https://doi.org/10.4337/9781788974875.00009>
- Ferragina, E. (2019). Does Family Policy Influence Women’s Employment?: Reviewing the Evidence in the Field. *Political Studies Review*, 17(1), 65–80. <https://doi.org/10.1177/1478929917736438>
- Ferragina, E., & Seeleib-Kaiser, M. (2015). Determinants of a Silent (R)evolution: Understanding the Expansion of Family Policy in Rich OECD Countries. *Social Politics: International Studies in Gender, State & Society*, 22(1), 1–37. <https://doi.org/10.1093/sp/jxu027>
- Firpo, S. P., Fortin, N. M., & Lemieux, T. (2018). Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, 6(2), Article 2. <https://doi.org/10.3390/econometrics6020028>

- Fischbach, R., & Eckert, R. (1979). Geschlechtsspezifische Arbeitsteilung im Vergleich von fünf Ländern unterschiedlicher politischer und wirtschaftlicher Ordnung. In *Geschlechtsrollen und Arbeitsteilung. Mann und Frau aus soziologischer Sicht*. (S. 202–233). Beck.
- Folbre, N., Shaw, L. B., & Stark, A. (2005). Introduction: Gender and Aging. *Feminist Economics*, 11(2), 3–5. <https://doi.org/10.1080/13545700500115803>
- Fortin, N., Lemieux, T., & Firpo, S. (2011). Chapter 1—Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Hrsg.), *Handbook of Labor Economics* (Bd. 4, S. 1–102). Elsevier. [https://doi.org/10.1016/S0169-7218\(11\)00407-2](https://doi.org/10.1016/S0169-7218(11)00407-2)
- Foster, L., & Ginn, J. (2018). Transmitting inequality: Pensions policy and the gendered life course. In *Handbook on gender and social policy*. Edward Elgar Publishing.
- Frericks, P. (2020). Die Absicherung familialer Sorgearbeit in der Rente. Entwicklungen und Konzepte im internationalen Vergleich. In *Neustart in der Rentenpolitik* (S. 139–156). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783748906889-139>
- Frericks, P., Maier, R., & De Graaf, W. (2007). European Pension Reforms: Individualization, Privatization and Gender Pension Gaps. *Social Politics: International Studies in Gender, State & Society*, 14(2), 212–237. <https://doi.org/10.1093/sp/jxm008>
- Frericks, P., Maier, R., & Graaf, W. D. (2006). Shifting the Pension Mix: Consequences for Dutch and Danish Women. *Social Policy & Administration*, 40(5), 475–492. <https://doi.org/10.1111/j.1467-9515.2006.00500.x>
- Frommert, D. (2020). Der Gender Pension Gap und soziale Elemente im Mehssäulensystem. *Deutsche Rentenversicherung*, 75(2/2020), 211–225.

- Frommert, D., & Lizon, M. (2017). Teilzeitarbeit neben Kindererziehung—Auswirkung auf die gesetzliche Rente. *Deutsche Rentenversicherung*, 3/2017, 273–291.
- Frommert, D., & Strauß, S. (2013). Biografische Einflussfaktoren auf den Gender Pension Gap – Ein Kohortenvergleich für Westdeutschland. *Journal for Labour Market Research*, 46(2), 145–166. <https://doi.org/10.1007/s12651-012-0125-7>
- Gabadinho, A., Ritschard, G., Müller, N. S., & Studer, M. (2011). Analyzing and Visualizing State Sequences in R with TraMineR. *Journal of Statistical Software*, 40(4), 1–37.
- Gangl, M., & Ziefle, A. (2009). Motherhood, Labor Force Behavior, and Women's Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States. *Demography*, 46(2), 341–369. <https://doi.org/10.1353/dem.0.0056>
- Gangl, M., & Ziefle, A. (2015). The Making of a Good Woman: Extended Parental Leave Entitlements and Mothers' Work Commitment in Germany. *American Journal of Sociology*, 121(2), 511–563. <https://doi.org/10.1086/682419>
- García-Sierra, A. (2023). *Learning by Parenting: How do Mothers Respond to Their Children's Developmental Declines?* Forthcoming as preprint.
- Gardiner, J., Robinson, A. M., & Fakhfakh, F. (2016). Exploring the private pension gender gap and occupation in later working life. *Work, Employment and Society*, 30(4), 687–707. <https://doi.org/10.1177/0950017015575868>
- Garrouste, C., & Paccagnella, O. (2011). Data quality: Three examples of consistency across SHARE and SHARELIFE data. *Retrospective Data Collection in the Survey of Health, Ageing and Retirement in Europe. SHARELIFE Methodology. Mannheim Research Institute for the Economics of Ageing (MEA): Mannheim*, 62–72.

- Gauthier, J.-A., Widmer, E. D., Bucher, P., & Notredame, C. (2010). 1. Multichannel Sequence Analysis Applied to Social Science Data. *Sociological Methodology*, 40(1), 1–38.
- Geyer, J., Grabka, M. M., & Haan, P. (2021). 20 Years of the Riester Pension—Personal Retirement Provision Requires Reform. *DIW Weekly Report*. https://doi.org/10.18723/DIW_DWR:2021-40-1
- Geyer, J., Haan, P., & Harnisch, M. (2020). *Zur Wirkung der Grundrente und der Mütterrente auf die Altersarmut*. 32.
- Geyer, J., Haan, P., & Wrohlich, K. (2015). The effects of family policy on maternal labor supply: Combining evidence from a structural model and a quasi-experimental approach. *Labour Economics*, 36, 84–98. <https://doi.org/10.1016/j.labeco.2015.07.001>
- Ginn, J. (2003). Parenthood, Partnership Status and Pensions: Cohort differences among women. *Sociology*, 37(3), 493–510.
- Ginn, J. (2004). European Pension Privatisation: Taking Account of Gender. *Social Policy and Society*, 3(2), 123–134. <https://doi.org/10.1017/S1474746403001581>
- Ginn, J., & Arber, S. (1996). Patterns of employment, gender and pensions: The effect of work history on older women's non-state pensions. *Work, employment and society*, 10(3), 469–490.
- Ginn, J., Daly, M., & Street, D. (2001). Engendering pensions: A comparative framework. *Women, work and pensions: International issues and prospects*. Buckingham: Open University Press, 1–10.
- Ginn, J., & MacIntyre, K. (2013). UK Pension Reforms: Is Gender Still an Issue? *Social Policy and Society*, 12(1), 91–103. <https://doi.org/10.1017/S1474746412000504>

- Grabka, V. M. M., Jotzo, B., Rasner, A., & Westermeier, C. (2017). *Der Gender Pension Gap verstärkt die Einkommensungleichheit von Männern und Frauen im Rentenalter*. 84(5), 12.
- Grady, J. (2015). Gendering Pensions: Making Women Visible. *Gender, Work & Organization*, 22(5), 445–458. <https://doi.org/10.1111/gwao.12096>
- Grech, A. G. (2013). Assessing the sustainability of pension reforms in Europe. *Journal of International and Comparative Social Policy*, 29(2), 143–162. <https://doi.org/10.1080/21699763.2013.836980>
- Gregorutti, B., Michel, B., & Saint-Pierre, P. (2017). Correlation and variable importance in random forests. *Statistics and Computing*, 27(3), 659–678. <https://doi.org/10.1007/s11222-016-9646-1>
- Gruijters, R. J., Van Winkle, Z., & Fasang, A. E. (2023). Life Course Trajectories and Wealth Accumulation in the United States: Comparing Late Baby Boomers and Early Millennials. *American Journal of Sociology*, 129(2), 530–569. <https://doi.org/10.1086/726445>
- Guillemard, A.-M., & Rein, M. (1993). Comparative Patterns of Retirement: Recent Trends in Developed Societies. *Annual Review of Sociology*, 19, 469–503.
- Haan, P., & Wrohlich, K. (2011). Can child care policy encourage employment and fertility?: Evidence from a structural model. *Labour Economics*, 18(4), 498–512. <https://doi.org/10.1016/j.labeco.2010.12.008>
- Haitz, N. (2015). Old-age poverty in OECD Countries and the Issue of Gender Pension Gaps. *CESifo DICE Report*, 13(2), 73–75.
- Halman, L., & Vloet, A. (1994). *Measuring and comparing values in 16 countries of the Western world: Documentation of the European Values Study 1981-1990 in Europe and North America*. Work and Organization Research Centre, Tilburg University.

- Halpern-Manners, A., Warren, J. R., Raymo, J. M., & Nicholson, D. A. (2015). The Impact of Work and Family Life Histories on Economic Well-Being at Older Ages. *Social Forces*, 93(4), 1369–1396. <https://doi.org/10.1093/sf/sov005>
- Halpin, B. (2017). SADI: Sequence Analysis Tools for Stata. *The Stata Journal*, 17(3), 546–572. <https://doi.org/10.1177/1536867X1701700302>
- Hamjediers, M., & Sprengholz, M. (2023). Comparing the Incomparable? Issues of Lacking Common Support, Functional-Form Misspecification, and Insufficient Sample Size in Decompositions. *Sociological Methodology*, 008117502311697. <https://doi.org/10.1177/00811750231169729>
- Hammerschmid, A., & Rowold, C. (2019a). Gender Pension Gaps in Europe Are More Explicitly Associated with Labor Markets than with Pension Systems. *DIW Weekly Report*, 9(25), 204–211. https://doi.org/10.18723/DIW_DWR:2019-25-1
- Hammerschmid, A., & Rowold, C. (2019b). Gender pension gaps—A problem in many European countries. *DIW Weekly Report*, 9(16/18), 156–157. https://doi.org/10.18723/DIW_DWR:2019-16-3
- Hänisch, C., & Klos, J. (2014). *A decomposition analysis of the German gender pension gap*. Discussion Paper Series, Wilfried-Guth-Stiftungsprofessur für Ordnungs-und Wettbewerbspolitik, Universität Freiburg.
- Harkness, S., Borkowska, M., & Pelikh, A. (2019). Employment pathways and occupational change after childbirth. In *Government Equalities Office: London, UK*. [Report]. Government Equalities Office. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/840062/Bristol_Final_Report_1610.pdf

- Hegewisch, A., & Gornick, J. C. (2011). The impact of work-family policies on women's employment: A review of research from OECD countries. *Community, Work & Family*, 14(2), 119–138. <https://doi.org/10.1080/13668803.2011.571395>
- Helske, S., Helske, J., & Chihaya, G. K. (2023). From Sequences to Variables: Rethinking the Relationship between Sequences and Outcomes. *Sociological Methodology*, 00811750231177026. <https://doi.org/10.1177/00811750231177026>
- Hinrichs, K. (2021). Recent pension reforms in Europe: More challenges, new directions. An overview. *Social Policy & Administration*, 55(3), 409–422. <https://doi.org/10.1111/spol.12712>
- Hlavac, M. (2014). *oaxaca: Blinder-Oaxaca Decomposition in R* (SSRN Scholarly Paper 2528391; S. 1–19). Social Science Research Network. <https://doi.org/10.2139/ssrn.2528391>
- Hofmeister, H., Blossfeld, H.-P., & Mills, M. (2006). Globalization, Uncertainty and Women's Mid-Career Life Courses: A Theoretical Framework. In H.-P. Blossfeld & H. Hofmeister, *Globalization, Uncertainty and Women's Careers* (S. 3–31). Edward Elgar Publishing. <https://doi.org/10.4337/9781781007495.00008>
- Horstmann, S., Hülsman, J., Corsi, M., Ginn, J., Leitner, S., & Leppik, L. (2009). *The socio-economic impact of pension systems on women*. Report prepared for European Commission, Directorate-General for Employment, Social Affairs, and Equal Opportunities, Brussels.
- Huebener, M., Müller, K.-U., Spieß, C. K., & Wrohlich, K. (2016). The parental leave benefit: A key family policy measure, one decade later. *DIW Economic Bulletin*, 6(49), 571–578.

- Ice, E., Ang, S., Greenberg, K., & Burgard, S. (2020). Women's Work-Family Histories and Cognitive Performance in Later Life. *American Journal of Epidemiology*, 189(9), 922–930. <https://doi.org/10.1093/aje/kwaa042>
- Jackson, J., & VanderWeele, T. (2018). Decomposition analysis to identify intervention targets for reducing disparities. *Epidemiology (Cambridge, Mass.)*, 29(6), 825–835. <https://doi.org/10.1097/EDE.0000000000000901>
- Jalovaara, M., & Fasang, A. E. (2020). Family Life Courses, Gender, and Mid-Life Earnings. *European Sociological Review*, 36(2), 159–178.
- Janitza, S., Celik, E., & Boulesteix, A.-L. (2018). A computationally fast variable importance test for random forests for high-dimensional data. *Advances in Data Analysis and Classification*, 12(4), 885–915. <https://doi.org/10.1007/s11634-016-0276-4>
- Jann, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *The Stata Journal*, 8(4), 453–479.
- Jędrzychowska, A., Kwiecień, I., & Poprawska, E. (2020). The Motherhood Pension Gap in a Defined Contribution Pension Scheme—The Case of Poland. *Sustainability*, 12(11), Article 11. <https://doi.org/10.3390/su12114425>
- Jefferson, T. (2009). Women and Retirement Pensions: A Research Review. *Feminist Economics*, 15(4), 115–145. <https://doi.org/10.1080/13545700903153963>
- Johnson, R. W., Sambamoorthi, U., & Crystal, S. (1999). Gender Differences in Pension Wealth: Estimates Using Provider Data. *The Gerontologist*, 39(3), 320–333. <https://doi.org/10.1093/geront/39.3.320>
- Joshi, H., & Davies, H. (2002). *Women's incomes over a synthetic lifetime*.

- Kapelle, N., & Vidal, S. (2021). Heterogeneity in Family Life Course Patterns and Intra-Cohort Wealth Disparities in Late Working Age. *European Journal of Population*, 38, 59–92. <https://doi.org/10.1007/s10680-021-09601-4>
- Kaufman, L., & Rousseeuw, P. (1990). *Finding groups in data; an introduction to cluster analysis*. J. Wiley.
- Kelle, N., Simonson, J., & Gordo, L. R. (2017). Is Part-Time Employment after Childbirth a Stepping-Stone into Full-Time Work? A Cohort Study for East and West Germany. *Feminist Economics*, 23(4), 201–224. <https://doi.org/10.1080/13545701.2016.1257143>
- Kitagawa, E. M. (1955). Components of a Difference Between Two Rates. *Journal of the American Statistical Association*, 50(272), 1168–1194.
- Klammer, U., Klenner, C., Ochs, C., Radke, P., & Ziegler, A. (2000). WSI-FrauenDaten-Report (Report on Women by the WSI). *Berlin: Ed. Sigma*.
- Komp-Leukkunen, K. (2019). Capturing the diversity of working age life-courses: A European perspective on cohorts born before 1945. *PLOS ONE*, 14(2), e0212400. <https://doi.org/10.1371/journal.pone.0212400>
- Komp-Leukkunen, K. (2021). Breadwinner Models Revisited: How a Couple's Combined Work Histories Influence the Retirement Transition. *Social Politics: International Studies in Gender, State & Society*, 28(2), 335–358. <https://doi.org/10.1093/sp/jxz008>
- König, S., Johansson, B. E. A., & Bolin, K. (2019). Invisible Scars or Open Wounds? The Role of Mid-career Income for the Gender Pension Gap in Sweden. *Frontiers in Sociology*, 4(84), 1–8. <https://doi.org/10.3389/fsoc.2019.00084>
- Korbmacher, J. M. (2014). Recall Error in the Year of Retirement. *SHARE Working Paper Series*, 21–2014, 1–41.

- Kreyenfeld, M. (2010). Uncertainties in Female Employment Careers and the Postponement of Parenthood in Germany. *European Sociological Review*, 26(3), 351–366. <https://doi.org/10.1093/esr/jcp026>
- Kreyenfeld, M. (2015). Maternal and Paternal Employment across the Life Course. In R. A. Scott & S. M. Kosslyn (Hrsg.), *Emerging Trends in the Social and Behavioral Sciences* (1. Aufl., S. 1–15). Wiley. <https://doi.org/10.1002/9781118900772.etrds0210>
- Kreyenfeld, M., & Bastin, S. (2016). Reliability of union histories in social science surveys: Blurred memory, deliberate misreporting, or true tales? *Advances in Life Course Research*, 27, 30–42. <https://doi.org/10.1016/j.alcr.2015.11.001>
- Kreyenfeld, M., & Geisler, E. (2006). Müttererwerbstätigkeit in Ost- und Westdeutschland: Mothers' Employment in East and West Germany. *Journal of Family Research*, 18(3), 333–360.
- Kreyenfeld, M., Mika, T., & Radenacker, A. (2018). Der Gender Pension Gap in Ost- und Westdeutschland: Welchen Einfluss hat eine Scheidung auf die Alterssicherung? *Sozialer Fortschritt*, 67(11–12), 973–996. <https://doi.org/10.3790/sfo.67.11-12.973>
- Kreyenfeld, M., Schmauk, S., & Mika, T. (2023). The gender pension gap in Germany: Is divorce a gender-equaliser? *Ageing & Society*, 43(11), 2700–2720. <https://doi.org/10.1017/S0144686X21001513>
- Krüger, H., & Levy, R. (2001). Linking Life Courses, Work, and the Family: Theorizing a Not So Visible Nexus between Women and Men. *Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 26(2), 145–166. <https://doi.org/10.2307/3341676>

- Kuitto, K., Kuivalainen, S., & Möhring, K. (2021). Inequalities in pensions and retirement from a life course perspective: An introduction. *Social Policy & Administration*, 55(3), 403–408. <https://doi.org/10.1111/spol.12663>
- Kuitto, K., Madia, J. E., & Podestà, F. (2023). Public Pension Generosity and Old-Age Poverty in OECD countries. *Journal of Social Policy*, 52(2), 256–275. <https://doi.org/10.1017/S0047279421000544>
- Kuivalainen, S., Nivalainen, S., Järnefelt, N., & Kuitto, K. (2020). Length of working life and pension income: Empirical evidence on gender and socioeconomic differences from Finland. *Journal of Pension Economics & Finance*, 19(1), 126–146. <https://doi.org/10.1017/S1474747218000215>
- Kunze, A. (2008). Gender wage gap studies: Consistency and decomposition. *Empirical Economics*, 35(1), 63–76. <https://doi.org/10.1007/s00181-007-0143-4>
- Kunze, A. (2018). The Gender Wage Gap in Developed Countries. In S. L. Averett, L. M. Argys, & S. D. Hoffman (Hrsg.), *The Oxford Handbook of Women and the Economy* (S. 369–394). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190628963.013.11>
- Kursa, M. B., & Rudnicki, W. R. (2010). Feature Selection with the Boruta Package. *Journal of Statistical Software*, 36(11). <https://doi.org/10.18637/jss.v036.i11>
- Lee, K. O., Smith, R., & Galster, G. (2017). Subsidized Housing and Residential Trajectories: An Application of Matched Sequence Analysis. *Housing Policy Debate*, 27(6), 843–874. <https://doi.org/10.1080/10511482.2017.1316757>
- Leitner, S. (2001). Sex and Gender Discrimination Within Eu Pension Systems. *Journal of European Social Policy*, 11(2), 99–115. <https://doi.org/10.1177/095892870101100201>

- Levine, P., Mitchell, O., & Phillips, J. (1999). *Worklife Determinants of Retirement Income Differentials Between Men and Women*. National Bureau of Economic Research. <https://doi.org/10.3386/w7243>
- Lewis, J. (1992). Gender and the development of welfare regimes. *Journal of European social policy*, 2(3), 159–173.
- Liao, T. F., Bolano, D., Brzinsky-Fay, C., Cornwell, B., Fasang, A. E., Helske, S., Piccarreta, R., Raab, M., Ritschard, G., Struffolino, E., & Studer, M. (2022). Sequence analysis: Its past, present, and future. *Social Science Research*, 107, 102772. <https://doi.org/10.1016/j.ssresearch.2022.102772>
- Lin, I.-F., & Brown, S. L. (2021). The Economic Consequences of Gray Divorce for Women and Men. *The Journals of Gerontology: Series B*, 76(10), 2073–2085. <https://doi.org/10.1093/geronb/gbaa157>
- Lis, M., & Bonthuis, B. (2019). Drivers of the gender gap in pensions: Evidence from EU-SILC and the OECD pension model. In R. Holzmann, E. Palmer, R. Palacios, & S. Sacchi (Hrsg.), *Progress and Challenges of Nonfinancial Defined Contribution Pension Schemes* (Bd. 2, S. 25–53). International Bank for Reconstruction and Development / The World Bank.
- Lundberg, I. (2022). The Gap-Closing Estimand: A Causal Approach to Study Interventions That Close Disparities Across Social Categories. *Sociological Methods & Research*, 00491241211055769. <https://doi.org/10.1177/00491241211055769>
- MacIndoe, H., & Abbott, A. (2004). *Sequence analysis and optimal matching techniques for social science data*. Thousand Oaks, CA: Sage.

- Madero-Cabib, I., & Fasang, A. E. (2016). Gendered work–family life courses and financial well-being in retirement. *Advances in Life Course Research*, 27, 43–60. <https://doi.org/10.1016/j.alcr.2015.11.003>
- Madero-Cabib, I., Le Feuvre, N., & König, S. (2021). Gendered retirement pathways across lifecourse regimes. *Ageing and Society*, 1–30. <https://doi.org/10.1017/S0144686X21001781>
- Markus Gangl & Andrea Ziefle. (2009). Motherhood, Labor Force Behavior, and Women’s Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States. *Demography*, 46(2), 341–369. <https://doi.org/10.1353/dem.0.0056>
- Meyer, T., & Pfau-Effinger, B. (2006). Gender Arrangements and Pension Systems in Britain and Germany: Tracing change over five decades. *International Journal of Ageing and Later Life*, 1(2), 67–110. <https://doi.org/10.3384/ijal.1652-8670.061267>
- Mika, T., & Czaplicki, C. (2017). Fertility and Women’s Old-Age Income in Germany. In M. Kreyenfeld & D. Konietzka (Hrsg.), *Childlessness in Europe: Contexts, Causes, and Consequences* (S. 331–349). Springer International Publishing. https://doi.org/10.1007/978-3-319-44667-7_16
- Mincer, J., & Polachek, S. (1974). Family investments in human capital: Earnings of women. *Journal of political Economy*, 82(2, Part 2), S76–S108.
- Möhring, K. (2015). Employment Histories and Pension Incomes in Europe: A multilevel analysis of the role of institutional factors. *European Societies*, 17(1), 3–26. <https://doi.org/10.1080/14616696.2014.934874>

- Möhring, K. (2018). Is there a motherhood penalty in retirement income in Europe? The role of lifecourse and institutional characteristics. *Ageing and Society*, 38(12), 2560–2589. <https://doi.org/10.1017/S0144686X17000812>
- Möhring, K. (2021). The consequences of non-standard working and marital biographies for old age income in Europe: Contrasting the individual and the household perspective. *Social Policy & Administration*, 55(3), 456–484. <https://doi.org/10.1111/spol.12720>
- Möhring, K., & Weiland, A. P. (2022). Couples' Life Courses and Women's Income in Later Life: A Multichannel Sequence Analysis of Linked Lives in Germany. *European Sociological Review*, 38(3), 371–388. <https://doi.org/10.1093/esr/jcab048>
- Mousteri, V., Daly, M., & Delaney, L. (2018). The scarring effect of unemployment on psychological well-being across Europe. *Social Science Research*, 72, 146–169. <https://doi.org/10.1016/j.ssresearch.2018.01.007>
- Muller, J. S., Hiekel, N., & Liefbroer, A. C. (2020). The Long-Term Costs of Family Trajectories: Women's Later-Life Employment and Earnings Across Europe. *Demography*, 57(3), 1007–1034. <https://doi.org/10.1007/s13524-020-00874-8>
- Müller, K.-U., & Wrohlich, K. (2020). Does subsidized care for toddlers increase maternal labor supply? Evidence from a large-scale expansion of early childcare. *Labour Economics*, 62, 101776. <https://doi.org/10.1016/j.labeco.2019.101776>
- Müller, N. S., Gabadinho, A., Ritschard, G., & Studer, M. (2008). Extracting Knowledge from Life Courses: Clustering and Visualization. In I.-Y. Song, J. Eder, & T. M. Nguyen (Hrsg.), *Data Warehousing and Knowledge Discovery* (S. 176–185). Springer. https://doi.org/10.1007/978-3-540-85836-2_17

- Musick, K., Bea, M. D., & Gonalons-Pons, P. (2020). His and Her Earnings Following Parenthood in the United States, Germany, and the United Kingdom. *American Sociological Review*, 85(4), 639–674. <https://doi.org/10.1177/0003122420934430>
- Netherlands Interdisciplinary Demographic Institute (distributor). (2020). *Generations and Gender Programme*. Generations and Gender Contextual Database. <https://www.ggp-i.org/data/ggp-contextual-database/>
- Neumark, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human resources*, 23(3), 279–295.
- Neyer, G. R. (2003). *Family policies and low fertility in Western Europe* (WP-2003-021; 0 Aufl., S. WP-2003-021). Max Planck Institute for Demographic Research. <https://doi.org/10.4054/MPIDR-WP-2003-021>
- Nieuwenhuis, R., Van Lancker, W., Collado, D., & Cantillon, B. (2020). Trends in Women's Employment and Poverty Rates in OECD Countries: A Kitagawa–Blinder–Oaxaca Decomposition. *Italian Economic Journal*, 6(1), 37–61. <https://doi.org/10.1007/s40797-019-00115-x>
- Nolan, A., Whelan, A., McGuinness, S., & Maitre, B. (2019). *Gender, Pensions and Income in retirement*. (Research series number RS87). Economic and Social Research Institute (ESRI).
- Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The review of economics and statistics*, 90(2), 290–299.
- Nutz, T., & Lersch, P. M. (2021). Gendered employment trajectories and individual wealth at older ages in Eastern and Western Germany. *Advances in Life Course Research*, 47, 100374. <https://doi.org/10.1016/j.alcr.2020.100374>
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*, 14(3), 693–709.

- Oaxaca, R. L., & Sierminska, E. (2023). *Oaxaca-Blinder Meets Kitagawa: What Is the Link?*
- OECD. (2013). *Pensions at a Glance 2013: OECD and G20 Indicators*. OECD Publishing.
https://doi.org/10.1787/pension_glance-2013-en
- OECD. (2015). *Pensions at a Glance 2015: OECD and G20 indicators*. OECD.
https://doi.org/10.1787/pension_glance-2015-en
- OECD. (2019). *Pensions at a Glance 2019: OECD and G20 Indicators*. OECD Publishing.
<https://doi.org/10.1787/b6d3dcfc-en>
- OECD. (2020). *Employment: Time spent in paid and unpaid work, by sex*.
<https://stats.oecd.org/index.aspx?queryid=54757>
- OECD. (2021). *Towards Improved Retirement Savings Outcomes for Women*.
https://www.oecd-ilibrary.org/finance-and-investment/towards-improved-retirement-savings-outcomes-for-women_f7b48808-en
- O’Rand, A. M. (1996). The Precious and the Precocious: Understanding Cumulative Disadvantage and Cumulative Advantage Over the Life Course. *The Gerontologist*, 36(2), 230–238. <https://doi.org/10.1093/geront/36.2.230>
- Orloff, A. S. (1993). Gender and the Social Rights of Citizenship: The Comparative Analysis of Gender Relations and Welfare States. *American Sociological Review*, 58(3), 303–328. <https://doi.org/10.2307/2095903>
- Overweg, C. (2023). *Female Work-family life courses in the Dutch and West German pension systems* [Master thesis]. Humboldt university.
- Palomäki, L.-M. (2017). Reference Groups and Pensioners’ Subjective Economic Well-Being in Europe. *Social Indicators Research*, 131(2), 509–525.
<https://doi.org/10.1007/s11205-016-1262-0>

- Pedersen, A. W. (2004). The Privatization of Retirement Income? Variation and Trends in the Income Packages of Old Age Pensioners. *Journal of European Social Policy*, 14(1), 5–23. <https://doi.org/10.1177/0958928704039782>
- Pelletier, D., Bignami-Van Assche, S., & Simard-Gendron, A. (2020). Measuring Life Course Complexity with Dynamic Sequence Analysis. *Social Indicators Research*, 152(3), 1127–1151. <https://doi.org/10.1007/s11205-020-02464-y>
- Pfau-Effinger, B. (1998). Gender cultures and the gender arrangement—A theoretical framework for cross-national gender research. *Innovation: The European Journal of Social Science Research*, 11(2), 147–166.
- Pfau-Effinger, B. (2005). WELFARE STATE POLICIES AND THE DEVELOPMENT OF CARE ARRANGEMENTS. *European Societies*, 7(2), 321–347. <https://doi.org/10.1080/14616690500083592>
- Piccarreta, R. (2017). Joint Sequence Analysis: Association and Clustering. *Sociological Methods & Research*, 46(2), 252–287. <https://doi.org/10.1177/0049124115591013>
- Piccarreta, R., & Struffolino, E. (2023). Identifying and Qualifying Deviant Cases in Clusters of Sequences: The Why and The How. *European Journal of Population*, 40(1), 1. <https://doi.org/10.1007/s10680-023-09682-3>
- Pollock, G. (2007). Holistic trajectories: A study of combined employment, housing and family careers by using multiple-sequence analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(1), 167–183. <https://doi.org/10.1111/j.1467-985X.2006.00450.x>
- Ponomarenko, V. (2016). Cumulative disadvantages of non-employment and non-standard work for career patterns and subjective well-being in retirement. *Advances in Life Course Research*, 30, 133–148. <https://doi.org/10.1016/j.alcr.2016.06.003>

- Ponomarenko, V., Leist, A. K., & Chauvel, L. (2019). Increases in wellbeing in the transition to retirement for the unemployed: Catching up with formerly employed persons. *Ageing & Society*, 39(2), 254–276. <https://doi.org/10.1017/S0144686X17000976>
- Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), e1301. <https://doi.org/10.1002/widm.1301>
- Quadagno, J. (1988). Women's Access to Pensions and the Structure of Eligibility Rules: Systems of Production and Reproduction. *The Sociological Quarterly*, 29(4), 541–558. <https://doi.org/10.1111/j.1533-8525.1988.tb01433.x>
- Raab, M. (2022). *ggseqplot: Render Sequence Plots using „ggplot2“*. <https://maraab23.github.io/ggseqplot/>
- Raab, M., & Struffolino, E. (2020). The Heterogeneity of Partnership Trajectories to Childlessness in Germany. *European Journal of Population*, 36(1), 53–70. <https://doi.org/10.1007/s10680-019-09519-y>
- Raab, M., & Struffolino, E. (2022). *Sequence Analysis* (Bd. 190). SAGE Publications, Inc. <https://us.sagepub.com/en-us/nam/sequence-analysis/book272086>
- Radl, J. (2013). Labour Market Exit and Social Stratification in Western Europe: The Effects of Social Class and Gender on the Timing of Retirement. *European Sociological Review*, 29(3), 654–668. <https://doi.org/10.1093/esr/jcs045>
- Rahimi, E., & Hashemi Nazari, S. S. (2021). A detailed explanation and graphical representation of the Blinder-Oaxaca decomposition method with its application in health inequalities. *Emerging Themes in Epidemiology*, 18(1), 12. <https://doi.org/10.1186/s12982-021-00100-9>

- Reimers, C. W. (1983). Labor Market Discrimination Against Hispanic and Black Men. *The Review of Economics and Statistics*, 65(4), 570–579.
<https://doi.org/10.2307/1935925>
- Riekhoff, A.-J., & Järnefelt, N. (2017). Gender Differences in Retirement in a Welfare State with High Female Labour Market Participation and Competing Exit Pathways. *European Sociological Review*, 33(6), 791–807.
<https://doi.org/10.1093/esr/jcx077>
- Riekhoff, A.-J., & Järnefelt, N. (2018). Retirement Trajectories and Income Redistribution Through the Pension System in Finland. *Social Forces*, 97(1), 27–54.
<https://doi.org/10.1093/sf/soy028>
- Riekhoff, A.-J., Kuitto, K., Möhring, K., & Palomäki, L.-M. (2023). *A long arm of adversity? Financial hardship during the life course and income in retirement.*
- Riekhoff, A.-J., & Vaalavuo, M. (2021). Health shocks and couples' labor market participation: A turning point or stuck in the trajectory? *Social Science & Medicine*, 276, 113843. <https://doi.org/10.1016/j.socscimed.2021.113843>
- Rios-Avila, F. (2020). Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal*, 20(1), 51–94.
<https://doi.org/10.1177/1536867X20909690>
- Ritschard, G. (2021). Measuring the Nature of Individual Sequences. *Sociological Methods & Research*, 00491241211036156. <https://doi.org/10.1177/00491241211036156>
- Ritschard, G., Liao, T. F., & Struffolino, E. (2023). Strategies for Multidomain Sequence Analysis in Social Research. *Sociological Methodology*, 53(2), 288–322.
<https://doi.org/10.1177/00811750231163833>

- Rosenfeld, R. A., Trappe, H., & Gornick, J. C. (2004). Gender and Work in Germany: Before and After Reunification. *Annual Review of Sociology*, 30(1), 103–124.
<https://doi.org/10.1146/annurev.soc.30.012703.110531>
- Rowold, C. (2022). Der Zusammenhang von Familien- und Arbeitsbiografien und Rentensystemen mit dem Gender-Pension-Gap – geschlechtsspezifische Rentenlücken in Europa. *Deutsche Rentenversicherung*, 2, 131–161.
- Rowold, C. (2023a). Differences in Gender Pension Gaps in public and private pensions in West Germany: Which role do work-family life courses play? *Forthcoming as working paper*.
- Rowold, C. (2023b, Januar 25). *Missing data in the SHARE Job Episodes Panel: A potential cause and solution*. Blogpost.
https://crowold.github.io/posts/2023/01/blogpost_jep/
- Rowold, C., Struffolino, E., & Fasang, A. (2023). *Life-course-sensitive analysis of group inequalities: Combining Sequence Analysis with the Kitagawa-Oaxaca-Blinder decomposition*. SocArXiv. <https://doi.org/10.31235/osf.io/7k4vt>
- Rowold, C., & Van Winkle, Z. (2023). Cohort Change in Family Life Course Complexity of Adults and Children. In R. Schoen (Hrsg.), *The Demography of Transforming Families* (S. 203–225). Springer International Publishing.
https://doi.org/10.1007/978-3-031-29666-6_10
- Rutten, A., van Vuuren, D., & Knoef, M. (2023). The effect of the statutory retirement age on spousal labor force participation. *Netspar*.
<https://www.netspar.nl/en/publication/the-effect-of-the-statutory-retirement-age-on-spousal-labor-force-participation/>
- Samek Lodovici, M., Crepaldi, C., Corsi, M., & Naaf, S. (2011). *The socio-economic impact of pension systems on the respective situations of women and men and the*

- effects of recent trends in pension reforms*. Report prepared for European Commission.
- Samek Lodovici, M., Druifuca, S., Patrizio, M., & Pesce, F. (2016). *The Gender Pension Gap: Differences Between Mothers and Women Without Children: Study*. Policy Department C: Citizen's Rights and Constitutional Affairs, European Parliament.
- Saraceno, C. (1994). The Ambivalent Familism of the Italian Welfare State. *Social Politics*, 1(1), 60–82. <https://doi.org/10.1093/sp/1.1.60>
- Schimmele, C. M., & Wu, Z. (2016). Repartnering After Union Dissolution in Later Life. *Journal of Marriage and Family*, 78(4), 1013–1031. <https://doi.org/10.1111/jomf.12315>
- Schmauk, S. (2023). *Divorce and economic wellbeing. A life course perspective on the interplay of work-family lives in West Germany and Sweden* [PhD Thesis]. Hertie School.
- Schmieder, J., & Wrohlich, K. (2021). Gender pay gap in a European comparison: Positive correlation between the female labor force participation rate and the gender pay gap. *DIW Weekly Report*, 11(9), 65–70. https://doi.org/10.18723/diw_dwr:2021-9-1
- Schmitt, L., & Auspurg, K. (2022). A Stall Only on the Surface? Working Hours and the Persistence of the Gender Wage Gap in Western Germany 1985–2014. *European Sociological Review*, jcac001. <https://doi.org/10.1093/esr/jcac001>
- Schmitz-Kießler, J. (2020). Die Rente als Bilanz des Lebens – Alterssicherung und Arbeit von Frauen auf dem Prüfstand. In *Neustart in der Rentenpolitik* (S. 119–138). Nomos Verlagsgesellschaft mbH & Co. KG. <https://doi.org/10.5771/9783748906889-119>

- Schneider, F. G. (2012). The Shadow Economy and Work in the Shadow: What Do We (Not) Know? *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2031951>
- Sefton, T., Evandrou, M., & Falkingham, J. (2011). Family Ties: Women's Work and Family Histories and their Association with Incomes in Later Life in the UK. *Journal of Social Policy*, 40(1), 41–69. <https://doi.org/10.1017/S0047279410000255>
- Sefton, T., Evandrou, M., Falkingham, J., & Vlachantoni, A. (2011). The relationship between women's work histories and incomes in later life in the UK, US and West Germany. *Journal of European Social Policy*, 21(1), 20–36. <https://doi.org/10.1177/0958928710380475>
- Smeeding, T. M., & Sandstrom, S. (2005). *Poverty and income maintenance in old age: A cross-national view of low income older women* (Working Paper 398). LIS Working Paper Series. <https://www.econstor.eu/handle/10419/95568>
- Sohier, L., Van Ootegem, L., & Verhofstadt, E. (2021). Well-Being During the Transition from Work to Retirement. *Journal of Happiness Studies*, 22(1), 263–286. <https://doi.org/10.1007/s10902-020-00228-6>
- Speiser, J. L., Miller, M. E., Tooze, J., & Ip, E. (2019). A comparison of random forest variable selection methods for classification prediction modeling. *Expert Systems with Applications*, 134, 93–101. <https://doi.org/10.1016/j.eswa.2019.05.028>
- Spiess, C. K., & Schneider, A. U. (2003). Interactions between care-giving and paid work hours among European midlife women, 1994 to 1996. *Ageing and Society*, 23(1), 41–68. <https://doi.org/10.1017/S0144686X02001010>
- Sprengholz, M., & Hamjediers, M. (2022). Intersections and Commonalities: Using Matching to Decompose Wage Gaps by Gender and Nativity in Germany. *Work and Occupations*, 0(0), online first. <https://doi.org/10.1177/07308884221141100>

- Sprengholz, M., & Hamjediers, M. (2023). *nopo. Stata-Module for Nopo Decompositions*, version 0.1 [Stata].
https://github.com/mhamjediers/nopo_decomposition/blob/main/README.md
- Sprengholz, M., Wieber, A., & Holst, E. (2022). Gender identity and wives' labor market outcomes in West and East Germany between 1983 and 2016. *Socio-Economic Review*, 20(1), 257–279. <https://doi.org/10.1093/ser/mwaa048>
- Statistisches Bundesamt (Destatis). (2023). *Armutsgefährdungsschwelle und Armutsgefährdung (monetäre Armut) in Deutschland*. Statistisches Bundesamt. <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Lebensbedingungen-Armutsgefaehrdung/Tabellen/armutsschwelle-gefaehrdung-silc.html>
- Stegmann, M. (2007). Effekte der Kindererziehung auf die Erwerbsprofile und Alterseinkommen von Frauen in West- und Ostdeutschland. *WSI-Mitteilungen*, 60(2), 86–93. <https://doi.org/10.5771/0342-300X-2007-2-86>
- Strittmatter, A., & Wunsch, C. (2021). The Gender Pay Gap Revisited with Big Data: Do Methodological Choices Matter? *WWZ Working Paper 2021/05*, 41.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8, 25. <https://doi.org/10.1186/1471-2105-8-25>
- Studer, M. (2013). *WeightedCluster Library Manual: A practical guide to creating typologies of trajectories in the social sciences with R*. LIVES. <https://doi.org/10.12682/LIVES.2296-1658.2013.24>
- Studer, M. (2018). Divisive property-based and fuzzy clustering for sequence analysis. In *Sequence analysis and related approaches* (S. 223–239). Springer, Cham.

- Studer, M. (2021). Validating Sequence Analysis Typologies Using Parametric Bootstrap. *Sociological Methodology*, 51(2), 290–318.
<https://doi.org/10.1177/00811750211014232>
- Studer, M., & Ritschard, G. (2016). What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(2), 481–511.
<https://doi.org/10.1111/rssa.12125>
- Sullivan, O. (2021). The gender division of housework and child care. *Research Handbook on the Sociology of the Family*.
<https://www.elgaronline.com/view/edcoll/9781788975537/9781788975537.00033.xml>
- Tambellini, E. (2023). Exploring the relationship between working history, retirement transition and women's life satisfaction. *Ageing & Society*, 43(8), 1754–1783.
<https://doi.org/10.1017/S0144686X2100132X>
- Tambellini, E., Danielsbacka, M., & Rotkirch, A. (2023). Do Working and Parenting Trajectories Influence Retirement Timing? Evidence From Spain, Using a Sequence Analysis Approach and Focusing on Women. *Work, Aging and Retirement*, waad018. <https://doi.org/10.1093/workar/waad018>
- Tinios, P., Bettio, F., Betti, G., & Georgiadis, T. (2015). *Men, women and pensions*. Publications Office of the European Union Luxembourg, UK.
- Toloşi, L., & Lengauer, T. (2011). Classification with correlated features: Unreliability of feature ranking and solutions. *Bioinformatics*, 27(14), 1986–1994.
<https://doi.org/10.1093/bioinformatics/btr300>

- Tophoven, S., & Tisch, A. (2016). Employment trajectories of German baby boomers and their effect on statutory pension entitlements. *Advances in Life Course Research*, 30, 90–110. <https://doi.org/10.1016/j.alcr.2016.04.003>
- Tosi, M., & Grundy, E. (2021). Work–family lifecourses and later-life health in the United Kingdom. *Ageing & Society*, 41(6), 1371–1397. <https://doi.org/10.1017/S0144686X19001752>
- Trappe, H., Pollmann-Schult, M., & Schmitt, C. (2015). The Rise and Decline of the Male Breadwinner Model: Institutional Underpinnings and Future Expectations. *European Sociological Review*, 31(2), 230–242. <https://doi.org/10.1093/esr/jcv015>
- Tyros, S., van Ewijk, C., & Stevens, R. (2023). *Shocks to future income and (lack of) consumption-smoothing: Evidence from the abolition of the state pension partner supplement in the Netherlands*.
- Uccheddu, D., Emery, T., Gauthier, A. H., & Steverink, N. (2022). Gendered work-family life courses and late-life physical functioning: A comparative analysis from 28 European countries. *Advances in Life Course Research*, 53, 1–20. <https://doi.org/10.1016/j.alcr.2022.100495>
- van Hedel, K., Mejía-Guevara, I., Avendaño, M., Sabbath, E. L., Berkman, L. F., Mackenbach, J. P., & van Lenthe, F. J. (2016). Work-Family Trajectories and the Higher Cardiovascular Risk of American Women Relative to Women in 13 European Countries. *American Journal of Public Health*, 106(8), 1449–1456. <https://doi.org/10.2105/AJPH.2016.303264>
- Van Winkle, Z. (2018). Family Trajectories Across Time and Space: Increasing Complexity in Family Life Courses in Europe? *Demography*, 55(1), 135–164. <https://doi.org/10.1007/s13524-017-0628-5>

- Van Winkle, Z. (2019). Family policies and family life course complexity across 20th-century Europe. *Journal of European Social Policy*, 30(3), 320–338.
- Van Winkle, Z. (2020). Early Family Life Course Standardization in Sweden: The Role of Compositional Change. *European Journal of Population*, 36(4), 765–798. <https://doi.org/10.1007/s10680-019-09551-y>
- Van Winkle, Z., & Fasang, A. (2017). Complexity in Employment Life Courses in Europe in the Twentieth Century—Large Cross-National Differences but Little Change across Birth Cohorts. *Social Forces*, 96(1), 1–30. <https://doi.org/10.1093/sf/sox032>
- Van Winkle, Z., & Fasang, A. (2021). The complexity of employment and family life courses across 20th century Europe: More evidence for larger cross-national differences but little change across 1916–1966 birth cohorts. *Demographic Research*, 44, 775–810. <https://doi.org/10.4054/DemRes.2021.44.32>
- Van Winkle, Z., & Fasang, A. E. (2020). Parenthood Wage Gaps Across the Life Course: A Comparison by Gender and Race. *Journal of Marriage and Family*, 82(5), 1515–1533. <https://doi.org/10.1111/jomf.12713>
- Vecgaile, L. (2023). *Changing Work-family Life Courses and the Future Gender Pension Gap in Germany: A Machine Learning Approach*.
- Veremchuk, A. (2020). *Gender Gap in Pension Income: Cross-Country Analysis and Role of Gender Attitudes* [SSRN Scholarly Paper]. Social Science Research Network. <https://doi.org/10.2139/ssrn.3662968>
- Vlachantoni, A. (2012). Financial inequality and gender in older people. *Maturitas*, 72(2), 104–107. <https://doi.org/10.1016/j.maturitas.2012.02.015>
- Waitkus, N., & Minkus, L. (2021). Investigating the Gender Wealth Gap Across Occupational Classes. *Feminist Economics*, 27(4), 114–147. <https://doi.org/10.1080/13545701.2021.1973059>

- Wiß, T. (2020). *Lebensstandardsicherung durch betriebliche Altersvorsorge? Deutschland im Vergleich*. 207–234. <https://doi.org/10.5771/9783748906889-207>
- Yabiku, S. T. (2000). Family history and pensions: The relationships between marriage, divorce, children, and private pension coverage. *Journal of Aging Studies*, 14(3), 293–312. [https://doi.org/10.1016/S0890-4065\(00\)08023-3](https://doi.org/10.1016/S0890-4065(00)08023-3)
- Yu, A., & Elwert, F. (2023). *Nonparametric Causal Decomposition of Group Disparities* (arXiv:2306.16591). arXiv. <http://arxiv.org/abs/2306.16591>
- Zagel, H., & Lohmann, H. (2020). Conceptual Approaches in Comparative Family Policy Research. In R. Nieuwenhuis & W. Van Lancker (Hrsg.), *The Palgrave Handbook of Family Policy* (S. 119–139). Springer International Publishing. https://doi.org/10.1007/978-3-030-54618-2_6
- Zagel, H., & Van Winkle, Z. (2022). Women’s Family and Employment Life Courses Across Twentieth-Century Europe: The Role of Policies and Norms. *Social Politics: International Studies in Gender, State & Society*, 29(2), 446–476. <https://doi.org/10.1093/sp/jxz056>
- Zhao, R., & Zhao, Y. (2018). The Gender Pension Gap in China. *Feminist Economics*, 24(2), 218–239. <https://doi.org/10.1080/13545701.2017.1411601>
- Ziefle, A., & Gangl, M. (2014). Do Women Respond to Changes in Family Policy? A Quasi-Experimental Study of the Duration of Mothers’ Employment Interruptions in Germany. *European Sociological Review*, 30(5), 562–581. <https://doi.org/10.1093/esr/jcu056>
- Zoch, G., & Heyne, S. (2023). The evolution of family policies and couples’ housework division after childbirth in Germany, 1994–2019. *Journal of Marriage and Family*, n/a(n/a). <https://doi.org/10.1111/jomf.12938>

8 Data, software, and funding

8.1 Data

This paper uses data from SHARE Waves 2, 3, 4, 5, 6, 7 (DOIs: [10.6103/SHARE.w2.710](https://doi.org/10.6103/SHARE.w2.710), <http://dx.doi.org/10.6103/SHARE.w3.710>, [10.6103/SHARE.w4.710](https://doi.org/10.6103/SHARE.w4.710), [10.6103/SHARE.w5.710](https://doi.org/10.6103/SHARE.w5.710), [10.6103/SHARE.w6.710](https://doi.org/10.6103/SHARE.w6.710), [10.6103/SHARE.w7.711](https://doi.org/10.6103/SHARE.w7.711)). The SHARE data collection has been funded by the European Commission, DG RTD through FP5 (QLK6-CT-2001–00360), FP6 (SHARE-I3: RII-CT-2006-062193, COMPARE: CIT5-CT-2005-028857, SHARELIFE: CIT4-CT-2006-028812), FP7 (SHARE-PREP: GA N°211909, SHARE-LEAP: GA N°227822, SHARE M4: GA N°261982, DASISH: GA N°283646) and Horizon 2020 (SHARE-DEV3: GA N°676536, SHARE-COHESION: GA N°870628, SERISS: GA N°654221, SSHOC: GA N°823782, SHARE-COVID19: GA N°101015924) and by DG Employment, Social Affairs & Inclusion through VS 2015/0195, VS 2016/0135, VS 2018/0285, VS 2019/0332, and VS 2020/0313. Additional funding from the German Ministry of Education and Research, the Max Planck Society for the Advancement of Science, the U.S. National Institute on Aging (U01_AG09740-13S2, P01_AG005842, P01_AG08291, P30_AG12815, R21_AG025169, Y1-AG-4553-01, IAG_BSR06-11, OGHA_04-064, HHSN271201300071C, RAG052527A) and from various national funding sources is gratefully acknowledged (see www.share-project.org).

I also apply data from the generated Job Episodes Panel release 7.1.0, which is based on SHARE Waves 3 and 7 (DOIs: [10.6103/SHARE.w3.710](https://doi.org/10.6103/SHARE.w3.710), [10.6103/SHARE.w7.710](https://doi.org/10.6103/SHARE.w7.710)), see Brugiavini et al. (2019) for methodological details. The Job Episodes Panel release 8.0.0 is based on SHARE Waves 3 and 7 (DOIs: [10.6103/SHARE.w3.800](https://doi.org/10.6103/SHARE.w3.800), [10.6103/SHARE.w7.710](https://doi.org/10.6103/SHARE.w7.710)).

I further used administrative data from SHARE-RV for background information on how to deal with the remaining missing information on retrospective employment data. Forschungsdatenzentrum der Rentenversicherung, Max-Planck-Institut für Sozialrecht und Sozialpolitik (2020). SHARE-RV. Release version: 7.1.0. SHARE-ERIC. Dataset. DOI: 10.6103/SHARE.SHARE-RV.710.

For details see Börsch-Supan et al. (2013), Börsch-Supan et al. (2018) and Börsch-Supan et al. (2020).

Börsch-Supan, A., C. Czaplicki, S. Friedel, I. Herold, J. Korbmacher, T. Mika (2020). *SHARE-RV: Linked Data to Study Aging in Germany*. In: Winker, P., T. v. Büttner, R. Riphahn, W. Smolny, J. Wagner (eds.): *Jahrbücher für Nationalökonomie und Statistik*, 240(1), 121–132. Berlin: De Gruyter. DOI: 10.1515/jbnst-2018-0034.

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2*. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w2.710](https://doi.org/10.6103/SHARE.w2.710)

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 3 – SHARELIFE*. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w3.710](https://doi.org/10.6103/SHARE.w3.710)

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4*. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w4.710](https://doi.org/10.6103/SHARE.w4.710)

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5*. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w5.710](https://doi.org/10.6103/SHARE.w5.710)

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6*. Release version: 7.1.0. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w6.710](https://doi.org/10.6103/SHARE.w6.710)

Börsch-Supan, A. (2020). *Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7*. Release version: 7.1.1. SHARE-ERIC. Data set. DOI: [10.6103/SHARE.w7.711](https://doi.org/10.6103/SHARE.w7.711)

Forschungsdatenzentrum der Rentenversicherung, Max-Planck-Institut für Sozialrecht und Sozialpolitik (2020). SHARE-RV. Release version: 7.1.0. SHARE-ERIC. Dataset. DOI: 10.6103/SHARE.SHARE-RV.710

8.2 Software and packages

The analyses were performed in stata and R.

The packages used for the decomposition analyses were *oaxaca* (Jann, 2008), *oaxaca rif* (Rios-Avila, 2020) and *nopo* (Sprengholz & Hamjediers, 2023)

The main packages used for the sequence analysis tools were *TraMineR* and *TraMineRextras* (Gabadinho et al., 2011), *WeightedCluster* (Studer, 2013) and *ggseqplot* (Raab, 2022).

8.3 Funding

This research was supported by funding from the Economic and Social Research Council grant ES/P0000649/1 as well as from Nuffield College.

A2 Appendix for Chapter 2

A2.1 Data and sample

I use cross-sectional data from the Survey of Health, Ageing and Retirement in Europe (SHARE, Bergmann et al., 2019; Börsch-Supan et al., 2013). I pool waves 2-6 and use wave 6 as baseline (as the most current, complete wave) and add individuals from waves 5, 4, and 2 if they were not surveyed in wave 6 to maximize the number of observations across cohorts while making use of the most recent data available. For wave 7 the regular panel questionnaire was only surveyed for individuals who were already included in the SHARELIFE survey in wave 3. The fieldwork for wave 8 was suspended due to the COVID pandemic which led to a switch of the interview mode. I thus use wave 6 as the most current data containing the full sample that was surveyed without any disruptions.

The sample contains all respondents aged 65 or older to demonstrate the implications based on a broad definition of the elderly.

A2.2 Survey of self-assessed retirement

The self-assessed employment state is surveyed in question ep005 in the core questionnaire of SHARE as follows: Please look at card 7. In general, which of the following best describes your current employment situation? The response options are 1. Retired, 2. Employed or self-employed (including working for family business), 3. Unemployed, 4. Permanently sick or disabled, 5. Homemaker, and 97. Other. In SHARE additionally, the interviewer is asked to refer to the following if respondents are in doubt about their current employment situation: ‘1. Retired (retired from own work, including semi-retired, partially retired, early retired, pre-retired). Retired refers to retired from own work only. Recipients of survivor pensions who do not receive pensions from their own work should not be coded

as retired. If they do not fit in categories 2 through 5, they should go into other.’⁶⁵ Women are less likely to receive pensions in general (Tinios et al., 2015) and are more likely to be in doubt about their situation in old age in the first place, given the gendered notion of retirement. Thus, the structure of the questions and the instructions for the interviewers are likely to level down the share of women self-reporting themselves as retired and thus exacerbate the gendered sample bias.

A2.3 Retrospective data and sample

Sample. Roughly 40% of the sample have no retrospective data and must be excluded. Table A2.1 shows the number of observations before and after the exclusion by gender and self-assessed retirement status. It reveals that the overall share of those who identify as retirees is almost the same (81.6 vs. 81.8%). Among women, a slightly higher share declare themselves retirees after the exclusion compared to the full sample (74.7 vs. 73.8%) whereas for men the reverse is evident, but to a lower degree (90.8 vs. 91%).

⁶⁵ Other surveys focussing on the old-age population, such as HRS, do not have such an interviewer instruction. In HRS the question is “Are you working now, temporarily laid off, unemployed and looking for work, disabled and unable to work, retired, a homemaker, or what?”.

Table A2.1 Summary table for retirement identification

	Full sample			Sample with retrospective data (after exclusion)		
	Male	Female	Total	Male	Female	Total
Does not identify as retired	1607 (9.0)	5653 (26.2)	7260 (18.4)	969 (9.2)	3374 (25.3)	4343 (18.2)
Identifies as retired	16295 (91.0)	15884 (73.8)	32179 (81.6)	9528 (90.8)	9975 (74.7)	19503 (81.8)
Total	17902 (100.0)	21537 (100.0)	39439 (100.0)	10497 (100.0)	13349 (100.0)	23846 (100.0)

Note: Relative shares in parenthesis. Own calculations based on SHARE release 7.1.0. Not weighted.

Pension receipt is a little less frequent for the matched sample (89.7% compared to 90.5% for the full sample) and the pattern is similar across genders (see Table A2.2).

Overall, the changes in the distribution of self-assessed retirement and pension receipt vary only to a minor degree across samples and do not substantially affect the gendered bias in both measurements.

Table A2.2 Summary table for pension receipt

Pension receipt	Full sample			Sample with retrospective data (after exclusion)		
	Male	Female	Total	Male	Female	Total
No	1050 (5.9)	2711 (12.6)	3761 (9.5)	659 (6.3)	1797 (13.5)	2456 (10.3)
Yes	16852 (94.1)	18826 (87.4)	35678 (90.5)	9838 (93.7)	11552 (86.5)	21390 (89.7)
Total	17902 (100.0)	21537 (100.0)	39439 (100.0)	10497 (100.0)	13349 (100.0)	23846 (100.0)

Note: Relative shares in parenthesis. Own calculations based on SHARE release 7.1.0. Not weighted.

Missing data. The retrospective data for the working trajectories comes from the Job Episodes Panel (JEP) which is a generated dataset based on the raw retrospective SHARELIFE (wave 3 and 7 of SHARE) data (Brugiavini et al., 2013, 2019). I use the *situation* variable which contains the self-reported situation or activity individuals were in for each year, covering multiple employment states but also categories out of the labor force. However, the *situation* variable contains missing information for sizeable shares of

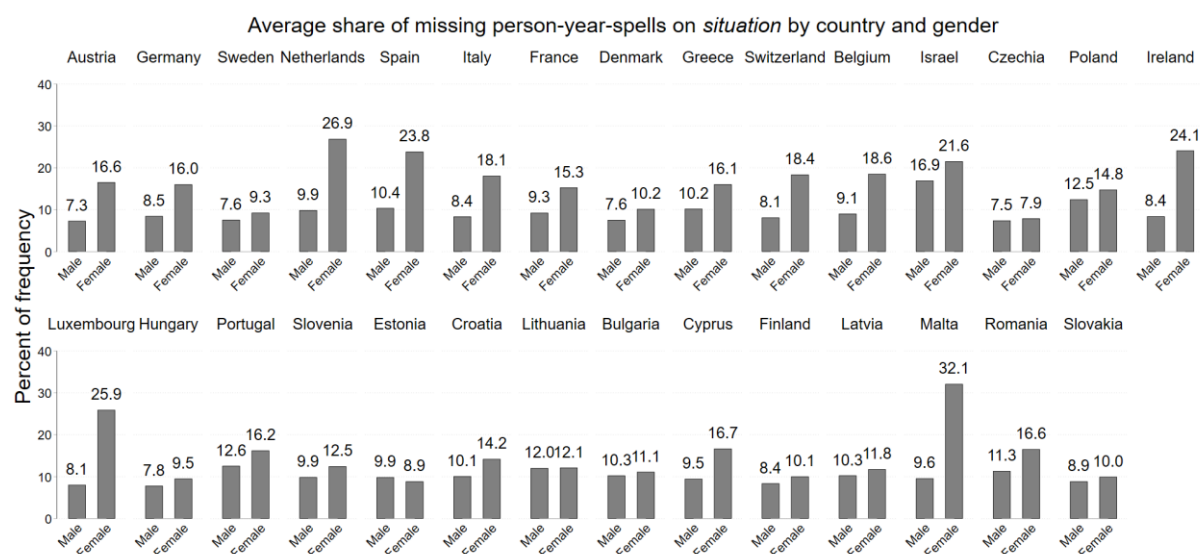
person-year spells. This is problematic particularly for this analysis because the level of missingness depends largely on country and gender (see Chapter 2.3). I filled in the missing data based on additional information from the raw SHARELIFE data that was not included in generating the *situation* variable (for a detailed description see Chapter 2.3).

This still leaves missing information for ages 59-65 of individuals who were not surveyed in wave 7 of SHARELIFE, so that information for the years between the survey of wave 3 and the survey year used for the cross-sectional data (i.e., self-defined employment situation and pension receipt) is missing. I forward-fill these spells based on their last observed work state.

Working categories considered. The work trajectories cover 10 mutually exclusive work states: education or training, full-time care work, part-time employment in the private sector, full-time employment in the private sector, civil service, self-employment, unemployed, other, missing, and retired. I differentiate between full-time and part-time employment only for employment in the private sector. Part-time employment is considered if any part-time employment occurs for that year and includes short-term employment without information on working hours. Care work is based on the response item ‘*Looking after home or family*’. This measure exclusively captures full-time and unpaid care work that was done for at least 6 months a year. Therefore, the actual degree of unpaid care work given per year is rather underestimated. For example, it is most likely that ‘part-time employment’ contains large shares of care work given that women often work part-time to reconcile care duties and paid work. ‘Other’ mostly contains activities that did not fit with any of the suggested options (57.84%) as well as activities such as military service (19.79%), leisure (7.53%), or volunteering (7.12%). Missing contains the remaining person-year spells with no information on the missing data.

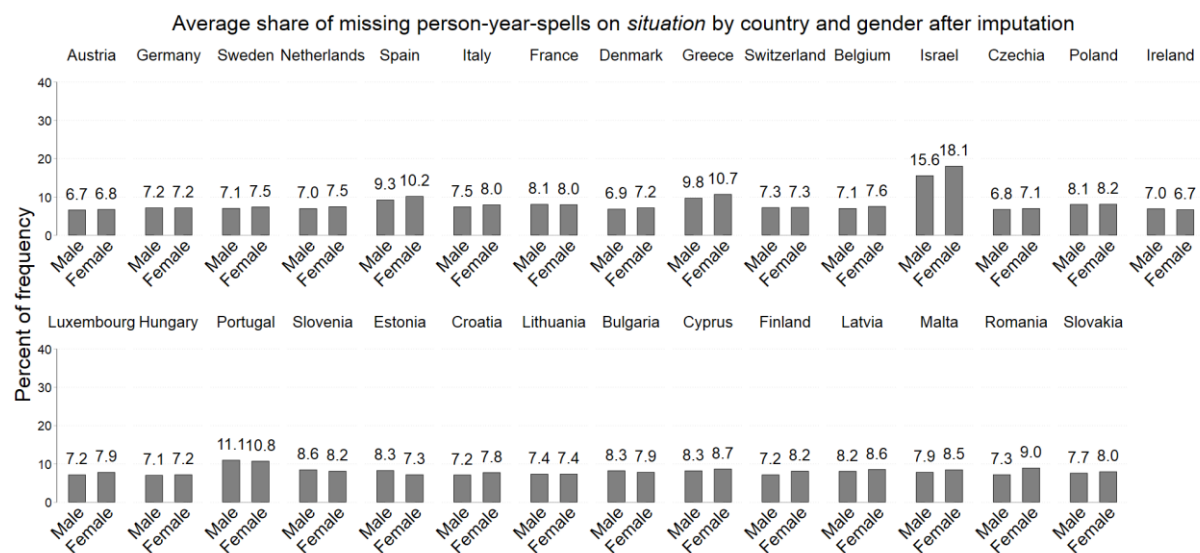
A2.4 Additional Figures

Figure A2.1 Share of missingness for all person-year spells before filling missing information



Notes: Own calculations based on whole sample of JEP release 7.1.0.

Figure A2.2 Share of missingness for all person-year spells after filling missing information



Notes: Own calculations based on whole sample of JEP and SHARELIFE release 7.1.0.

A3 Appendix for Chapter 3

A3.1 Additional Tables

Table A3.1 Norms and family policies in Italy and Germany

		Italy	West Germany	Sources
Institutions	Publicly funded and regulated childcare available for children within 1 year of childbirth since (until 1980)	-	-	Van Winkle 2019
	Publicly funded and regulated childcare available for children by age 3	2003	1996	Van Winkle 2019
	Childcare rate in day care centers /public childcare for children aged 0–2	1997: 6%	1991, 1996: 4%	Del Boca 2002; Kreyenfeld and Geisler 2006
	Childcare rate in day care centers/public childcare for children aged >2	1997: 91% (age 3–6)	1991: 51% (age 3), 75% (age 4–6) 1996: 54% (age 3), 76% (age 4–6)	Del Boca 2002; Kreyenfeld and Geisler 2006
Norms	Traditional family patterns 1981	0.269	0.238	Halman and Vloet 1994
	Rejection of traditional women's roles	-0.595	-0.009	Halman and Vloet 1994
	Equal roles for men and women	0.099	-0.111	Halman and Vloet 1994
	Percentage of people that agree or agree strongly with the statement that a pre-school child is likely to suffer if his or her mother works (1990)***	78%	83%	Atlas of European Values 2020
Outcomes	Female labor force participation rate in 1990	35%	45%	<u>Netherlands</u> <u>Interdisciplinary Demographic Institute (distributor)</u> , 2020.
	Gender wage gap (median) for full-time employees in 1998	7%	21%	Hammerschmid and Rowold 2019a

Table A3.2 Pension systems in Italy and Germany

	Pension characteristics in 2012 ⁶⁶		Italy	West Germany	Source
Public/state pension	Access	Minimum contribution length	20 years ⁶⁷	5 years	OECD 2013
		Age for pension qualification/ min pension age	Pensionable age: 66 for men, 62 for women Av. Effective age of labor-market exit: 61.1 for men, 60.5 for women	Pensionable age: 65 for men, 65 for women Av. Effective age of labor-market exit: 62.1 for men, 61.6 for women	OECD 2013
		Share of 65+ population receiving no independent pension income	12% of men, 35% of women	2% of men, 9% of women	Own calculation based on sample
	Amount	Level of linkage to earnings (Progressivity Index) ⁶⁸	1.4	25.4	OECD 2013
		Maximum amount as % of average earnings	332%	150%	OECD 2013
		Childcare benefits	More generous transformation coefficient for mothers depending on n of children (+1 year if 1–2 children, +2 years if 3–4 children); only for employment interruptions; earnings-related	1 pension point per children born before 1992; employment allowed; standard level (average earnings)	OECD 2013
		Other care benefits	Max. 2 years	Increase of pension by €21.22–65.28 for elderly care	Horstmann et al. 2009; Samek Lodovici et al. 2011
		Divorce benefit	No splitting	Pension rights added up and split for years of marriage equally between partners	Horstmann et al. 2009; Samek Lodovici et al. 2011
		Gender Gap in replacement rate ⁶⁹	-18	6.5	European Institute for Gender Equality 2018; Eurostat 2018

⁶⁶ 2012 is used as the reference year for the pension systems, since the pension income of the majority of sample was surveyed in 2012 (see Table A3.2). No relevant changes took place in either pension system from 2006–2013.

⁶⁷ Since the Prodi reform in 2007 the minimum contribution has been age dependant: age + years must sum up to at least 95 (from 2011 = 96 and from 2013 = 97), which equals 30 years at age 65. However, this means that even at age 80, individuals must have at least 15 years of contribution and thus much more than the 5 years in Germany.

⁶⁸ Measures the relationship between inequality in employment income and pension income (only mandatory components, not including childcare benefits); higher score = the more progressive the mandatory pension system (maximum value: 100).

⁶⁹ Absolute difference in the aggregate replacement ratio between women and men. Positive gap = women have a higher aggregate replacement ratio in comparison to men; women benefit on average more from the pension

system's redistribution. Aggregated replacement ratio = relationship between median gross pension income of 65 to 74-year-olds and median gross earnings of 50 to 59-year-olds.

Occupational & private pensions	Access	Share of occupational pension	2% 3% of men, 1% of women	30% 41% of men, 20% of women	Own calculation based on sample
		Share of private pension	0.2% 0.2% of men, 0.1% of women	4.9% 6.6% of men, 3.2% of women	Own calculation based on sample

Table A3.3 Waves, survey years and number of observations

Wave	Survey year	Italy		West Germany	
		Men	Women	Men	Women
2	2006–2007	71	67	56	68
4	2011	125	139	87	91
5	2013	907	1,007	565	543
6	2015	254	296	96	113

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.4 Overview of sample selection and dropouts.

		Men	Women	Total
Base sample age cutoff	Italy	2141	2308	4449
	West Germany	1475	1475	2950
Non-response pension	Italy	1966	2169	4135
	West Germany	1274	1302	2576
Employment	Italy	1845	2083	3928
	West Germany	1182	1241	2423
Missing pension income	Italy	1844	2080	3924
	West Germany	1182	1241	2423
Final sample incl. retro data	Italy	1357	1509	2866
	West Germany	804	815	1619

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.5 Pension types included in the analysis

Public pension	Occupational pension	Private Pension
1. Old age pension	1. Old age pension from the last, a second or a third job	1. Average payments of regular life insurance
2. Old age supplementary pension or public old age second pension	2. Early retirement pension	2. Regular private annuity or private personal pension
3. Early retirement or pre-retirement pension	3. Disability or invalidity insurance	
4. Main or secondary disability insurance pension, or sickness benefits	<i>(For wave 6: sum of all occupational pensions)</i>	

Table A3.6 Options for reference coefficients, interpretation, benefits, and specifics for SA-KOB decomposition

Reference coefficient option	I.a) Group 1 (here: men)	I.b) Group 2 (here: women)	II) Pooled	III) Cluster-specific majority group
Concept	Reference coefficients retrieved from a regression model for group 1 (e.g., men): $\hat{\beta}_M$	Reference coefficients retrieved from a regression model for group 2 (e.g., women): $\hat{\beta}_W$	Reference coefficients retrieved from a regression model pooled over both groups (controlling for the group indicator): $\hat{\beta}^*$	Coefficients of group 1 or group 2 chosen as reference for each cluster separately. - For group-specific clusters: coefficients of the majority group. - For group-neutral clusters: coefficients of the theoretically nondiscriminatory group.
Interpretation explained part	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_M$ Change of gap if the prevalence of women would be equal to the one of men in the specific life-course cluster and if women received the same reward as men for this life course	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_W$ Change of gap if the prevalence of women would be equal to the one of men in the specific life-course cluster and if men received the same reward as women for this life course	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}^*$ Change of gap if the prevalence of women would be equal to the one of men in the specific life-course cluster and if women received the return from a pooled model for this life course	$(\bar{X}_M - \bar{X}_W)' \hat{\beta}_M$ <u>or</u> $(\bar{X}_M - \bar{X}_W)' \hat{\beta}_W$ - Change of gap if the prevalence of the minority group in the specific life course would be equal to the majority group and would be rewarded for this life course equivalent to the majority group. - Interpretation equivalent to options I.a) and I.b), depending on the life-course cluster and coefficients chosen.
Interpretation returns component from unexplained part	$\bar{X}'_W (\hat{\beta}_M - \hat{\beta}_W)$ Change of gap if the observed share of women in the respective life-course cluster received the same returns for the life-course cluster as men	$\bar{X}'_M (\hat{\beta}_M - \hat{\beta}_W)$ Change of gap if the observed share of men in the respective life-course cluster received the same returns for the life-course cluster as women	$\bar{X}'_M (\hat{\beta}_M - \hat{\beta}^*) + \bar{X}'_W (\hat{\beta}^* - \hat{\beta}_W)$ Sum of change of gap if the observed share of men and the share of women in the respective life-course cluster received the returns of the pooled sample for this life-course cluster respectively	$\bar{X}'_W (\hat{\beta}_M - \hat{\beta}_W)$ <u>or</u> $\bar{X}'_M (\hat{\beta}_M - \hat{\beta}_W)$ - Change of gap if the underrepresented group received the same reward for the respective life-course cluster as the majority group. - Interpretation equivalent to options I.a) and I.b), depending on the life-course cluster and coefficients chosen.
Benefits	- Straightforward interpretation - Reflects assumption that one group is not discriminated against		- Sensitive to the group composition of the covariates and the implications for the rewarding structure	- Straightforward interpretation - Sensitive to the group composition of the covariates and the implications for the rewarding structure (e.g., group-specific (positive) discrimination for group-specific life-course clusters)
Limitations	- Nondiscrimination assumption unlikely to hold (esp. for all covariates of interests) - Disregards group composition of the clusters		- Artificial rewarding structure - Unclear which group drives coefficient - Ambiguous interpretation	- Necessary to define thresholds for group-specific and neutral clusters & choice of majority group; application-specific
Specifics SA+KOB	for - Concept theoretically and empirically challenged in case of high level of group-specific life-course clusters		- In case of highly group-specific life-course cluster, the less typical group might skew the nondiscriminatory coefficients in the pooled regression model	- Addresses group-specific life-course clusters and its theoretical and empirical implications - Addresses some issues of group-specific within-cluster heterogeneity as majority group more likely to be coherent with overall cluster characteristics

Table A3.7 Characteristics of life-course clusters for Italy

	Cluster 1: Ft Employed, 2+ Children		Cluster 2: Civil Servant, 2+ Children		Cluster 3: Self- employed, 2+ Children		Cluster 4: Pt Employed, 2+ Children		Cluster 5: Ft Care and Other, 2+ Children		Cluster 6: Ft Employed Mostly, 1 Child		Cluster 7: Ft Employed Mostly, Married No Children		Cluster 8: Ft Employed Mostly, Single	
Share on working states	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Full-Time Employed	76.85	69.49	7.78	3.87	13.56	9.56	29.58	13.29	27.38	6.49	61.02	28.13	56.36	34.95	52.01	38.99
(Any) Part-Time Employed	0.37	1.07	0.53	0.00	1.13	0.94	13.50	14.17	1.07	0.60	0.43	5.03	0.44	1.45	3.36	1.52
Self-Employed	1.88	1.29	0.85	0.72	71.79	66.00	8.35	1.11	6.73	1.11	19.85	11.51	15.60	5.17	12.79	5.60
Civil Servant	0.22	0.22	70.23	67.43	0.27	0.00	1.81	1.50	2.08	0.01	0.00	0.22	7.47	8.33	5.23	7.12
Care Work And Other	1.61	7.44	1.76	3.66	1.87	9.42	3.42	15.71	53.10	87.17	1.66	42.22	1.93	31.44	6.98	26.19
Retired	16.74	17.35	14.61	16.79	10.22	12.25	31.75	21.60	6.31	3.96	14.78	11.29	15.34	15.89	16.12	16.18
Education Or Training	1.75	1.87	3.41	5.40	0.79	0.69	0.77	3.77	3.33	0.39	1.98	0.82	2.15	1.82	2.01	2.49
Unemployed	0.60	1.26	0.83	2.13	0.36	1.14	10.81	28.85	0.00	0.28	0.29	0.78	0.71	0.95	1.50	1.92
Share on family states																
Single, No Child	18.91	10.76	20.41	13.10	18.36	8.58	20.96	11.80	22.38	10.89	21.00	13.71	24.76	21.02	99.34	99.95
Single, 1+ Children	0.53	4.64	0.00	0.00	0.17	0.00	0.81	0.14	4.76	0.21	0.26	0.47	0.00	0.00	0.00	0.00
Married/Cohab., No Children	3.09	3.06	3.69	4.08	3.19	3.03	10.17	4.24	16.67	3.04	5.93	6.18	74.22	78.85	0.66	0.05
Married/Cohab., 1 Child	7.93	8.40	21.03	27.72	8.28	7.22	18.00	9.42	10.65	7.54	70.95	79.28	0.98	0.02	0.00	0.00
Married/Cohab., 2+ Children	68.88	69.60	53.61	53.57	69.94	80.97	48.17	74.41	45.54	78.26	1.05	0.28	0.00	0.00	0.00	0.00
Divorced w/o Children	0.67	3.54	1.26	1.53	0.07	0.19	1.90	0.00	0.00	0.06	0.80	0.07	0.04	0.11	0.00	0.00
Cohort																
1909–1929	8.72	4.33	4.82	2.04	12.38	13.33	15.00	10.10	11.43	12.95	8.95	13.39	9.57	8.04	8.77	12.99
1930–1939	40.44	31.73	27.71	16.33	39.11	25.33	33.00	33.65	45.71	37.41	34.74	30.36	30.85	36.61	31.58	35.06
1940–1944	24.66	28.37	28.92	36.73	29.70	29.33	22.00	29.33	25.71	20.86	27.89	28.13	27.66	23.21	19.30	27.27
1945–1950	26.17	35.58	38.55	44.90	18.81	32.00	30.00	26.92	17.14	28.78	28.42	28.13	31.91	32.14	40.35	24.68
Wave																
2	5.37	2.88	0.00	0.00	7.43	2.67	4.00	4.33	0.00	5.04	6.32	5.80	5.32	7.14	5.26	1.30
4	11.58	7.69	1.20	2.04	7.43	5.33	9.00	10.58	8.57	10.07	8.42	11.16	10.64	4.46	3.51	12.99
5	65.10	65.38	63.86	63.27	73.27	68.00	66.00	71.15	77.14	66.73	65.79	63.39	65.96	67.86	66.67	67.53
6	17.95	24.04	34.94	34.69	11.88	24.00	21.00	13.94	14.29	18.17	19.47	19.64	18.09	20.54	24.56	18.18
Interview Year																
2006	0.84	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.54	1.05	0.89	0.00	0.00	0.00	0.00
2007	4.53	2.88	0.00	0.00	6.44	2.67	4.00	4.33	0.00	4.50	5.26	4.91	5.32	7.14	5.26	1.30
2011	11.58	7.69	1.20	2.04	7.43	5.33	9.00	10.58	8.57	10.07	8.42	11.16	10.64	4.46	3.51	12.99
2013	65.10	65.38	63.86	63.27	73.27	68.00	66.00	71.15	77.14	66.73	65.79	63.39	65.96	67.86	66.67	67.53
2015	17.95	24.04	34.94	34.69	11.88	24.00	21.00	13.94	14.29	18.17	19.47	19.64	18.09	20.54	24.56	18.18
Education																

Max. lower secondary (Isced 0–2)	74.19	72.06	54.32	28.57	87.31	88.00	85.71	80.30	82.86	94.73	71.51	84.23	75.00	70.91	78.95	73.68
Max. post-secondary (Isced 3, 4)	19.83	21.57	28.40	40.82	9.64	10.67	9.18	17.24	5.71	4.36	20.43	13.96	15.22	20.91	14.04	18.42
Tertiary education (Isced 5,6)	5.98	6.37	17.28	30.61	3.05	1.33	5.10	2.46	11.43	0.91	8.06	1.80	9.78	8.18	7.02	7.89
Born In Country Of Interview																
Not Born In Country	0.85	1.47	1.23	2.04	0.51	1.33	0.00	1.48	0.00	1.27	1.61	1.35	2.17	1.80	0.00	0.00
Native - Born In Country	99.15	98.53	98.77	97.96	99.49	98.67	100.00	98.52	100.00	98.73	98.39	98.65	97.83	98.20	100.00	100.00
Zero pension																
Any Pension Income	89.09	83.65	86.75	91.84	88.61	80.00	82.00	72.60	77.14	44.24	93.16	74.55	87.23	67.86	85.96	88.31
No Pension Income	10.91	16.35	13.25	8.16	11.39	20.00	18.00	27.40	22.86	55.76	6.84	25.45	12.77	32.14	14.04	11.69
Mean pension income	13,855.01	9,114.86	14,897.98	13,186.51	9,206.65	6,600.11	10,206.97	6,209.10	9,360.51	3,061.79	14,148.63	6,719.03	12,735.11	7,290.58	10,387.80	9,278.45
Mean age at time of interview	73.11	71.43	71.41	70.22	74.09	73.27	73.98	72.84	75.00	73.49	72.66	73.32	72.33	72.51	72.44	73.83

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.8 Characteristics of life-course clusters for West Germany

	Cluster 1: Ft Employed, 2+ Children		Cluster 2: Civil Servant, 2+ Children		Cluster 3: Self- employed, 2+ Children		Cluster 4: Pt Employed, 2+ Children		Cluster 5: Ft Care and Other, 2+ Children		Cluster 6: Ft Employed Mostly, 1 Child		Cluster 7: Ft Employed Mostly, Married No Children		Cluster 8: Ft Employed Mostly, Single	
Share on working states	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Full-Time Employed	83.60	64.47	5.93	1.76	23.90	10.58	44.44	17.56	35.09	12.65	76.19	35.98	64.05	55.30	60.57	48.54
(Any) Part-Time Employed	0.08	9.19	0.14	1.69	0.00	7.83	5.27	37.96	0.00	2.57	0.12	18.78	0.00	9.42	0.62	1.46
Self-Employed	1.35	1.82	0.02	0.39	66.26	62.83	4.92	0.17	4.61	0.22	9.21	3.00	5.93	3.76	9.26	3.13
Civil Servant	0.65	0.45	73.96	73.76	0.16	0.75	5.38	0.31	0.33	0.77	0.38	0.22	15.04	4.98	3.40	5.83
Care Work And Other	0.86	8.14	0.39	2.47	0.61	8.83	2.31	21.79	47.70	78.88	1.14	29.81	0.49	11.46	10.73	19.17
Retired	9.64	12.29	9.17	8.46	4.53	6.67	26.97	16.38	8.22	3.09	9.09	9.62	9.68	10.78	10.42	15.21
Education Or Training	3.51	2.77	10.40	11.46	4.53	2.50	5.56	2.66	4.06	1.63	3.35	1.65	4.48	3.22	3.55	6.67
Unemployed	0.31	0.87	0.00	0.00	0.00	0.00	5.15	3.18	0.00	0.18	0.52	0.94	0.32	1.09	1.47	0.00
Share on family states																
Single, No Child	15.70	8.56	17.42	12.57	16.99	11.17	17.65	8.15	18.75	10.96	15.60	11.32	28.60	17.53	98.30	94.79
Single, 1+ Children	0.89	4.60	0.20	0.00	0.08	0.00	3.82	0.29	3.95	0.68	0.68	1.04	0.00	0.00	0.00	0.00
Married/Cohab., No Children	4.26	4.28	5.14	5.79	3.31	3.08	17.25	3.37	5.48	3.21	6.98	6.09	69.24	80.48	1.70	5.21
Married/Cohab., 1 Child	8.05	7.35	19.25	13.61	5.76	9.00	13.37	8.07	19.30	7.30	71.72	78.51	0.67	0.00	0.00	0.00
Married/Cohab., 2+ Children	67.23	57.90	56.28	65.23	71.98	76.00	43.34	79.59	50.77	77.25	1.46	0.66	0.42	0.00	0.00	0.00
Divorced w/o Children	3.88	17.31	1.71	2.80	1.88	0.75	4.57	0.54	1.75	0.60	3.55	2.38	1.06	1.99	0.00	0.00
Cohort																
1909–1929	9.51	9.09	7.38	3.13	9.80	16.00	16.67	8.33	15.79	13.97	11.19	20.13	3.39	8.70	11.11	10.00
1930–1939	38.04	31.82	37.70	3.13	50.98	36.00	25.00	36.76	21.05	40.61	28.67	24.53	33.90	28.26	29.63	50.00
1940–1944	31.70	30.00	33.61	40.63	19.61	16.00	25.00	26.96	31.58	31.88	31.47	23.27	35.59	30.43	29.63	10.00
1945–1950	20.75	29.09	21.31	53.13	19.61	32.00	33.33	27.94	31.58	13.54	28.67	32.08	27.12	32.61	29.63	30.00
Wave																
2	8.07	10.91	3.28	3.13	7.84	0.00	0.00	5.39	15.79	11.35	6.99	8.81	6.78	8.70	11.11	0.00
4	9.80	4.55	13.93	12.50	5.88	8.00	19.44	11.76	10.53	10.48	12.59	17.61	6.78	8.70	7.41	0.00
5	71.76	66.36	72.13	56.25	78.43	72.00	72.22	68.14	68.42	71.62	62.24	58.49	69.49	65.22	70.37	80.00
6	10.37	18.18	10.66	28.13	7.84	20.00	8.33	14.71	5.26	6.55	18.18	15.09	16.95	17.39	11.11	20.00
Interview Year																
2006	1.73	0.91	0.00	0.00	0.00	0.00	0.00	1.47	5.26	0.87	2.10	1.26	0.00	0.00	3.70	0.00
2007	6.34	10.00	3.28	3.13	7.84	0.00	0.00	3.92	10.53	10.48	4.90	7.55	6.78	8.70	7.41	0.00
2011	9.51	4.55	13.93	12.50	5.88	8.00	19.44	11.76	10.53	10.48	11.89	16.35	6.78	8.70	7.41	0.00
2012	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70	1.26	0.00	0.00	0.00	0.00
2013	71.76	66.36	72.13	56.25	78.43	72.00	72.22	68.14	68.42	71.62	62.24	58.49	69.49	65.22	70.37	80.00
2015	10.37	18.18	10.66	28.13	7.84	20.00	8.33	14.71	5.26	6.55	18.18	15.09	16.95	17.39	11.11	20.00

Education																
Max. lower secondary (Isced 0–2)	4.32	27.27	3.28	0.00	3.92	28.00	20.59	26.96	10.53	41.92	9.79	24.68	8.47	13.04	11.11	30.00
Max. post-secondary (Isced 3, 4)	65.42	58.18	25.41	12.50	60.78	60.00	50.00	62.75	57.89	49.34	67.13	61.39	62.71	63.04	62.96	20.00
Tertiary education (Isced 5, 6)	30.26	14.55	71.31	87.50	35.29	12.00	29.41	10.29	31.58	8.73	23.08	13.92	28.81	23.91	25.93	50.00
Born In Country Of Interview																
Not Born In Country	15.56	14.55	13.11	6.25	7.84	4.00	27.78	17.65	21.05	10.92	10.49	16.98	18.64	6.52	22.22	10.00
Native - Born In Country	84.44	85.45	86.89	93.75	92.16	96.00	72.22	82.35	78.95	89.08	89.51	83.02	81.36	93.48	77.78	90.00
Zero pension																
Any Pension Income	98.85	98.18	96.72	93.75	96.08	92.00	97.22	94.12	100.00	86.46	98.60	87.42	98.31	95.65	100.00	90.00
No Pension Income	1.15	1.82	3.28	6.25	3.92	8.00	2.78	5.88	0.00	13.54	1.40	12.58	1.69	4.35	0.00	10.00
Mean pension income	20,011.36	10,316.90	31,798.87	22,671.13	13,888.74	6,030.56	15,878.52	7,897.04	16,523.29	3,738.53	20,305.42	8,649.98	21,592.14	12,759.74	14,254.80	10,367.60
Mean age at time of interview	72.58	71.68	72.33	68.41	73.73	73.64	73.03	71.94	71.95	73.75	71.97	72.90	71.32	71.00	71.52	74.40

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.9 Gender-specific distribution on clusters and gender gaps therein, before and after reassigning

	Italy							
	Pre-reassignment				Post-reassignment			
	Men	Women	Abs. difference	Rel. difference	Men	Women	Abs. difference	Rel. difference
1 FT employed, 2+ chi.	45.39	13.78	31.6	69.6	43.92	13.78	30.1	68.6
2 Civil servant, 2+ chi.	6.48	3.25	3.2	49.9	6.12	3.25	2.9	46.9
3 Self-employed, 2+ chi.	15.40	4.97	10.4	67.7	14.89	4.97	9.9	66.6
4 PT employed, 2+ chi.	3.61	13.78	10.2	73.8	7.37	13.78	6.4	46.5
5 FT care/other, 2+ chi.	1.18	36.85	35.7	96.8	2.58	36.85	34.3	93.0
6 FT empl./care, 1 child	15.48	14.84	0.63	4.3	14.00	14.84	0.8	5.7
7 FT employed, married	8.25	7.42	0.83	10.1	6.93	7.42	0.5	6.7
8 FT employed, single	4.20	5.10	0.9	21.5	4.20	5.10	0.9	17.7
	West Germany							
	Pre-reassignment				Post-reassignment			
	Men	Women	Abs. difference	Rel. difference	Men	Women	Abs. difference	Rel. difference
1 FT employed, 2+ chi.	45.15	13.50	31.7	70.1	43.16	13.50	29.7	68.7
2 Civil servant, 2+ chi.	15.55	3.93	11.6	74.7	15.17	3.93	11.2	74.1
3 Self-employed, 2+ chi.	6.59	3.07	3.5	53.5	6.34	3.07	3.3	51.6
4 PT employed, 2+ chi.	1.12	25.03	23.9	95.5	4.48	25.03	20.6	82.1
5 FT care/other, 2+ chi.	1.12	28.10	27.0	96.0	2.36	28.10	25.7	91.6
6 FT empl./care, 1 child	18.78	19.51	0.7	3.7	17.79	19.51	1.7	8.8
7 FT employed, married	8.33	5.64	2.7	32.3	7.34	5.64	1.7	23.1
8 FT employed, single	3.36	1.23	2.1	63.5	3.36	1.23	2.1	63.5

Notes: The pre-reassigned values refer to the original MSA solution (used in M11, see Supplementary Materials). The post-reassigned values were used in the main SA-KOB model (see Supplementary Materials E6 for details on the reassignment procedure). The pre-assigned relative differences were used to identify group-specific life courses and the majority groups respectively. Relative difference: the respective majority group (group with higher row percentages) used for the denominator. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.10 Robustness checks, set I: different reference coefficients used for KOB.

	Italy								West Germany							
	Main: MAJORITY group ref. (M1)		MEN's ref. (M2)		WOMEN's ref. (M3)		POOLED ref. (M4)		Main: MIXED ref. (M1)		MEN's ref. (M2)		WOMEN's ref. (M3)		POOLED ref. (M4)	
Mean pension inc. men	12,660**		12,660**		12,660**		12,660**		21,119**		21,119**		21,119**		21,119**	
	-57.31		-57.31		-57.31		-57.56		-53.62		-53.62		-53.62		-53.95	
Mean pension inc. women	6,009**		6,009**		6,009**		6,009**		8,030**		8,030**		8,030**		8,030**	
	-37.13		-37.13		-37.13		-37.26		-28.52		-28.52		-28.52		-28.68	
Absolute GPG	6,651**		6,651**		6,651**		6,651**		13,089**		13,089**		13,089**		13,089**	
	-24.29		-24.29		-24.29		-24.39		-27.03		-27.03		-27.03		-27.2	
Explained, total	2,054**		1,288**		2,112**		1,875**		3,514**		2,830**		3,418**		3,261**	
	-10.07		-2.686		-11.33		-11.25		-8.725		-3.802		-9.076		-9.675	
Unexplained, total	4,597**		5,364**		4,540**		4,777**		9,575**		10,260**		9,671**		9,828**	
	-14.52		-9.925		-14.61		-15.52		-17.65		-12.29		-18.37		-18.8	
	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns
1 FT employed, 2+ chi.	621.2**	92.16	621.2**	92.16	419.7**	293.7	568.4**	145	214.9	99.19	214.9	99.19	-3.073	317.2	202.3	111.8
	-4.88	-1.178	-4.88	-1.178	-3.487	-1.181	-6.072	-0.93	-1.04	-0.741	-1.04	-0.741	(-0.0147)	-0.742	-1.439	-0.51
2 Civil servant, 2+ chi.	79.72*	-84.40*	79.72*	-84.40*	154.3**	-159.0*	109.9**	-114.6*	1,401**	2.913	1,401**	2.913	1,392**	11.26	1,417**	-13.21
	-2.49	(-2.258)	-2.49	(-2.258)	-3.238	(-2.308)	-3.168	(-2.278)	-6.738	-0.0499	-6.738	-0.0499	-6.323	-0.0499	-7.029	(-0.106)
3 Self-employed, 2+ chi.	-245.2**	-69.06+	-245.2**	-69.06+	-107.4+	-206.8+	-209.4**	-104.9+	-181.2*	-33.92	-181.2*	-33.92	-145.0*	-70.15	-162.7**	-52.41
	(-3.929)	(-1.655)	(-3.929)	(-1.655)	(-1.766)	(-1.675)	(-5.084)	(-1.771)	(-2.499)	(-0.598)	(-2.499)	(-0.598)	(-2.321)	(-0.600)	(-2.624)	(-0.643)
4 PT employed, 2+ chi.	92.08**	-15.25	105.4*	-28.53	92.08**	-15.25	78.94**	-2.114	479.3**	-58.72	748.8*	-328.2	479.3**	-58.72	395.8**	24.78
	-3.069	(-0.247)	-2.061	(-0.247)	-3.069	(-0.247)	-2.82	(-0.0256)	-3.784	(-0.791)	-2.327	(-0.796)	-3.784	(-0.791)	-3.383	-0.164
5 FT care/other, 2+ chi.	1,568**	58.67+	788.1+	838.2+	1,568**	58.67+	1,370**	256.4**	1,646**	88.07	686.7	1,047+	1,646**	88.07	1,426**	307.7*
	-13.68	-1.778	-1.931	-1.858	-13.68	-1.778	-12.81	-2.91	-9.03	-1.625	-1.295	-1.74	-9.03	-1.625	-9.86	-2.468
6 FT empl./care, 1 child	-19.3	477.6**	-19.3	477.6**	7.817	450.5**	-6.16	464.4**	-19.99	518.0*	-19.99	518.0*	25.77	472.3*	-0.599	498.6*
	(-0.633)	-4.48	(-0.633)	-4.48	-0.62	-4.447	(-0.615)	-4.426	(-0.735)	-2.394	(-0.735)	-2.394	-0.833	-2.389	(-0.0591)	-2.18
7 FT employed, married	-4.059	87.04	-4.059	87.04	1.748	81.24	-1.965	84.95	41.67	-2.24	41.67	-2.24	42.34	-2.912	45.51	-6.083
	(-0.463)	-1.275	(-0.463)	-1.275	-0.414	-1.273	(-0.442)	-1.247	-1.134	(-0.0251)	-1.134	(-0.0251)	-1.217	(-0.0251)	-1.264	(-0.0562)
8 FT employed, single	13.84	-160.1**	13.84	-160.1**	-14.47	-131.8**	-4.483	-141.8**	-105.6*	-58.09	-105.6*	-58.09	-4.669	-159.0+	-75.96*	-87.70+
	-0.936	(-2.677)	-0.936	(-2.677)	(-1.056)	(-2.635)	(-0.761)	(-3.023)	(-2.020)	(-1.574)	(-2.020)	(-1.574)	(-0.112)	(-1.713)	(-2.049)	(-1.745)
Constant		4,259**		4,259**		4,259**		4,259**		8,651**		8,651**		8,651**		8,651**
		-8.649		-8.649		-8.649		-9.341		-10.37		-10.37		-10.37		-10.25
n men	1357	1357	1357	1357	1357	1357	1357	1357	804	804	804	804	804	804	804	804
n women	1509	1509	1509	1509	1509	1509	1509	1509	815	815	815	815	815	815	815	815

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.11 Robustness checks, set II: excluding individuals poorly assigned to cluster based on low silhouette widths.

	Italy						West Germany					
	Main model (M1)		Excluded < silhouette p5 (M5)		Excluded < silhouette p10 (M6)		Main model (M1)		Excluded < silhouette p5 (M5)		Excluded < silhouette p10 (M6)	
Mean pension inc. men	12,660**		12,681**		12,635**		21,119**		21,402**		21,501**	
	-57.31		(57.95)		(62.17)		-53.62		(57.64)		(50.46)	
Mean pension inc. women	6,009**		5,884**		5,868**		8,030**		7,992**		7,854**	
	-37.13		(43.24)		(32.17)		-28.52		(28.07)		(22.39)	
Absolute GPG	6,651**		6,797**		6,766**		13,089**		13,410**		13,648**	
	-24.29		(30.45)		(25.37)		-27.03		(29.07)		(23.77)	
Explained, total	2,054**		2,060**		2,254**		3,514**		3,418**		3,362**	
	-10.07		(10.24)		(9.008)		-8.725		(8.804)		(8.579)	
Unexplained, total	4,597**		4,737**		4,513**		9,575**		9,992**		10,286**	
	-14.52		(16.93)		(11.76)		-17.65		(16.54)		(15.97)	
	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns
1 FT employed, 2+ chi.	621.2**	92.16	624.0**	44.37	673.4**	58.94	214.9	99.19	282.4	74.41	307.4	101.5
	-4.88	-1.178	(4.351)	(0.599)	(4.835)	(0.666)	-1.04	-0.741	(1.440)	(0.696)	(1.383)	(1.068)
2 Civil servant, 2+ chi.	79.72*	-84.40*	77.61*	-80.15+	84.57*	-88.02*	1,401**	2.913	1,456**	-24.84	1,440**	-28.29
	-2.49	(-2.258)	(2.160)	(-1.935)	(2.455)	(-2.517)	-6.738	-0.0499	(6.978)	(-0.345)	(7.287)	(-0.363)
3 Self-employed, 2+ chi.	-245.2**	-69.06+	-256.4**	-70.23*	-280.5**	-77.02*	-181.2*	-33.92	-176.1*	-66.04	-152.8*	-41.65
	(-3.929)	(-1.655)	(-4.340)	(-2.171)	(-4.211)	(-2.079)	(-2.499)	(-0.598)	(-2.308)	(-0.979)	(-2.252)	(-0.612)
4 PT employed, 2+ chi.	92.08**	-15.25	112.0**	5.765	115.3**	22.47	479.3**	-58.72	362.1*	-94.71	316.1*	-115.6
	-3.069	(-0.247)	(3.412)	(0.113)	(3.535)	(0.371)	-3.784	(-0.791)	(2.534)	(-1.192)	(2.398)	(-1.608)
5 FT care/other, 2+ chi.	1,568**	58.67+	1,599**	60.53	1,744**	71.38	1,646**	88.07	1,606**	82.86+	1,594**	83.06+
	-13.68	-1.778	(12.49)	(1.390)	(18.44)	(1.603)	-9.03	-1.625	(9.782)	(1.948)	(9.161)	(1.843)
6 FT empl./care, 1 child	-19.3	477.6**	-36.19	486.9**	-19.94	414.6**	-19.99	518.0*	-54.12	601.9**	-60.88	524.9*
	(-0.633)	-4.48	(-0.965)	(4.241)	(-0.625)	(4.088)	(-0.735)	-2.394	(-0.992)	(2.579)	(-1.569)	(2.317)
7 FT employed, married	-4.059	87.04	-6.535	94.98	-5.658	102.4	41.67	-2.24	22.88	-41.25	10.44	-29.07
	(-0.463)	-1.275	(-0.380)	(1.337)	(-0.451)	(1.400)	-1.134	(-0.0251)	(0.875)	(-0.412)	(0.352)	(-0.291)
8 FT employed, single	13.84	-160.1**	12.56	-164.1**	15.69	-162.6*	-105.6*	-58.09	-109.4*	-16.48	-111.4*	-21.28
	-0.936	(-2.677)	(0.765)	(-2.758)	(1.050)	(-2.412)	(-2.020)	(-1.574)	(-2.173)	(-1.006)	(-2.412)	(-0.824)
Constant		4,259**		4,496**		4,335**		8,651**		9,021**		9,568**
		-8.649		(10.90)		(9.685)		-10.37		(10.62)		(12.26)
n men	1357	1357	1324	1324	1294	1294	804	804	767	767	747	747
n women	1509	1509	1441	1441	1350	1350	815	815	769	769	711	711

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.12 Robustness checks, set III: Including more controls.

	Italy								West Germany							
	Main, control sample (M7)		Cohort more detailed (M8)		Education (M9)		Nativity (M10)		Main, control sample (M7)		Cohort more detailed (M8)		Education (M9)		Nativity (M10)	
Mean pension inc. men	12,717**		12,717**		12,717**		12,717**		21,148**		21,148**		21,148**		21,148**	
	(69.30)		(55.02)		(62.94)		(63.15)		(50.58)		(52.58)		(64.92)		(50.47)	
Mean pension inc. women	6,013**		6,013**		6,013**		6,013**		8,029**		8,029**		8,029**		8,029**	
	(31.21)		(35.08)		(32.33)		(45.20)		(23.79)		(27.46)		(31.00)		(28.79)	
Absolute GPG	6,705**		6,705**		6,705**		6,705**		13,119**		13,119**		13,119**		13,119**	
	(26.00)		(20.61)		(24.69)		(31.29)		(23.37)		(27.18)		(30.19)		(24.65)	
	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns
1 FT employed, 2+ chi.	652.8**	93.45	652.7**	92.68	633.1**	86.41	632.9**	86.31	203.3	93.85	182.6	82.78	189.9	46.71	187.1	44.55
	(4.349)	(1.160)	(5.884)	(1.263)	(5.733)	(1.073)	(5.397)	(1.058)	(1.228)	(0.694)	(0.861)	(0.617)	(1.010)	(0.386)	(1.137)	(0.324)
2 Civil servant, 2+ chi.	82.08*	-80.45+	82.03**	-80.41*	45.12	-40.28	45.02	-40.39	1,400**	1.236	1,391**	-1.169	1,130**	-40.17	1,124**	-42.17
	(2.495)	(-1.943)	(2.795)	(-2.001)	(1.528)	(-1.060)	(1.628)	(-1.207)	(7.545)	(0.0162)	(6.775)	(-0.0142)	(6.385)	(-0.459)	(5.712)	(-0.458)
3 Self-employed, 2+ chi.	-238.3**	-67.91+	-238.7**	-67.93*	-175.3**	-74.47*	-175.1**	-74.36*	-183.0**	-34.99	-183.0*	-34.79	-193.6*	-55.01	-196.6**	-57.42
	(-3.715)	(-1.869)	(-4.402)	(-2.327)	(-4.219)	(-2.561)	(-3.899)	(-2.246)	(-2.683)	(-0.587)	(-2.170)	(-0.640)	(-2.492)	(-0.910)	(-3.362)	(-0.932)
4 PT employed, 2+ chi.	91.86**	-16.21	92.13**	-15.71	67.30*	0.257	67.29*	0.668	485.7**	-43.05	485.2**	-37.87	410.7**	-17.07	408.1**	-14.90
	(3.591)	(-0.248)	(3.080)	(-0.200)	(2.246)	(0.00371)	(2.286)	(0.0119)	(3.825)	(-0.626)	(3.927)	(-0.481)	(3.340)	(-0.242)	(3.559)	(-0.278)
5 FT care/other, 2+ chi.	1,579**	59.79	1,576**	59.42	1,217**	35.13	1,216**	35.35	1,648**	87.18*	1,641**	89.54+	1,505**	81.14+	1,505**	82.61
	(13.91)	(1.231)	(12.67)	(1.592)	(10.48)	(1.133)	(9.935)	(1.133)	(9.102)	(1.994)	(7.727)	(1.703)	(10.70)	(1.790)	(7.727)	(1.644)
6 FT empl./care, 1 child	-21.44	487.1**	-21.43	487.7**	-19.30	368.8**	-19.24	367.8**	-17.64	505.8*	-16.73	497.3*	-27.89	586.7*	-26.92	572.9*
	(-0.688)	(4.431)	(-0.813)	(4.691)	(-1.029)	(3.613)	(-0.616)	(3.858)	(-0.711)	(2.287)	(-0.471)	(2.329)	(-0.910)	(2.361)	(-0.673)	(2.461)
7 FT employed, married	-2.966	73.67	-2.974	73.14	-2.721	82.00	-2.670	81.19	41.22	-4.608	39.74	-9.571	45.32	12.90	45.75	14.75
	(-0.231)	(0.893)	(-0.271)	(1.123)	(-0.333)	(1.346)	(-0.395)	(1.138)	(1.020)	(-0.0557)	(1.218)	(-0.109)	(1.021)	(0.128)	(1.234)	(0.158)
8 FT employed, single	12.79	-163.7**	12.73	-162.9**	10.42	-145.5**	10.38	-145.2**	-106.8*	-58.67	-105.9*	-58.09	-94.10+	-46.75	-92.63*	-45.88
	(0.800)	(-2.688)	(0.844)	(-3.260)	(0.821)	(-3.235)	(0.893)	(-3.441)	(-2.426)	(-1.354)	(-2.429)	(-1.436)	(-1.942)	(-1.319)	(-1.965)	(-1.271)
Cohort 1909–1929	3.559	-2.588	5.870	-20.92	3.543	-34.05	3.503	-33.64	-23.65	63.29	-32.92	100.3	-35.00	82.67	-35.83	85.28
	(0.425)	(-0.0468)	(0.414)	(-0.278)	(0.438)	(-0.501)	(0.307)	(-0.545)	(-0.721)	(0.471)	(-0.616)	(0.612)	(-0.899)	(0.528)	(-1.166)	(0.657)
Cohort 1930–1939	-21.17	-216.5+	-29.56	-272.3*	-9.023	-137.6	-9.290	-139.9	-7.078	-78.51	0.779	14.34	6.551	58.28	7.138	66.46
	(-1.286)	(-1.900)	(-1.596)	(-2.150)	(-0.566)	(-1.062)	(-0.542)	(-1.057)	(-0.281)	(-0.318)	(0.0428)	(0.0583)	(0.376)	(0.221)	(0.398)	(0.310)
Cohort 1940–1948	-24.38	373.7							-1.762	-148.8						
	(-1.155)	(1.528)							(-0.0987)	(-0.397)						
Cohort 1940–1944			2.279	101.1	2.187	109.8	2.184	109.5			6.847	173.4	3.381	176.5	4.256	184.5
			(0.216)	(0.855)	(0.189)	(0.947)	(0.191)	(0.925)			(0.296)	(0.763)	(0.147)	(0.716)	(0.230)	(1.012)
Cohort 1945–1950			0.773	164.4	12.98	79.43	12.99	80.55			-0.397	-347.2	4.063	-347.5	4.333	-365.0
			(0.109)	(1.154)	(1.261)	(0.519)	(1.536)	(0.593)			(-0.0177)	(-1.600)	(0.186)	(-1.336)	(0.188)	(-1.466)

Continuation of Table A3.12

	Italy								West Germany							
	Main, control sample (M7)		Cohort more detailed (M8)		Education (M9)		Nativity (M10)		Main, control sample (M7)		Cohort more detailed (M8)		Education (M9)		Nativity (M10)	
Mean pension inc. men	12,717**		12,717**		12,717**		12,717**		21,148**		21,148**		21,148**		21,148**	
	(69.30)		(55.02)		(62.94)		(63.15)		(50.58)		(52.58)		(64.92)		(50.47)	
Mean pension inc. women	6,013**		6,013**		6,013**		6,013**		8,029**		8,029**		8,029**		8,029**	
	(31.21)		(35.08)		(32.33)		(45.20)		(23.79)		(27.46)		(31.00)		(28.79)	
Absolute GPG	6,705**		6,705**		6,705**		6,705**		13,119**		13,119**		13,119**		13,119**	
	(26.00)		(20.61)		(24.69)		(31.29)		(23.37)		(27.18)		(30.19)		(24.65)	
	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns
Max. lower secondary edu.					310.9**	-566.0	310.9**	-564.9					1,045**	-893.1**	1,033**	-878.7**
					(3.242)	(-1.080)	(4.198)	(-0.851)					(6.805)	(-4.806)	(7.745)	(-4.158)
Max. post-secondary edu.					29.71	41.88	30.00	42.82					-14.49	-209.1	-15.78	-235.0
					(1.142)	(0.362)	(0.996)	(0.386)					(-0.651)	(-0.585)	(-0.549)	(-0.725)
Tertiary education					113.5**	14.44	113.2**	14.11					1,026**	535.1**	1,025**	534.6**
					(2.771)	(0.287)	(2.731)	(0.271)					(8.419)	(3.876)	(5.570)	(3.897)
Not native							0.133	3.670							-1.118	-64.58
							(0.0502)	(0.291)							(-0.117)	(-0.716)
Native							0.133	-403.4							-1.118	374.4
							(0.0502)	(-0.294)							(-0.117)	(0.730)
Wave 2	-5.681	-6.751	-5.874	-6.108	-3.771	-17.64	-3.769	-17.34	9.134	-53.13	8.192	-80.96	5.835	-39.46	5.867	-44.00
	(-0.535)	(-0.158)	(-0.629)	(-0.162)	(-0.393)	(-0.424)	(-0.542)	(-0.412)	(0.667)	(-0.655)	(0.585)	(-0.992)	(0.363)	(-0.455)	(0.399)	(-0.510)
Wave 4	0.0480	21.47	0.103	22.98	0.0530	12.37	0.0529	12.37	5.906	189.6+	-0.735	154.2	0.0854	134.5	-0.0997	139.2
	(0.00924)	(0.362)	(0.0167)	(0.388)	(0.00807)	(0.193)	(0.00564)	(0.192)	(0.271)	(1.910)	(-0.0467)	(1.360)	(0.00571)	(1.092)	(-0.00825)	(1.099)
Wave 5	-0.325	-200.3	-0.330	-194.3	-0.191	-28.83	-0.194	-30.16	50.08	677.6	41.81	490.9	32.07	264.5	32.08	265.1
	(-0.0466)	(-0.666)	(-0.0498)	(-0.643)	(-0.0201)	(-0.0842)	(-0.0275)	(-0.109)	(1.164)	(1.271)	(1.163)	(0.974)	(0.932)	(0.554)	(1.261)	(0.571)
Wave 6	-6.241	37.42	-6.519	29.88	-7.241	46.72	-7.204	45.99	7.716	-278.3	-7.862	-131.5	-2.194	-142.4	-2.671	-139.3
	(-0.498)	(0.286)	(-0.546)	(0.270)	(-0.397)	(0.313)	(-0.658)	(0.389)	(0.281)	(-1.503)	(-0.267)	(-0.713)	(-0.0479)	(-0.551)	(-0.0843)	(-0.668)
Constant		4,211**		4,375**		4,644**		5,046**		8,690**		8,768**		7,894**		7,633**
		(9.641)		(8.522)		(6.016)		(3.378)		(10.26)		(11.97)		(9.782)		(9.093)
n men	1331	1331	1331	1331	1331	1331	1331	1331	802	802	802	802	802	802	802	802
n women	1489	1489	1489	1489	1489	1489	1489	1489	814	814	814	814	814	814	814	814

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.13 Robustness checks, set IV: using original cluster and cross-sectional weights.

	Italy						Germany					
	Main model (M1)		Not manipulated (M11)		Weighted (M12)		Main model (M1)		Not manipulated (M11)		Weighted (M12)	
Mean pension inc. men	12,660**		12,660**		12,559**		21,119**		21,119**		20,690**	
	-57.31		(59.99)		(54.36)		-53.62		(52.43)		(49.05)	
Mean pension inc. women	6,009**		6,009**		5,750**		8,030**		8,030**		7,940**	
	-37.13		(33.76)		(33.51)		-28.52		(32.23)		(23.70)	
Absolute GPG	6,651**		6,651**		6,808**		13,089**		13,089**		12,750**	
	-24.29		(23.37)		(23.66)		-27.03		(26.15)		(23.67)	
Explained, total	2,054**		2,404**		2,112**		3,514**		3,718**		3,217**	
	-10.07		(11.17)		(9.665)		-8.725		(10.08)		(7.303)	
Unexplained, total	4,597**		4,247**		4,696**		9,575**		9,371**		9,533**	
	-14.52		(13.94)		(13.57)		-17.65		(15.50)		(15.57)	
	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns	explained	returns
1 FT employed, 2+ chi.	621.2**	92.16	750.8**	135.4+	662.0**	108.4	214.9	99.19	207.8	90.01	121.5	162.4
	-4.88	-1.178	(5.563)	(1.810)	(4.715)	(1.374)	-1.04	-0.741	(0.989)	(0.753)	(0.649)	(1.146)
2 Civil servant, 2+ chi.	79.72*	-84.40*	94.50*	-79.84*	62.30*	-69.60**	1,401**	2.913	1,433**	-1.914	1,485**	29.97
	-2.49	(-2.258)	(2.375)	(-2.115)	(2.211)	(-2.621)	-6.738	-0.0499	(7.416)	(-0.0279)	(6.419)	(0.526)
3 Self-employed, 2+ chi.	-245.2**	-69.06+	-204.9**	-43.75	-255.2**	-70.65*	-181.2*	-33.92	-197.9*	-36.41	-181.2*	-71.86
	(-3.929)	(-1.655)	(-3.477)	(-1.215)	(-4.715)	(-2.038)	(-2.499)	(-0.598)	(-2.358)	(-0.640)	(-2.200)	(-1.034)
4 PT employed, 2+ chi.	92.08**	-15.25	146.0**	-7.044	76.01*	-1.908	479.3**	-58.72	557.6**	-9.089	348.0*	-61.68
	-3.069	(-0.247)	(2.911)	(-0.143)	(2.191)	(-0.0311)	-3.784	(-0.791)	(3.403)	(-0.310)	(2.061)	(-0.762)
5 FT care/other, 2+ chi.	1,568**	58.67+	1,632**	-0.544	1,531**	52.46	1,646**	88.07	1,725**	39.18	1,589**	80.31+
	-13.68	-1.778	(19.04)	(-0.0332)	(12.55)	(1.449)	-9.03	-1.625	(10.88)	(1.527)	(7.661)	(1.785)
6 FT empl./care, 1 child	-19.3	477.6**	17.12	540.4**	-22.70	481.8**	-19.99	518.0*	-8.952	531.6*	-29.75	459.8
	(-0.633)	-4.48	(0.507)	(4.706)	(-0.813)	(3.881)	(-0.735)	-2.394	(-0.251)	(2.432)	(-0.642)	(1.586)
7 FT employed, married	-4.059	87.04	10.35	118.6	2.072	145.9*	41.67	-2.24	59.23	-16.73	55.13	-6.885
	(-0.463)	-1.275	(0.679)	(1.546)	(0.161)	(2.057)	-1.134	(-0.0251)	(1.024)	(-0.187)	(1.153)	(-0.0707)
8 FT employed, single	13.84	-160.1**	9.351	-134.7*	-1.987	-198.6**	-105.6*	-58.09	-101.6*	-55.79	-170.1*	-55.84
	-0.936	(-2.677)	(0.646)	(-2.241)	(-0.123)	(-2.824)	(-2.020)	(-1.574)	(-2.187)	(-1.521)	(-2.468)	(-1.396)
Constant		4,259**		3,780**		4,462**		8,651**		8,420**		8,186**
		-8.649		(9.004)		(9.483)		-10.37		(8.396)		(9.266)
n men	1357	1357	1357	1357	1357	1357	804	804	804	804	804	804
n women	1509	1509	1509	1509	1509	1509	815	815	815	815	815	815

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Models M1 and M11 not weighted.

Table A3.14 Comparing the SA-KOB results to standard KOB (summary measures).

	Italy								West Germany							
	Main model (SA-KOB)				Standard model				Main model (SA-KOB)				Standard model			
Absolute GPG	6,651**				6,651**				13,089**				13,089**			
	-24.29				(21.27)				-27.03				-26.33			
Explained, total	2,054**				1,682**				3,514**				4,499**			
	-10.07				(6.781)				-8.725				-6.457			
Unexplained, total	4,597**				4,970**				9,575**				8,590**			
	-14.52				(12.09)				-17.65				-12.55			
	explained	shares	returns	shares	explained	shares	returns	shares	explained	shares	returns	shares	explained	shares	returns	shares
1 FT employed, 2+ chi.	621.2**	9.34%	92.16	1.39%					214.9	1.64%	99.19	0.76%				
	-4.88		-1.178						-1.04		-0.741					
2 Civil servant, 2+ chi.	79.72*	1.20%	-84.40*	-1.27%					1,401**	10.70%	2.913	0.02%				
	-2.49		(-2.258)						-6.738		-0.0499					
3 Self-employed, 2+ chi.	-245.2**	-3.69%	-69.06+	-1.04%					-181.2*	-1.38%	-33.92	-0.26%				
	(-3.929)		(-1.655)						(-2.499)		(-0.598)					
4 PT employed, 2+ chi.	92.08**	1.38%	-15.25	-0.23%					479.3**	3.66%	-58.72	-0.45%				
	-3.069		(-0.247)						-3.784		(-0.791)					
5 FT care/other, 2+ chi.	1,568**	23.58%	58.67+	0.88%					1,646**	12.58%	88.07	0.67%				
	-13.68		-1.778						-9.03		-1.625					
6 FT empl./care, 1 child	-19.3	-0.29%	477.6**	7.18%					-19.99	-0.15%	518.0*	3.96%				
	(-0.633)		-4.48						(-0.735)		-2.394					
7 FT employed, married	-4.059	-0.06%	87.04	1.31%					41.67	0.32%	-2.24	-0.02%				
	(-0.463)		-1.275						-1.134		(-0.0251)					
8 FT employed, single	13.84	0.21%	-160.1**	-2.41%					-105.6*	-0.81%	-58.09	-0.44%				
	-0.936		(-2.677)						(-2.020)		(-1.574)					
Years FT employment					1,368**	20.56%	-706.7**	-10.63%					3,136**	23.96%	206.7	1.58%
					(6.491)		(-3.644)						-5.452		-0.431	
Years PT employment					-68.74**	-1.03%	-116.2**	-1.75%					-1,282**	-9.79%	-34.23	-0.26%
					(-2.964)		(-3.268)						(-6.136)		(-0.973)	
Years civil service					79.66*	1.20%	-311.1**	-4.68%					1,939**	14.82%	-134.9	-1.03%
					(2.224)		(-4.926)						-8.523		(-1.340)	
Years education					109.7	1.65%	629.9**	9.47%					940.9**	7.19%	916.5**	7.00%
					(1.141)		(3.931)						-3.169		-2.879	
Years unemployment					322.6**	4.85%	-362.1*	-5.44%					156.3	1.19%	-280.2	-2.14%
					(2.798)		(-2.450)						-0.843		(-1.066)	
Years divorced					3.471	0.05%	30.66	0.46%					27.32	0.21%	-514.3**	-3.93%
					(0.299)		(1.078)						-0.621		(-3.764)	
Years married, 2+ch.					145.3**	2.19%	-776.8	-11.68%					55.69	0.43%	517.6	3.95%
					(2.782)		(-1.324)						-0.51		-0.541	
Years married, no children					-1.068	-0.02%	-60.36	-0.91%					44.46	0.34%	87.09	0.67%
					(-0.0634)		(-0.423)						-1.226		-0.386	
Years single, no children					-254.3**	-3.82%	-856.5**	-12.88%					-558.2**	-4.26%	-829.2**	-6.33%
					(-3.080)		(-3.685)						(-3.278)		(-2.840)	
Constant		4,259**					7,648**			8,651**					8,369**	
		-8.649					(6.353)			-10.37					-4.163	
n male	1357		1357		1357		1357		804		804		804		804	
n female	1509		1509		1509		1509		815		815		815		815	

Notes: All models with absolute independent pension income as the dependent variable and controlled for birth cohort and wave. Main model based on normalized linear regression models, standard model on linear regression models. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A3.15 Results from Kitagawa-Oaxaca-Blinder decompositions, full table.

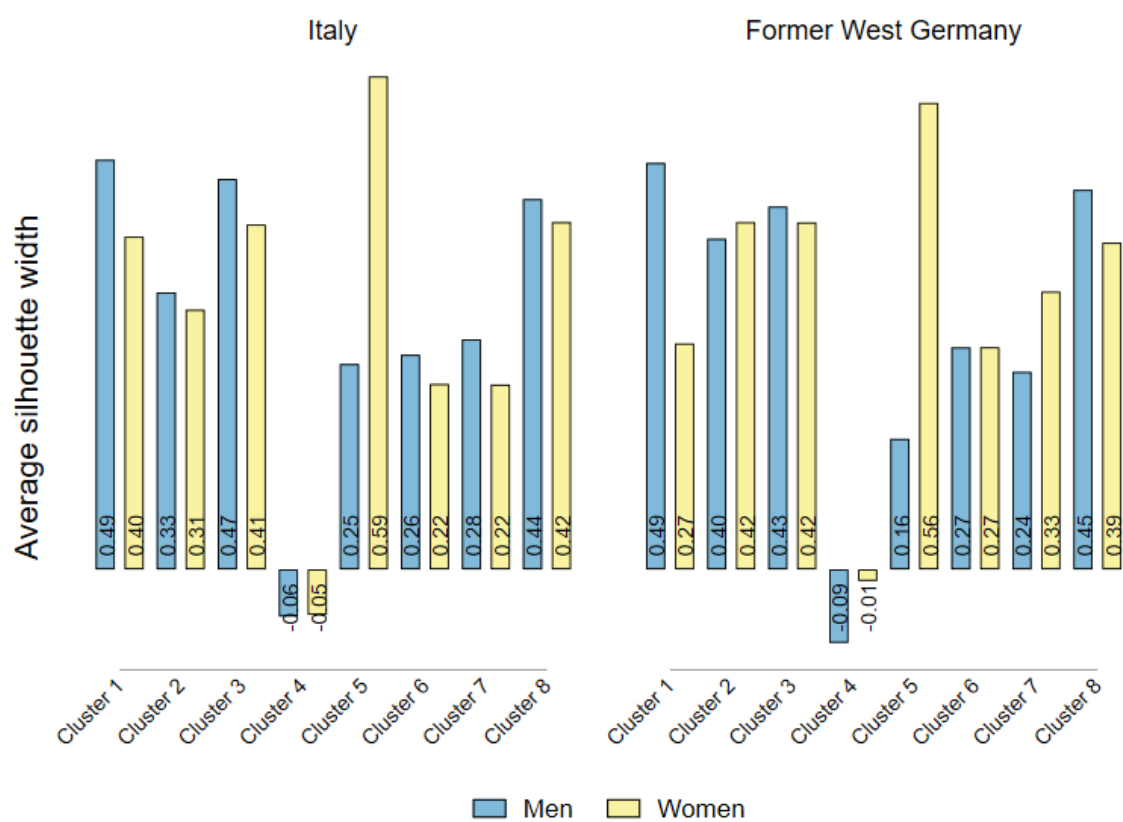
	Italy				West Germany			
Mean pension inc. men	12,660**				21,119**			
	-57.31				-53.62			
Mean pension inc. women	6,009**				8,030**			
	-37.13				-28.52			
Absolute GPG	6,651**				13,089**			
	-24.29				-27.03			
Explained, total	2,054**				3,514**			
	-10.07				-8.725			
Unexplained, total	4,597**				9,575**			
	-14.52				-17.65			
	explained	shares	returns	shares	explained	shares	returns	shares
Sum life courses	2106.281	31.67%	386.66	5.81%	3476.08	26.56%	555.203	4.24%
1 FT employed, 2+ chi.	621.2**	9.34%	92.16	1.39%	214.9	1.64%	99.19	0.76%
	-4.88		-1.178		-1.04		-0.741	
2 Civil servant, 2+ chi.	79.72*	1.20%	-84.40*	-1.27%	1,401**	10.70%	2.913	0.02%
	-2.49		(-2.258)		-6.738		-0.0499	
3 Self-employed, 2+ chi.	-245.2**	-3.69%	-69.06+	-1.04%	-181.2*	-1.38%	-33.92	-0.26%
	(-3.929)		(-1.655)		(-2.499)		(-0.598)	
4 PT employed, 2+ chi.	92.08**	1.38%	-15.25	-0.23%	479.3**	3.66%	-58.72	-0.45%
	-3.069		(-0.247)		-3.784		(-0.791)	
5 FT care/other, 2+ chi.	1,568**	23.58%	58.67+	0.88%	1,646**	12.58%	88.07	0.67%
	-13.68		-1.778		-9.03		-1.625	
6 FT empl./care, 1 child	-19.3	-0.29%	477.6**	7.18%	-19.99	-0.15%	518.0*	3.96%
	(-0.633)		-4.48		(-0.735)		-2.394	
7 FT employed, married	-4.059	-0.06%	87.04	1.31%	41.67	0.32%	-2.24	-0.02%
	(-0.463)		-1.275		-1.134		(-0.0251)	
8 FT employed, single	13.84	0.21%	-160.1**	-2.41%	-105.6*	-0.81%	-58.09	-0.44%
	-0.936		(-2.677)		(-2.020)		(-1.574)	
Cohort 1909–1929	2.916		1.009		-21.41		52.94	
	(0.261)		(0.0173)		(-0.733)		(0.387)	
Cohort 1930–1939	-21.11		-220.0		-3.367		-29.85	
	(-1.359)		(-1.628)		(-0.179)		(-0.120)	
Cohort 1940–1948	-21.62		359.2+		-2.317		-183.7	
	(-1.151)		(1.728)		(-0.166)		(-0.453)	
Wave 2	-5.787		-0.0227		9.153		-50.22	
	(-0.667)		(-0.000561)		(0.545)		(-0.564)	
Wave 4	0.000238		24.26		4.454		159.1	
	(4.24e-05)		(0.371)		(0.213)		(1.561)	
Wave 5	-0.354		-247.0		52.63		714.2	
	(-0.0390)		(-0.875)		(1.228)		(1.458)	
Wave 6	-5.012		21.05		4.550		-252.6	
	(-0.444)		(0.205)		(0.207)		(-1.438)	
Constant			4,259**				8,651**	
			-8.649				-10.37	

n men	1357	1357	804	804
n women	1509	1509	815	815

Notes: Based on normalized linear regression models with absolute independent pension income as the dependent variable. Controlled for birth cohort and wave. z-statistics in parentheses. ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

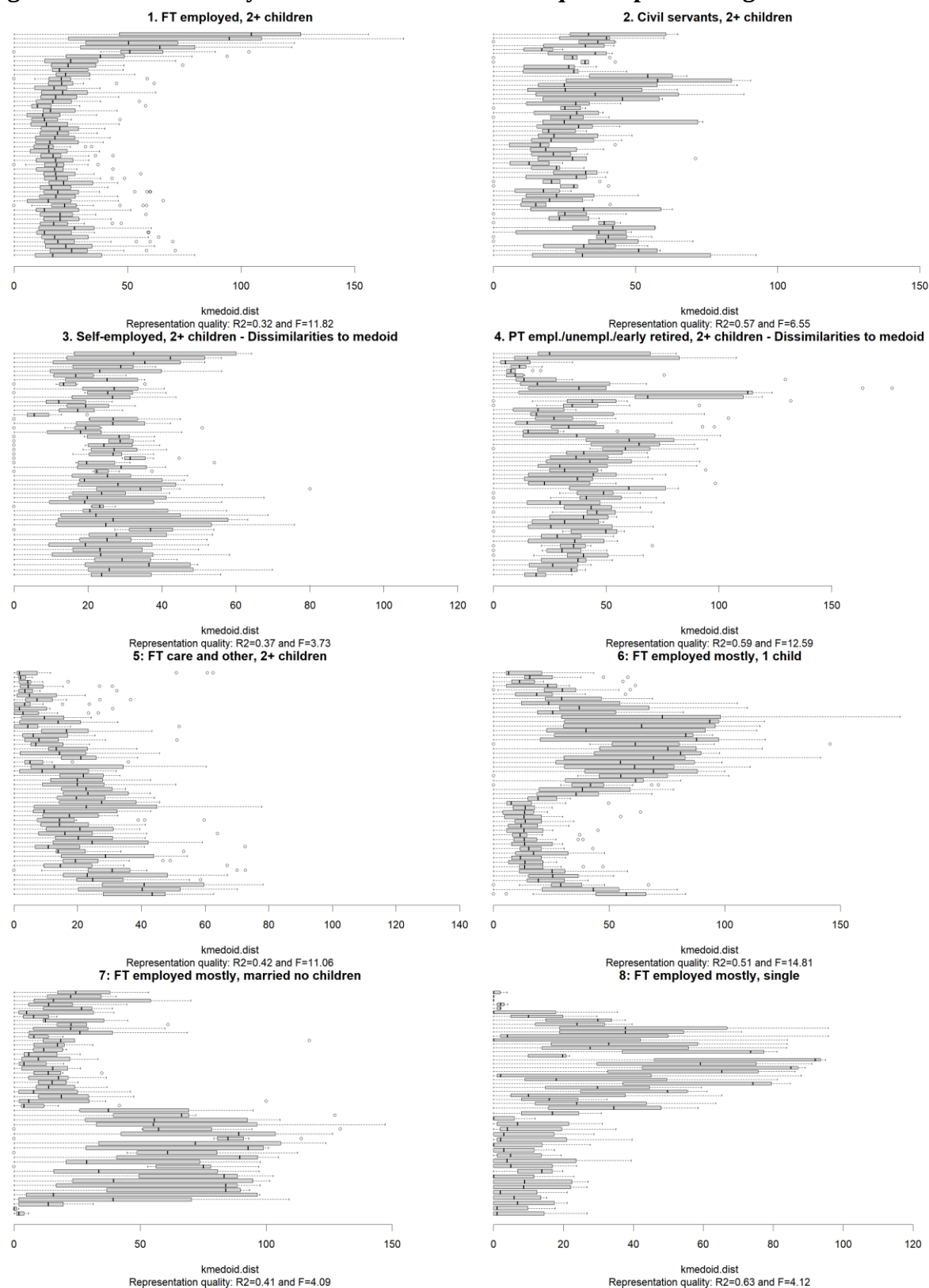
A3.2 Additional Figures

Figure A3.1 Average silhouette width (ASW) by cluster, country & gender



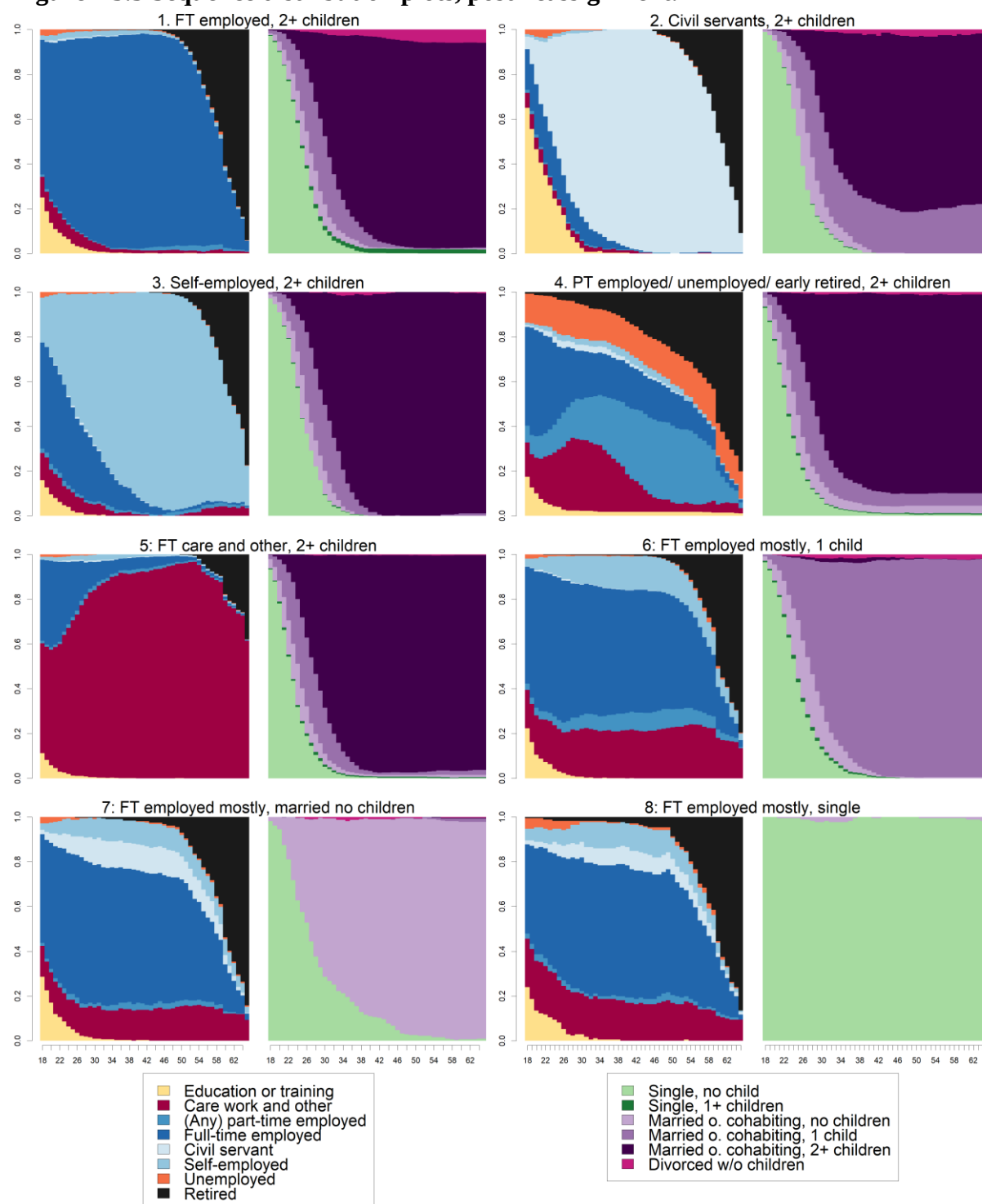
Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted

Figure A3.2 Dissimilarity from the medoid for the sequence plots in Figure 3.1



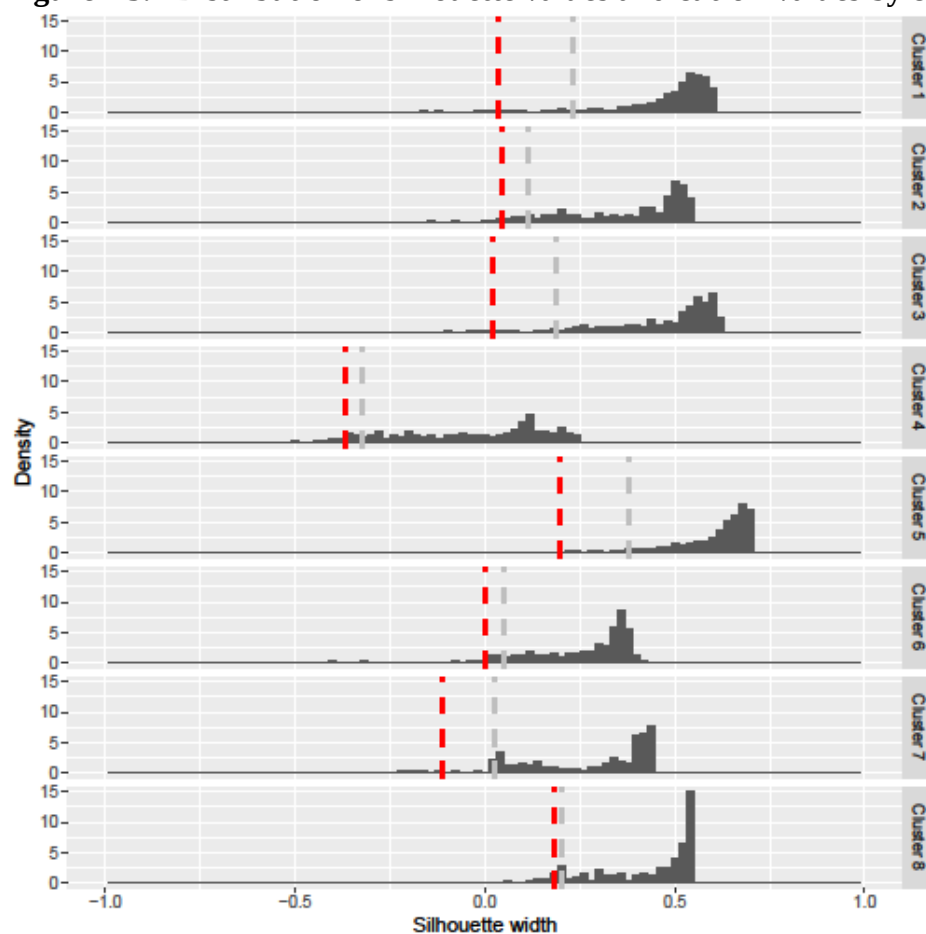
Notes: The dissimilarity from the medoid for each cluster plot in Figure 3.1. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A3.3 Sequence distribution plots, post-reassignment.



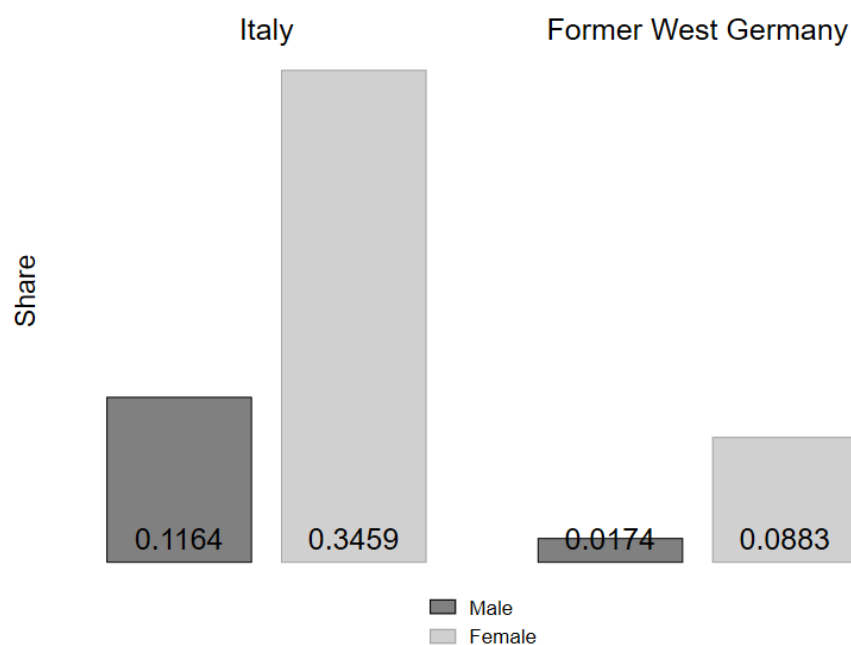
Notes: Cluster membership after the reassignment of single trajectories (see A3.5.5 for details). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A3.4 Distribution of silhouette values and cut-off values by cluster



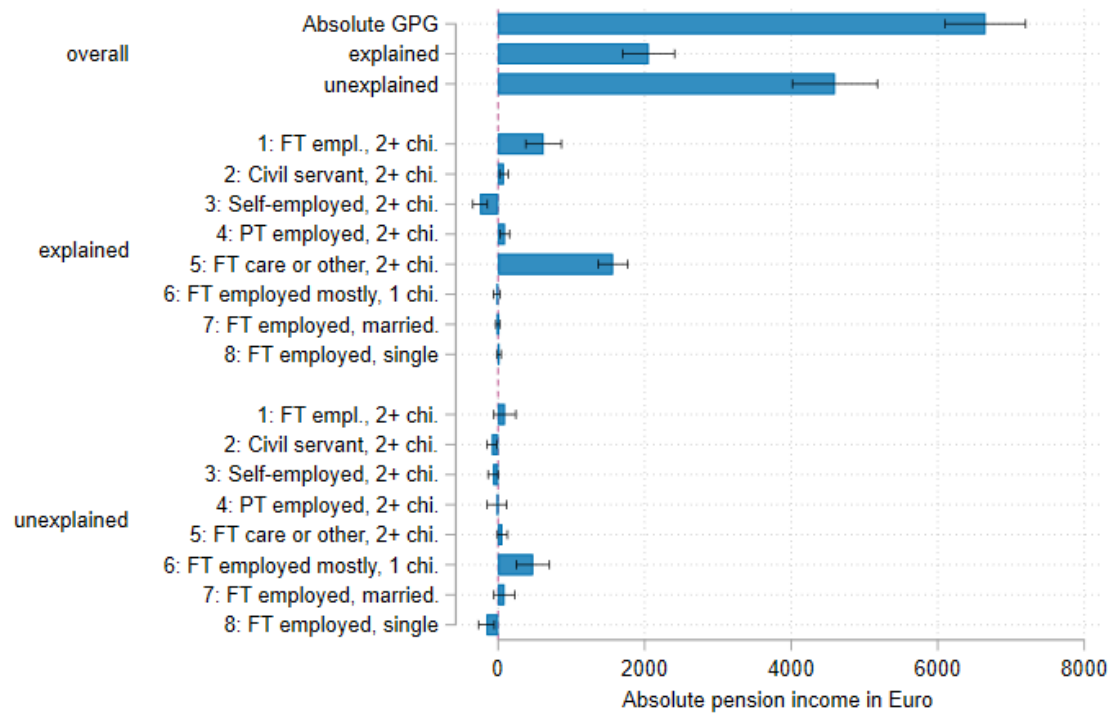
Notes: Red line indicates cluster-specific cut-offs based on the 5th percentile of the silhouette distribution and the grey line based on the 10th percentile. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A3.5 Share of individuals with no pension income by country and gender.



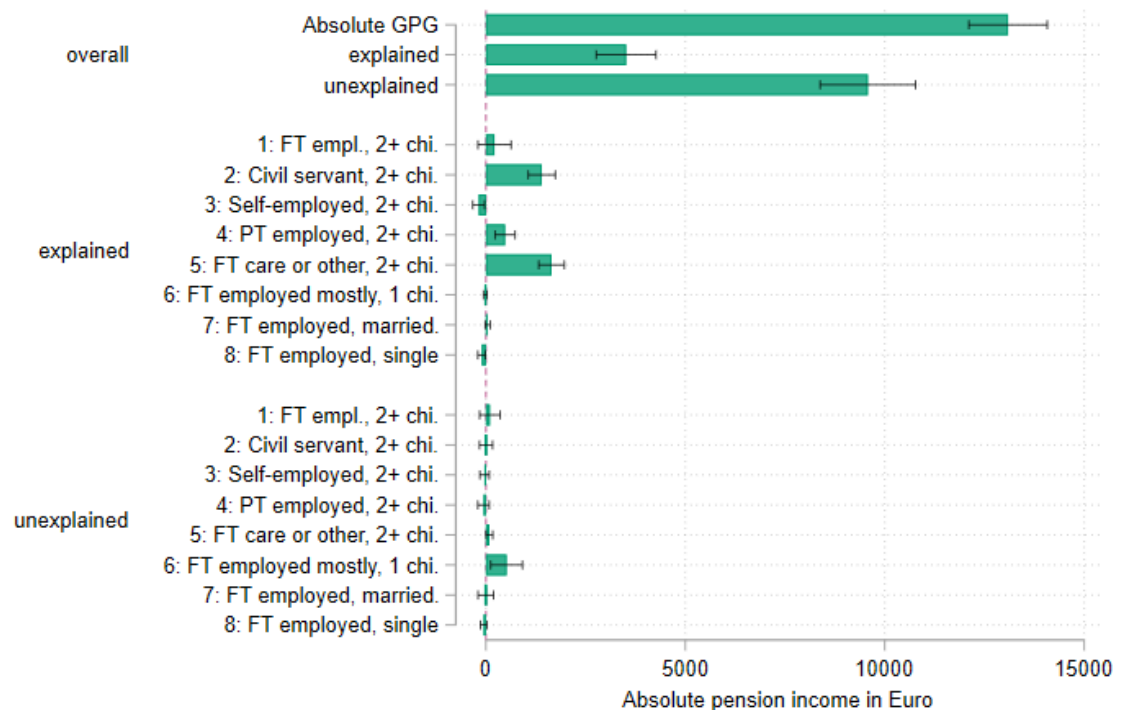
Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A3.6 Confidence intervals (95%) for the decomposition of the GPG in Italy



Notes: Based on normalized linear regression models with absolute independent pension income as dependent variable. Controlled for birth cohort and wave. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A3.7 Confidence intervals (95%) for the decomposition of the GPG in West Germany



Notes: Based on normalized linear regression models with absolute independent pension income as dependent variable. Controlled for birth cohort and wave. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A3.3 Data and sample

By using a threshold of 65 years for the sample definition, we covered the average effective age of labor-market exit (Table A3.2). Additionally, literature in the overall field of old-age income inequality (e.g., Ebbinghaus, 2021), but also on GPGs (e.g., Hänisch & Klos, 2014) has chosen the same sample cut-off. This makes our results comparable to recent literature.

Table A3.4 summarizes all steps of the sample creation and tracks how many respondents were excluded per step.

Imputations of missing retrospective working life courses

The information on the working life courses is based on the self-reported situation or activity in the generated Job Episodes Panel data (Brugiavini et al., 2019). The variable of interest had missing information on 11.56% of all person-year spells, which was highly driven by women's entries. We first filled in the missing information based on several variables in the raw SHARELIFE data that was not used for generating the Job Episodes Panel data. This allowed us to fill most missing person-year spells and left us with 2.17% of missing person-year spells. We then forward filled the last observed situation for those individuals who were only surveyed in the first retrospective survey (Wave 3) up to their latest interview year. Lastly, we imputed few spells on *retired* for missing person-year spells in the reported year of retirement and above and imputed *care work* for those with missing information who stated that they had never done any paid work in their lives. This reduced the share of missing information on the work life courses to 1.1%.

For Germany, we then matched the survey data to register data on pension relevant activities (SHARE-RV) to consider whether there was any pattern for the remaining missing information. We found that for all missing spells that were matched, being employed is the most prevalent state for men (56.8%) and women (40.6%). However, for

both men and women, 29% of the missing spells in the survey data were attributed to noncontributory periods in the register data, which covers any pension-irrelevant activity, most likely homemaking, or unregistered unemployment. Further, in 22.6% of the person-year spells, women were registered as being active in unpaid care and child-raising/homemaking.

Based on this heterogeneous evidence, we assigned all remaining missing spells to the working state covering care work and other states, namely “full-time care work/other/missing” (4.61 % of person-year-spells in this state).

A3.4 Dependent variable

In Wave 6, the survey of occupational pension income did not differentiate between different types of occupational pensions. Wave 5 is the most recent wave with a detailed differentiation between occupational survivor pensions and occupational pensions based on individual’s own jobs. Therefore, we used Wave 5 as the basis for our sample.

The inclusion of individuals aged 65 or older without any own pension income is especially relevant in light of country comparisons, as the unequal access to pensions is highly country-specific and gendered. Whereas 12% (2%) of male respondents in Italy (West Germany) have a pension income of 0, this is the case for 35% (9%) of women (Figure A3.5).

A3.5 Robustness checks and extensions

In this section, we present robustness checks and extensions to our main analyses in which we elaborate on the specificities of the SA-KOB decomposition.

A3.5.1 Different reference coefficients in KOB.

As discussed in the main text, the decomposition results depend greatly on the reference coefficients chosen. When combining KOB with sequence analysis tools, this may lead to further substantive implications, which we discuss with reference to single examples in the following.

Men's coefficients. As expected, using the coefficients from the models for men as a reference for all clusters (M2, Table A3.10) leads to differences in the explained share and returns components for clusters 4 and 5, for which the women's reference coefficients are used in the main model M1. In both Germany and Italy, the explained shares shrink while the returns component increases for cluster 5, thus rendering differential returns more relevant than compositional differences. The interpretation would thus be that a higher share of the GPG is based on differences in returns in the care-work-dominated life courses of parents rather than on the fact that there are much fewer fathers with such poorly rewarded biographies compared to mothers. In substantive terms, we do not believe that this shift in the interpretation provides any relevant added value. Moreover, empirically, it is rather meaningless to choose the men's coefficient as a nondiscriminatory return for the life-course cluster dominated by long-term unpaid care work. This life-course cluster is experienced by very few men. In both countries, men in this cluster differ from their female peers in exactly the characteristics that are typical of the pattern as they spend a much larger share of their life courses in full-time employment and a much lower share in unpaid care work (or other states covered in the category, see Table A3.7 and Table A3.8 for cluster 5). This is also mirrored in the coefficients. Despite the care benefits implemented in both

pension systems, which women in this life course should benefit more from, the higher share of full-time employment apparently leads to much better returns for men (Table 3.6). Additionally, the gender-specific ASW for cluster 5 reveals that women's life courses are, on average, much more consistent with cluster 5 compared to men's (see Figure A3.1). We thus argue that the women's coefficient is a more representative return for this life-course cluster. Otherwise, the share of the gap that is due to the dominance of women dedicating their lives to unpaid care work and the ways in which pension systems reward such lives (including child-care benefits) would be underestimated and skewed by unrealistic returns to such female-typical life courses. When using the men's coefficient as a reference, the difference in returns is weighted by the high prevalence of women in the cluster (the mean) and thus the returns component is inflated.

Women's coefficients. Using the female coefficients as the reference for all clusters leads to an increase in the returns component, while the explained parts shrink for most male-majority or gender-neutral clusters (cluster 1–3, 8). This is because the returns for women differ greatly from those of men for these life courses. Again, this is most likely largely driven by gender-specific within-cluster heterogeneity. For example, the returns of continuously full-time employed mothers are much lower than those of fathers with similar careers (see Table 3.6), and the share of full-time employment among mothers in this cluster is lower than that of fathers (see Table A3.7 and Table A3.8). Even women who managed to combine motherhood with continuous full-time employment spend on average 7–8% of their observed life course on unpaid care work. The average life-course characteristics for men are more in line with the overall characteristics of life-course clusters in cluster 1. Indeed, the gender-specific ASW for cluster 1 reveals that men's life courses fit on average better with cluster 1 (see Figure A3.1). The same pattern is visible in the other life-course clusters, especially for cluster 3 and Italy more generally: Men's average life-course

characteristics more closely resemble the overall life-course cluster in the respective cluster (see Table A3.7 and Table A3.8). As such, their returns to the respective life courses are less likely to be biased by other characteristics therein and thus closer to a nondiscriminatory return for the respective cluster.

From a policy-intervention perspective, it is more interesting to consider which parts of the GPG are due to a lack of women in male-typical life courses that are rewarded based on the male reward structure and not based on a female reward structure that is systematically lower *because* women in this life-course cluster have a higher share of poorly rewarded characteristics. If we assume that as many mothers as fathers of two or more children experienced continuous full-time employment (cluster 1) and mothers were rewarded just like fathers (as we do in our main model), the GPG would drop by €621 in Italy and by €215 (nonsignificant difference) in West Germany (see Table 3.5). The recommendation for policy makers would then be, as we discuss in the main text, to implement policies and labor-market cultures that make the combination of parenthood with continuous full-time employment as accessible for mothers as fathers, if the reward structure for men remains unchanged. However, if the reward structure for women is adopted, the results would only suggest a drop in the gap by €420 in Italy, and no drop at all in West Germany (see Table A3.10). Thus, we would underestimate the explained share, at least partly because women's observed characteristics differ systematically from men's in continuous full-time employment clusters.

Using the female-typical reward structure would thus imply that a higher share of the GPG could be reduced by creating the same pension rewards for similar life courses. In our case, this may partly be achieved by offsetting the observed remaining life-course gender differences within the group of individuals with continuous full-time employment. But we cannot estimate directly how much of this part is due to gender differences in unobserved

characteristics such as pay or gender segregation on occupations and industries. We thus believe that using the female coefficients is less helpful for policy makers because the implications with regard to the larger returns component are less clear.

Pooled coefficients. The decomposition results using coefficients from a regression model pooled over the decomposition groups (M4 in Table A3.10) are close to our preferred option of using cluster-specific reference coefficients based on the majority group (M1 in Table A3.10). This is also the case for clusters with a high level of gender-specific segregation (especially clusters 1, 4, and 5 for both countries), for which the results differ to a larger extent when comparing the models using the male and female coefficients only (M2 and M3 in Table A3.10). This is because the dominant group is likely to also drive the coefficients in the pooled model. The coefficients from the pooled model are very close to the male coefficients for the male majority clusters 1, 2, and 3, whereas they are closer to the female coefficients for the female majority clusters 4 and 5 (see Table 3.6). However, when using the pooled coefficients, it is not clear which group drives the rewarding structure. The strategy of using the coefficients from the majority group in case of high levels of group-specific segregation can be understood as a clearer version compared to using the coefficients from a pooled model. In the pooled model, the reader might be less clear about which group drives the coefficients and what the reward structure means, whereas in our preferred option, it is clear how to interpret the returns since we know what the returns mean.

Overall, this robustness check reveals the importance of the reference coefficients as parameter for the interpretation of the decomposition results. Even though our main results are robust generally, the magnitude of explained and returns components differ across different options for the reference coefficients used. When combining SA with the KOB, we argue to select the reference coefficients based on theoretical and empirical

considerations. We suggest opting for the coefficients of the majority group, identified based on the suggested guidelines in section 3.3.2.1 (especially in the cases of extremely group-specific clusters) to achieve the most empirically robust and substantively relevant interpretation of the KOB.

A3.5.2 Replicating standard KOB: Using summary proxies of life courses.

To highlight the theoretical, empirical, and methodological contribution of the SA-KOB decomposition compared to the standard decomposition approach in the literature, we replicate the KOB decomposition using one-dimensional summary indicators as proxies for family and work life courses (the “standard” model). For comparability reasons, we use the same life-course states as used for the SA but focus on the life-course states that are used most often as main covariates of interest in the standard decomposition approach (see Table 3.1): the number of years spent in full-time employment, part-time employment, self-employment, the public sector, and unemployment. Given that one of our main results is that continuous engagement in unpaid care-work of women drives the GPG, we also would like to include the duration of care work, which has been only included in one of the previous decomposition studies (Frommert & Strauß, 2013). However, including the duration of care-work increases the multicollinearity in the gender and country specific regression models underlying the KOB (variance influence factor of up to 15.8). We thus must exclude it and focus on the other work life-course summary measures. Given that multicollinearity still reaches critical levels (variance influence factor of up to 7.8), we further exclude the duration of self-employment. This leads to more acceptable levels of multicollinearity, but they still slightly exceed the level of the regression models using the clusters of our main model containing all life-course states (2.59 compared to a maximum of 2.35 for the regressions underlying the SA-KOB). We proceed similarly for the selection of family life-course summary indicators and include the duration spent in the following

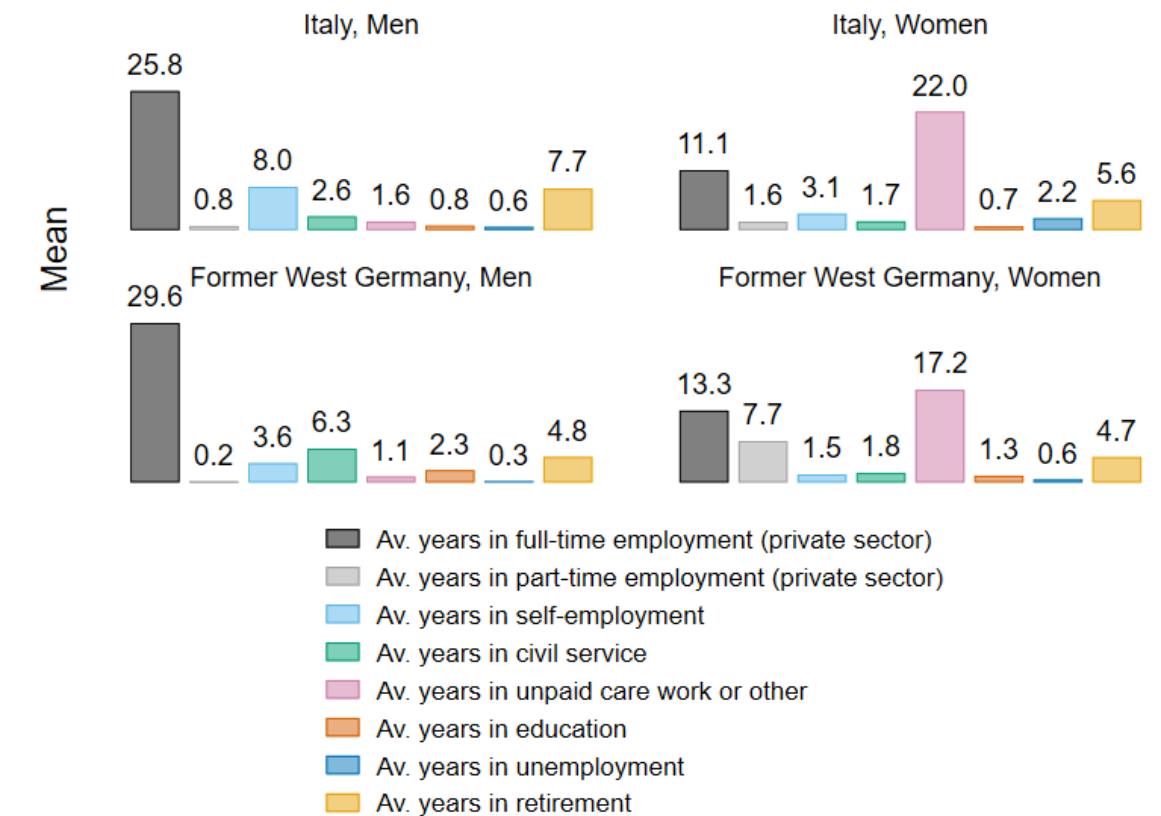
states: being married and having 2 children, being married without having children, being divorced, and being single (without having children) without increasing the multicollinearity further. We refrain from including any interactions between work and family life courses to ensure our replication is as similar as possible to the standard decomposition model and does not exacerbate the multicollinearity concerns even further.

For the KOB decomposition, we apply the same parameters as in our main analysis to ensure comparability (choosing reference coefficients based on the majority group and using the coefficient from the women's regression for the duration in part-time employment and the coefficients from the men's regressions otherwise and controlling for wave and cohort). The standard approach confirms many findings from the main model. In the next paragraph, we focus on describing the most relevant differences between the results of both modeling approaches.

By using continuous variables for the duration in different states, we explore more heterogeneity in the data in the standard decomposition than when using the categorical life-course clusters as covariates. As a result, the explained shares of the implemented life-course states are at times higher (Table 3.4). For example, the duration spent in full-time employment explains almost one quarter of the GPG in West Germany, whereas the explained shares of continuous full-time employment careers parallel to parenthood in the SA-KOB main model (Table 3.5, cluster 1) is not significant. Thus, the standard model suggests that women's pension income would increase substantially if they spent on average the same years in full-time employment as men, while the SA-KOB decomposition points out that this is most likely rather driven by mothers who spent particularly few years in full-time employment before taking up unpaid care work for a long period, if not the rest of their life courses (clusters 4 and 5). As such, the SA-KOB highlights in a more accurate way the dynamics behind women's fewer years in full-time employment. The continuous

unpaid care-work provided by these mothers upon family formation prevented them reaching as many years in full-time employment as fathers after family formation.

Figure A3.8 Average years spent in work life-course states from age 18–65.

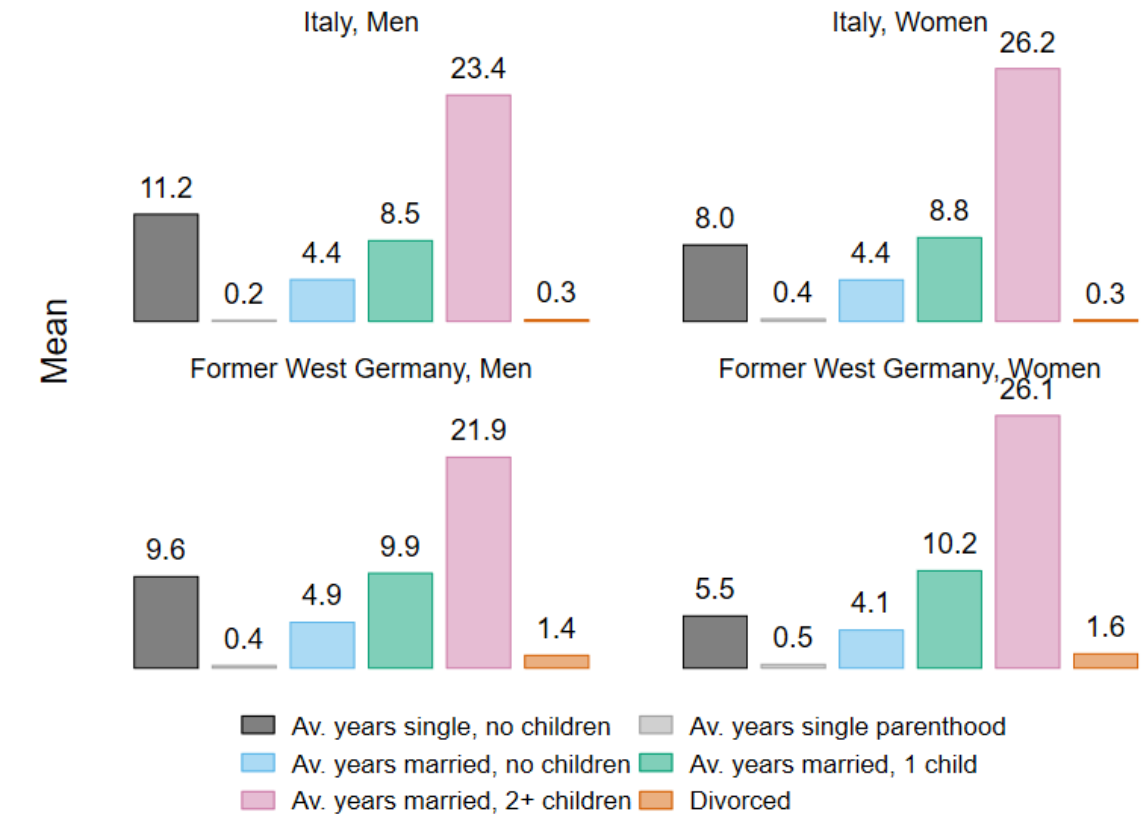


Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

However, the standard approach reveals associations between less common life-course states (e.g., unemployment, or divorce) and the GPG that were concealed in the SA-KOB decomposition. In Italy, a share of the GPG is due to the greater number of years women have spent on average in unemployment (Table A3.7, Figure A3.8). However, this share is cancelled out by better pension income returns for women per year being unemployed. In West Germany, an analysis of the duration in divorce reveals that the GPG is reduced due to better pension income accruing to women per year being divorced compared to men. This is an important finding, which confirms recent studies on divorce being a gender-equalizer in Germany (Kreyenfeld et al., 2023) that is not visible in the SA-KOB due to

small shares of being divorced spread over several clusters (1, 2, and 6 predominantly, see Figure A3.3).

Figure A3.9 Average years spent in family life-course states from age 18–65.



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

In both countries, similar shares of the GPG are due to lower pension returns for each additional year spent in education by women compared to men. In West Germany, a substantial part of the GPG is also due to women’s fewer years spent in education after age 17. Even though the life-course clusters differ regarding the length of educational periods, these differences by years in education are masked by more dominant patterns (for example, cluster 2 contains most of the life courses with long educational periods, while the main characteristic of cluster 2 is continuous civil service careers of parents). On the other hand, however, the SA-KOB shows visually that long educational spells are often associated with continuous civil service. Therefore, it captures the synergy of these two favorable life course aspects (long education and continuous civil service) that generate particularly high

pension income (especially in West Germany) and large gender differences in pension income following the very limited access of mothers to these favorable life courses. In West Germany, the standard approach explains higher shares of the gaps in pension income. This is not surprising because the metric variables pick up more heterogeneity than the categorical life-course clusters. The life-course clusters still include some within-cluster heterogeneity, particularly with regard to the duration in the dominant states within the clusters. Indeed, some of the associations revealed by the standard approach are of substantial interest and do not emerge in the SA-KOB (e.g., differences in returns to educational years). However, in some cases such associations are likely to obscure more complex mechanisms that only become visible in the SA-KOB (e.g., the association with i) long education spells and civil service, and ii) short full-time employment spells and long unpaid care work). Additionally, in Italy, the standard decomposition approach explains less of the GPG than the SA-KOB, because we could not consider the duration in care work due to multicollinearity.

However, the aim of this study is not to maximize the share of the GPG that can be explained by the data but to reveal how life-course dynamics in different domains are associated with the GPG. In this vein, explaining a higher share of the GPG might not help identify the association with life-course complexities relevant for the outcomes. In this respect, the standard approach fails in some respects where the SA-KOB contributes substantively instead. The benefits and limitations of both approaches are listed in Table 3.7.

First, the standard approach does not uncover the role of the interaction between family and work life courses. Only the SA-KOB decomposition (using MSA) reveals that the GPG is almost exclusively driven by the life courses of parents of at least 2 children. Additionally, the standard approach does not capture the theoretically important notion of linked lives,

realized through the temporal intersection of work and family life courses – an important characteristic of female life course clusters – and its negative relation with GPGs. The standard approach does not make the link between life-course patterns from different domains visible and show how these are rewarded (by pension systems in our illustration). Second, the standard approach is more prone to multicollinearity, meaning that not all life-course states can be considered in the model. It is simply impossible to consider durations in all life-course states that are relevant at the same time. As a result, research using standard KOB must fall back on models that will obviously miss some of the associations. SA-KOB, by contrast, allows researchers to include a multitude of different (categorical) life-course states and thus efficiently circumvents multicollinearity issues. As a result, standard decomposition conceals the important finding that gender differences in engaging in unpaid care work strongly drive the GPG. This would most likely have been visible if we would have included the duration spent in care work, but due to severe multicollinearity issues, we would have to exclude the duration in full-time employment at the very least, which has been the covariate used most often in the standard approach. In Italy, this results in a lower total explained share of the gap due to differences in life-course characteristics compared to our main model—even though we exploit much more heterogeneity in the data when using continuous measures instead of the eight categorical clusters.

Third, related to the previous point, the selection and combination of the life-course aspects (and potentially their interactions) to be considered in the model are not straightforward and depend largely on the researcher. For example, in this application we decided to drop the duration in care work (and self-employment) instead of the duration in full-time employment due to high multicollinearity. SA-KOB does not easily lead to such difficulties because more life-course states (up to a limit) can be considered in all their complexities.

Fourth, life course summary measures are likely to conceal how these different life-course states unleash their associations with inequalities in certain interactions and synergies across the life course. For example, in West Germany, a strong GPG driver is the privileged life-course pattern characterized by long educational spells followed by continuous civil service work. In the standard approach, we see that both the years men spent in education and working in the civil service are associated with the GPG, but not that their inequality driving mechanisms most likely largely unfold jointly.⁷⁰ Another aspect that the standard model obscures is that part-time employment often follows family-related employment interruptions (cluster 4). It thus does not show the association of the GPG with mothers' life courses, in which part-time employment mostly starts after childbirth, because mothers need to reconcile family care and their careers. This is especially relevant in the German case, given the increasing share of female (particularly maternal) part-time employment (Schmitt & Auspurg, 2022).

A3.5.3 Sensitivity check for the impact of the heterogeneity within the same life-course cluster

Following Jalovaara & Fasang (2020), we apply a robustness check that excludes the poorly assigned individuals from each cluster respectively and therefore leads to higher cluster coherence. First, this check reveals whether the results are overall robust to within-cluster heterogeneity. Second, it uncovers whether the heterogeneity is group specific and in which direction the decomposition results might be skewed by it. This can help to better understand the returns component of the unexplained shares. For the explained shares,

⁷⁰ Still, the sum of the explained shares due to the years spent in education and the civil service in the standard approach more than doubles the share due to cluster 2 (continuous civil service and long education) in the SA-KOB decomposition. Again, this is largely because the continuous duration proxies explore more heterogeneity in the data (e.g., differences in duration of educational spells in cluster 2) and consider any years spent in these states independently of family life and how continuous or disruptive the states are. We believe that this is a good example for highlighting that explaining more variance does not always mean that the most important mechanisms related to the outcome of interest are revealed.

excluding poorly assigned individuals potentially results in “more representative” reference coefficients with regard to the cluster characteristics used for weighting the mean differences in characteristics. For the returns component, it aims to reduce the gender-specific heterogeneity within the same life-course cluster, making the “endowment” with the same life-course cluster more similar between groups and thus revealing whether parts of the returns component of the unexplained shares are due to group-specific differences within the cluster.

We exclude individuals with a low silhouette value, which is a proxy for each life courses’ fit with the overall cluster (Kaufman & Rousseeuw, 1990). So far, SA applications have used absolute cut-offs, for example by excluding all life courses with negative silhouettes (e.g., Jalovaara & Fasang, 2020). However, we suggest applying relative instead of absolute cut-off values. In applications with clusters that are differently heterogeneous, the distribution of silhouette values will differ largely by cluster. In our case, cluster 4 merges different atypical employment patterns, which is of high theoretical relevance in this illustrative application but leads to a low average silhouette width (ASW) value of this cluster (Figure A3.4). If we apply an absolute cut-off of 0, we exclude about half of the cluster from our analyses. The selection of which cut-off to use depends on the level of heterogeneity in the cluster-specific ASW in each application. Where there are large differences in the ASW between clusters, we suggest applying cut-offs based on percentiles of the cluster-specific silhouette distribution.

Excluding all life courses with silhouettes below the 5th percentile of each cluster respectively, we find that the results remain the same for both countries (Table A3.11).⁷¹ Even when using a more conservative cut-off (10th percentile), the main results are robust.

⁷¹ We do not exclude those trajectories that have been reassigned to the cluster post-MSA to avoid systematically worsening the common support problem in the robustness check.

Overall, the returns components are not largely driven by poorly assigned individuals. However, in our application, the gender-specific within-cluster heterogeneity regarding the work dimension only slightly decreases in Italy (not shown, available upon request). The average women's characteristics slightly approach the mean characteristics of men in cluster 6 (higher share of full-time employment, lower share of care work) after excluding poorly assigned individuals, and the size of the returns component decreases. This suggests that minor parts of the returns component are due to the gender differences regarding work life characteristics in cluster 6. Because the average gender differences in (employment) life-course characteristics are still large between men and women in both countries after excluding poorly assigned individuals, the results highlight the notion of persistent gender differences in life-course aspects, even when experiencing *similar* life-course clusters (returns component of unexplained parts). It is likely that the gender-specific within-cluster heterogeneity drives parts of the returns component.

Excluding poorly assigned life courses leads to larger explained shares for most clusters, because the reference coefficients have larger effect sizes (not shown, available upon request)—they are less impacted by life courses that do not share the main characteristics of the cluster. Thus, the main model tends to underestimate the explained parts but only to a marginal degree.

A3.5.4 Additional variables in the decomposition models.

We adjust the main models by adding variables to ease the potential omitted variable bias. The results for the baseline model differ slightly from our main model, due to the exclusion of individuals with missing information on the controls (Table A3.12, M7).

First, we use a variable categorized by using a more fine-grained definition of birth-cohorts to account for significant shifts in the pension system, most importantly the privatization of pensions. The results do not change.

Second, we adjust for the educational level (categorized based on ISCED). However, it is important to note that to some extent educational differences might be endogenous to the work-family life courses as they are somewhat accounted for in the life-course clusters themselves. For example, civil servants have longer periods in education than other life courses. Indeed, we find that the explained and returns components due to civil servant careers in both countries are partly driven by educational differences (cluster 2). In Italy, both shares for civil servant careers become insignificant when controlling for education. In West Germany, the explained share drops from €1,391 to €1,130. Additionally, in both countries, parts of the explained shares due to parents with working life courses dominated by self-employment (cluster 3) and unpaid care work (cluster 5) are explained by differences in education. However, all these shares remain significant, and the main substantive interpretation is unchanged. Indeed, the educational differences between men and women are likely to unfold over the life courses and are thus captured in our covariate of interests, the work-family life-course clusters. This seems to be the case, as educational differences capture some of the variation in the GPG associated with life-course clusters. 4.6% (8%) of the GPG in Italy (West Germany) is due to more women only having up to lower secondary education and 1.7% (7.8%) is due to more men having tertiary education. In West Germany, the returns components for educational level are likewise significant. Women would receive €893 pension income less per year if having up to lower secondary education would result in as low pension income rewards for women compared for men.

Third, we adjust the model by using a variable that accounts for whether the respondents were born in Italy or Germany, respectively. Next to other inequalities that migration background might create, pension income could differ between natives and those with migration background since the latter might start contributing to the Italian or German pension system later in life and thus for a shorter period. However, we find no significant

association with migration background and the GPG, and our results do not change, except for the returns component of cluster 5 in West Germany, which becomes insignificant.

Other potentially omitted variables cannot be adjusted for in the models due to data limitations, such as age at retirement and early childhood circumstances.

A3.5.5 Common support.

If the SA reveals high group-specific clusters, the lack of common support for the KOB analysis might become an issue. In the MSA analysis, we find strong gender-specific segregation for some of the life-course clusters. To some extent, such segregation is expected, especially in the cohorts of individuals in our sample. Yet, the degree of segregation is an important empirical finding and points to extreme inequalities between male- and female-typical life courses. In particular, the care- and part-time dominated life courses (clusters 4 and 5) are almost exclusively experienced by women. At the same time, this important empirical finding raises challenges for KOB decomposition in the second step of our analysis. Because the indicators of clusters 4 and 5 take the value of zero for almost all men, we run into a lack of variation for men in this variable. This empirical finding is a special finding for our illustrative application. Applications with a lower level of group-specific life courses will not run into this empirical challenge when using SA-KOB.

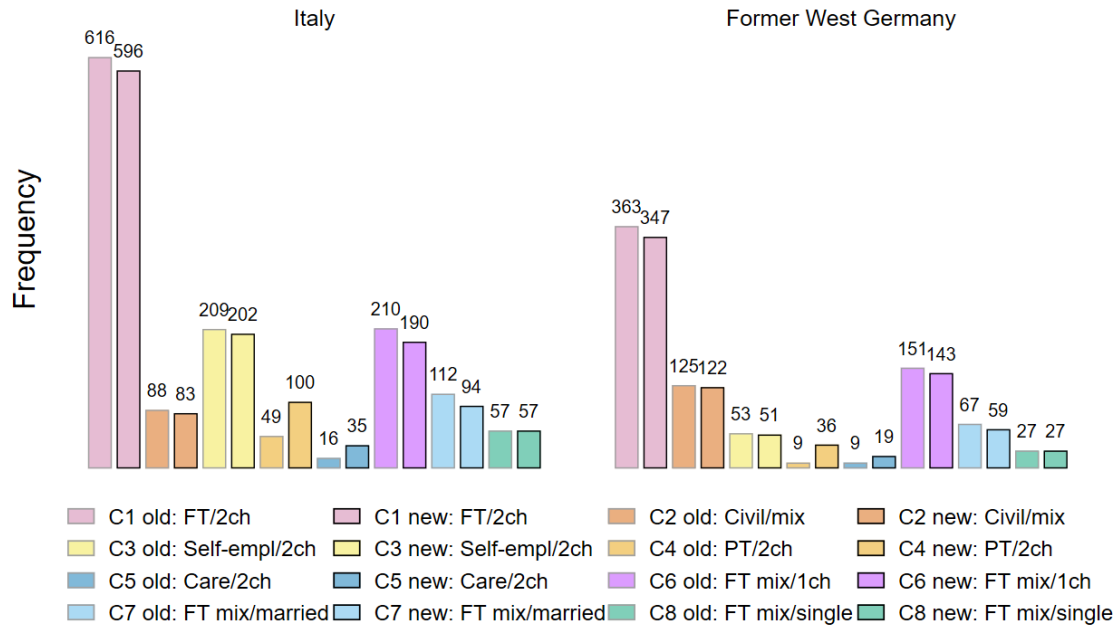
There is no obvious methodological solution for the lack of common support in our case, since the underlying gender-specific segregation systematically leads to a lack of data points for men in the care cluster. Among some candidates for (second-best) work-around strategies, we opted for the following, which enables us to perform a substantively meaningful KOB decomposition despite the strong segregation patterns found in the

MSA.⁷² This strategy implies a manual relocation of some men's life courses from other clusters in the female majority clusters after having identified the clusters in the MSA step. For this approach to be meaningful, we need to assume that the relocated observations are a sufficiently good comparison group for the female care dominated life courses. First, we identify men that are not well assigned to their cluster using low relative silhouette values per cluster (lowest 10% of the country-specific within-cluster silhouette distribution). We then combine this information with whether the life courses of the men contain characteristics that are close to the working characteristics of the female majority clusters. For the care cluster 5, we consider a minimum sum of years of having had care, other or missing periods. For the cluster 4, the part-time cluster, a minimum duration in part-time work, unemployment, or retirement (as an indicator for early retirement) is considered.

However, this combined approach alone does not identify enough cases for both clusters. Thus, we also consider men in the other six clusters who spent most the years in 'care work/other/missing' or part-time employment (i.e., we identified men within the highest 2–3 percentiles of years in the specific state, e.g., care work). We further did not reassign men from the singlehood cluster, cluster 8, given the low number of observations in this cluster. With both approaches combined, we identify almost all men that were assigned to the women-majority life-course clusters originally by the cluster algorithm.

⁷² Other options were (i.) joining the care cluster with a very similar cluster, (ii.) duplicating the very few male observations in the care cluster to reach a sufficiently stable base for estimation or (iii.) dropping the observations of very segregated clusters (4, 5) in the KOB analysis. Whereas the first option was not empirically feasible in our application, the third option would have required us to correct the selection bias. The application of a Heckman model would have resulted in further difficulties. We advise against the second option, because it would artificially increase the weight of rather selective observations in the original data. Moreover, this approach also leads to strong dependencies in the sample of duplicated men.

Figure A3.10 Distribution on work-family life courses for men by country, before and after reassignment (frequencies)



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

As a robustness check we show the results using the original clusters, that is before the reassignment of men's cases. The model including the original cluster, (Table A3.13, M11), shows that the main results remain highly consistent. However, in the clusters most cases were reassigned to or taken from (clusters 1, 2, 3, 4, and 5, see Figure A3.10), the explained shares increase and the returns components decrease slightly. The explained shares increase because we reassign some men's life courses from the male-majority clusters 1 and 3 to the female-majority clusters 4 and 5, which leads to an overall decrease of the gender-specific segregation over these life courses. The returns components increase due to more gender-specific differences in the life courses that are the result of the reassignment of life courses from other clusters too and thus drive gender differences in returns for the same cluster. Two returns components in Italy only become significant after reassignment: the one for the self-employed cluster 3 as well as the female majority cluster 5.

Using weights (M12 in Table A3.13) does not change the results substantially, with the exception that the returns component for cluster 5 in Italy becomes insignificant once

weights are applied. In West Germany, the returns component for cluster 5 only becomes significant when using weights, whereas the one for cluster 6 becomes insignificant.

A4 Appendix for Chapter 4

A4.1 Additional Tables

Table A4.1 Family policies, gender norms and outcomes across West Germany

		West Germany	Sources
Family policies	Publicly funded and regulated childcare available for children by age 3	1996	(Van Winkle, 2019)
	Childcare rate in daycare centres for children aged 0-2 in 1991	4%	(Kreyenfeld & Geisler, 2006)
	Childcare rate in daycare centres for children aged 3 in 1991	51%	(Kreyenfeld & Geisler, 2006)
Norms	Percentage of people that agree or agree strongly that a toddler is likely to suffer if his or her mother works (1992)	74%	(Kreyenfeld & Geisler, 2006)
Outcomes	Female labour force participation rate 1973	34.3%	(Fischbach & Eckert, 1979)
	Female labour force participation rate 1989	89%	(Rosenfeld et al., 2004) based on (Klammer et al., 2000)
	Gender wage gap (median) for full-time employees (1998)	21%	(Hammerschmid & Rowold, 2019a)

Table A4.2 Pension policies and indicators for West Germany

Pension characteristics in 2012		West Germany	Source
Public/state pension			
Access	Minimum contribution length	5 years	OECD 2013
	Age for pension qualification/ min pension age	Pensionable age: 65 for men, 65 for women Av. Effective age of labour market exit: 62.1 for men, 61.6 for women	OECD 2013
	Share of 65+ population receiving no independent pension income	1% of men, 7% of women	Own calculation
	Level of linkage to earnings (Progressivity Index)**	25.4	OECD 2013
	Maximum amount as % of average earnings	150%	OECD 2013
Amount	Child care benefits	1 pension point per child born before 1992; parallel employment allowed; standard level (average earnings)	OECD 2013
	Other care benefits	Increase of pension by 21.22-65.28€ for elderly care	Lodovici et al. 2010
	Divorce benefit	Pension rights added up and split for years of marriage equally between partners	Horstmann et al. 2009; Kreyenfeld 2018
	Gender Gap in replacement rate***	6.5	Eurostat (2018); European Institute for Gender Inequality (2018)
Occupational & personal pensions			
Access	Share of occupational pension	31.19% of total 39.76 of men 21.69 of women	Own calculation
	Share of personal pension	5.03% of total 6.88% of men 2.98% of women	Own calculation

Table A4.3. Sample selection

	Men		Women	
Base sample age cutoff	1475	(50.0)	1475	(50.0)
Non-response pension	1274	(49.5)	1302	(50.5)
Employment	1182	(48.8)	1241	(51.2)
Sample incl. retro data	804	(49.7)	815	(50.3)
Sample incl. SHARE-RV data	670	(52.6)	604	(47.4)

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A4.4 Waves, survey years and n of observations (final sample)

Wave	Survey year	Men	Women
2	2006-2007	26	26
4	2011	62	41
5	2013	500	435
6	2015	82	102
<i>Total</i>		670	604

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A4.5 Pension types included and definition of dependent variables.

DV I: Public pensions only	DV II: Private pensions only	
<i>Public pension</i>	<i>Occupational pension</i>	<i>Personal Pension</i>
1. Old age pension 2. Old age supplementary pension or public old-age second pension 3. Early retirement or pre-retirement pension 4. Main or secondary disability insurance pension, or sickness benefits	1. Old age pension from the last, second or third job 2. Early retirement pension 3. Disability or invalidity insurance	1. Average payments of regular life insurance 2. Regular private annuity or private personal pension

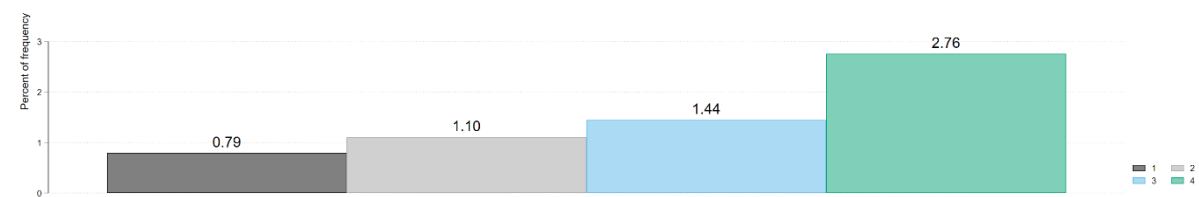
Table A4.6 Baseline coefficients from linear regression models for public and private pension income

	Public Pensions		Private Pensions	
	Men	Women	Men	Women
Cluster 1: High-income FT employed, 2+ chi.	2961.2	3463.7	3481.3	5535.5
Cluster 2: Low-medium income FT employed, 2+ chi.	-1606.2	-227.6	-1439.4	-321.6
Cluster 3: Civil servant, 2+ chi.	15139.8	13413.1	-3037.1	-1312.9
Cluster 4: Self-employed, 2+ chi.	-4327.5	-5674.2	352.1	-389.7
Cluster 5: Pt or non-contributory employed, 2+ chi.	-2895.6	-2754.9	-747.6	-487.2
Cluster 6: FT care or other, 2+ chi.	-4435.7	-6226.5	81.50	-1263.8
Cluster 7: FT employed mostly, 1 chi.	1524.1	-1257.8	778.3	-673.6
Cluster 8: FT employed mostly, married	149.0	988.1	1417.6	-578.8
Cluster 9: FT employed mostly, single	-3336.7	-3921.7	-154.0	-465.0
Cluster 10: FT employed mostly, divorced	-3172.4	2197.8	-732.8	-42.98
co2038	734.0	269.9	434.5	29.79
co3944	378.0	-305.1	196.4	-236.1
co4550	-1112.0	35.18	-630.9	206.4
N	1274		1274	

Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

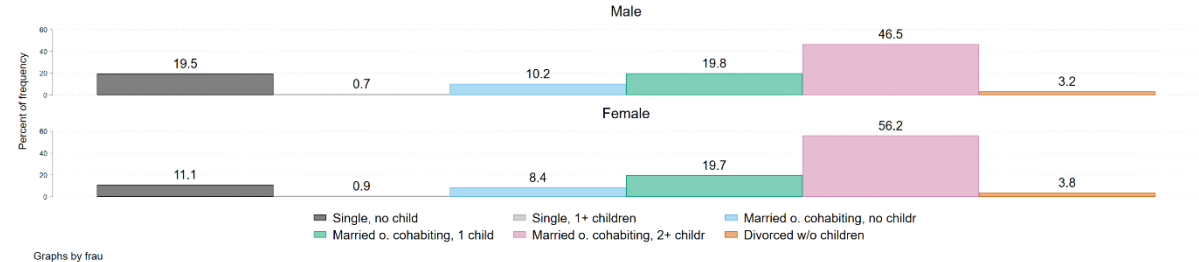
A4.2 Additional Figures

Figure A4.1 Earning point quartiles



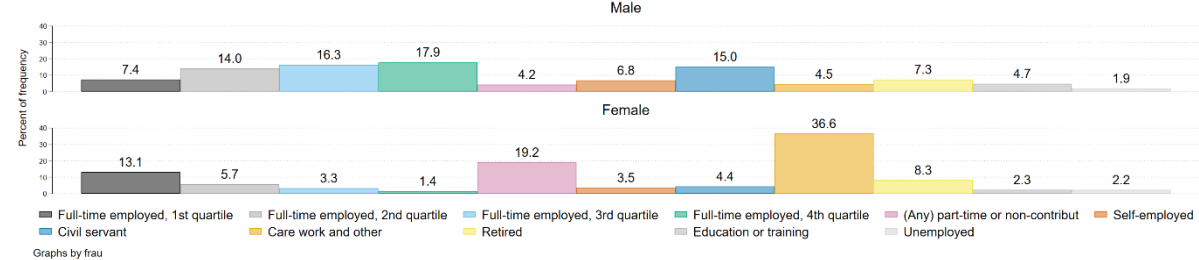
Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.2 Distribution on family states by gender, person-year spells



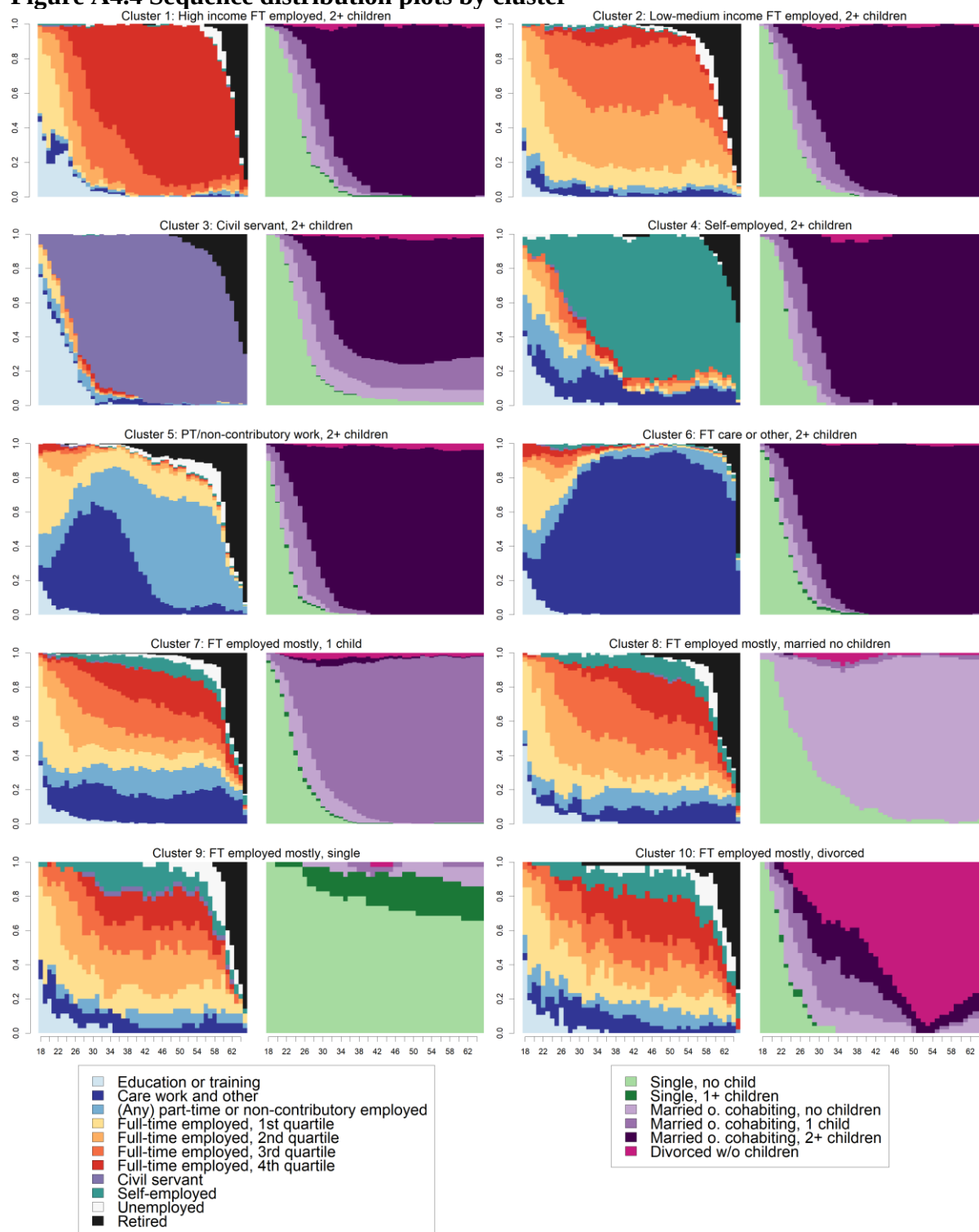
Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.3 Distribution on working states by gender, person-year spells



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.4 Sequence distribution plots by cluster



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.5 Original distribution on clusters by gender



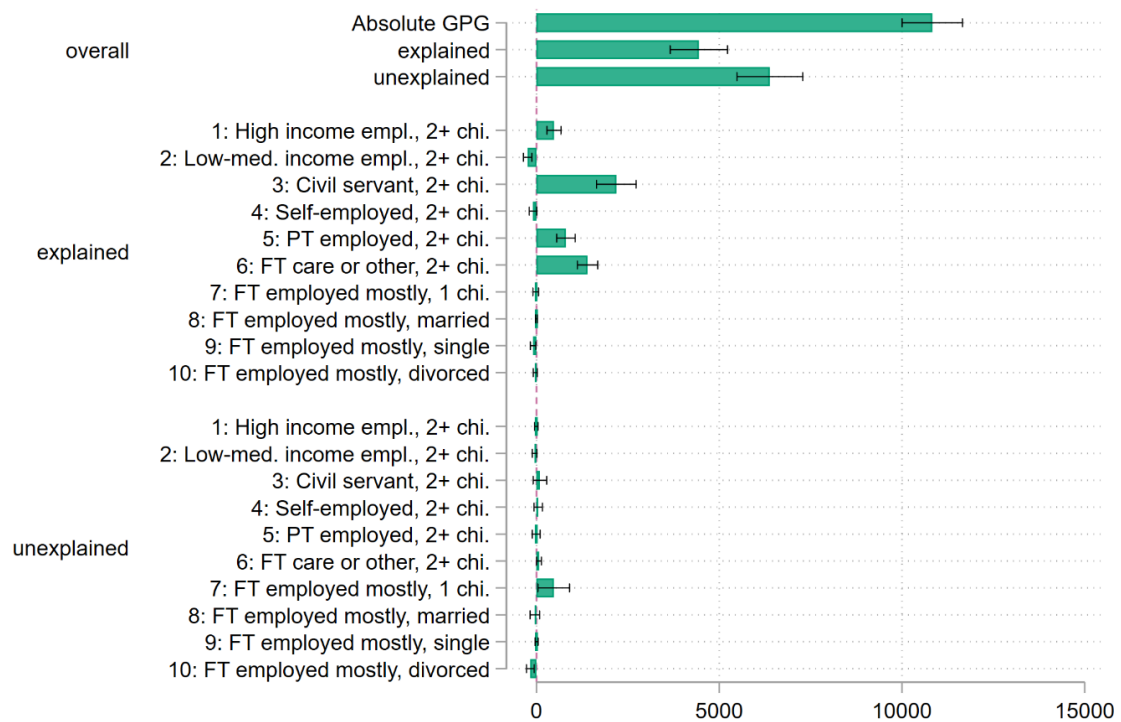
Notes: Not used for decomposition. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted

Figure A4.6 Final distribution on clusters by gender used for decomposition



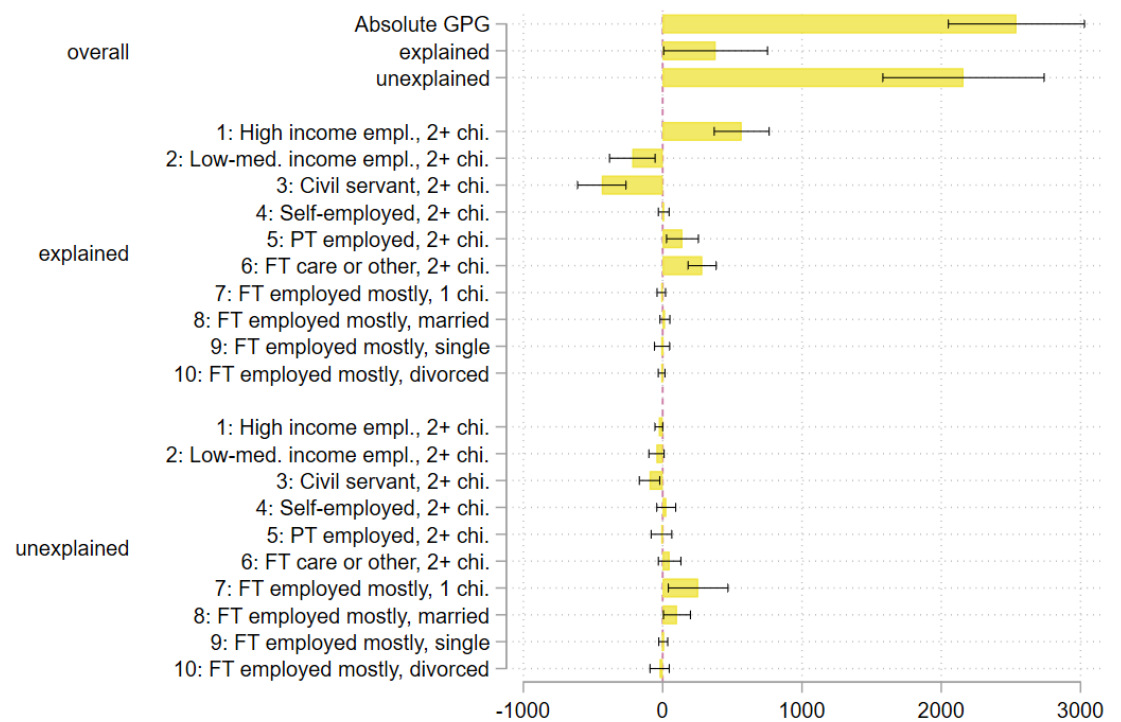
Notes: Used for decomposition. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted

Figure A4.7 Decomposition results with confidence intervals, public pension income



Notes: Based on normalized linear regression models with absolute independent pension income as dependent variable. Controlled for birth cohort. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.8 Decomposition results with confidence intervals, private pension income



Notes: Based on normalized linear regression models with absolute independent pension income as dependent variable. Controlled for birth cohort. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.9 Gender-specific distribution on work-family life course clusters by private pension receipt



Notes: Right panel used for decomposition M4, Table 4.3. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted

A4.3 Details on analytical extensions

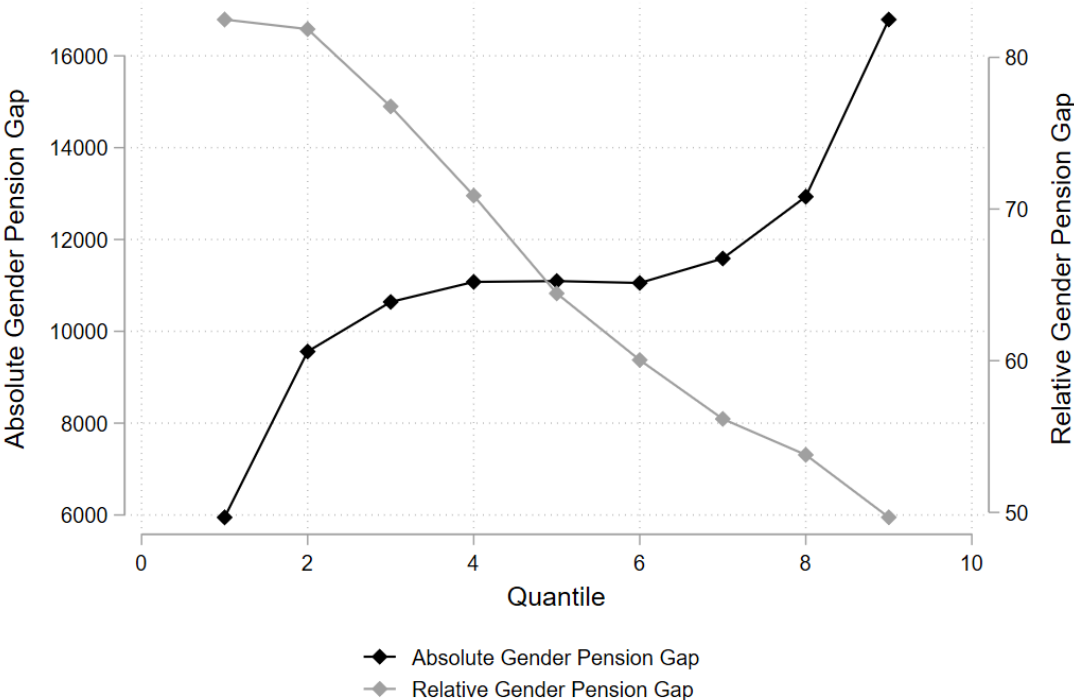
A4.3.1 Decomposition along the distribution of public pension income

To assess whether the association of the life course clusters and the GPG differs across the pension distribution I extend the analysis at the mean with a decomposition of public pension along the distribution in the main text. I apply the RIF (recentered influence function) decomposition which reveals the explained and returns shares for each characteristic separately similar to the KOB (Firpo et al., 2018). I use the strategy closest to the KOB decomposition applied for the main analysis that estimates separate RIF regressions for each group and decomposes the GPG on the selected percentiles based on these (Rios-Avila, 2020). I apply the decomposition parameters from the main model where possible.

In the main part of the manuscript, I use quartiles instead of deciles for the RIF decomposition to limit the bias due to the low number of observations per gender and cluster. Here, I show the results using deciles to get a better sense of how the associations differ beyond the quartiles. However, it is important to keep in mind that the number of observations per cluster is partly based on a few single observations and is not sufficient for presenting robust results.

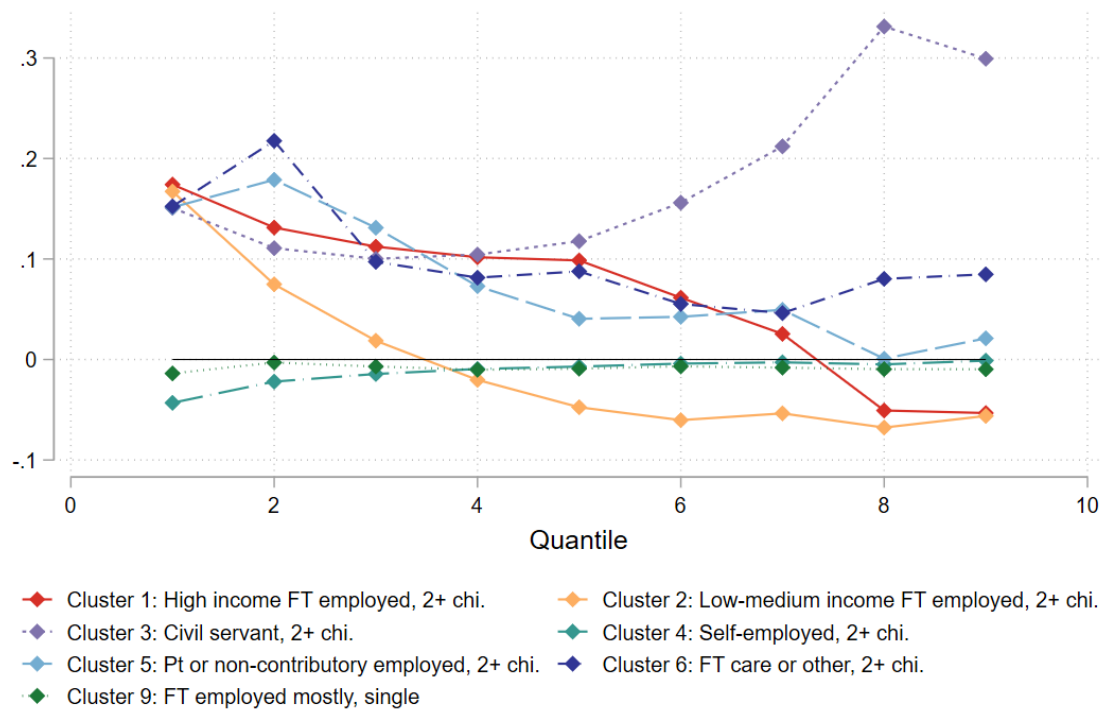
Overall, the results described in the main manuscript are confirmed when using deciles instead of quartiles. While the absolute gap between men’s and women’s pension income increases along the distribution from €5,949 to €16,791, the relative gap decreases from 82 at the first decile to 50% at the 9th decile (see Figure A4.10).

Figure A4.10 Gender gap in public pensions over the distribution, quantiles



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

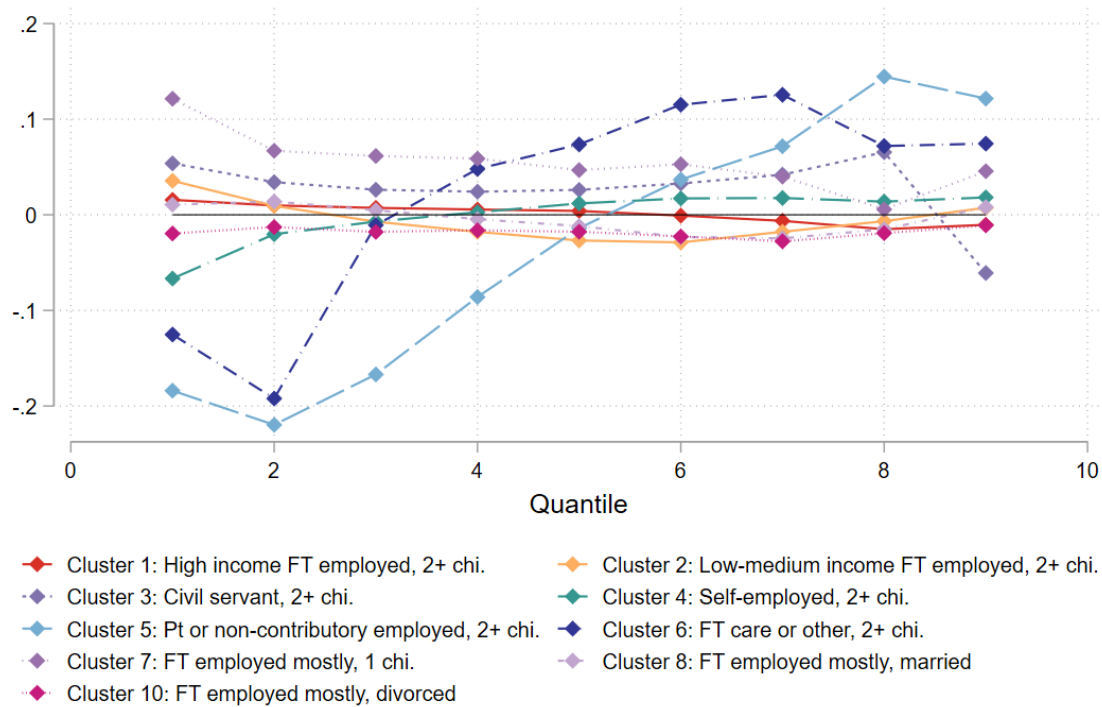
Figure A4.11 Decomposition over the deciles: selected explained shares



Notes: Results from RIF decomposition. Based on normalized models with absolute independent pension income as dependent variable. Only life course clusters with significant explained shares for at least one decile are depicted. Explained shares are shown in relative terms. Coefficients from male baseline models were used as reference. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The share of the gap that is explained by the dominance of mothers in life courses characterized by unpaid care work or care-related employment interruptions is highest at the first and second deciles (Figure A4.11). Whereas cluster 6 (continuous care work) remains important along all deciles, part-time employment of mothers after care-work-related employment interruptions does not play a major role at the very top of the distribution (8th and 9th decile). Additionally, similar shares of mothers compared to fathers in stable, but low-income employment careers (cluster 2) would be particularly important for decreasing the GPG at the very bottom of the distribution (17% of the GPG at the 1st decile). Access to civil service for women and similar returns for it compared to their male peers leads to the largest decrease of the GPG at the top of the distribution (explained and returns component jointly: 31% of the GPG, not shown),

Figure A4.12 Decomposition over the deciles: selected shares due to returns component

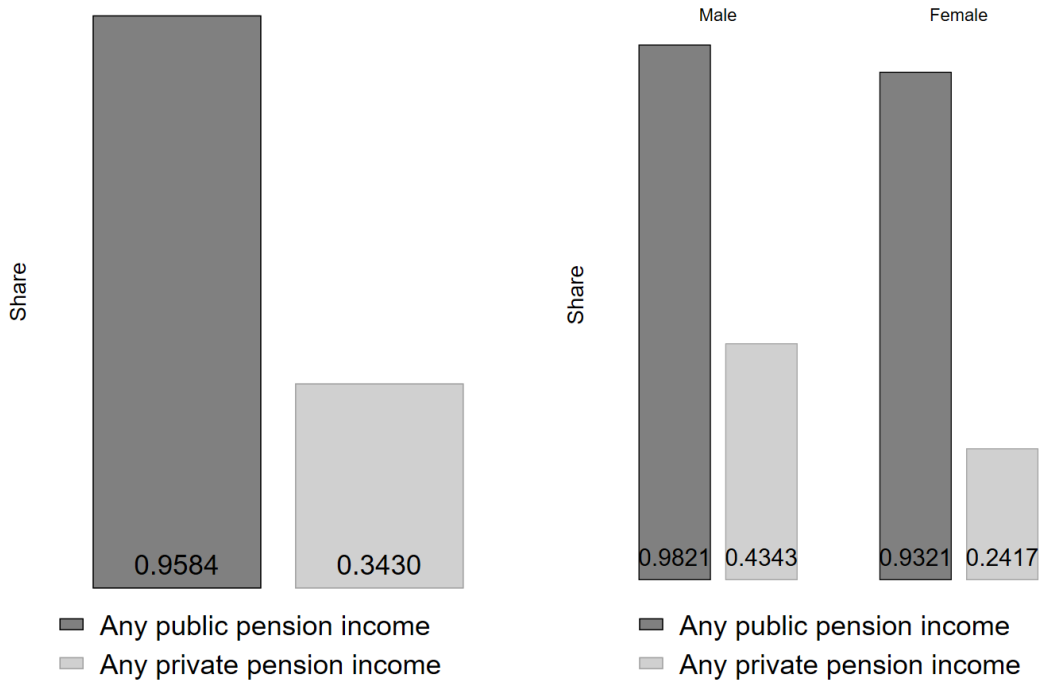


Notes: Results from RIF decomposition. Based on normalized models with absolute independent pension income as dependent variable. Only life course clusters with significant explained shares for at least one decile are depicted. Explained shares are shown in relative terms. Coefficients from male baseline models were used as reference. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A4.3.2 Dealing with zero pension income

Both public and private pensions include individuals with 0 pension income but for the private pension distribution the zero inflation is more severe (Figure A4.14 and Figure A4.15). Almost 96% receive an independent public pension, but only 34% receive their own personal pension (Figure A4.13).

Figure A4.13 Share with any independent public or private pension income overall and by gender.



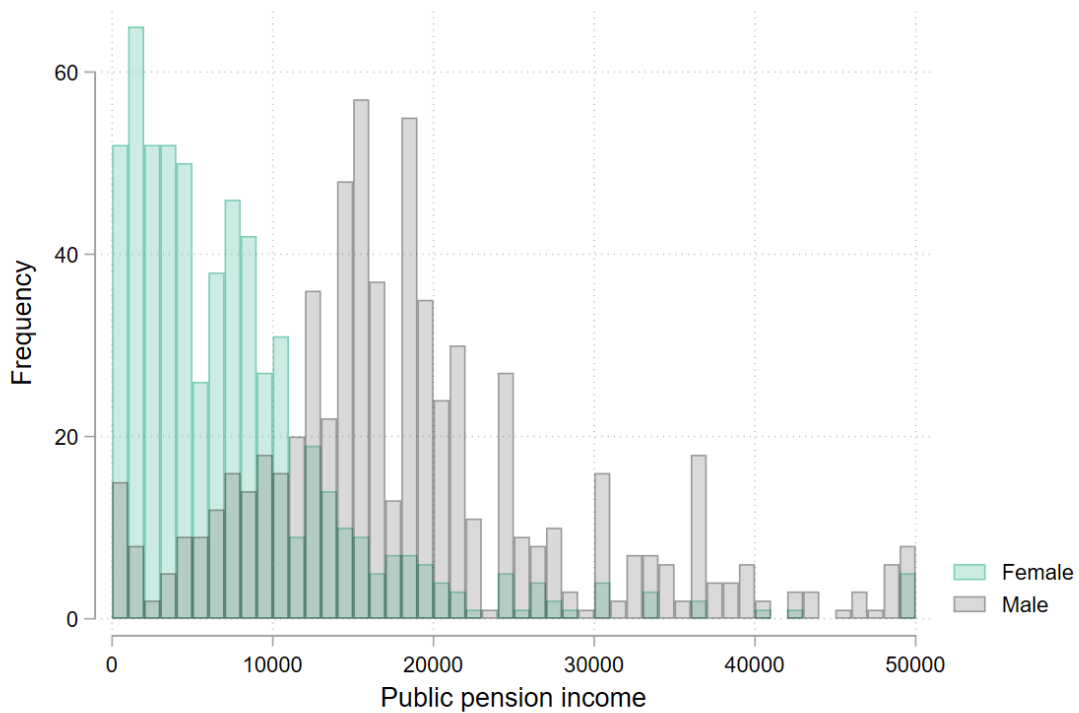
Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

There are two main options to deal with zero-inflated distributions.

Option I: Alternative baseline model

The first option would be to use modelling approaches in the decomposition that address zero-inflated distributions more effectively than linear regressions (e.g., zero-inflated negative binomial or Poisson regressions). However, there are several difficulties. First, such models are tailored to count data. However, private pension income is not an ideal example of count data, with observations of up to €36000. Thus, the data structure does not align perfectly with these alternative modelling approaches either. Secondly, as of my knowledge, the software used for the decomposition analyses in the main analysis does not currently support zero-inflated negative binomial or Poisson regressions.

Figure A4.14 Distribution of public pension income by gender.



Notes: 1000Euro bins. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Consequently, I decided to apply option II, the two-part model, which additionally contributes to the literature with a decomposition of the Gender Coverage Gap.

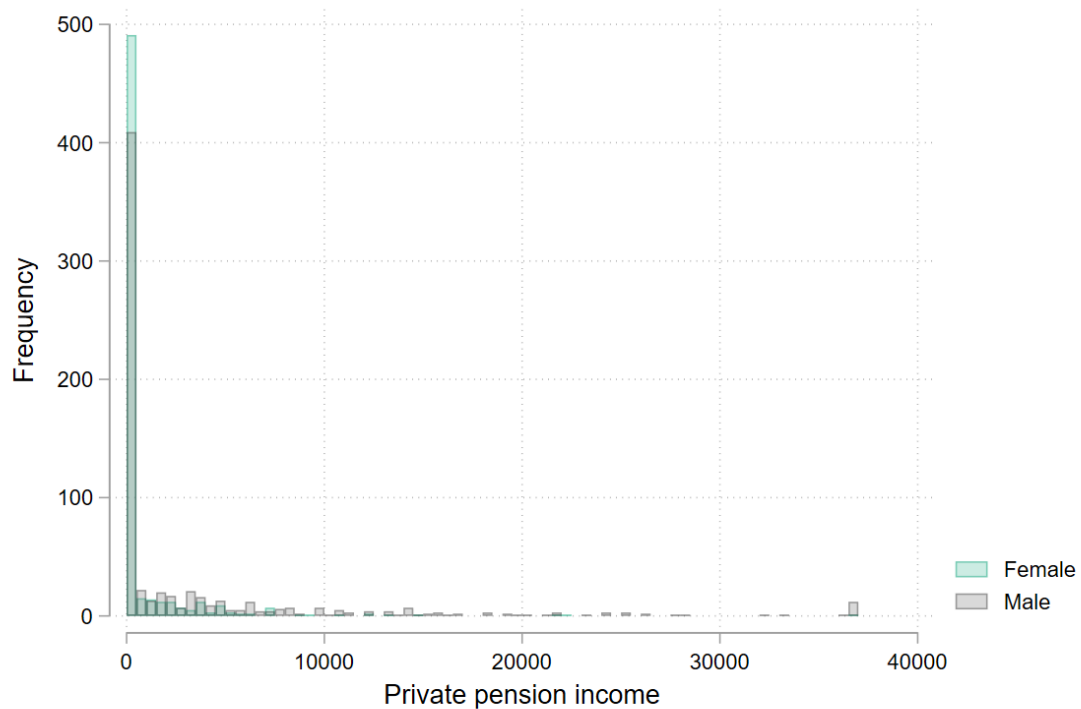
Option II: Two-part model

The second option addresses the zero-inflated continuous outcome by applying a two-part model (Belotti et al., 2015; Boulton & Williford, 2018, see for more detail Chapter 2.1.3). If zero pension income was included and methodologically addressed at all (see Table 2.1), the literature has applied the two-part model (Bardasi & Jenkins, 2010; Even & Macpherson, 2004; Ezeyi & Vujic, 2017; Gardiner et al., 2016; Johnson et al., 1999; Zhao & Zhao, 2018).

The idea is to split the outcome into two variables, i) a dummy for the coverage (whether receipt of any pension income) and ii) a continuous variable including only nonzero values (i.e., only these individuals with pension receipt) and then run separate models for both

outcomes. For the second model, it is advisable to account for the selection into pension receipt, e.g., using the Heckman selection model or by conditioning on the pension receipt in the sense of the model suggested by Belotti et al. (2015).

Figure A4.15 Distribution of personal pension income by gender.



Notes: 500Euro bins. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

However, two-part models pose some difficulties for the analyses applied in this dissertation. First, it requires a sufficient number of observations of individuals with and without pension income for a meaningful and empirically feasible analysis. Focussing on individuals with private pension receipts is challenging for the cohorts and data used in this study due to the low number of observations with private pensions. Second, the Heckman selection model needs an instrument or selection variable that determines access to pensions (i.e., the selection equation) but not the amounts of pensions (i.e., the outcome equation). However, finding an instrument is particularly challenging for the application in this study.

Bardasi and Jenkins (2010) used several family life course proxies and transitions from private employment to self-employment as instruments since self-employees are usually not covered by occupational pensions. While the latter works well for men, they do not find a variable meeting the exclusion restriction for women. Gardiner et al. (2016) used company size and proxies for financial awareness. However, such variables include lots of missing data in SHARE if available at all. Most importantly, it is advisable to consider only characteristics as instruments that occur before the start of the trajectories. But most of the plausible instruments will occur during the work and family life courses from age 18-65 that are used as main predictors in the outcome equation of the main analysis. For example, transitions to self-employment as well as family life courses are covered in the SA clusters. Company size is likely to correlate with the life course clusters and is shown to also impact earnings and thus pensions. In general, I assume that pension receipt is also determined by work and family life courses in pension systems without basic pensions (see Chapter 2.2). Alternative instruments occurring before age 18 are less straightforward and will be less reliable because the SHARE data was surveyed at old age (age 50+). Such measures (e.g., childhood circumstances) contain large degrees of missingness and are likely to suffer from recall bias.

I conclude that the Heckman correction model is not feasible for this analysis because life course trajectory groups starting from age 18 are the main predictors for the outcome equation and already incorporate many of the possible exclusion restriction variables. The data source does not provide sufficient and reliable alternative instruments prior to the start of the trajectories. I thus do not correct for selection.⁷³

⁷³ The two-part modelling approach as described by Belotti et al. (2015) might be an alternative to the Heckman model but still has to be proved to be applicable with the KOB decomposition.

Methods. For the decomposition of the Gender Coverage Gap, I used logistic regressions as the baseline models instead of linear regressions as in the main models. The results did largely not differ when using probit regression or linear probability models (the effect sizes tend to be slightly larger for the logistic regressions, results available upon request). Otherwise, I apply the same parameters as for the main model (i.e., regarding the choice of reference coefficients, and normalization of clusters).

A4.4 Combining life course data from the survey and the administrative records

The data on the working life courses, or pension-relevant activities, are available on a monthly basis in the administrative records of the German Public Pension Insurance (SUF SHARE-RV - source: FDZ-RV). However, the retrospective survey data provides family and working life courses on annual basis. I thus transform the more fine-grained monthly data of the pension records to annual data, using the mode per year, that is the situation the individuals were most in during the respective year, as has been done before (Bühler et al., 2022).

Any employees or other groups (such as apprentices, persons raising children, non-commercial caregivers, persons with disabilities, persons in military or voluntary service, persons who receive unemployment or sickness benefits, and students with marginal jobs) who are subject to mandatory public pension insurance are covered in the administrative data. Self-employees (with some exceptions) and civil servants as well as judges, professional militaries and potentially some other groups of occupations are normally not covered in the public pension system and thus in the data. However, as soon they had at least one month of an insured activity they are included in the data.

Such periods of non-insured activity are recorded as a ‘non-contributory period’ as their social income situation. On top of activities not insured in the public pension scheme,

activities declared as non-contributory may contain unemployment without being registered unemployed, homemaking or other pension-relevant activities that have not been reported (e.g., education). For further information on the data please consult the SHARE-RV User Guide and the VSKT User Information

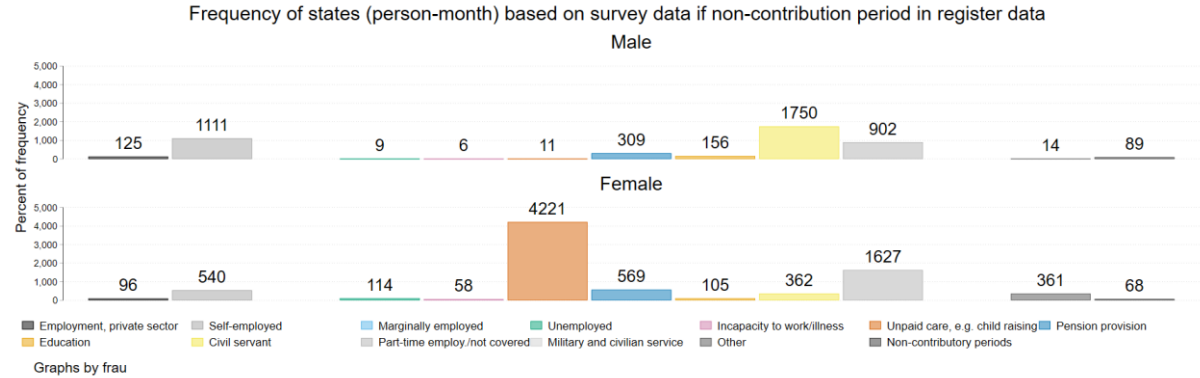
If any employment or short-term employment was reported in the survey while non-contributory periods were recorded in the register data, I grouped these spells with part-time employment if they did not report full-time employment and had earning points collected in this year. These spells might also capture full-time employment in occupations not covered in the public pension system. However, since I do not have a proxy for their earnings due to the lack of EPs for such employment types, I decided to combine them with part-time employment to have them reported as any employment. Employment spells that were originally recorded as non-contributory periods in the records, but reported full-time and earning points were recorded, I defined as full-time employment.

Imputation of non-contributory period

I impute non-contributory periods with the survey information from SHARELIFE since the ‘non-contributory period’ in the register data is likely to be very heterogenous but makes up for a large share of person-year spells in the register data for my sample (23.03%). I only impute (full-time) employment spells if full-time employment was reported in the survey data and earning points were available in the register data. That is because I need the earning points for differentiating between different earning quartiles in the private sector. The rest of the reported employment spells in the survey data without information on earnings points in the administrative data (including short-term employment) are coded on part-time employment.

Figure A4.16 shows for the ‘non-contributory periods’ in the register data the number of person-year spells that were imputed by the different working states from survey data. Quite large gender differences become visible: whereas for men, the non-contributory periods mostly entail civil service spells followed by self-employment and any non-contributory employment (for which no EPs were recorded), for women the by far most prevalent working state is unpaid care work. In general, the Figure reveals the heterogeneity of the social situations combined in non-contributory periods and the importance of differentiating between them. This is even more the case for the analyses with a focus on gender, as is the case here.

Figure A4.16. States replacing non-contributory periods in administrative records



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The procedure for civil servants and self-employees

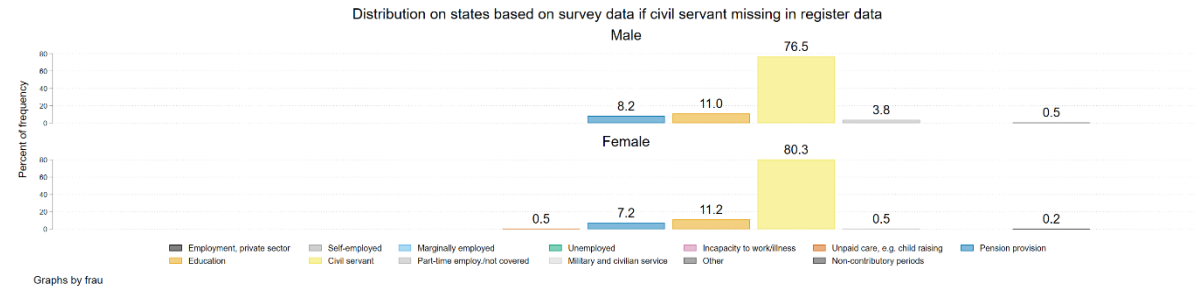
Since self-employees and civil servants are not covered in the administrative records, the information on them is solely based on the survey data. I impute both working states in two different ways.

First, there are individuals not covered at all in the administrative records who have reported for most of their working lives (defined by more than 19 years) to be self-employed or civil servants. These individuals are most likely not included as these employment types are not covered in the administrative records or because they did not agree to link the survey data with the administrative records. For civil servants, more than

half of the reported civil service person-month spells are not linkable to the administrative records, for self-employees about one-quarter of person-month spells were not matched.

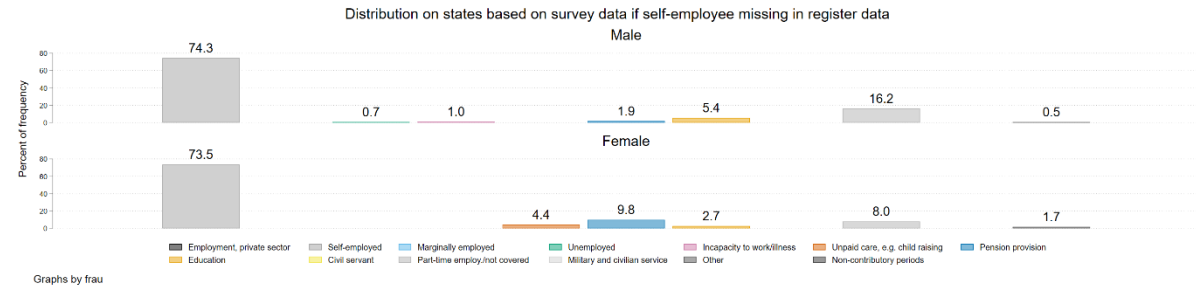
I include individuals who have reported in the survey to have worked for at least 19 years as civil servants or self-employed respectively. The working states of these individuals are thus fully based on the survey data. As information on the earning points for the reported years of (full-time) employment is unavailable for these individuals, I code these years as part-time employment (up to 16% of all person-year spells for those individuals, see Figure A4.18). For the civil servants that are not covered in the administrative data, 76.5-80.3% of their person-year spells are attributed to civil service (Figure A4.17). For self-employees, it is 73.5-74.3% of all person-year spells (Figure A4.18).

Figure A4.17 Distribution on states (based on survey data only) if civil servants missing in register data



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.18 Distribution on states (based on survey data only) if self-employee missing in register data

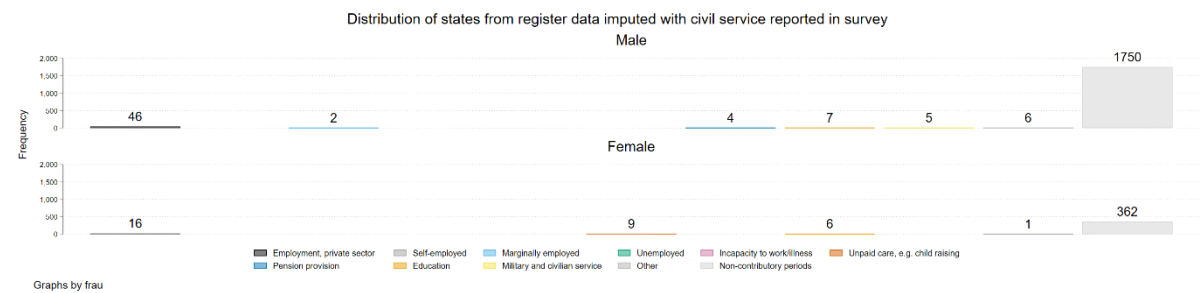


Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Secondly, even if individuals are matched with the administrative records, I mostly lack information on whether these individuals were self-employed or civil servants. In most

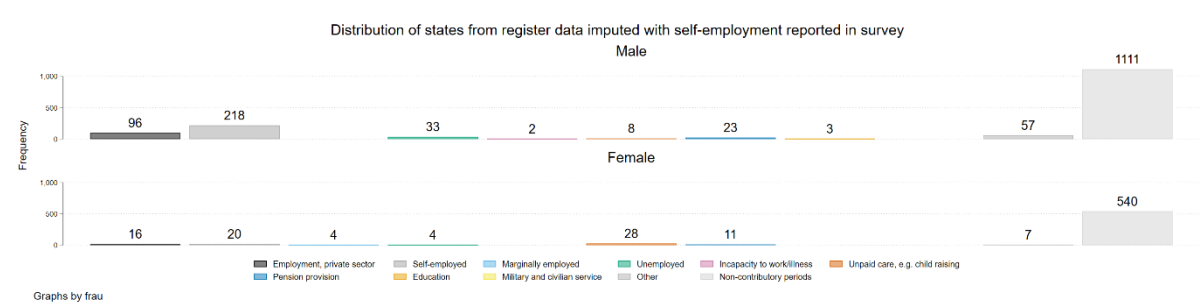
cases, such periods are coded as ‘non-contributory periods’. Thus, I replaced all ‘non-contributory periods’ with civil service or self-employment if these were reported in the survey data. Furthermore, I overwrite administrative records for person-year spells that responded reported to be civil servants or self-employees if the records recorded another working state for less than 12 months for the specific year. However, most of these self-employment and civil service spells overwrite non-contributory periods (Figure A4.19 and Figure A4.20).

Figure A4.19 Frequency of states from register data imputed with civil service reported in the survey



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A4.20 Frequency of states from register data imputed with self-employment reported in the survey



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A4.5 Multichannel Sequence Analysis: analytical choices and robustness

Sequence analysis is an exploratory approach. Still, its analytical choices should be guided by theoretical and, where appropriate, statistical measures (Aisenbrey & Fasang, 2010; Raab & Struffolino, 2022).

For Multichannel sequence analysis, that is when several life course domains are considered in the Sequence Analysis jointly, Piccarreta (2017) suggest first considering the correlation between the domains of interest. The idea is that the domains have to be interrelated to some extent to justify a joint MSA over them. Using the Mantel coefficient, I observe an overall weak but statistically significant positive relationship among the domains (.03). Since I theoretically assume that the linkage between the working and family spheres differs largely between genders and that women's working lives are stronger linked to their family life, I conduct the analysis by gender (as suggested by Raab & Struffolino (2022)). As assumed, I find that the linear correspondence of family dissimilarities to working dissimilarities is indeed much more positively associated for women (.165) compared to men (insignificant and slightly negative).

In such cases, Raab & Struffolino (2022) suggest considering separate MSAs by subgroups, here gender, to prevent concealing gender-specific joint life course patterns. However, in combination with the KOB decomposition, I need the *same* covariates, i.e. cluster, per group to decompose the outcome of interest based on the second step (see Chapter 3). Additionally, as in most applications of the SA-KOB decomposition, I assume that life-course patterns differ by the group of interests, and this might be specifically also the case for the linkage between domains. Thus, when combining MSA with the KOB, one must still go for a joint sequence analysis, even if the linkage between domains is driven by only one of the groups.

To take advantage of the exploratory nature of SA, I assess different parameters along the analytical choices. First, I apply different versions of optimal matching (OM) to calculate the distance matrix for individual sequences which is used in the next step, the cluster analysis. I apply two different types of data-driven substitution costs, based on transition rates as well as a common future (holding the indel costs constant at $\max(sm)/2$). I then use

substitution costs based on transition rates for Hamming distances as well as Dynamic Hamming Distances (DHD), which are more time-sensitive approaches.⁷⁴

As second parameter I use the two most applied clustering approaches in the field, hierarchical clustering based on Ward's Linkage, as well as Partitioning Around Medoids (PAM). Lastly, for the hierarchical clustering outcomes, I guide my choice of the number of clusters retrieved based on standard Cluster Quality Indices (CQI) such as the Average Silhouette Width (ASW), the pseudo R² as well as Hubert's Gamma (HG). I further disregard solutions suggesting less than 6 or more than 12 clusters for empirical and theoretical reasons.

This leaves me with a total of 18 cluster solutions which I visually examine to assess the analytical and theoretical fit with the research question and assumptions.

The best-performing solution (ASW of .282) is a cluster outcome with 11 clusters based on OM with substitution cost calculated based on a common future and clustered by Ward. However, the only remarkable difference to the cluster solution using the same parameters but deriving 10 life course patterns only, is that the 11-cluster solution reveals two instead of one cluster dominated by stable careers in civil service. One of the clusters is characterized by being married and having at least 2 children over the life course, while the individuals in the other cluster have different family life courses, most dominantly being married and having one child. This does not seem to provide added theoretical value important enough to justify the disadvantages with regard to the number of observations per group that each additional cluster brings. The 10-cluster solution contains only life course patterns that are theoretically highly relevant while still achieving the third-highest ASW among all cluster solutions examined in more detail.

⁷⁴ For a detailed overview of these parameters, see Studer & Ritschard (2016).

All cluster solutions considered that were clustered by PAM do not include a family life course that is dominated by divorce periods. This is, however, highly relevant concerning the German pension system, since pension claims are split after divorce and therefore pension claims up to the divorce are equalized.

Apart from that, the chosen cluster solution is robust against the cost specification applied. I find very similar clusters when using costs based on transition rates, as well as Dynamic Hamming Distances or normal hamming distances (correlation between different cluster solutions .92-93).

A4.6 Reassigning single trajectories to address the common support problem.

To improve the robustness of the decomposition, I systematically select and assign single trajectories to the low-n clusters for each group as previously suggested (Rowold et al. 2022). I do the reassignment based on three options.

For the first option, I consider the other four well-performing cluster solutions suggesting 10 clusters (see description in S2). I re-group individual trajectories if they were assigned to one of the low-n clusters in one of the alternative cluster solutions. The low-n clusters for men and women are the high-income cluster (1) for women, the part-time and non-contributory cluster (5) and care work cluster (6) for men. I do not re-group trajectories from other small clusters (most importantly clusters 9 and 10) to not create more common support issues.

For the reassignment of male trajectories, this first procedure succeeds in reassigning a few male trajectories to clusters 5 and 6. All reassigned trajectories indeed have a rather poor fit with the life course patterns they were previously grouped with in the main cluster solution (measured by silhouette values, not shown), which validates this procedure.

However, no other cluster solution has assigned additional women to the high-income cluster.

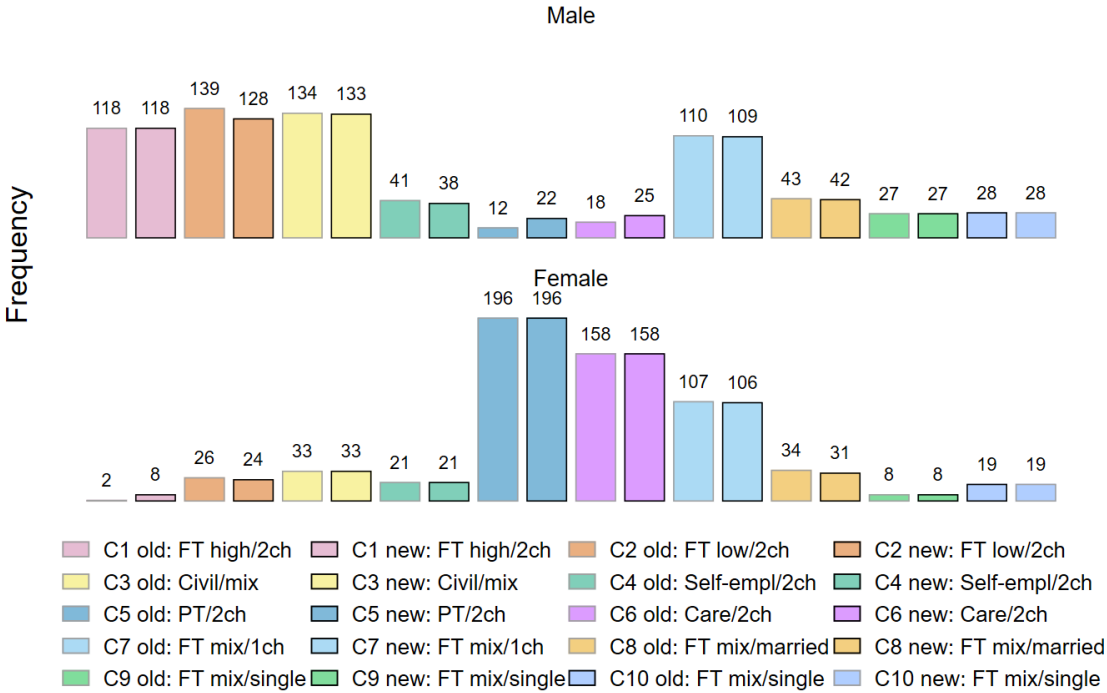
As second option, I combine information on how coherent the individual trajectories are with the overall cluster (silhouette value distribution, using the 10th percentile as the cut-off for identifying poorly assigned individuals) with the sum of years individuals have spent in the life course states that are important characteristics for the cluster. For men, I only do this for cluster 5 since I already have n=25 men in the care cluster (6).

For women, this option does not identify additional women for the high-income cluster.

Thus, for women only, I apply option 3, in which I only consider the sum of years in high (4th quartile) or medium-high (3rd quartile) earning full-time jobs)as well as the time spent in part-time and care work. I aim to identify women with many years spent in high-earning full-time jobs and few years with care work and part-time employment. Doing this, I identify 6 more women who broadly share the characteristics of life course pattern 1.

The absolute gender-specific distribution on the cluster before and after the reassignment is visible in Figure A4.21 and the final relative distribution is given in Figure A4.5.

Figure A4.21 Frequency of individuals in clusters before and after reassignment



Notes: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A4.7 Robustness checks

The results of the robustness checks are available upon request.

a) Adding controls

The results do not change when controlling for survey waves and nativity. Controlling for the highest educational level though marginally changes the magnitude of some results. For both income types, parts of the gender gaps that are due to more men in life course patterns 1-3 (parents with high or low-income careers in the private sectors or civil servants) are being picked up by educational differences between men and women (results available upon request). This is because educational differences are partly inherent in the life course clusters (e.g., clusters 1 and 3 have a higher share of educational periods after age 18, Figure A4.4, Table 4.1). While the parts of the explained shares for clusters 1 and 3 due to educational differences are only marginal, the negative association of more fathers having stable but low-income employment careers is almost fully absorbed by the educational gradient between men and women. For private pensions, educational differences also pick up most of the explained share due to careers characterized by employment interruptions (cluster 5).

The higher share of women with maximal lower secondary as well as the higher share of men with tertiary education (not shown) explain together 13% of the gender gap in public pensions and 28% of the gender gap in private pensions. The mediating role of education in the association between life course patterns and the GPG suggests that the educational level also influences the (gendered) odds for realizing different life course patterns.

Furthermore, women benefit from higher public and private pension returns for a low educational level compared to men. For public pension, this might be due to upgrades of

female caretakers' pension income. On the other hand, however, parts of the GPGs are due to higher pension returns for men with tertiary education compared to their female peers.

b) Using different reference coefficients

Using different reference coefficients (male, female or pooled for all covariates) the magnitude of the explained parts changes only slightly, and particularly the results using the reference coefficients from a pooled model are very similar to the main results (similar to Chapter 3). For public pension, the returns components are not sensitive to the reference coefficient chosen but for private pension the negative returns share for cluster 2 becomes significant, suggesting that mothers with low-income employment receive better pension rewards than fathers with similar life courses.

c) Using the original MSA solution

As expected and in line with Rowold et al. (2023), the explained shares are larger when using the original, more unequal gendered distribution on clusters. This is particularly the case for the life courses of parents characterised by high-income jobs, part-time employment and care work (clusters 1, 4 and 5) for private pensions. The explained shares in the main analyses are thus slightly underestimated.

A5 Appendix for Chapter 5

A5.1 Additional Tables

Table A5.1 Pension types included and definition of dependent variables.

Dependent Variable		
<i>Public pension</i>	<i>Occupational pension</i>	<i>Private Pension</i>
1. Old age pension	1. Old age pension from the last, a second or a third job	1. Average payments of regular life insurance
2. Old age supplementary pension or public old-age second pension	2. Early retirement pension	2. Regular private annuity or private personal pension
3. Early retirement or pre-retirement pension	3. Disability or invalidity insurance	
4. Main or secondary disability insurance pension, or sickness benefits		

Table A5.2 Sample selection

		Male		Female	
Base sample age cutoff	Netherlands	1372	(48.8)	1442	(51.2)
	West Germany	1475	(50.0)	1475	(50.0)
Non-response pension	Netherlands	1144	(48.0)	1241	(52.0)
	West Germany	1276	(49.5)	1304	(50.5)
Employment	Netherlands	1073	(47.2)	1201	(52.8)
	West Germany	1184	(48.8)	1243	(51.2)
Final sample incl. retro data	Netherlands	500	(45.9)	589	(54.1)
	West Germany	806	(49.7)	817	(50.3)

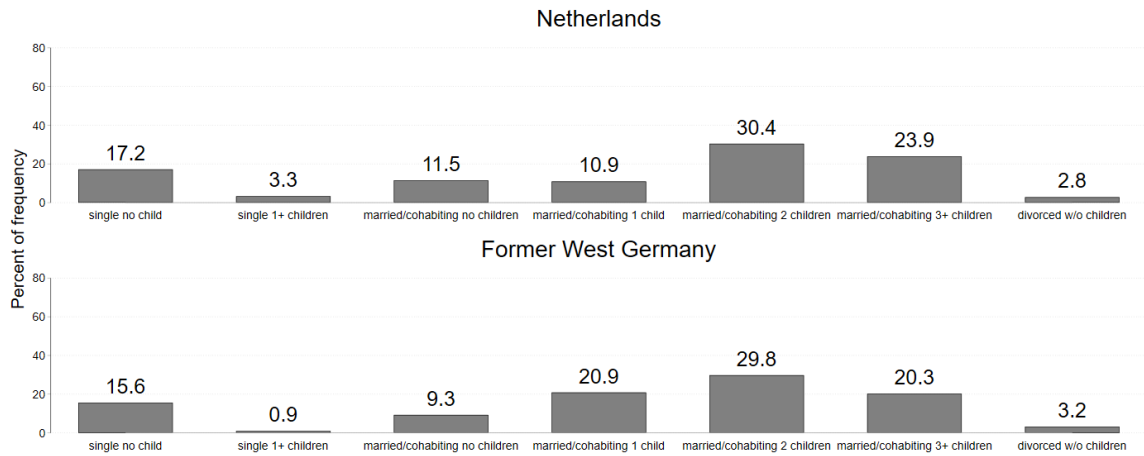
Table A5.3. Full results of Boruta Life Course Feature Selection, main model

West Germany						Netherlands					
features	Mean imp.	Min. imp.	Median imp.	Max. imp.	norm Hits	features	Mean imp.	Min. imp.	Median imp.	Max. imp.	norm Hits
duration.CIV	32.47	19.28	32.81	35.26	1.00	duration.CW	29.04	12.63	29.33	32.60	1.00
duration.CW	29.73	15.87	30.01	31.37	1.00	duration.EDU	28.76	11.85	28.91	36.75	1.00
pattern..CW	18.50	12.63	18.56	23.61	0.99	AFtransition..OTHFT	18.33	4.82	18.77	20.44	0.95
spell.CW_age_1	18.45	9.09	18.55	23.76	0.99	Complexity.Entrop_work	16.37	7.88	17.00	22.03	0.93
AFtransition..CIVRET	18.23	16.47	18.26	20.13	0.99	duration.Ret	14.95	10.43	14.92	16.82	0.90
AFtransition..FTRET	17.96	10.38	18.20	22.09	0.98	duration.SELF	13.41	5.31	13.54	17.41	0.85
duration.FT	16.03	12.52	16.02	22.88	0.97	pattern..CW..OTH	13.21	3.36	13.20	15.00	0.83
AFtransition..CW.PT..P						duration.SNC	12.24	2.12	12.70	18.07	0.77
TRET	15.87	3.84	16.32	16.67	0.97	spell.EDU_age_1	10.84	7.13	10.83	14.49	0.76
spell.CIV_age_1	14.56	12.75	14.49	17.47	0.96	Age23_27.End_EDU	10.10	6.14	10.03	13.83	0.71
pattern..CIV	13.55	11.58	13.12	16.56	0.94	duration.CIV	9.57	5.11	9.68	10.75	0.68
AFtransition.CWFT..FTC						spell.CW_age_1	9.40	7.33	9.26	13.49	0.68
W	12.76	3.10	12.90	13.96	0.92	pattern..OTH..FT	9.08	3.56	9.22	10.81	0.65
pattern..EDU..CIV	11.10	10.06	11.10	13.54	0.88	pattern..CW	8.72	7.14	8.62	12.84	0.63
duration.EDU	10.44	6.59	10.49	12.40	0.85	AFtransition..EDUCIV	8.60	6.56	8.63	10.87	0.61
duration.Ret	9.59	4.72	9.71	13.13	0.79	pattern..UNE..FT	7.96	3.28	7.84	10.38	0.56
pattern..FT..CW	9.32	7.55	9.22	17.75	0.78	duration.FT	7.87	3.77	7.95	11.74	0.55
spell.PT_age_1	8.80	3.33	8.83	11.82	0.75	pattern..FT..CW	7.10	4.44	6.95	11.80	0.49
AFtransition..MNCM1C	8.77	4.00	8.78	13.70	0.74	pattern..EDU	6.45	1.27	6.38	11.36	0.43
pattern...PT	8.40	2.56	8.56	11.30	0.71						
Complexity.Entropy_w											
ork	8.19	3.25	8.30	10.67	0.70						
AFtransition..EDUFT	7.83	2.69	7.80	11.66	0.68						
Age23_27.End_EDU	7.62	4.19	7.69	9.71	0.67						
duration.SELF	7.51	3.11	7.58	9.90	0.64						
Complexity.Turbulence											
_work	7.25	3.15	7.35	8.72	0.62						
Age58_62.Start_Ret	6.80	0.62	6.85	12.22	0.56						
AFtransition.EDUFT..F											
TRET	6.61	4.37	6.60	11.30	0.56						
duration.PT	5.86	3.91	5.84	9.61	0.46						
Age58_62.End_CIV	5.09	3.30	5.07	8.27	0.35						
AFtransition..FTCW	4.42	3.11	4.22	10.56	0.25						
AFtransition..CWRET	3.85	1.94	3.96	7.76	0.18						
AFtransition..EDUCIV	3.82	2.36	3.79	8.29	0.18						
pattern..EDU	3.40	0.96	3.41	9.94	0.12						
AFtransition.EDUCIV.											
CIVRET	3.00	-0.92	3.05	10.46	0.09						
spell.EDU_age_1	2.48	0.10	2.01	9.84	0.07						
Age63_65.End_CW	1.97	-0.83	1.95	8.44	0.03						
Age63_65.End_CIV	0.74	-1.95	0.71	9.97	0.01						

Notes: Complete Boruta results, including all importance values and hits.

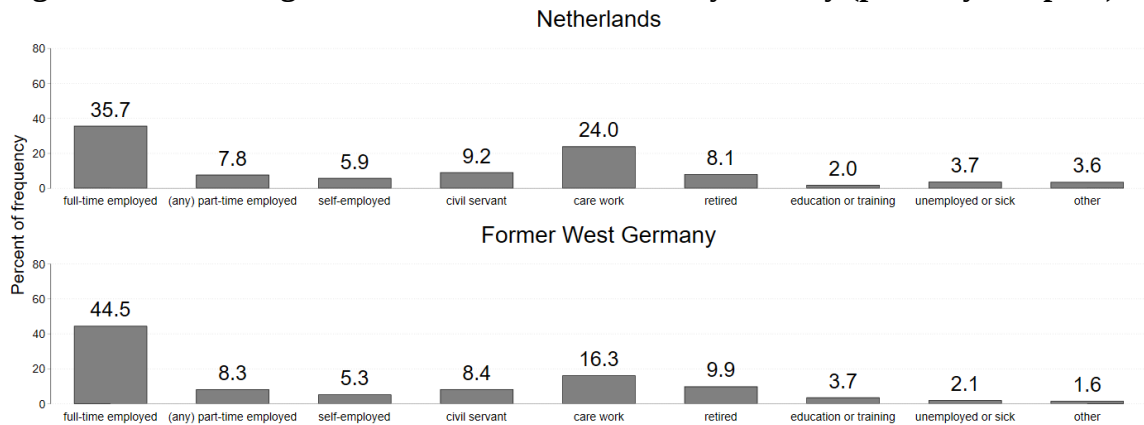
A5.2 Additional Figures

Figure A5.1 Family life course states considered by country (person-year spells)



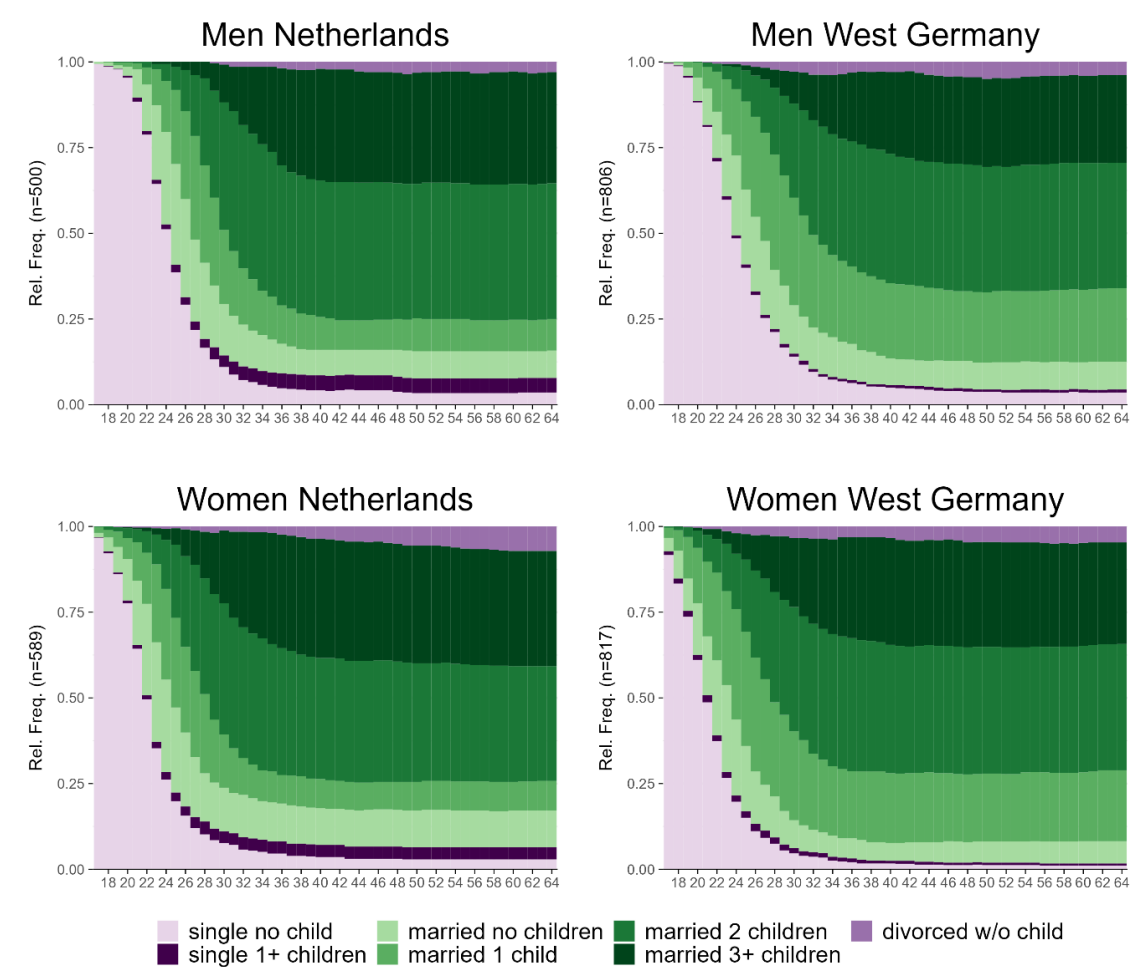
Note: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.2 Working life course states considered by country (person-year spells).



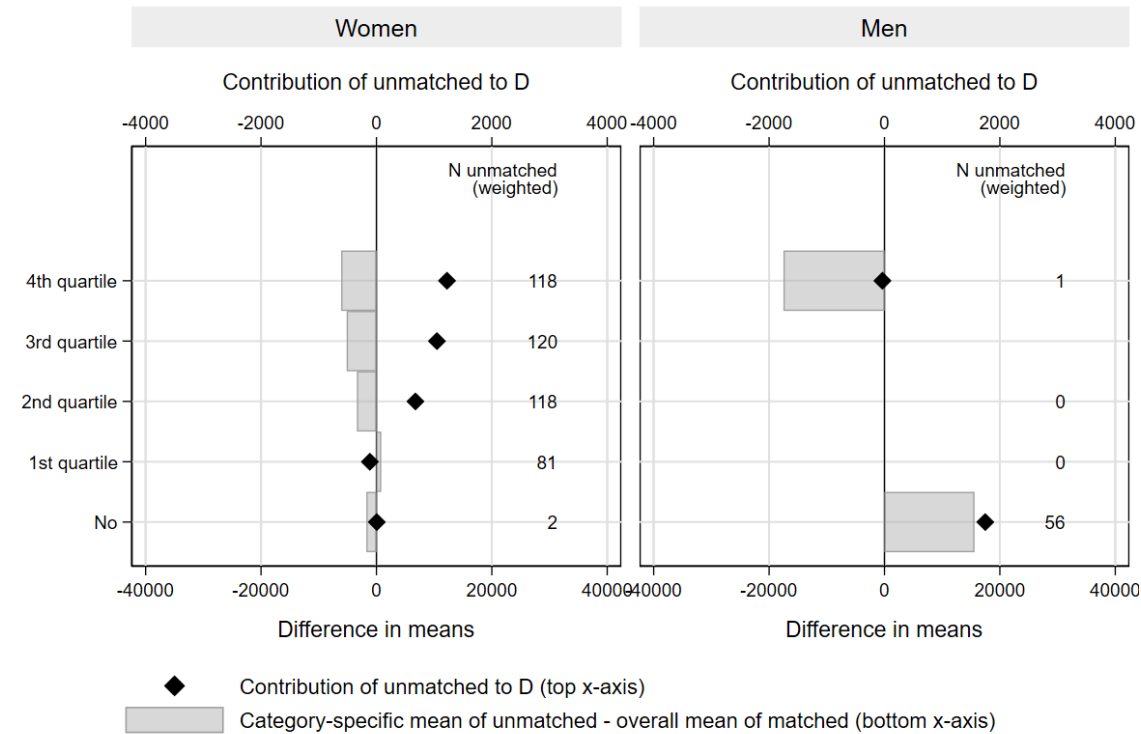
Note: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.3 Family life courses across genders in West Germany and the Netherlands



Note: Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.4 Contributions of years spent in care work (to GPG (DA/DB), Netherlands (model 2)



Notes: Bars depict the difference between the mean pension income of the unmatched individuals per variable category and the overall pension income of matched individuals by decomposition groups. The squares depict the contribution of the category-specific pension difference to the overall GPG (D) and thus to DA/DB by weighting it by the number of unmatched individuals in each category relative to all individuals for each group. Based on Sprengholz & Hamjediers (2022) and the Ñopo package in stata (Sprengholz & Hamjediers, 2023). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.5 Contributions of years spent in civil service (to GPG (DA/DB), West Germany (model 3)

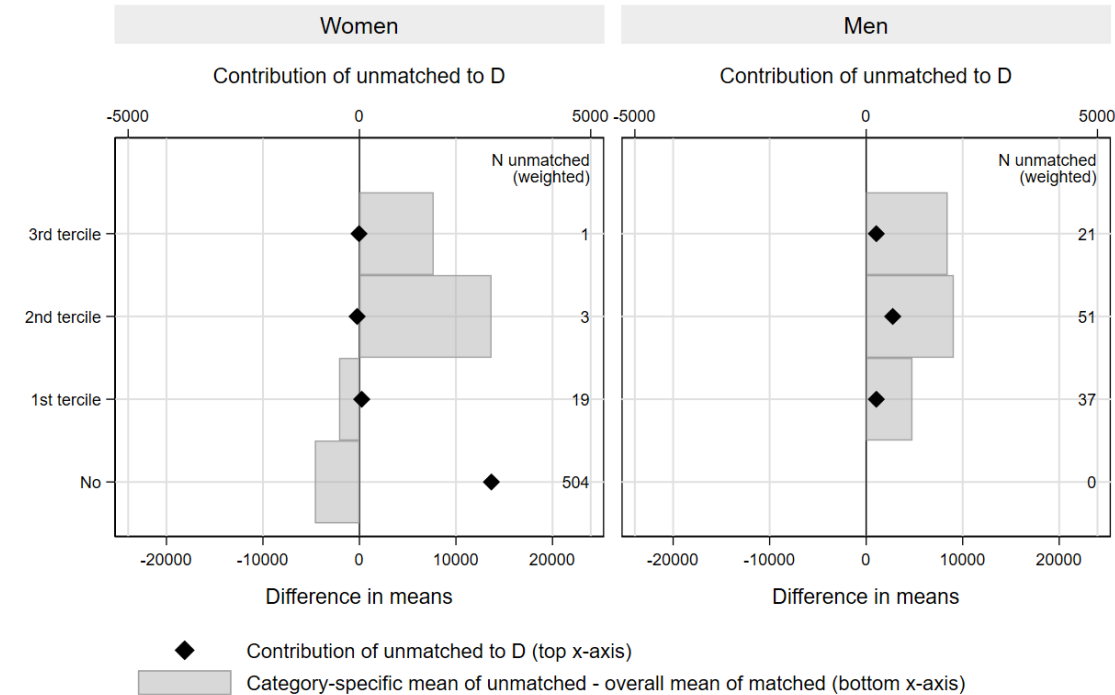
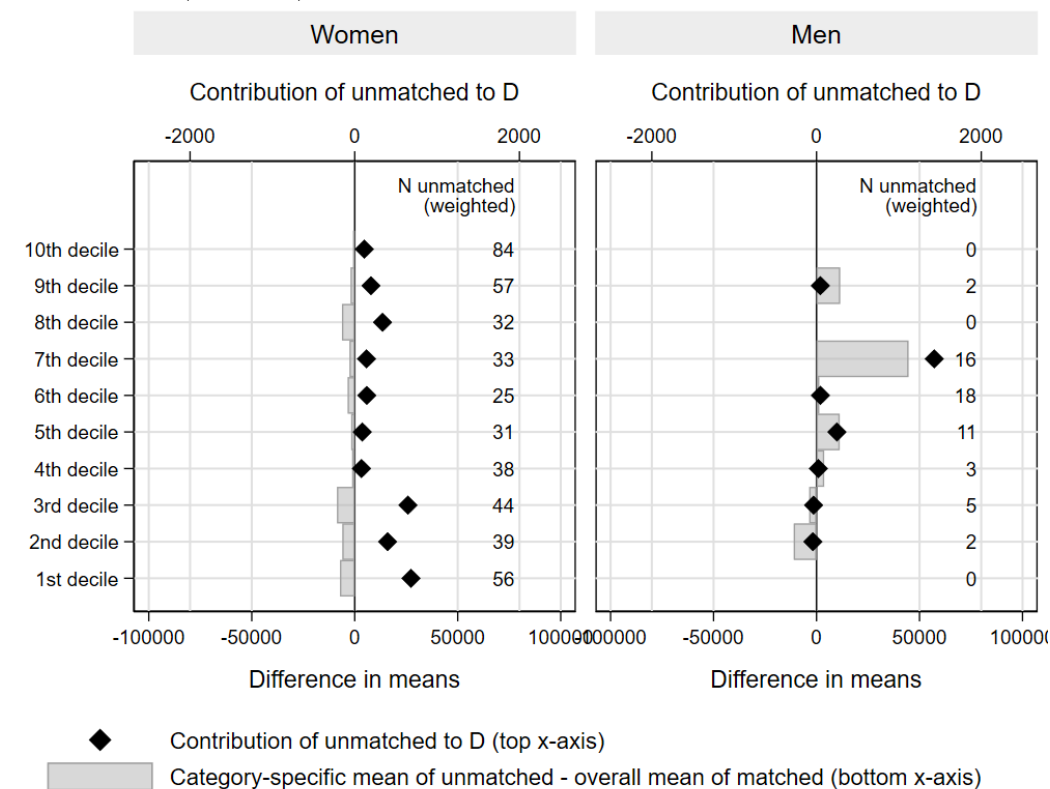


Figure A5.6 Contributions of age at transition from civil service to retired (to GPG (DA/DB), West Germany (model 3))



Notes: Bars depict the difference between the mean pension income of the unmatched individuals per variable category and the overall pension income of matched individuals by decomposition groups. The squares depict the contribution of the category-specific pension difference to the overall GPG (D) and thus to DA/DB by weighting it by the number of unmatched individuals in each category relative to all individuals for each group. Based on Sprengholz & Hamjediers (2022) and the *Ñopo* package in stata (Sprengholz & Hamjediers, 2023). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.7 Contributions of working life complexity categories to GPG (DA/DB), Netherlands (model 2)



A5.3 Gender-specific pension predictors and predictors' importance for GPG

This extension is to explore whether the overall pension predictors selected in the main model of the Life Course Feature Selection are also relevant predictors of the Gender Pension Gap. Additionally, it aims to assess whether there are additional important predictors for gender-specific pension inequality. These predictors might be relevant for the Gender Pension Gap as well but could be concealed in the feature selection for overall inequality.

In the main analysis, I focus on the most important predictors of *overall* pension income inequality, i.e., the inequality pooled over gender. While gender inequality is inherent in overall pension inequality, and it is likely to constitute a large component given the high level of the Gender Pension Gap, it is important to recognize that other inequality aspects exist in overall pension inequality. Examples might be educational, occupational, or nativity-related pension inequalities. However, the focus of this study is *gendered* pension inequality. Thus, the feature selection for overall pension inequality risks selecting variables that are i) important for these other inequality aspects primarily while ii) concealing or underestimating proxies that might be relevant for the Gender Pension Gap. For example, relying only on the feature selection pooled over gender risks concealing variables that have contrary associations for men and women with pension income. The associations of such variables might be close to 0 for the pooled model over both samples. However, variables that constitute highly gender-specific associations with pensions might be an important mechanism for the explanation of gendered pension inequality.

To my knowledge, there are no feature selection approaches that can directly identify the determinants of group gaps in outcomes. For selecting relevant features for decomposing the gender wage gap, Bonaccolto-Töpfer and Briel (2022) suggest applying a post-double-LASSO procedure which consists of two selection steps. In the first step and comparable

to the feature selection used in the main analysis, the most important predictors for wage are selected using the LASSO algorithm. In a second feature selection they then additionally identify which variables have pronounced gender differences, using a second LASSO model with the gender dummy as an outcome variable. The selected features from both feature selection models are then used to decompose the gender wage gap.

However, through the second step, this approach highly stresses the compositional component of the decomposition technique and is thus likely to overestimate this part of the gender gap. Other predictors might be of relevance as well but unfold their associations with the gender gap through differences in returns (returns component). I believe such predictors would have to be systematically identified as well to result in a balanced set of selected features to decompose the gender gap based on.

Another approach involves adjusting the outcome (pension income) to align with the GPG. However, changing the outcome to the Gender Pension Gap is better applicable for couple-level analyses (for example, using the intra-couple gender gap as outcome) than for the individual-level analysis. A workaround could involve creating artificial (heterosexual) couples through matching and using the GPG at the couple level as the outcome for the feature selection. This, however, would halve the number of observations.

Instead, I run the Life Course Feature Selection separately by gender.

As a second robustness check, I use single variables in the decomposition to assess whether the most important pension predictors of overall pension inequality are substantially related to the Gender Pension Gap.

A) Gender-specific Life Course Feature Selection

For the gender-specific Life Course Feature Selection I use the same parameters for Boruta as in the main model.

First, I compare the importance of gender-specific pension predictors with those of overall pension inequality to obtain an indication of which predictors are particularly important for gender inequality in pensions and which might be driven by gender. The gender-specific pension inequality does not capture gender inequality in pensions but identifies pension predictors that are important for pension inequality among men and among women only. Thus, comparing the ranks of variables across the pooled, male, and female samples reveals the features that are foremost important for gendered pension inequality as this dimension of inequality is exclusively integrated into the pooled model. The underlying assumption is that other inequality dimensions, for example educational, occupational, or nativity-related pension inequalities, similarly influence both the overall and gender-specific pension predictors.⁷⁵ Consequently, pension predictors that are much more important for *overall* pension inequality compared to pension inequalities among men or women indicate which life course aspects are particularly important predictors of *gendered* pension inequalities.

Additionally, it provides a sense of which pension predictors might be driven by gender. Gender-specific predictors of pensions might be concealed in the pooled model if they have contrary associations with pension income by gender. This might be particularly the case for family life course aspects because family experiences have been shown to have opposing associations for men and women with pensions (for example, for divorce, Kreyenfeld et al., 2023, see also Chapter 4). Such gendered associations might cancel each other out in the feature selection for pension income pooled over gender. Consequently, related variables might not be selected as important predictors. However, these extreme ‘differences in returns’ might be an important source of Gender Pension Gaps and would be captured in the returns component of the standard KOB decomposition (i.e., differences in returns).

⁷⁵ A challenge to this assumption are interactions of other inequality dimensions with gender.

Thus, second, I briefly discuss life course aspects, especially those related to family lives, that are important for gender-specific pension inequalities.

I. Differentiation between pension predictors driven by one gender and by gendered pension inequality.

As discussed in the main model, comparing the relative importance across the pooled and the gender-specific model (via the ranks) indeed reveals several pension predictors that are only or much more important in the pooled model. In addition to the life course aspects mentioned in the main text, some other life course variables stand out as particularly important for gender inequalities, since they emerge as important predictors of overall pensions but not of women's or men's pensions.

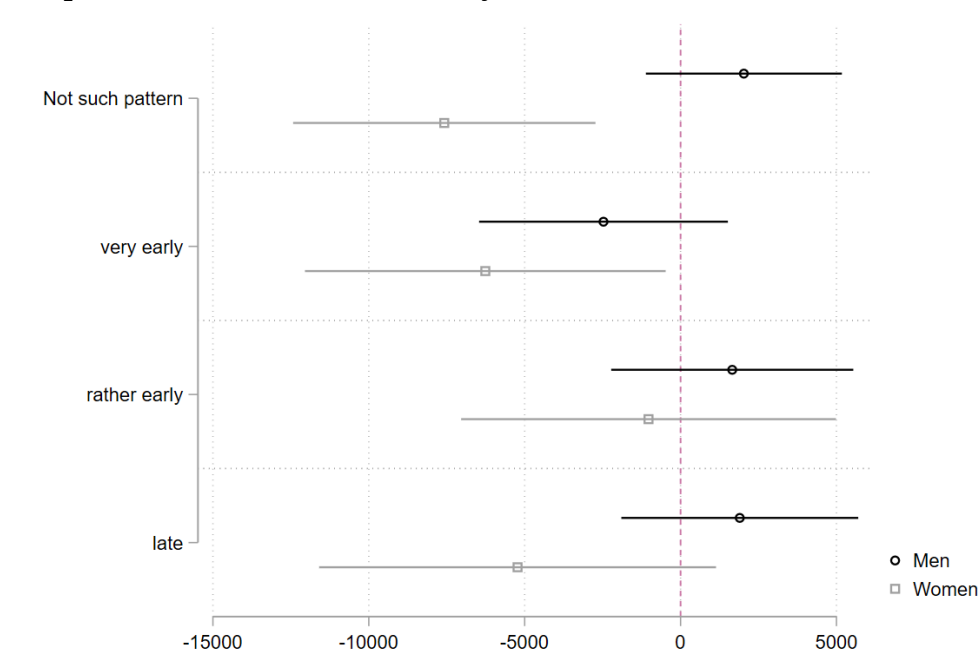
In West Germany, the age at the transition from full-time employment to retirement as well as the duration of retirement are both much more important for the pooled pension inequality (see Table 5.2 and Figure A5.6). As discussed in the main text, this probably traces back to the finding of the gender-specific retirement notion and reporting in Chapter 2.2. Chapter 2.2 reveals that women who do not report retirement and consequently do not have a transition to retirement, disproportionally comprise those with extensive years of caregiving responsibilities. Given that such a selective sample of women is unlikely to report retirement and these women are likely to receive low pensions (see section 5.4.2.1), variables related to retirement might be good proxies for gendered pension inequalities, while being less important for overall pension income. Indeed, especially women who do not experience any retirement or transition to retirement by age 65, receive a much lower pension income (see Figures A5.10-12). Thus, it is likely that the life course aspects related to retirement are proxies for specific life courses with a large share of care work, which drive the effects.

Figure A5.8 Gender-specific pension predictors West Germany



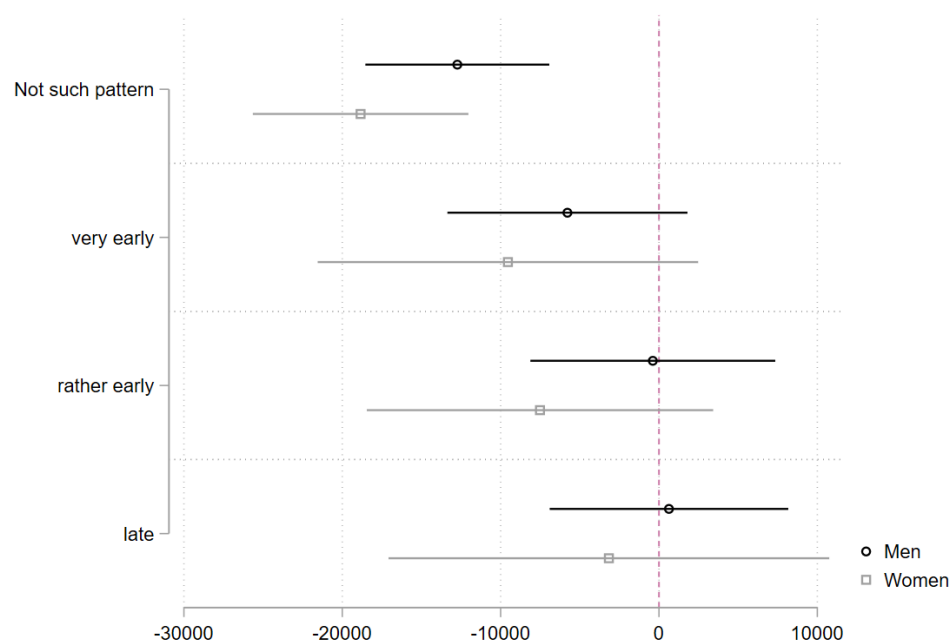
Notes: Results for all features selected in the Netherlands or West Germany based on the Boruta algorithm. Mean importance as z scores. See Figure 5.3 for the state alphabet. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.9 Association between age at transition from FT employment to retirement and pension income, West Germany



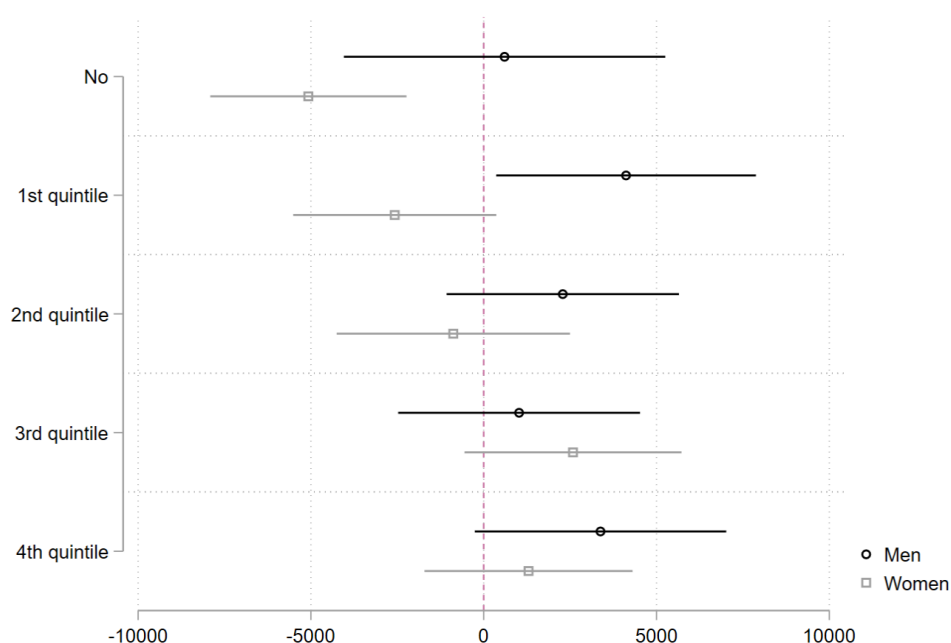
Notes: Reference category: 'very late'. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.10 Association between age at transition from civil service to retirement and pension income, West Germany



Notes: Reference category: 'very late'. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.11 Association between duration of retirement and pension income, West Germany

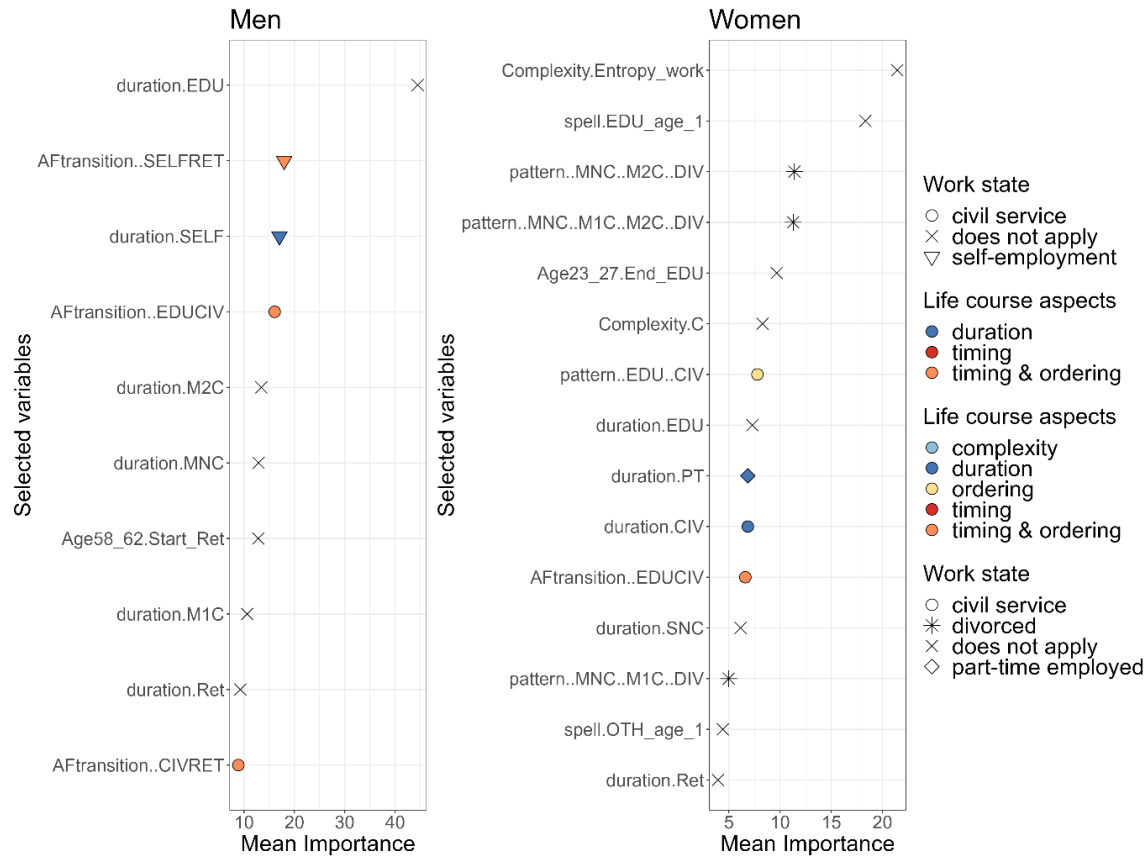


Notes: Reference category: 5th quintile (longest duration in retirement). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Additionally, the timing and ordering of part-time employment as well as the age at the transition from care work to part-time employment if a direct retirement from part-time

employment follows later in the life course, seem to be exclusively important for gendered pension inequalities. In West Germany, the timing and ordering of part-time employment is much more important for the overall pension income inequality, and thus presumably for the gendered inequality, than for the pension inequality among men or women. However, in both pension systems, the duration of part-time employment is an important pension predictor for women only (rank 9, see Table 5.2, Figure A5.12). This highlights that the importance of different life course aspects severely differs across different inequality dimensions of the same outcome.

Figure A5.12 Gender-specific pension predictors the Netherlands



Notes: Results for all features selected in the Netherlands or West Germany based on the Boruta algorithm. Mean importance as z scores. See Figure 5.3 for the state alphabet. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Other life course aspects stand out to be only important for the gender-typical pension inequality. Apart from the number of full-time employment periods in West Germany, this is also the case for the duration of self-employment which is only important for pension

inequality among men. Additional life course aspects that seem to be driven by one gender (i.e., they are important for pooled inequality and one gender) are the complexity of work life course and the timing of education in the Netherlands as well as the duration of care work in West Germany (all potentially driven by women).

II. Family life course aspects concealed in pooled income inequality model

Second, as expected, particularly family life course aspects stand out as being exclusively important for gender-specific pension inequalities. These might not have been detected as relevant for the overall pension inequality due to opposing associations for men and women with pensions. Different associations with pension income by gender have been shown for divorce or never been married (e.g., Kreyenfeld et al., 2023; Möhring, 2015).

In addition to life course aspects related to divorce, for men in West Germany the age at the transition from being married but having no children to fatherhood as well as at the transition from singlehood without having children to being a married father of one child (if later the transition to a second child follows) is more important than for the overall pension inequality. Thus, contrary to my expectations the timing of parenthood seems to be more important for father's than for mother's pension incomes, respectively. Instead, for women, besides a life course proxy related to divorce, the only other family life course aspect among the top-10 pension predictors is the years spent as a single mother. In the Netherlands, the duration of being married without children, with one child as well as with two or more children are among the top-10 pension predictors of men. For Dutch women, exclusively transitions of divorce in different constellations appear to be important family life course predictors of pensions.

Consequently, the sensitivity check confirms that many pension predictors related to family life courses come to light when focusing on gender-specific than on overall pension inequalities. The results further suggest that family life course proxies are not particularly

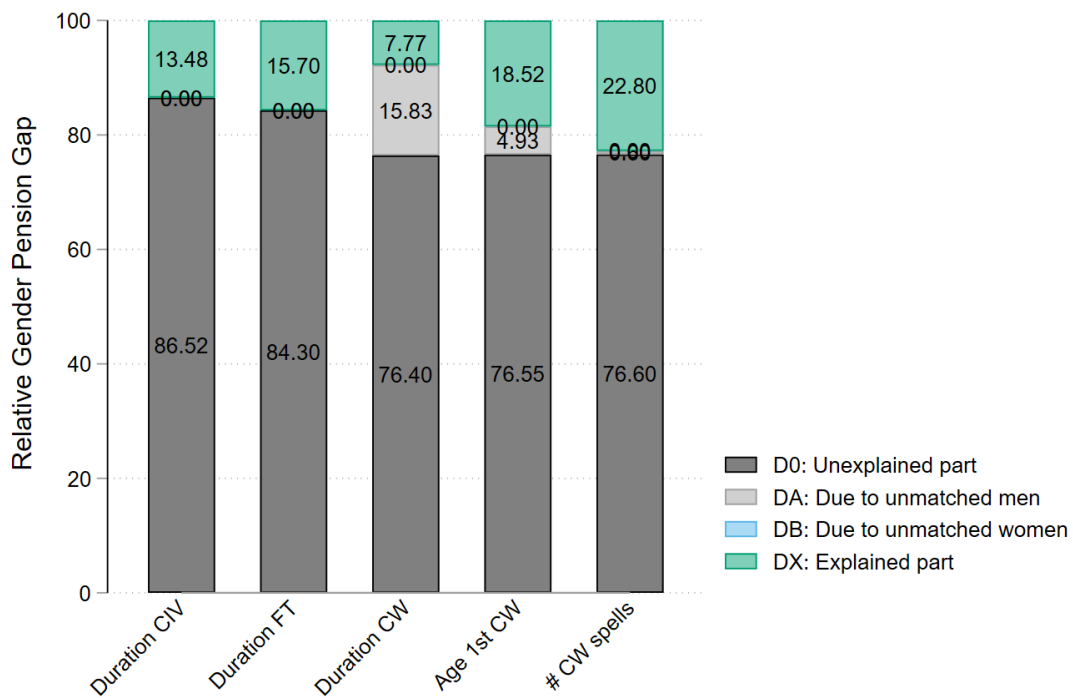
relevant for women's pensions, but even slightly more important predictors of men's compared to women's pension inequalities. However, this indeed points to gender differences in returns for the same family life course characteristics (see Chapter 3). The role of these gender-specific returns of certain life course aspects for the Gender Pension Gap must be assessed further in future research.

B) Bivariate association of most important pension predictors and GPG

As the second sensitivity analysis I assess whether the most important pension predictors also play a role for the gender gap in pension income. I run separate Ñopo decompositions for the duration of full-time employment, as well as the most important pension income predictors.

Overall, the findings reveal that the result of the Life Course Feature Selection highlighting the importance of care work over full-time employment life course aspects for *overall* pension inequality holds when focussing on *gender* inequality in pension income. The duration, timing, and ordering of care work contribute higher shares of the Gender Pension Gap compared to the duration of full-time employment in West Germany (Figure A5.13). However, they also explain a higher share of the GPG compared to the years spent as a civil servant (the most important pension predictor), suggesting that the ranking of the Gender Pension Gap predictors deviates at least slightly from those of overall pension income inequality.

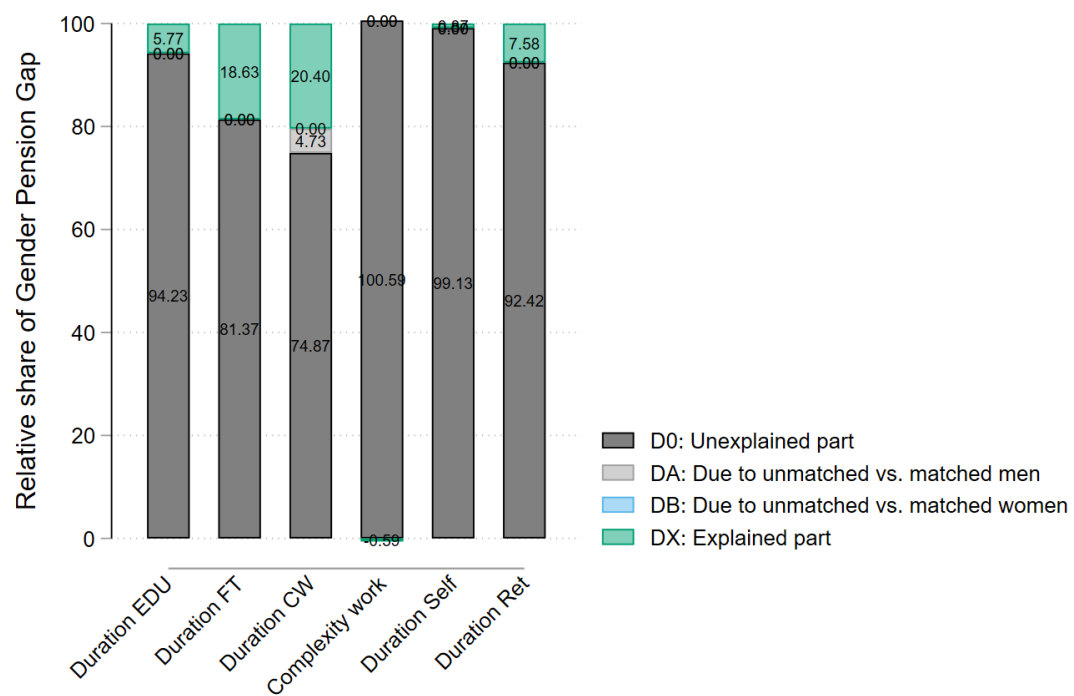
Figure A5.13 Ñopo decomposition of Gender Pension Gap in West Germany



Notes: Ñopo decomposition of the gender gap in total pension income using single variables as covariates. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

The finding for the Netherlands is similar (Figure A5.14). The duration in care work attributes most to the GPG (25%) followed by the duration in full-time employment (19%). However, other important pension predictors are much less related to the GPG (duration in education and retirement) or not at all (complexity of working lives and duration of self-employment). Furthermore, in contrast to West Germany, apart from the duration of care work, no significant DA or DB components occur when considering single variables. That underlines that the association with the GPG is often more complex and cannot be captured in bivariate associations. Rather, interactions with other variables seem to be important, which is implemented well by matching based on several variables in the main model.

Figure A5.14 Ñopo decomposition of Gender Pension Gap in the Netherlands



Notes: Ñopo decomposition of the gender gap in total pension income using single variables as covariates. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A5.4 Additional analysis: Kitagawa-Oaxaca-Blinder decomposition over common support

The Kitagawa-Oaxaca-Blinder decomposition (Blinder, 1973; Kitagawa, 1955; R. Oaxaca, 1973) relies on one linear OLS regression for each group (i.e., men and women) separately. Thus, other than for the nonparametric $\tilde{N}$ opo decomposition, multicollinearity issues are more severe. For fitting the KOB models over common support for the matching samples from the $\tilde{N}$ opo decomposition, I, therefore, exclude highly multicollinear variables. Furthermore, I do not categorize the continuous life course variables as for the $\tilde{N}$ opo decomposition but model them as linear effects, following the conventional approach in the literature. Otherwise, the KOB models use the same variable sets as the stepwise $\tilde{N}$ opo models.

I apply the same choices for the KOB parameters as in Chapters 3 and 4: normalisation of the clusters, choice of reference coefficient based on the majority group, and coefficients of men, if no clear majority group.

The results are documented in Table A5.4 and Table A5.5 and discussed in section 5.4.2.

Table A5.4 KOB decomposition over the matched sample for West Germany, stepwise models

	Model 1: top 2 variables				Model 2: top 4 variables				Model 3: top 9 variables			
Men	22,115**				21,758**				21,100**			
	(48.74)				-36.07				-33.83			
Women	10,714**				11,568**				11,507**			
	(20.63)				-15.88				-15.62			
GPG	11,401**				10,190**				9,592**			
	(16.75)				-11.07				-9.926			
Total explained	2,305**	20.21%			2,272**	22.30%			1,859*			
	(7.024)				-3.453				-2.5			
Total unexplained	9,097**	79.79%			7,918**	77.70%			7,734**			
	(10.75)				-8.672				-8.056			
	explained		returns effect		explained		returns effect		explained		returns effect	
Duration CIV	1,279**	11.22%	-33.13	-0.29%	820.2**	8.05%	11.93	0.12%	-463.4	-4.83%	332.2	3.46%
	(6.043)		(-0.299)		-2.996		-0.0487		(-1.443)		-1.132	
Duration CW	1,025**	8.99%	-40.36*	-0.35%	504.5	4.95%	-40.65	-0.40%				
	(3.526)		(-2.191)		-1.162		(-1.451)					
Age at first CW spell					-534.7	-5.25%	-16.22	-0.16%				
					(-0.895)		(-0.0742)					
Age start of FT->Ret					1,482**	14.54%	-258	-2.53%				
					(3.493)		(-0.351)					
No transition FT->Ret									757.4+	7.90%	797.9	8.32%
									-1.701		-0.571	
Early transition FT->Ret									-28.62	-0.30%	206.9	2.16%
									(-0.613)		-0.818	
Rather early transition FT->Ret									121.6	1.27%	-44.23	-0.46%
									-1.207		(-0.117)	
Rather late transition FT->Ret									-28.05	-0.29%	-300.6	-3.13%
									(-0.216)		(-0.896)	
Late transition FT->Ret									241.4	2.52%	119.5	1.25%
									-1.086		-0.896	
Duration FT									1,258	13.12%	1,274	13.28%
									(1.491)		(0.864)	
Constant			9,170**				8,221**				5,348+	55.75%
			(10.11)				(6.424)				(1.724)	
n male	806		806	806	768	768	768		697		697	
n female	526		526	526	331	331	331		290		290	

Notes: Variables related to the duration and timing are included as linear effects (in contrast to the Nopo decomposition). Single variables were excluded due to multicollinearity. z-statistics in parentheses, ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A5.5 KOB decomposition over the matched sample for the Netherlands

	Model 1: top 2 variables			
Men	26,932**			
	(23.75)			
Women	14,526**			
	(16.23)			
GPG	12,406**			
	(9.530)			
Total explained	4,931**	39.75%		
	(4.712)			
Total unexplained	7,475**	60.25%		
	(4.441)			
	explained		returns effect	
Duration EDU	2,304**	18.57%	1,577**	12.71%
	(3.183)		(2.745)	
Duration CW	2,628**	21.18%	34.06	0.27%
	(3.638)		(0.147)	
Constant			5,864**	
			(3.437)	
n male	500	500	500	500
n female	425	425	425	425

Notes: Variables related to the duration and timing are included as linear effects (in contrast to the Ñopo decomposition). Single variables were excluded due to multicollinearity. z-statistics in parentheses, ** p<0.01, * p<0.05, + p<0.10. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A5.5 Robustness checks

A5.6.1 Alternative Feature Selection Algorithms

To verify the results of the feature selection, I apply complementary algorithms to the Boruta.

First, the variable importance measures of Random Forest algorithms, such as Boruta, are biased in the case of correlated features (e.g., Gregorutti et al., 2017; Tološi & Lengauer, 2011). Even though I exclude highly correlated variables (.9), the variable importance of remaining correlated variables might still be biased downwards. Gregorutti et al. (2017) have shown that the recursive feature elimination (RFE), a wrapper algorithm based on Random Forest, reduces the correlation bias on the variable importance. Another recently developed feature selection approach, that has been suggested as a more stable and robust alternative to Boruta for high-dimensional data (Degenhardt et al., 2019), is the vita algorithms (Janitza et al., 2018).

Second, other feature selection approaches, other than the ‘all-relevant’ algorithms like Boruta, aim at selecting only the most efficient set of features for predicting the outcome (Bolano & Studer, 2020). This is helpful, especially for confirming the top of the most important life course variables from the larger set resulting from the Boruta. LASSO is one of these more conservative feature selection approaches and a very popular algorithm.

I use the recursive feature elimination, and the lasso algorithms (both ‘minimal set’ approaches) as well as the vita algorithm (‘all-relevant’ approach) to assess the robustness of the main findings.

Overall, the main finding persists: none of the other algorithms suggests that the duration of full-time employment performs better than the duration of care work for predicting pensions. For West Germany, the duration of unpaid care work is even more important than

the duration of civil service (identified as the most important pension predictor by Boruta) and for RFE and Vita, it is the most important pension predictor.⁷⁶ For the Netherlands, Lasso and RFE do not even select the duration of full-time employment as the important predictor at all.

However, some differences stand out (Table A5.6). For Germany, life course aspects related to retirement (particularly the duration of retirement, but also AFtransition..CIVRET, AFtransition..FTRET) tend to perform poorer in the other algorithms. Lasso identifies late education as the most important, and late marriage and having one child as the second most important pension predictor. The overall complexity of working lives scores much higher in the RFE.

For the Netherlands, RFE identified two complexity measures for family life as the most important pension predictors that were not selected as relevant predictors by Boruta. Just like for West Germany, Lasso identifies a late end of education as the most important pension predictor. Additionally, the age at the first care work spell is assessed as much more important by RFE and Vita.

In conclusion, apart from several differences across the relevant pension predictors identified by the different algorithms, the main finding of the first analysis step is confirmed. Life course aspects related to care work are more important pension predictors compared to the duration of full-time employment. Furthermore, the variable sets for the decomposition would be similar when taking into account the results of all algorithms.

⁷⁶ Given that the RFE iteratively removes least important variables until only the most important ones remain, several variables may share a rank (i.e., they were removed at the same round or remained as most important).

Table A5.6 Important pension predictors by identified by different algorithms, ranks

West Germany					Netherlands				
features	rank Boruta	rank Lasso	rank RFE	rank Vita	features	rank Boruta	rank Lasso	rank RFE	rank Vita
duration.CIV	1	4	7	4	duration.CW	1	4	3	1
duration.CW	2	3	1	1	duration.EDU	2	11	2	4
pattern..CW	3		1	3	AFtransition..OTHFT	3	25	10	21
spell.CW_age_1	4	12	2	2	Complexity.Entropy_work	4		6	11
AFtransition..CIVRET	5	8	11	8	duration.Ret	5		9	5
AFtransition..FTRET	6	18	12	16	duration.SELF	6		10	7
duration.FT	7	5	8	5	pattern..CW..OTH	7		10	14
AFtransition..CW.PT...PTRET	8		9	13	duration.SNC	8			52
spell.CIV_age_1	9	10	9	10	spell.EDU_age_1	9	8	8	8
pattern..CIV	10		10	9	Age23_27.End_EDU	10	1		12
AFtransition..CWFT..FTCW	11			90	duration.CIV	11		9	13
pattern..EDU..CIV	12		11	14	spell.CW_age_1	12	19	4	2
duration.EDU	13	26	5	7	pattern..OTH..FT	13			38
duration.Ret	14			31	pattern..CW	14		7	6
pattern..FT..CW	15		9	6	AFtransition..EDUCIV	15	6		18
spell.PT_age_1	16	32	12	17	pattern..UNE..FT	16			40
AFtransition..MNCM1C	17		11	26	duration.FT	17			25
pattern...PT	18		12	34	pattern..FT..CW	18	7	5	3
Complexity.Entropy_work	19		6	38	AFtransition..UNEFT	19			47
AFtransition..EDUFT	20			25	pattern..EDU	20			10
Age23_27.End_EDU	21	1	6	11	Complexity.C	21		7	9
duration.SELF	22	27		46	pattern..EDU..CIV	22	9		17
Complexity.Turbulence_work	23	22	3	58	AFtransition..FTCW	23			15
Age58_62.Start_Ret	24	17		87	AFtransition..SELFRET	24	5		16
AFtransition..EDUFT..FTRET	25	11	10	18	Age58_62.Start_Ret	25	2		20
duration.PT	26		10	65	Age58_62.End_FT	26			28
Complexity.C	27		7	67	AFtransition..CIVFT	27	15		44
AFtransition..FTCW	31		10	12	pattern..MNC..M1C..M2C..DIV	28	3		56
Age28_32.End_MNC	36	7		52	spell.OTH_age_1	29	20		29
pattern..EDU	43	15	11	20	pattern..MNC..M1C..DIV	33			66
spell.FT_age_1	44		12	43	AFtransition..EDUFT..FTRET	34	18		24
duration.OTH	52	13			Age18_22.Start_CIV	40	12		60
Complexity.Trans_work	54		8	84	AFtransition.MNCM1C..M1CM2C	44	16		51
pattern..FT...PT	56		4	41	pattern..MNC..M2C..DIV	45	26		55
Age23_27.Start_CW	59		11	30	Complexity.Trans_work	49		8	36
AFtransition..CW.PT	61		12	15	pattern..FT...PT	52		9	72
Age28_32.End_EDU	63	6			Complexity.Entropy_family	63		1	23
Age18_22.End_FT	85	14		97	Complexity.Turbulence_family	68		1	34
Complexity.Entropy_family	87		12		duration.PT	69		10	49
Age18_22.Start_M2C	93	16			AFtransition...PTRET	87	17		
Age38_42.End_M1C	104	9			pattern..SNC	108	10		
Age38_42.Start_M1C	109	2		85	AFtransition..CIVFT..FTRET	117	14		62
					Age23_27.End_CIV	118	13		

Notes: Results for all features selected in the Netherlands or West Germany ranked by mean importance based on the Boruta algorithm. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A5.5.2 Social policy relevant sensitivity

I further run some social-policy-relevant sensitivity analyses, most importantly the inclusion of survivor pension in the pension definition and the exclusion of individuals without own pensions (see Table A5.7). This is of relevance to assess whether the predictors of pensions change when two of the major conceptual parameters change (see Chapter 2.1), given that many studies on gender pension inequalities have considered survivor pension or have focussed on gender pension inequalities among pension recipients only (see Table 2.1 for an overview). Consequently, these analyses assess whether the importance of pension predictors, specifically the duration of full-time employment, varies based on the two conceptual choices in question.

I. Including survivor pensions.

First, I include survivor pensions in the outcome variable pension income.

Including survivor pensions does not lead to remarkable differences in the main results. However, in West Germany, the rank of the duration of full-time employment increases from 7 to 3. At the same time, the duration of care work remains the second most important pension predictor, and the distance in mean importance between both predictors remains the same (not shown).

In the Netherlands, the duration of full-time employment remains a relatively irrelevant predictor for pension inequality. The relative importance of the duration of care work decreases slightly, while, for example, the years spent single without having children increases.

Table A5.7 Important pension predictors for different pension definitions, ranks

West Germany				Netherlands			
features	rank main	rank survivor pension	rank pension recipient s	features	rank main	rank survivor pension	rank pension recipients
duration.CIV	1	1	1	duration.CW	1	3	1
duration.CW	2	2	2	duration.EDU	2	1	2
pattern..CW	3	10	8	AFtransition..OTHFT	3	7	3
spell.CW_age_1	4	6	7	Complexity.Entropy_work	4	6	4
AFtransition..CIVRET	5	7	3	duration.Ret	5	5	8
AFtransition..FTRET	6	4	10	duration.SELF	6	4	6
duration.FT	7	3	9	pattern..CW..OTH	7	20	9
AFtransition..CW.PT...PTRET	8	5	4	duration.SNC	8	2	21
spell.CIV_age_1	9	11	13	spell.EDU_age_1	9	10	5
pattern..CIV	10	8	15	Age23_27.End_EDU	10	8	7
AFtransition..CWFT..FTCW	11	12	5	duration.CIV	11	13	17
pattern..EDU..CIV	12	9	6	spell.CW_age_1	12	15	12
duration.EDU	13	14	11	pattern..OTH..FT	13	9	13
duration.Ret	14	16	12	pattern..CW	14	18	16
pattern..FT..CW	15	17	23	AFtransition..EDUCIV	15	11	10
spell.PT_age_1	16	21	24	pattern..UNE..FT	16	21	11
AFtransition..MNCM1C	17	23	34	duration.FT	17	17	18
pattern...PT	18	32	16	pattern..FT..CW	18	14	14
Complexity.Entropy_work	19	25	14	pattern..EDU	19	16	19
AFtransition..EDUFT	20	22	22	Complexity.C	20	19	15
Age23_27.End_EDU	21	13	17	spell.OTH_age_1	21	12	20
duration.SELF	22	15	36				
Complexity.Turbulence_work	23	28	21				
Age58_62.Start_Ret	24	18	33				
AFtransition..EDUFT..FTRET	25	26	19				
duration.PT	26	30	25				
duration.SNC	27	24	28				
Age23_27.Start_FT	28	20	26				
Age58_62.End_CIV	29	27	35				
AFtransition..FTCW	30	34	32				
AFtransition..CWRET	31	31	18				
AFtransition..EDUCIV	32	35	31				
pattern..EDU	33	29	29				
AFtransition..EDUCIV..CIVRET	34	36	39				
pattern..FT..FT	35	38	20				
spell.EDU_age_1	36	33	30				
AFtransition..SELFRET	37	19	37				
Age63_65.End_CW	38	39	27				
Age63_65.End_CIV	39	37	38				

Notes: Results for all features selected in the Netherlands or West Germany ranked by mean importance based on the Boruta algorithm. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

II. Excluding individuals without pension income.

As a second major sensitivity analysis, I exclude individuals with no own pension receipt. Consequently, this robustness check focuses only on pension recipients and ignores the gender inequalities in access to pensions in the first place.

Excluding individuals with 0 pensions, removes 86 individuals in West Germany and 20 in the Netherlands.

Again, the main patterns do not change. The duration of care work remains as important as in the main model and the relative importance of the duration of full-time employment for overall pension income even decreases slightly and is not even confirmed as a relevant pension predictor in the Netherlands (not shown).

In West Germany the relative importance of ordering and timing of care work decreases slightly when considering only pension recipients as well as when including survivor pensions. Instead, more complex patterns that capture the timing and ordering of care work in synergy with other life course aspects are more relevant (e.g., AFtransition..CW.PT...PTRET for including survivor pensions, and AFtransition..CWFT..FTCW for excluding individuals without pension receipt).

In conclusion, the main findings do not change when adjusting the two main conceptual choices concerning the definition of pension income. Even though the duration of full-time employment reaches a higher relative importance when including survivor pensions in the outcome, care work indicators are still more or just as important. As a result, future literature should consider care-work-related life course aspects besides more standard work characteristics – independent of the conceptual choices for defining pension income.

III. Adjusting for single households.

To take into account that the public pension payments in the Netherlands depend on the household size, I control for whether individuals live in a single household. The ranking of almost all variables, especially the most important ones, proves to be very robust (results are available upon request).

A5.5.3 Life course variable operationalisation and functional forms

I differentiated between three different life course variable operationalisations for the feature selection. The first option incorporates duration and complexity as continuous variables. This aligns with how the duration in life course states is typically considered in literature stream 1 and is consistent with the aim of this study to evaluate whether the duration in life course states, especially in employment, is indeed the most important pension predictor. Life course aspects III (age at the nth spell) and V (age at the transition at the start of a certain subsequence) are categorised as quartiles because of the distributions and the zero-inflation. For the life course aspect II, the number of frequent sub-sequences of states ('pattern'), I merge small categories for each country separately. If the number of observations in any category is lower than 15, I merge this category to the next smallest category. This operationalisation is used in the main model.

Option II categorises all variables, including the duration and complexity. The duration is categorised in terciles, quartiles or quintiles, depending on the distribution. Complexity measures are categorised in deciles. This is the operationalisation of variables that are used in the $\tilde{N}$ opo decomposition since I need categorical data for the exact matching.

Option III exploits all heterogeneity in the data but ignores any data structure and considers the data as it is. Highly correlated variables were not excluded. It thus also contains more variables compared to options I and II.

To check the robustness of the results using functional form option I, I run the analysis using variable sets applying the operationalisations of options II and II (results are available upon request).

For West Germany, with two exceptions (AFtransition..CW.PT...PTRET, AFtransition..CWFT..FTC not selected) the top-15 pension predictors remain almost the same and the duration of care work ranks most important (results are available upon request). However, when considering the duration of full-time employment in quintiles, it ranks second most important, slightly more important than the timing and ordering of care work, but still with some distance to the importance of care work duration.

For the Netherlands, categorising the duration and complexity variables leads to much lower importance of the complexity of working lives as well as the duration of self-employment. However, the duration of care work is still the most important pension predictor while the duration of full-time employment remains relatively unimportant.

I conclude that the results of the Life Course Feature Selection are slightly sensitive to the functional form of the variables, but the main results do not change.

A5.5.4 Variable selection for $\tilde{N}opo$ decomposition

I. Overall variable selection

In the main model, I select the variables for the stepwise $\tilde{N}opo$ decompositions based on the variable importance ranks resulting from the feature selection pooled over gender. For choosing the cutoffs between different variable sets for the stepwise models, I usually relied on particularly large differences in variable importance. However, I added highly influential variables as single variables since they largely impacted the matchable samples (e.g., complexity work lives or the duration of retirement for the Netherlands which both

increased the number of matching strata dramatically). This approach allows for a meaningful and comprehensible interpretation of the stepwise models.

Additionally, given that the Ñopo decomposition cannot consider many variables simultaneously due to the curse of dimensionality, I exclude single variables that might not add to the understanding of the GPGs due to empirical reasons or lack of substantive relevance. For the Netherlands, I exclude variables related to the state “other” since these seem not to have a substantive meaning (this applies to ‘AFtransition__OTH_FT’ and ‘pattern__CW__OTH’). Including them in the Ñopo decomposition does not change the results and thus does not contribute anything substantively (not shown). Both variables were also not robust as main predictors in several sensitivity analyses (e.g., using alternative algorithms).

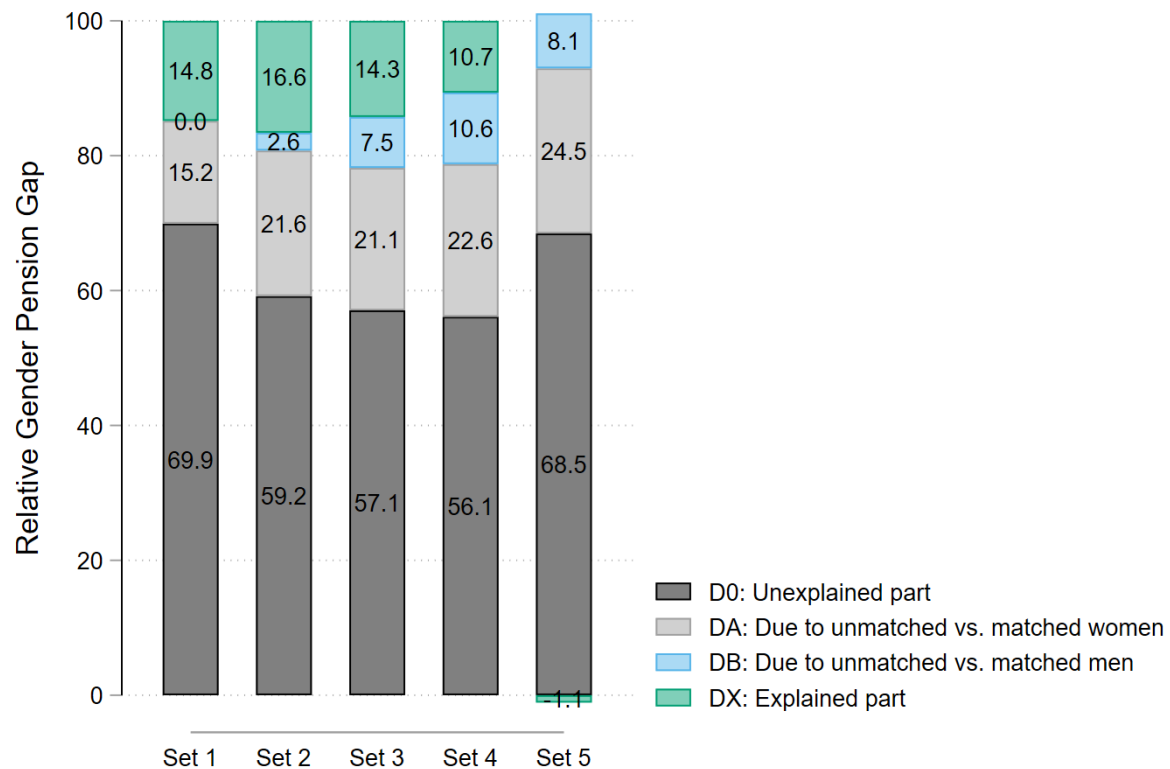
For both countries, I excluded highly multicollinear variables. Even though these are less severe for matching compared to regression analyses, multicollinear variables might conceal which associations are important for the GPG. For West Germany, I excluded the number of care work spells (ordering) in set 2 and subsequently, and the ordering of civil service from set 3 onwards. In set 4, I excluded the duration of retirement and the frequency of the ordering of civil service following education spells over the life course. In set 5, I excluded the number of part-time spells (ordering). In the Netherlands, multicollinearity is less of an issue. For variable sets 3 onwards, I exclude the duration in childless singlehood.

II. Choosing the maximal number of variables in the main models

Due to the exact matching approach, only a limited number of variables is feasible to be included until the ‘curse of dimensionality’ is reached. With the increasing number of characteristics in the models, the size of the matched samples decreases as well as the explained share in West Germany. Particularly in model 5 in West Germany and model 4 in the Netherlands the ‘curse of dimensionality’ becomes visible: including more variables

in the model turns the explained share negative. However, with only 11% (14%) of women and 23% (49%) of men being matched in West Germany (Netherlands), the explained share is based on highly selective samples and its meaningfulness is questionable. As a result, I focus on models 1-4 in West Germany and models 1-3 in the Netherlands for a meaningful interpretation.

Figure A5.15 Ñopo decomposition of Gender Pension Gap in West Germany



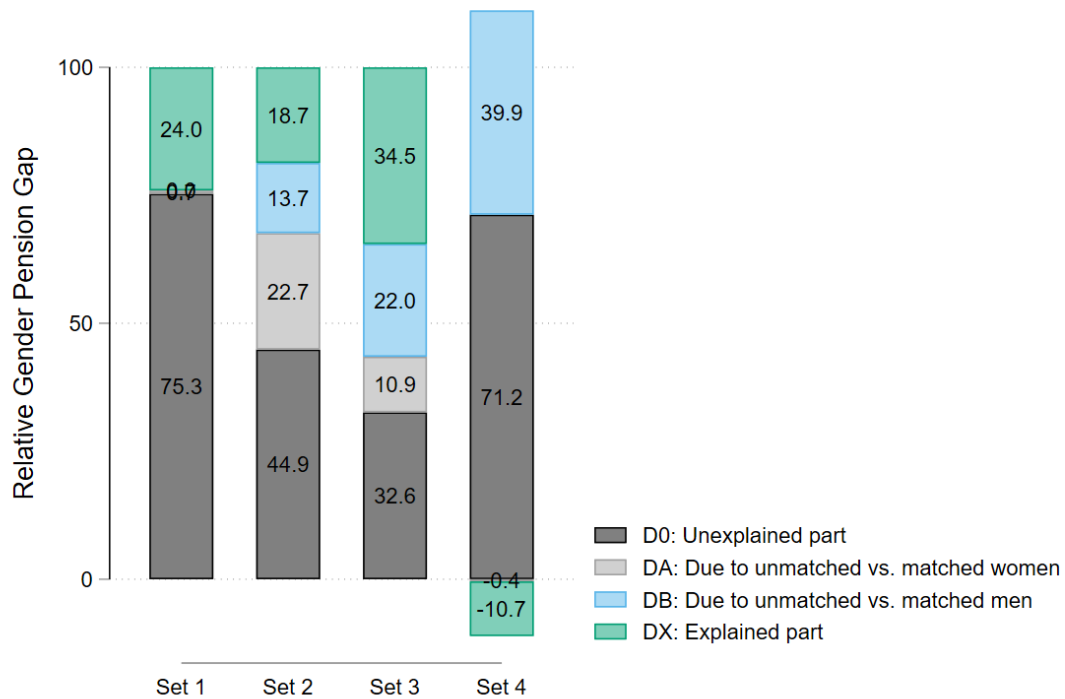
Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.3). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A5.8 Ñopo decomposition of Gender Pension Gap in West Germany

Models	Variable set 1	Variable set 2	Variable set 3	Variable set 4	Variable set 5
	Duration care work, duration civil services	set 1+ age at first CW spell, age at start of CIV->Ret & FT->Ret subsequences	set 2 + duration FT, age at start of CW->PT_PT->RET and CW->FT_FT->CW subsequences, age at 1st civil service	set 3 + duration EDU, number of FT- subsequence >CW	SET 4+ age at first PT spell, age at transition of MNC->M1C, Complexity work life courses, age at transition of EDU->FT
Variables					
% matched women	64	41	35	30	11
% matched men	100	95	86	72	23
N of strata	13	79	280	475	1148
N of strata matched	6	19	46	62	57
D0	9405	7959	7674	7547	9207
DA	2048	2902	2842	3043	3294
DB	0	356	1015	1426	1095
DX	1997	2232	1918	1433	-147

Note: Ñopo decomposition of the Gender Pension Gap into the four components presented in equation 2. % matched refers to the share of women and men respectively matched for the respective characteristics controlled for.

Figure A5.16 Ñopo decomposition of Gender Pension Gap in the Netherlands.



Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.4). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Table A5.9 Ñopo decomposition of Gender Pension Gap in the Netherlands

Models	Variable set 1	Variable set 2	Variable set 3	Variable set 4
Variables	Duration care work, duration education	Set 1 + Complexity work life courses (entropy)	Set 2 + duration self- employment	Set 3 +, duration_Re t
% matched women	72	25	20	14
% matched men	100	89	75	49
N of strata	24	116	188	391
N of strata matched	10	36	37	45
D0	9403	5603	4071	8894
DA	82	2841	1357	-53
DB	0	1712	2749	4981
DX	3003	2332	4311	-1334

Notes: Ñopo decomposition of the Gender Pension Gap into the four components presented in equation 2. % matched refers to the share of women and men respectively matched for the respective characteristics controlled for.

A5.5.5 Parameters and robustness of $\tilde{\text{Nopo}}$ decomposition

Besides some more general limitations, the results of the $\tilde{\text{Nopo}}$ decomposition might be subject to some parameters.

I approach two of these issues empirically via sensitivity analyses.

I. Reference group

Due to the nonparametric nature of the $\tilde{\text{Nopo}}$ decomposition, the references cannot be chosen per variable category (as suggested for the KOB decomposition in Chapter 3) but must be the same for the whole model. Switching the reference group affects only the explained and the unexplained part of the $\tilde{\text{Nopo}}$ decomposition.

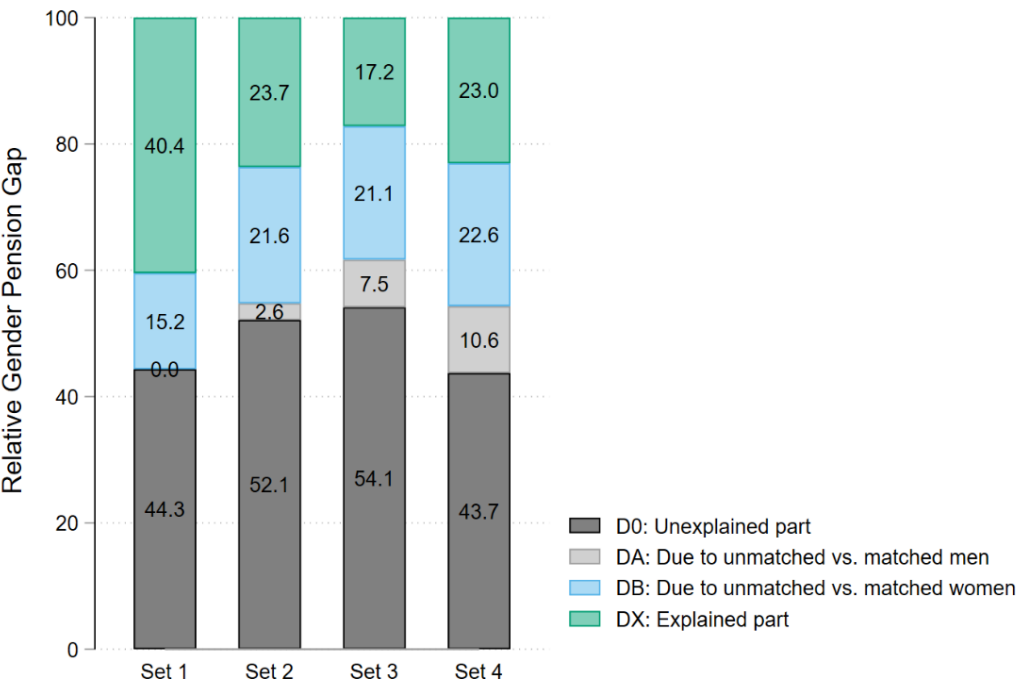
With women as group A in the main model (and thus as a reference in terms of returns to characteristics, $Dx = \bar{Y}_{AB,m} - \bar{Y}_{A,m}$), DX is the part due to differences in characteristics when the returns of women to these characteristics are applied. When using men as group A we have the part due to differences in characteristics weighted by the returns to these characteristics of men. The latter is the most common choice in KOB decompositions in the field.

In the main model, $D0$ ($D0 = \bar{Y}_{B,m} - \bar{Y}_{AB,m}$) refers to the part of the gap due to differences in returns if both, men and women, had the characteristics of men. If we switch the groups or references, we have the reverse, i.e., the differences in returns are weighted by the characteristics of women.

In West Germany, switching references leads to larger explained shares, which is most likely due to higher returns of men for the same characteristics (Figure A5.17). In the Netherlands, the explained part increases for model 1 only (see Figure A5.18). For all other models, it decreases to very small or even negative shares, suggesting that the explained

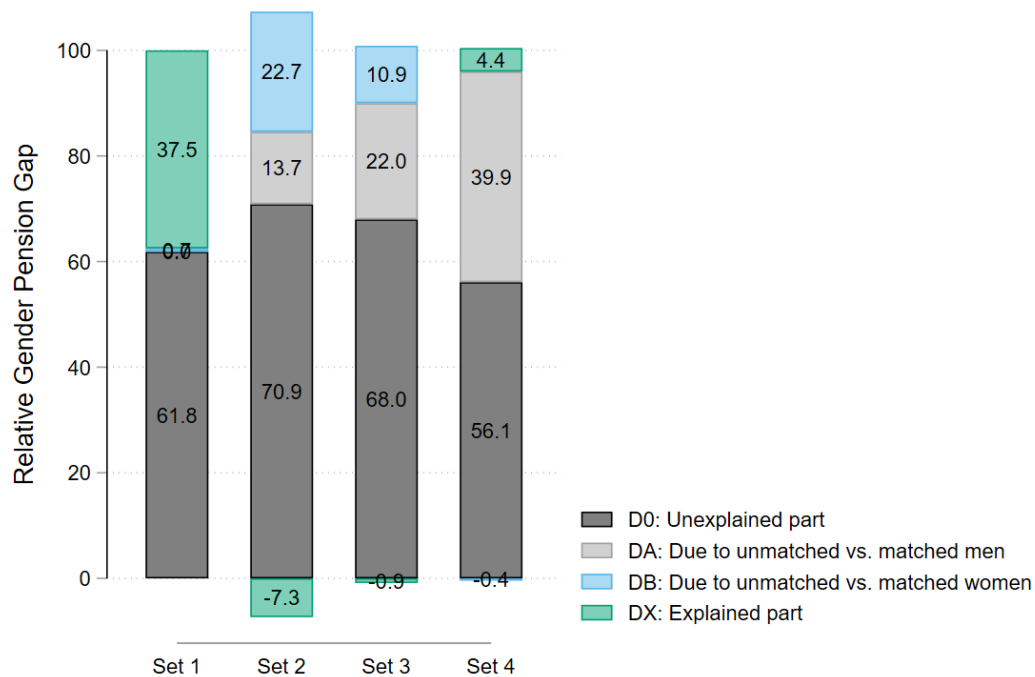
shares in the Netherlands are not robust regarding the references and the pension returns between matched men and women differ substantially.

Figure A5.17 Ñopo decomposition with men as group A (‘returns reference’), West Germany



Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.3). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.18 Ñopo decomposition with men as group A, Netherlands



Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.4). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

II. Matching approach and robustness of DA and DB

The results of the main analysis highlight that large shares of the GPGs in both countries are due to the very high segregation of men and women on different life course aspects. However, the analysis also shows the limits of the Ñopo decomposition when using a large number of characteristics given that only small shares of women are matchable after matching based on a few variables („curse of dimensionality“, Hamjediers & Sprengholz, 2023, S. 12). Particularly the low share of matched women raises the question whether DA and DB arise because less frequent combinations of life course characteristics do not appear in the small sample whereas they do occur in the overall population.

Hamjediers and Sprengholz (2023) suggest two ways of assessing this data-driven challenge of the Ñopo decomposition. First, using the less exact Propensity Score Matching to ease the curse of dimensionality and see whether the DA and DB are still significant (Hamjediers & Sprengholz, 2023). Second, applying a larger dataset to rule out that the

lack of support is only observable in the study sample. Given that I do not have access to such data, the latter has to be assessed by future research, e.g., using register data containing life course data on the work and family categories. Still, it is highly unlikely that the share of men who spent many years doing unpaid care work would increase in larger samples – at least for the studied cohorts in West Germany and the Netherlands.

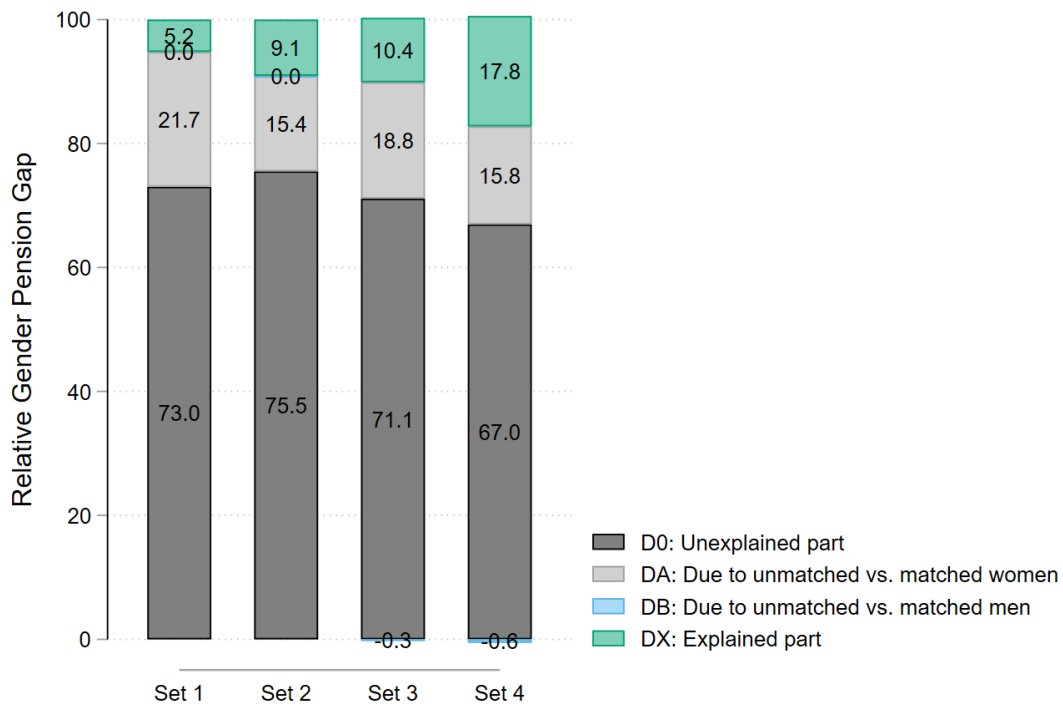
Propensity Score Matching

As expected, using a less strict matching approach like Propensity Score Matching leads to overall higher shares of matched individuals (e.g., a minimum of 51% of women matched in West Germany, and 72% in the Netherlands; not shown). This confirms the significant lack of common support for life course experiences across genders in both countries. For West Germany the large DA components are clearly confirmed, reaching similar but partly slightly lower levels (Figure A5.19). For the Netherlands, however, the DA component is much stronger reduced to 5% but still substantial (Figure A5.20, compared to 23% and 11% in the main model).

On the contrary, nearly all men are successfully matched in both countries when using Propensity Scores (minimum 99%, not shown). Consequently, the DB components of the GPG are not confirmed. This suggests that if the matching is relaxed, women with comparable life course aspects to men are available.

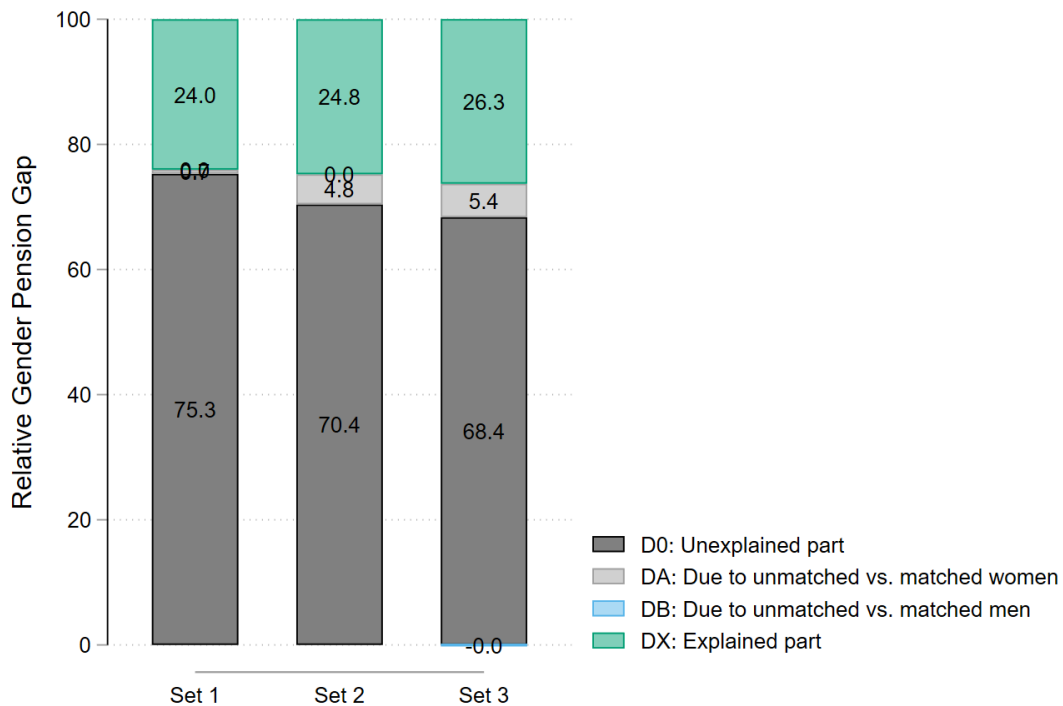
Overall, the sensitivity analyses reinforces the main result highlighting large parts of the GPG resulting from gender-exclusive life course experiences that are differently rewarded – though to a lesser degree in the Netherlands. Even when relaxing the matching, for up to half of the women no comparable men are identifiable, particularly in West Germany. As the main results suggest, the lack of common support is most likely attributable to the limited care work engagement of men.

Figure A5.19 Ñopo decomposition using PS Matching, West Germany



Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.2). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

Figure A5.20 Ñopo decomposition using PS Matching, the Netherlands



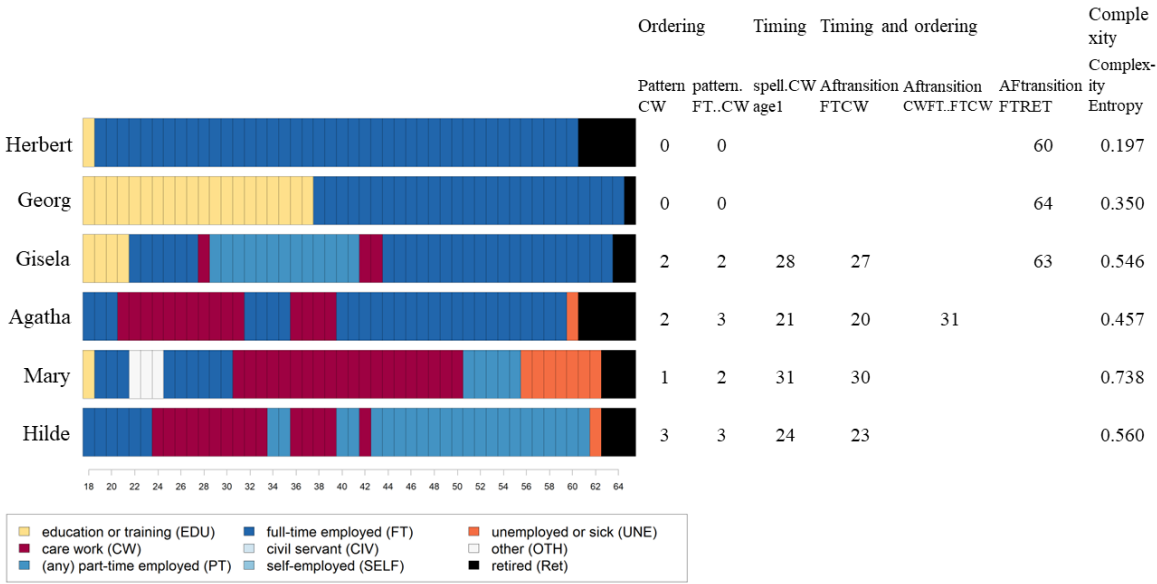
Notes: Ñopo decomposition of the gender gap in total pension income using the analysis sample and different sets of variables as covariates, chosen in descending order by their variable importance as estimated by the Boruta algorithm (see Table 5.2). Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

A5.6 Further documentation

A5.6.1 Life course variable description

Figure A5.21 depicts the example work trajectories as well as example variables for the different life course aspects.

Figure A5.21 Example trajectories and life course variables.



Note: Example working life courses selected from the analysis sample. The colour code shows which working life states the individuals experienced for each year, spanning ages 18-65.

The variable for *ordering* ('pattern') captures the times an individual experienced a certain life course state period or subsequence along the observed life course. For example, 'pattern CW' captures how often the individuals have had care work periods. 'Pattern FT..CW' captures how often full-time periods were followed by care work periods. For example, Mary had only one care work period, but before that two distinct full-time employment periods, and consequently, the subsequence FT->CW occurred twice in her life course.

The *timing* variables cover the age period when transitions between states occurred as dummies (not depicted in Figure A5.21) as well as the age at which each nth spell per state occurred along the trajectory. For example, the variable spell.CW.age1 captures the age at the first care work spell, which is 21 for Agatha and 31 for Mary.

Variables with the prefix ‘Afrtransition’ capture the age at the first transition of a certain subsequence, and consequently are proxies for the life course aspects of *timing* and *ordering*. For example, Afrtransition FTCW returns the age at the first transition from full-time employment to care work. Afrtransition CWFT..FTCW captures the age at the first transition from care work to full-time employment if later in the life course a transition from full-time employment to care work occurs. Among the example trajectories, only Agatha experiences such a combination of transitions. Afrtransition FTRET returns the age at the transition from full-time employment to retirement. Only Herbert, Georg and Gisela experience this transition, as the other individuals retire directly from a period of unemployment or sickness.

And lastly, sequence complexity indices measure the instability and unpredictability of the whole observed life course (for an overview see Pelletier et al., 2020; Ritschard, 2021). Among the examples in Figure A5.21, Mary has the most unpredictable life course.

A5.6.2 Boruta Algorithm

In the following, I briefly summarise the steps of the Boruta algorithms based on Kursa and Rudnicki (2010).

The Boruta algorithm first creates shadow features, that is copies, of all variables. After eliminating some correlations of the shadow features, it runs a random forest over the whole set of variables, comprised of the original as well as the shadow variables. It then compares the importance of all original variables to the best-performing shadow variable. Variables with significantly higher importance are deemed important, and those with significantly lower importance as unimportant. The variables deemed unimportant are removed together with all shadow variables.

The whole procedure is repeated until all variables are categorised as either important or unimportant or the limited number of random forest runs is reached. For more details, please refer to Kursa and Rudnicki (2010).

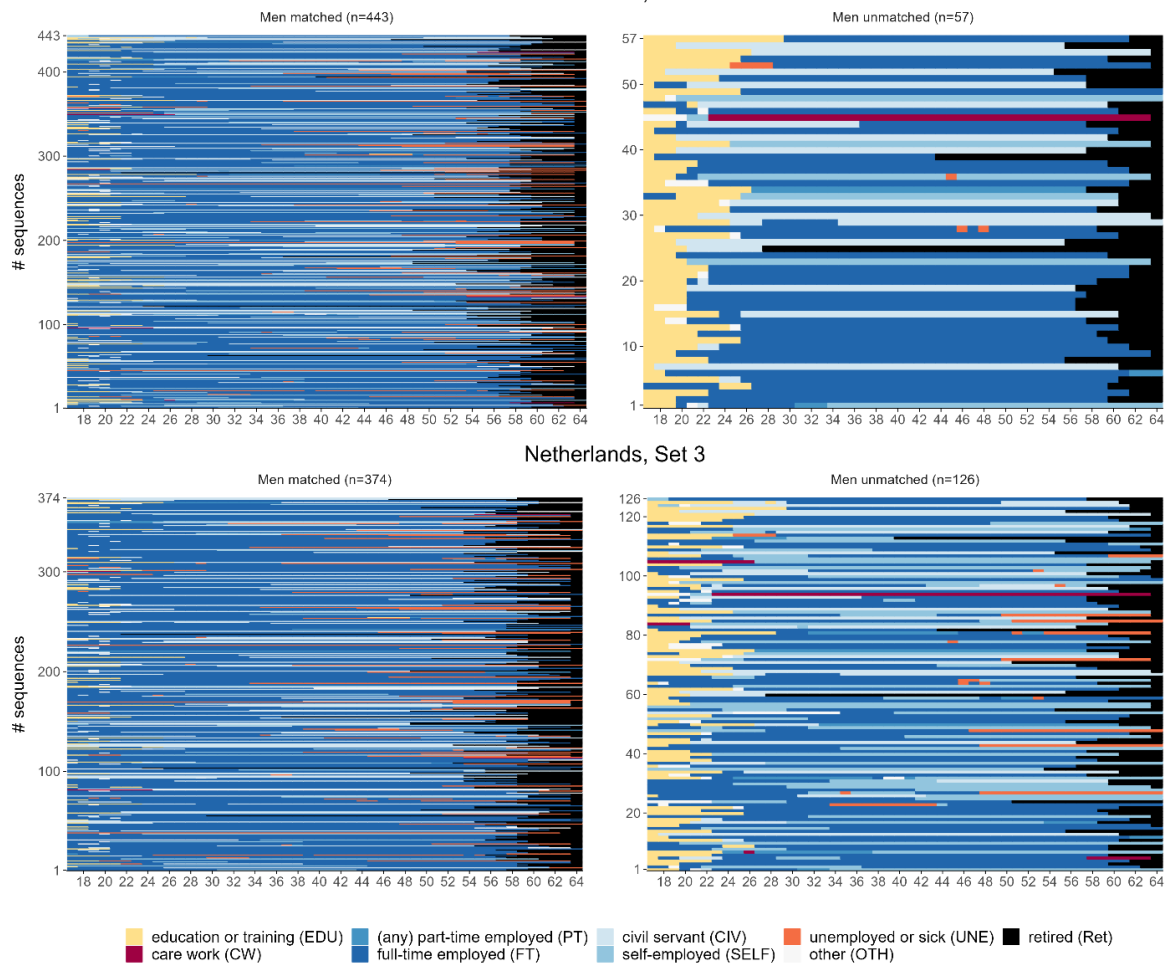
A5.6.3 Parameter tuning for Boruta

For the Random Forests underlying the Boruta algorithm, I tune the number of variables considered at each split of the trees, m_{try} , the tree complexity (minimal node size) and the sampling (fraction of sample used for each tree). I focus on these three parameters as these are the most important parameters for RFs considering variable importance (Probst et al., 2019). As suggested, I use subsampling without replacement (Strobl et al., 2007) and do not tune the number of trees but rather increase the number of trees for the final model (to 10 times the number of features) to win some stability for variable importance (Boehmke & Greenwell, 2019). For the other parameters, I use the default values. To find the optimal parameter values, I use a Cartesian search and a random grid search (Boehmke & Greenwell, 2019) as well as a tuning approach specialised in Random Forests (Probst et al., 2019, tuneRanger).

A5.6.4 Life course visualisation of different samples by gender and matching status for men

For Dutch men, including the duration of self-employment to the overall complexity and educational years (model 3) sorts the more privileged self-employees (those starting self-employment mostly after long educational periods) in the group of unmatchable men (Figure A5.22).

Figure A5.22 Work life courses of men by matching status and model, Netherlands
Netherlands, Set 2



Notes: Samples based on $\tilde{\text{Nopo}}$ decomposition of the Gender Pension Gap presented in equation 2. Own calculation based on the analysis sample and SHARE waves 2–6, v7.1.0. Not weighted.

This seems to be the explanation for the increases in DB, i.e., the part of the GPG due to pension income differences between unmatched and matched men, from models 2 to 3. Privileged self-employed individuals, despite lacking access to occupational pensions, may possess the means to secure exceptionally generous personal pensions. Conversely, less privileged self-employed individuals may find themselves without access to any form of pension, be it public, occupational, or personal.