


ORIGINAL ARTICLE

Open Access



# Harmonising French and German administrative DADS-EDP and SIAB data

A practical guide on how to create comparable data sets on maternal employment

Andreas Filser<sup>1\*</sup> , Pascal Achard<sup>2</sup>, Inga Marie Amend<sup>1,3</sup>, Corinna Frodermann<sup>1</sup>, Ida Gaede<sup>4</sup> and Sander Wagner<sup>5,6</sup>

## Abstract

Administrative data are increasingly used in social science research, including for studying the labour market consequences of motherhood. Yet despite the growing use of such data within countries, cross-national comparative research remains limited because administrative data require substantial harmonisation before meaningful comparisons are possible. In this paper, we harmonise two rich administrative data sets suitable for comparative research on maternal employment: the French DADS-EDP and the German SIAB. France and Germany present a compelling case for such comparisons given their differing norms and policies regarding childcare and maternal employment. We document our harmonisation decisions on data structure, sample selection, employment measurement, and key socio-economic variables, and provide the full Stata code as supplementary material. We illustrate the potential of the harmonised data by examining how maternal employment trajectories evolve over time in France, West Germany, and East Germany. The analysis reveals substantial cross-national and temporal variation in employment trajectories after childbirth, highlighting the value of harmonised administrative data for comparative research on maternal employment.

**Keywords** Motherhood penalty, Administrative data, Data harmonisation, Comparative research, Harmonisation, France, Germany

**JEL classification** C81, J16, J13, J21, J22

## 1 Introduction

Large-scale administrative registry data are increasingly used in social science research, yet comparative cross-national studies based on such data remain scarce. This is despite administrative social security data having numerous advantages over traditional surveys: they avoid panel attrition and recall bias, generally cover longer time spans and larger populations, and often allow linkage to surveys or other administrative data sources, such as firm-level data or health records. The increasing use of administrative data in single-country studies demonstrates the importance of these advantages for addressing many of our most pressing research questions.

\*Correspondence:

Andreas Filser  
andreas.filser2@iab.de

<sup>1</sup>Institute for Employment Research, Nuremberg, Germany

<sup>2</sup>Paris Dauphine University - PSL, Paris Cedex 16, France

<sup>3</sup>Friedrich-Alexander University Erlangen-Nürnberg (FAU), Nuremberg, Germany

<sup>4</sup>Sciences Po, Centre for Research on Social Inequalities (CRIS), CNRS, Paris, France

<sup>5</sup>Leverhulme Centre for Demographic Science, University of Oxford, Oxford OX13UQ, UK

<sup>6</sup>Nuffield College, University of Oxford, Oxford OX1 1NF, UK

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

However, the increased use of administrative data in within-country studies has not been matched by a comparable rise in cross-national comparisons. A key challenge for cross-national research using administrative data lies in the heterogeneity of data structures. Among other things, differences in sampling frames, population coverage, and variable coding and definitions require extensive harmonisation before meaningful comparisons can be made. While promising initiatives have emerged (e.g., Bassanini et al. 2024; Bertheau et al. 2023; Kleven et al. 2025; Tomaskovic-Devey et al. 2020; van der Vleuten et al. 2023), the analytical potential of harmonised administrative data for cross-national comparisons remains largely inaccessible to the broader research community. The substantial upfront investment required to harmonise administrative data across countries often prevents researchers from leveraging the unique insights such data could offer when it comes to comparing welfare regimes, as well as national and regional cultural, social, and policy contexts.

Among European countries, France and Germany stand out as especially suitable cases for comparative research based on harmonised administrative data. They are the European Union's two largest economies, which regularly compare and reference each other's social and economic policies. In both countries, rich administrative data, including linked employer–employee information, are available to researchers in comparably structured forms: the French DADS-EDP and the German SIAB are based on individual-level social security notifications and provide individual employment histories, which can be augmented with firm-level information. Despite some differences, both sources enable longitudinal analyses of employment trajectories dating back to 1967 (DADS-EDP) and 1975 (SIAB). Their large sample sizes allow for socio-economic sub-group analyses, the exploration of small-scale regional differences, and the evaluation of long-term temporal trends and policy changes.

France and Germany present a compelling case for comparative labour market research, with questions around maternal employment emerging as a particularly relevant area of research. Despite many similarities in their welfare states (Esping-Andersen 1990), the two countries have taken distinct paths in integrating women into the workforce, designing family policies, and supporting dual-earner households (Mazur 2003). France provides extensive all-day childcare services, enabling mothers to re-enter the labour market quickly and has considerably lower motherhood earnings and employment penalties than Germany (Budig et al. 2016; Hook and Li 2025).

Despite their potential, administrative employment data from France and Germany have not been readily accessible for comparative research. Comparative

analyses require resolving differences in episode structures, variable definitions, categorisation schemes, and the sample coverage of the two original data sets. Among the most important differences to resolve are SIAB's recording of employment periods as precise employment episodes with exact daily start and end points compared to the annually aggregated employment information in DADS. Additionally, part-time employment is recorded categorically in SIAB, while DADS captures the hours worked. Addressing these differences is essential to unlocking the full potential of both data sets for comparative research.

In this paper, we document our harmonisation of the French DADS and German SIAB administrative data sets and provide the relevant code as Stata DoFiles. This will enable other researchers to conduct comparative research on employment outcomes across the two countries. The documentation outlines our decisions on variable construction, sample selection, and data structure alignment. As part of this effort, we provide a set of harmonised variables specifically tailored to the study of motherhood and its labour market consequences. To facilitate and encourage future comparative research using both data sets, we share the full Stata code in the supplementary material and explicitly indicate the corresponding DoFiles for each harmonisation step throughout our description. Despite the focus on maternal employment, our approach is adaptable for broader research on employment dynamics. The code can be easily used and adapted to construct additional harmonised variables for other research questions.

Throughout this paper, we distinguish between two types of harmonisation. Mechanical harmonisation refers to structural alignment of the two data sets; for instance, aggregating SIAB's daily employment episodes into an annual format to match the structure of DADS-EDP. Conceptual harmonisation, by contrast, aims at functional equivalence: ensuring that variables measure the same underlying concept across both data sets, even where the technical operationalisation necessarily differs. Where full equivalence cannot be achieved, we make remaining differences explicit. Researchers should therefore interpret harmonised variables as functionally comparable rather than identically operationalised.

We demonstrate the analytical potential of the harmonised dataset via an illustrative analysis of maternal employment trajectories in the five years following the birth of the first child. Leveraging the long observation time span and large sample size, we analyse how these trajectories evolved between 1997 and 2014. While our analysis highlights one application of the harmonised data sets, their linked employer–employee structure opens up further avenues for research on topics such as maternal sorting across firms or the role of firm

characteristics in shaping gendered employment outcomes of motherhood.

The remainder of the paper is structured as follows: Sect. 2 introduces the DADS and SIAB data sets and outlines sample construction. Section 3 details the harmonisation steps of the underlying data structure, while Sect. 4 describes our approach to harmonising information on employment volume (full- and part-time). Section 5 provides a brief overview of further harmonisation decisions for selected variables. In Sect. 6, we illustrate the analytical potential of the harmonised data through an example analysis of how mothers' employment trajectories evolved over time in both countries. Finally, Sect. 7 discusses remaining limitations of the harmonisation, and Sect. 8 outlines data access modalities. The supplementary materials include a complete list of harmonised variables together with the corresponding Stata code.

## 2 Data sources and harmonisation of sample construction

Comparable administrative data with linked employer-employee information is available to researchers in both France and Germany. These data sets are based on individual-level social security declarations and are complemented by employer information - and, in the French case, additional demographic data. This section outlines the sampling structure of each dataset and discusses key differences in sample coverage.

On the German side, our harmonisation draws upon the Sample of Integrated Labour Market Biographies version 1975–2021 (SIAB - Graf et al. 2023), which is a 2% sample drawn from the population of individuals characterised by at least one of the following situations: employment subject to social security, marginal part-time employment, or receipt of social welfare benefits. SIAB derives from employer notifications to the social security system and from unemployment and social security benefits processing by the Federal Employment Agency (BA). Additionally, SIAB contains information on registrations for job seeking and participation in employment as well as on training measures run or sponsored by the BA. Individuals are traceable across these data sources, allowing for a comprehensive reconstruction of employment biographies. Establishment-level information from the Establishment History Panel (BHP - Ganzer et al. 2022) can be linked to individual records to obtain a linked employer-employee data set. SIAB covers approximately 80% of the working population in Germany, with the 1975–2021 version comprising the employment biographies of 1,977,295 individuals (Schmucker et al. 2023). In SIAB, data coverage for East Germany is only considered complete from 1993 onward (Schmucker et al. 2023). Berlin is coded as part of East Germany. Researchers analysing periods prior to 1993 should therefore consider

reclassifying Berlin as West Germany, as East German observations in these years effectively reflect West Berlin only.

The French data are based on the Déclarations Annuelles de Données Sociales (DADS), which combines annual payroll declarations by employers and the state payroll files (INSEE, 2025a, 2025b; Robert-Bobée and Gualbert 2021). The identification of births leverages combining this data with the Echantillon Démographique Permanent (EDP), which contains socio-demographic information based mostly on the French census. The panel DADS-EDP has been running since 1967 and collects information on individuals born on the first four days of October. Before 2002, labour market data was only collected for individuals born in October of an even year.

The DADS-EDP sample is subject to two independent restrictions that shape our analytical sample: one on the coverage of birth certificates (relevant for identifying first births), and one on the coverage of the DADS payroll data (relevant for observing labour market histories).

Regarding birth identification, the EDP covers individuals born on the first four days of October, April, and July, as well as between January 2nd and 5th (Meurs and Pora 2025). Birth certificate coverage for the children of these individuals varies across subgroups. For individuals born on October 1st or 4th, birth certificates are complete and first births can be identified reliably throughout the observation period. For individuals born on October 2nd or 3rd, birth certificates are sparse for children born in the 1980s and 1990s; following Meurs and Pora (2025), missing birth years from 1982 to 1997 can be reconstructed using census data from 1990 to 1999. For individuals born in January, April, or July - who were added to the EDP at a later stage - no birth certificate information is available before 2004, so any birth observed in the data could in principle be a second or higher-order birth. Given that parity cannot be reliably established for this last group, we exclude them from our sample.

Regarding labour market histories, DADS data before 2002 are only available for individuals born in an even year; from 2002 onward, odd-year births are included as well. Combining both restrictions, our final sample consists of October-born individuals with reliable parity information, restricted to even-year births before 2002 and including odd-year births from 2002 onward. This amounts to approximately 0.5% to 1% of French employees. Additional firm-level information can be added from the DADS-postes and DADS-établissements data sets.

Due to differences in the underlying administrative sources, sample coverage related to civil servants varies between SIAB and DADS. In Germany, SIAB does not include individuals who do not pay contributions to the social security system: self-employed, students, and

civil servants with special employment status (*Beamte*). *Beamte* are public sector employees who have special loyalty duties, job security, as well as rights, and make up 3.8% of the total workforce (destatis 2022). Most notably, *Beamte* enjoy high employment security and can only be dismissed under specific circumstances involving a severe violation of rules and regulations following a formal investigation. Consequently, they are not part of the unemployment insurance system and do not appear in the SIAB data. Yet, all other public sector workers - making up 62% of the public sector workforce in Germany - are covered in SIAB (destatis 2022). An important caveat is that the SIAB data do not allow for clearly identifying these workers or their employers as being part of the public sector. In contrast, the DADS covers all public service employees in France, who are classified into three types: *Fonction Publique d'État*, *Fonction Publique Territoriale*, and *Fonction Publique Hospitalière*. The most notable data restriction is that employees in the *Fonction Publique d'État* do not have firm identifiers before 2009, which prevents these individuals from being included in employer-employee linked analyses. They can still be used in individual-level models. To harmonise the two data sets, we exclude the *Fonction Publique d'État* from the French data, as they are the closest equivalent to the German *Beamte*, who are excluded from the German SIAB.

To harmonise our sample for the analysis of motherhood penalties, we include two additional sample criteria. Researchers focusing on other labour market outcomes should adjust these criteria to suit their specific research questions. First, we focus on women who held an employment contract at the onset of maternity leave, in line with how maternity is recorded in the SIAB data. In the SIAB, motherhood is identified through employment interruptions linked with maternity leave, which are reported either by employers or the unemployment benefits system, depending on the woman's employment status (see Sect. 3 and Müller et al. 2022). Restricting the sample to women who were in active employment before childbirth ensures comparable labour market attachments across both countries. Second, women in our harmonised sample have to be at least 18 years old at the time of their first childbirth. This age threshold maintains consistency and comparability in the analysis (see scripts `FR_edp_data.do` and `DE_10b_annual_data.do`).

Finally, we do not restrict the sample based on work hours, but researchers may refine their analytical samples as needed. For example, they may focus only on full-time employees at the time of the first birth. To this end, we provide variables indicating work intensity to researchers in the harmonised data: `share_fulltime` (DADS-EDP) and `fulltime_pct` (SIAB) (see `DE_10b_annual_data.do` for SIAB and `FR_add_share_employment.do` for DADS).

### 3 Harmonising the data structure

In addition to differences in sample coverage, the SIAB and DADS-EDP data sets do not structure employment episodes in the same way. This section outlines these differences and describes our harmonisation approach.

SIAB is an episode-based dataset in which episodes of (un-)employment are recorded with daily precision, including exact start and end dates. Continuous employment across multiple years is split into separate annual episodes, each running from January 1 to December 31 (Schmucker et al. 2023). Periods in which individuals are neither employed nor registered for benefits or unemployment appear as gaps in the episode data. This includes periods of maternity protection and maternity leave. For mothers, a typical trajectory begins with employers filing a 'deregistration due to health benefits' at the onset of maternity leave, six weeks before the expected delivery date. This indicates the start of maternity protection, during which wages are not paid and partially replaced by health insurance payments. For women registered as unemployed, a similar notification is filed by the employment agency. Based on these deregistrations, we calculate the expected delivery date as an approximation of the birth date of the child (see `DE_05_Mother_identification.do` adapted from Müller et al. 2022). SIAB thus captures transitions into motherhood for women who are employees or registered as unemployed when entering maternity leave. Births that occur in phases when women are inactive are not traceable. After maternity leave, women reappear once they return to active employment and their employer files a new registration. A mother's re-entry into the labour market is also visible when she claims unemployment or social security benefits. Thus, the SIAB enables detailed tracking of employment interruptions related to childbirth and provides insight into the timing of mothers' re-entries into the workforce (see e.g., Frodermann et al. 2013, 2022, 2023).

The DADS data, by contrast, structures episodes around the calendar year for each combination of individual, firm, and year. Note that DADS records are aggregated at the person-firm (SIREN) level, whereas SIAB episodes are structured at the person-establishment level. In the German social security reporting system, establishments belonging to the same firm are assigned separate establishment identifiers if they are located in different municipalities, while those within the same municipality and of the same economic activity are typically reported under a single identifier (Statistik der Bundesagentur für Arbeit 2023). As a result, a woman working across multiple establishments of the same firm in different municipalities may appear as a single record in DADS but as multiple records in SIAB. If a person has more than one contract with the same firm in a given year, those employment episodes are grouped into

a single record. This record shows when a job at a given firm started and ended, along with how many hours the person was paid for during that time. The start and end dates in these records are not suitable for tracking the length of employment interruptions. For example, in the case of multiple contracts with the same firm in the same year, the records contain only the combined first begin and last end dates. Similarly, if one person holds more than one part-time job at the same time, the beginning and end dates of different employers can overlap. Additionally, some episodes appear to cover the entire year when they correspond to relatively few working hours. Researchers should therefore not use start and end dates in the DADS-EDP to derive measures of employment duration or intensity; for this purpose, the variable *durée de paie* (DP) should be used instead (see Sect. 4).

We mechanically harmonise the two data sets by aggregating the daily SIAB episodes into an annual person-year format aligned with the DADS structure. When multiple employment states (employed, unemployed, inactive) or work-intensity states (full-time, part-time) occur within a year, we classify the year according to the state in which the individual spent the most days. For example, if someone is both employed and inactive in the same year, we assign the annual state based on the greater number of days in either category (see Sect. 4). This procedure is implemented in

`DE_10b_annual_data.do` and can be adapted by researchers to reflect alternative definitions. A more detailed explanation of the aggregation rules is provided in Sect. 4.

Finally, we restrict the DADS sample to women who were employed in the year before birth and held an active contract, ensuring comparability with the SIAB-based identification of mothers. Given that unemployment registrations in DADS are only available from 2008 onward, we exclude women registered as unemployed in the year before birth from the harmonised sample. The same restriction is applied to the SIAB sample, where we only keep women who were actively employed in the year before birth (the corresponding `do`-files are `FR_cleaning_final.do` for DADS-EDP and `DE_10b_annual_data.do` for SIAB).

#### 4 Employment status and work volume

Establishing an annual structure for the harmonised data set involves establishing aggregation rules for employment status and work volume. Years in which individuals are either continuously employed or entirely inactive/unemployed are straightforward to classify. However, many person-years involve both employment and non-employment episodes. Since unemployment registrations in the DADS are incomplete before 2008, we do not

distinguish between unemployment and inactivity in the harmonised dataset (in `FR_cleaning_final.do`).

We define a woman as actively employed in a given year if she spent the majority of the calendar year in paid employment, operationalised as days with wage payments exceeding 50% of the days in that year.

In SIAB, this corresponds to employment episodes with a wage > 0 covering at least 50% of days. The denominator accounts for the actual length of the calendar year, including leap years (365 or 366 days; `DE_10b_annual_data.do`).

In DADS-EDP, we use the variable *durée de paie* (DP), which records the number of days in paid employment across all firms a woman worked for in a given year, to apply the same 50% threshold (`FR_add_share_employment.do`). Note that DP is top-coded at 360 days per person-year. The 50% employment threshold is therefore applied against a denominator of 360 days, ensuring consistency between the maximum possible value of the numerator and the denominator (`FR_add_share_employment.do`).

The two data sets similarly capture work volume by using information provided by the employer on whether the contract is part-time or full-time. In DADS-EDP, we define full-time employment based on that specific variable (variable CE for “Condition d’Emploi”, see `FR_add_share_employment.do`) that reports whether someone is employed part-time or full-time. In SIAB, work volume is indicated by the variable *teilzeit*, which flags employment below the “usual working hours in the establishment” (Graf et al. 2023). Based on this information, we categorise calendar years into full-time or part-time. To improve accuracy, we apply two correction procedures to the originally reported part-time variable. First, we adjust for the structural break in the part-time indicator in 2011, following Fitzenberger and Seidlitz (2020 - see `DE_06a_parttime_Fitzenberger_Seidlitz.do`). Second, we check post-birth employment for potential overreporting of full-time employment for mothers who return to their previous employer after their child-related employment interruption. Specifically, we assume that employment episodes that yield less than 90% of pre-birth inflation-adjusted daily earnings and are notified as full-time are, in fact part-time employments that employers failed to update (Frodermann et al. 2013 - see `DE_06b_return_to_employment.do`).

Based on these contract- and episode-based part-time variables, we apply a majority-rule approach to distinguish full-time and part-time employment. Specifically, we classify a year based on the state that dominates at least 50% of the time spent in employment. This implementation can be found in `DE_10b_annual_data.do` for the SIAB, and `FR_cleaning_final.do` for the DADS.

**Table 1** Recoding scheme of SIAB education categories to ISCED 2011

Ausbildung_imp	Professional training (imputed)	ISCED code	ISCED category name	Final education variable
1	Lower secondary school qual. w/o vocational qual	24	Lower secondary general education	Basic education
2	Lower secondary school qual. with vocational qual	35	Upper secondary vocational education	Secondary education
3	Entrance qual. for FH or University w/o vocational qual	34	Upper secondary general education	Secondary education
4	Entrance qual. for FH or University with vocational qual	34 + 35	Upper secondary general and vocational education	Secondary education
5	University of Applied Sciences (FH)	64 or 74	Bachelor's/Master's or equivalent level, academic	Tertiary education
6	University	64 or 74	Bachelor's/Master's or equivalent level, academic	Tertiary education

**Table 2** Recoding of DADS education categories to ISCED 2011

Dip_tot	Entry	Original entry	Final education variable
1	No diploma	1 - Aucun diplôme déclaré (aucun diplôme ou pas présent au recensement)	Basic education
2	Primary School	2 - CEP, DFEO	Basic education
3	Junior High School	3 - BEPC, BE, BEPS	Basic education
4	Vocational Training	4 - CAP, BEP, EFAA, BAA, BPA, FPA 1er	Secondary education
5	High School	5 - Baccalauréats technique et professionnel, brevet professionnel, autres brevets (BEA, BEC, BEH, BEI, BES, BATA), baccalauréat général, brevet supérieur, CFES	Secondary education
6	College	6 - BTS, DUT, DEST, DEUL, DEUS, DEUG, diplôme des professions sociales ou de la santé	Tertiary education
7	Master and above	7 - Diplôme universitaire de 2ème ou 3ème cycle, diplôme d'ingénieur, grande école	Tertiary education

## 5 Variable overview

Finally, we harmonise key socio-economic variables that are present in both data sources but differ in their original coding. Our approach is guided by international statistical framework standards commonly used to structure socio-economic data. In the following subsections, we describe the harmonisation of education (ISCED 2011), occupations (ISCO-88 and ISCO-08), employment sectors (NACE Rev. 2), and the construction of earnings quintiles to classify mothers' pre-birth earnings. Beyond these individual-level socio-economic variables that require explicit harmonisation, both DADS-EDP and SIAB provide firm-level information that can be used for comparative analyses. Table A1 lists these firm-level variables side by side for both data sets alongside the respective do files.

### 5.1 Education (ISCED 2011)

In SIAB, education is reported via employer notifications, which can be enhanced using the imputation routine developed by Fitzenberger et al. (2006) and later adapted versions by Thomsen et al. (2018) for more recent versions of the SIAB. We recode education into the first-digit codes of ISCED 2011, following the IAB-FDZ's official recoding scheme (DE\_09c\_ISCED.do; FDZ-IAB, 2023). Table 1 presents the mapping.

In DADS-EDP, education is recorded for 83% of women. The raw classification established by EDP is displayed in Table 2, and its recoding into ISCED 2011 is implemented in DoFile FR\_cleaning\_final.do.

In both data sets, we distinguish four categories: (1) missing, (2) basic, (3) secondary, and (4) tertiary education. We retain missing education as an additional category to avoid excluding these cases from our analytical sample.

### 5.2 Occupations (ISCO-88 & ISCO-08)

In the SIAB, occupations are coded according to two classification systems: the *Klassifikation der Berufe 1988* (KldB 1988) and the *Klassifikation der Berufe 2010* (KldB 2010). The KldB 1988 scheme applies to employment episodes ending before November 30, 2011, while the KldB 2010 scheme has been in use since that date (Matthes & Paulus, 2013, Schmucker et al. 2023). To enhance longitudinal consistency, the IAB-FDZ has converted the KldB 2010 into the old KldB 1988 classifications beyond 2011 and, conversely, backdated the KldB 2010 classifications to years prior to the classification change (Schmucker et al. 2023). We recode the KldB 1988 classifications into ISCO-88 categories and the KldB 2010 classifications, in conjunction with the niveau variable, into ISCO-08 categories (FDZ IAB 2023). Both ISCO codes are available for all years. Nevertheless, users should be aware of

**Table 3** ISCO-88 & ISCO-08 classifications in final data

ISCO-88	ISCO-08
01 Armed forces	01 Commissioned armed forces officers
11 Legislators and senior officials	02 Non-commissioned armed forces officers
12 Corporate managers	03 Armed forces occupations, other ranks
13 Managers of small enterprises	11 Chief executives, senior officials and legislators
21 Physical, mathematical and engineering science professionals	12 Administrative and commercial managers
22 Life science and health professionals	13 Production and specialized services managers
23 Teaching professionals	14 Hospitality, retail and other services managers
24 Other professionals	21 Science and engineering professionals
31 Physical and engineering science associate professionals	22 Health professionals
32 Life science and health associate professionals	23 Teaching professionals
33 Teaching associate professionals	24 Business and administration professionals
34 Other associate professionals	25 Information and communications technology professionals
41 Office clerks	26 Legal, social and cultural professionals
42 Customer services clerks	31 Science and engineering associate professionals
51 Personal and protective services workers	32 Health associate professionals
52 Models, salespersons and demonstrators	33 Business and administration associate professionals
61 Skilled agricultural and fishery workers	34 Legal, social and cultural associate professionals
71 Extraction and building trades workers	35 Information and communications technicians
72 Metal, machinery and related trades workers	41 General and keyboard clerks
73 Precision, handicraft, craft printing and related trades workers	42 Customer services clerks
74 Other craft and related trades workers	43 Numerical and material recording clerks
81 Stationary plant and related operators	44 Other clerical support workers
82 Machine operators and assemblers	51 Personal services workers
83 Drivers and mobile plant operators	52 Sales workers
91 Sales and services elementary occupations	53 Personal care workers
92 Agricultural, fishery and related labourers	54 Protective services workers
93 Labourers in mining, construction, manufacturing and transport	61 Market-oriented skilled agricultural workers
	62 Market-oriented skilled forestry, fishery and hunting workers
	63 Subsistence farmers, fishers, hunters and gatherers
	71 Building and related trades workers (excluding electricians)
	72 Metal, machinery and related trades workers
	73 Handicraft and printing workers
	74 Electrical and electronic trades workers
	75 Food processing, woodworking, Garment and other craft and related trades workers
	81 Stationary plant and machine operators
	82 Assemblers
	83 Drivers and mobile plant operators
	91 Cleaners and helpers
	92 Agricultural, forestry and fishery labourers
	93 Labourers in mining, construction, manufacturing and transport
	94 Food preparation assistants
	95 Street and related sales and services workers
	96 Refuse workers and other elementary workers

the discontinuities inherent in the German occupational coding framework: since the KldB 2010 provides greater detail than the KldB 1988, conversions between both schemes inevitably involve a degree of uncertainty. In contrast, ISCO-08 is finer-grained than ISCO-88, introducing additional imprecision when aligning data across the two systems.

In DADS, occupations are classified using INSEE's *Catégories Socio-Professionnelles (CSP)*. Using the R package *SocialPosition* (Falcon 2015), we first convert CSP into ISCO-88 categories and then apply a transition matrix from ISCO-88 to ISCO-08 to harmonise classifications. Researchers using the CSP information should be aware of the 2003 update to the classification scheme,

**Table 4** Sector categories based on NACE rev.2 as available in the harmonised data set

Code	Economic area	Final coding
A	Agriculture, forestry and fishing	<i>Other (A, B, R, F, T &amp; U)</i>
B	Mining and quarrying	<i>Other (A, B, R, F, T &amp; U)</i>
C	Manufacturing	Manufacturing (C)
D	Electricity, gas, steam and air conditioning supply	Electricity & Water Supply (D&E)
E	Water supply; sewerage, waste management and remediation activities	Electricity & Water Supply (D&E)
F	Construction	<i>Other (A, B, R, F, T &amp; U)</i>
G	Wholesale and retail trade; Repair of motor vehicles and motorcycles	Wholesales (G)
H	Transportation and storage	Transportation (H)
I	Accommodation and food service activities	Accommodation and Food (I)
J	Information and communication	Information (J)
K	Financial and insurance activities	Financial Activities & Real Estate (K&L)
L	Real estate activities	Financial Activities & Real Estate (K&L)
M	Professional, scientific and technical activities	Professional Scientific (M)
N	Administrative and support service activities	Administrative Activities (N)
O	Public administration and defence; Compulsory social security	Public Administration (O)
P	Education	Education (P)
Q	Human health and social work activities	Human Health (Q)
R	Arts, entertainment and recreation	<i>Other (A, B, R, F, T &amp; U)</i>
S	Other service activities	Other Service (S)
T	Activities of households as employers; Undifferentiated goods and services producing activities of households for own use	<i>Other (A, B, R, F, T &amp; U)</i>
U	Activities of extraterritorial organisations and bodies	<i>Other (A, B, R, F, T &amp; U)</i>

which might affect their findings when analysing changes over time (INSEE, 2024).

For analytical consistency, both ISCO code variants are made available at the two-digit level, i.e., the sub-major groups in both data sets. However, researchers should be aware that analyses at the two-digit level may yield small case numbers in some sub-groups; combining categories may be necessary depending on the sample size and level of detail required. More generally, the two ISCO versions differ in their fit across the observation period: ISCO-88 tends to be more reliable for earlier years, while ISCO-08 is better suited for more recent periods. In the SIAB, this reflects the imputation of KldB 2010 classifications prior to November 2011 (see above); in DADS-EDP, the accuracy of ISCO codes may also vary across the observation period. For detailed coding of the ISCO classifications, see DE\_09b\_ISCO.do (SIAB) and FR\_add\_isco.do (DADS). Based on the two-digit ISCO-88 and ISCO-08 codes (27 sub-major groups for ISCO-88 and 43 for

**Table 5** Number of observed women by year of first childbirth and country

Year of first childbirth	France	West Germany	East Germany
1997–2001	3,230	17,681	3,168
2002–2006	6,469	15,338	3,325
2007–2011	8,507	16,149	3,946
2012–2014	4,528	11,453	2,904
Total	22,734	60,621	13,343

ISCO-08), we create the following occupation groups (Table 3):

### 5.3 Sector (NACE rev.2)

In SIAB, sector information is available through linkage to the Establishment History Panel (BHP; Schmidtlein et al. 2020). Establishments are coded according to the 3-digit Classification of Economic Activities 2008 scheme, which can be recoded into NACE Rev. 2 codes using tables provided by FDZ-IAB (destatis 2025; FDZ-IAB, 2023).

In DADS, firm-level sector codes are based on the NAF classification, which can be directly converted into NACE codes. We use the section-level (single-letter) codes of the NACE Rev. 2 classification (Eurostat, 2008). To avoid small case numbers, we group some sectors into broader categories. The recoding is implemented in DE\_09a\_NACE\_sectors.do for SIAB and FR\_add\_nace.do for DADS. Table 4 presents the harmonised sector categories.

### 5.4 Earnings

Both data sets capture earnings from employment subject to social security contributions, but differ in an important respect regarding top-coding.

In SIAB, earnings are reported in the variable *tentgeld*, which captures either gross daily earnings (for employment episodes) or daily benefits (for benefit receipt episodes). However, since earnings above the social security contribution ceiling are not relevant for social security contributions and benefit claims, earnings information in the underlying data is capped at that ceiling. This affects approximately 5.2% of all employment episodes for workers between 1975 and 2021, of which only 18% are from women – i.e., only about 0.9% of women's employment episodes are top-coded at the contributions ceiling. Nevertheless, we address this issue by applying the established imputation procedure for earnings information of Stüber et al. (2023; see also Card et al. 2013; Dustmann et al. 2009; Gartner, 2005) in DoFile DE\_04b\_wage\_imputation.do. In newer SIAB versions (7523 or later), researchers can alternatively use the newly provided variable *tentgeld\_imp*, which provides imputed earnings information. Note, however, that *tentgeld\_imp* is not deflated, whereas we applied deflation when constructing

annual\_earnings, using 2015 as the baseline. The annual\_earnings variable created in DE\_10a\_annual\_wage\_prep\_data.do provides the cumulated annual earnings from employment(s) subject to social security.

In DADS-EDP, earnings are reported via the variable sbr, which corresponds to the gross (real) earnings received by workers and is not subject to top-coding. As a result, earnings in the DADS-EDP are fully observed across the entire distribution, in contrast to the SIAB.

This asymmetry has implications for comparative earnings analyses. While the top-coding correction applied to the SIAB substantially reduces this gap, the imputation cannot prevent uncertainty at the top of the earnings distribution. Researchers should therefore exercise caution when using the harmonised data to study earnings or wage inequality among workers in the upper earnings distribution, as remaining differences in measurement precision between the two data sets may affect comparability in this range.

Linked to our example of calculating motherhood penalties, we also provide code to calculate earnings quintiles based on pre-birth income. To this end, we define a running variable  $t$  capturing years before and after birth. We then calculate the earnings quintiles for each motherhood cohort in  $t = -1$  (see DE\_10b\_annual\_data.do for SIAB and FR\_cleaning\_final.do for DADS). Annual earnings might appear lower than a woman's average labour income for those giving birth early in the following year ( $t = 0$ ), as their earning in  $t = -1$  are affected by the onset of maternity leave. To compensate, we adjust the earnings of women giving birth within the first 6 weeks of  $t = 0$  by reweighting the earnings in  $t = -1$  by  $d_m$ , capturing the days spent on maternity leave in  $t = -1$ :

$$earnings_{t-1} \times \frac{365}{365 - d_m} \text{ if } d_m > 0$$

The variable  $d_m$  is calculated as follows, given that pre-birth maternity leave is 6 weeks in both countries:

$$d_m = \text{date of birth} - 6 \times 7 \text{ days} - 01.\text{January}_{t=0}$$

This adjustment ensures women are not misclassified as low earners due to earnings reductions attributable to maternity leave. Researchers should note that the quintile calculations are sensitive to the final sample restriction criteria. Those applying different restrictions should re-use our code to define new quintiles.

## 6 Example: mothers' employment trajectories over time

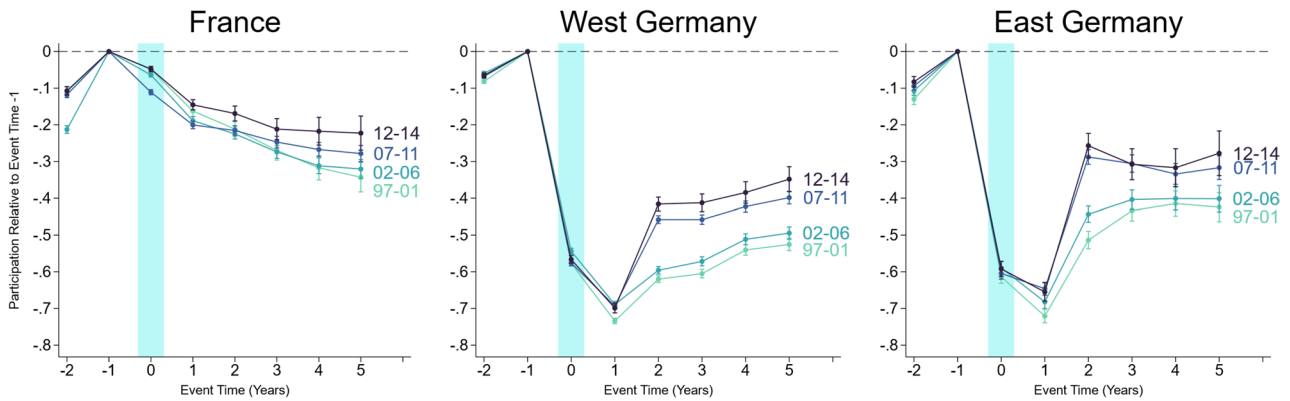
To demonstrate the potential of the harmonised DADS and SIAB data sets, we analyse the evolution of mothers' employment trajectories following their first childbirth

in France and Germany. Leveraging the large sample size and long observation period of both data sets, we examine mothers' labour market participation over the five years after childbirth, covering 18 years of entries into motherhood (1997–2014). In addition to the French-German cross-national comparison, we distinguish between West and East Germany. This separation reflects well-established approaches in the literature, which highlight persistent disparities between the two former German states in their labour market behaviour, childcare availability, gender norms, and attitudes toward maternal employment (Lietzmann and Frodermann 2021; Rosenfeld et al. 2004; Zoch 2021; Zoch and Hondralis 2017). Note that Berlin is coded as part of East Germany in our data preparation of SIAB; however, since our analysis starts in 1997 – after complete East German coverage is established – this coding does not affect our results.

To capture changes in maternal employment over time, we divide the samples from the three regional entities into four maternity cohorts based on the year of first childbirth: 1997–2001, 2002–2006, 2007–2011, and 2012–2014. This design accounts for the significant impact of the 2007 parental leave reform in Germany, which fundamentally altered mothers' employment patterns. Before the reform, compensation for parental leave was available for up to two years, but only for mothers in households below the median income. The 2007 reform introduced a universal compensation scheme, providing parents with 65% of their previous net income for up to fourteen months, with a maximum of twelve months per parent and two months reserved for the 'other parent' (the *daddy quota*). Although later reforms further expanded flexibility (e.g., the option to convert full-time leave benefits into part-time work months introduced in 2015), the 2007 reform is widely recognised as a turning point, leading to longer leave durations, especially among higher-income families who had previously lacked access to parental leave benefits (Frodermann et al. 2022; Kluge and Tamm 2013). By structuring our cohorts around this reform, our illustrative analysis can capture how maternal employment trajectories shifted around a major policy change.

### 6.1 Analytical sample

Our example analysis uses women from both data sets who gave birth between 1997 and 2014 and were actively employed for the majority of the year before birth ( $t = -1$ ). This gives us a total sample of 22,734 women in France, 60,621 in West Germany, and 13,343 in East Germany. Table 5 provides an overview of sample sizes broken down by the four maternity cohorts: 1997–2001, 2002–2006, 2007–2011, and 2012–2014. The increase in French cases after 2002 reflects the expansion of the DADS-EDP sample (see Sect. 2).



**Fig. 1** Labour market participation relative to the year before first childbirth

### 6.2 Modelling approach

We estimate the effects of entry into motherhood (via first birth) on labour market participation for each motherhood cohort in France as well as East and West Germany. Specifically, we trace women employed in the year before they had their first child for five years after birth. Our approach measures the motherhood penalty in employment among women who have been previously employed rather than among all mothers. Moreover, we estimate the effects of the onset of motherhood (which might include subsequent births) rather than the effect of a given child. The event study approach we follow is similar to Kleven et al. (2019). In sum, our estimates capture year-by-year changes in labour market participation following the transition to motherhood. As discussed, our time variable  $t \in \{-2, 5\}$  reflects the years since the first birth in  $t=0$ . We run separate regressions for each cohort and region (France, West Germany, and East Germany):

$$E_{iyt}^{cohort} = \sum_{j \neq -1} \alpha_j^{cohort} I_{j=t} + \sum_k \beta_k^{cohort} I_{k=age} + \sum_y \gamma_k^{cohort} I_{year=y} + \epsilon_{iyt}^{c,s} \tag{1}$$

Here,  $E_{iyt}^{cohort}$  is a binary indicator equal to 1 if individual  $i$  (who is part of the *cohort* that is being estimated) was actively employed for the majority of year  $y$  at event time  $t$  (i.e., spent more than 50% of days in paid employment), and 0 otherwise. The right-hand side of the equation includes dummies capturing the effect of running time  $t \in \{-2, \dots, 5\}$  via the coefficients  $\alpha_j^{cohort}$ , age via  $\beta_k^{cohort}$  and year via  $\gamma_y^{c,n}$ . The  $\alpha$  coefficients capture the percent change in labour market participation mothers experience relative to the years before the first birth ( $t=-1$ ) and hence captures the motherhood penalty. Following the approach of Kleven et al. (2019), we obtain our final estimates by calculating the average difference between the employment outcome predicted by all regressors and the predicted employment when omitting the contribution of the  $\alpha$  coefficients capturing the effects of motherhood.

### 6.3 Results

Figure 1 displays employment trajectories from two years before to five years after the first childbirth by cohort and region (France, West Germany, and East Germany). The lines trace predicted percent changes in employment participation relative to the year before childbirth (event time = -1) as a consequence of childbirth.

Across all three contexts, employment declines sharply in the first year after childbirth ( $t=1$ ), though the magnitude of this drop varies substantially. The largest drop occurs in West Germany (-73.6 to -67.9%), followed by East Germany (-70.3 to -61.0%). In France, the decline is far smaller (-20.0 to -16.2%).

Similarly, the effect of childbirth on employment trajectories continues to differ by country and by cohorts in subsequent years. In France, employment continues to decline gradually through year five.<sup>1</sup> The earliest cohort (1997–2001) reaches -34.3% by year five, while the most recent cohort (2012–2014) reaches -22.3%. Among mothers from both East and West Germany, the sharp initial decline is followed by a gradual recovery, which, however, remains below the levels of French mothers. In West Germany, maternal employment trajectories are characterised by a sharp initial decline and a partial recovery, which becomes larger for later cohorts. In the earliest cohort (1997–2001), employment remains 52.5% below pre-birth levels by year five. However, the 2002–2006 cohort reaches -49.4%, the 2007–2011 cohort -39.8%, and the 2012–2014 cohort -34.8% by year five, indicating a clear upward trend. In East Germany, the year of birth and one year after birth declines are of similar magnitude to West Germany, but employment then rebounds more quickly and to higher levels. The 1997–2001 cohort records a decline of -72.1% in year one and -42.4% in year five. Subsequent cohorts rebound

<sup>1</sup> Additional analyses revealed that the gradual decline can be attributed to second births: focussing on mothers with a single birth within the five years, yields a gradual increase in these mothers' employment participation in the later years after the first and only childbirth.

to successively higher levels five years after birth:  $-40.1\%$  (2002–2006),  $-31.7\%$  (2007–2011), and  $-27.8\%$  (2012–2014). Thus, the overall V-shaped trajectory among East German mothers is similar to that of their West German peers; however, cohort differences are somewhat less pronounced.

Overall, all three regions are marked by shrinking employment penalties associated with the onset of motherhood. Cohort differences are more modest in France but substantial in both parts of Germany. From the earliest to the latest cohorts, declines in year five diminish by 18% points in West Germany, 15% points in East Germany, and 12% points in France. The gains in employment participation among German mothers emerge primarily two years after birth, which is in line with the important effects that the 2007 parental leave benefits reform had in changing German mothers' employment trajectories: generous first-year benefits combined with shorter maximum durations created increasingly V-shaped employment trajectories. The importance of the reform is further supported by the jump in post-birth employment between the 2002–2006 and the 2007–2011 cohorts. However, the reform only mildly reduced inner-German differences in maternal employment participation. Return to employment remains substantially higher in East Germany, as it was before 2007, pointing to the long-lasting persistence of normative and institutional differences. These patterns illustrate how the harmonised data can capture both policy-driven and longer-term secular trends in maternal employment across countries.

This example analysis illustrates the potential of harmonised SIAB and DADS-EDP data sets. By combining large-scale, long-running administrative records from two countries, these data sets enable fine-grained, functionally comparable analyses of maternal employment trajectories that would be difficult to achieve with survey data, given typical limitations in sample size, measurement consistency, and longitudinal depth. For future research, the harmonised data sets provide a foundation for comparative research on family policies, labour market inequalities, and broader life-course dynamics (see e.g., Filser et al. 2026; Wagner et al. 2025).

## 7 Remaining harmonisation limitations

The harmonisation steps described in this paper bring the DADS-EDP and SIAB data sets to a high degree of functional comparability for the study of maternal employment. Nevertheless, several limitations remain that researchers should be aware of when interpreting results. These limitations reflect the distinction between mechanical and conceptual harmonisation: while the

data sets have been structurally aligned, functional equivalence cannot always be fully achieved, and the following points document where residual non-equivalences remain.

The two data sets differ in their coverage of public sector workers. SIAB excludes civil servants (Beamte,  $\sim 3.8\%$  of the German workforce), and we exclude the closest French equivalent (*Fonction Publique d'État*) for comparability (see Sect. 2). Furthermore, since unemployment registrations are only available in DADS-EDP from 2008 onward, the harmonised data does not distinguish between unemployment and inactivity in either country (see Sects. 3 and 4).

Additionally, both samples are restricted to women with active employment before childbirth. Births occurring during phases of inactivity are not captured in SIAB, and the same restriction is applied to DADS-EDP to ensure comparability (see Sect. 2). The harmonised data therefore represent mothers with an established employment attachment rather than all mothers.

Another asymmetry concerns earnings measurement. SIAB earnings are top-coded at the social security contribution ceiling ( $\sim 0.9\%$  of women's episodes), which we address through imputation (Stüber et al. 2023), whereas DADS-EDP earnings (sbr) are fully observed across the distribution (see Sect. 5.4). Researchers should therefore exercise caution when comparing earnings at the top of the distribution, as residual differences in measurement precision may affect comparability in this range.

Finally, the two data sets also differ in sample size (see Table 5), which has implications for statistical power in comparative analyses. In research settings where true effect sizes are small, estimates based on the French data may lack the precision to detect effects that are identifiable with the larger German sample. Researchers should account for this asymmetry when drawing cross-national comparisons, particularly for subgroup analyses.

## 8 Data availability

Both SIAB and DADS-EDP are available to the research community following a formal data application process. Researchers interested in analysing SIAB and DADS-EDP data in a harmonised way are required to apply for and obtain access to each data set separately. Although the data sets are hosted on distinct secure computing environments, researchers can access both environments through the same set of access points provided by the IDAN cooperation network at CASD in Malakoff (France), the FDZ-IAB in Nuremberg (Germany), GESIS in Cologne (Germany), and the Institute for Social

Research at the University of Michigan in Ann Arbor (USA).<sup>2</sup>

SIAB data are available to the international research community free of charge and are provided by the Research Data Centre (FDZ) at the IAB. Three access modes are offered according to the degree of anonymization. A highly anonymised version offering aggregated variable coding and a more limited scope of variables is available as Scientific Use File (SUF, referred to as *regional file* (Frodermann et al. 2021) and can be analysed within the institutional computing environment of the researcher. The detailed SIAB version is available only via remote execution via JoSuA or at one of the secure on-site data access points (see <https://fdz.iab.de/> for a full list; Eberle et al. 2017). In either case, users are required to sign a Data Use Agreement with the FDZ and must comply with further requirements according to the access mode. Further information is available on the homepage of the FDZ, which also provides related documentation, e.g., a detailed description of the dataset and frequency tables.<sup>3</sup>

Access to the DADS–EDP data is managed through the CASD (Centre d'Accès Sécurisé aux Données) under strict legal and technical safeguards. The first step is to submit an application through the Confidential Data Access Portal (CDAP). This requires creating an account, signing a confidentiality agreement, and providing details about your project, the research team, and the specific data sets you intend to use. Once the application is submitted, it must be reviewed by the *Comité du secret statistique*. This committee evaluates whether the project complies with the requirements of statistical confidentiality. Only after the Comité gives a favourable opinion can the project proceed. Following approval, researchers must finalise contracts with their institution and register with CASD. Registration involves attending a formal enrolment session, after which researchers receive the secure authentication equipment needed to connect to CASD servers. All work with the data takes place remotely on CASD's secure infrastructure. Access to the data is subject to fees. While the *Comité du secret statistique* does not charge for its review, there are fees associated with CASD itself. Descriptions of the DADS–EDP dataset, including details on variables and documentation, can be found directly on the CASD

website, where a dedicated page is available for the Panel DADS–EDP.<sup>4</sup>

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12651-026-00430-8>.

Supplementary Material 1.

### Acknowledgements

This paper has benefitted from the support and input of several individuals whom we would like to thank. At the IAB, we thank Dana Müller for leading the project that made this collaboration possible and for her continuous support throughout, including valuable guidance on data-related questions. We are grateful to Ann-Christin Bächmann for sharing code templates and for her input on the occupations data in SIAB. We also thank Kevin Ruf for providing the translation matrix from KldB 1988 to ISCO-88, which substantially facilitated the harmonisation of occupational classifications. We further thank Pierre Pora for sharing code and advice relating to the French DADS–EDP data. Finally, we thank the two anonymous reviewers for their exceptionally thorough and constructive reviews. Their comments helped us clarify the employment measurement in the DADS–EDP data, which has substantially improved the paper.

### Author contributions

All authors read and approved the final manuscript.

### Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported in this paper was supported by the grant “ANR-DFG French-German Collaboration for Joint Projects in Social Sciences” from the German Research Foundation (DFG grant FR3840/4 – 1) and the French Research Foundation (ANR-20-FRAL-0007) as well as by the Leverhulme Trust grant RC-2018-003.

### Declarations

#### Competing interests

The authors declare that they have no competing interests.

Received: 12 November 2025 / Accepted: 8 May 2026

Published online: 11 June 2026

### References

- Bassanini, A., Bovini, G., Caroli, E., Ferrando, J.C., Cingano, F., Falco, P., Felgueroso, F., Jansen, M., Martins, P.S., Melo, A., Oberfichtner, M., Popp, M.: Labor Market Concentration, Wages and Job Security in Europe. *J. Hum. Resour.* 0223–12757R1 (2024). <https://doi.org/10.3368/jhr.0223-12757R1>
- Bertheau, A., Acabbi, E.M., Barceló, C., Gulyas, A., Lombardi, S., Saggio, R.: The Unequal Consequences of Job Loss across Countries. *Am. Economic Review: Insights.* 5(3), 393–408 (2023). <https://doi.org/10.1257/aeri.20220006>
- Budig, M.J., Misra, J., Boeckmann, I.: Work–Family Policy Trade-Offs for Mothers? Unpacking the Cross-National Variation in Motherhood Earnings Penalties. *Work Occup.* 43(2), 119–177 (2016). <https://doi.org/10.1177/0730888415615385>
- Card, D., Heining, J., Kline, P.: Workplace Heterogeneity and the Rise of West German Wage Inequality. *Q. J. Econ.* 128(3), 967–1015 (2013). <https://doi.org/10.1093/qje/qjt006>
- destatis. Staff by type of employment contract. Federal Statistical Office. (2022), August 4 <https://www.destatis.de/EN/Themes/Government/Public-Service/Tables/public-service-personnel-type-employment-contract.html>

<sup>2</sup> Note however, that for all FDZ-IAB data including SIAB, additional data restrictions apply for data access from the United States. See [https://fdz.iab.de/en/appointment-locations-and-fdz-online-calendar/#Tutorial\\_Anonymization\\_process\\_for\\_on-site\\_use\\_for\\_locations\\_outside\\_the\\_European\\_Union](https://fdz.iab.de/en/appointment-locations-and-fdz-online-calendar/#Tutorial_Anonymization_process_for_on-site_use_for_locations_outside_the_European_Union).

<sup>3</sup> See <https://fdz.iab.de/en/our-data-products/individual-and-household-data/siab/>.

<sup>4</sup> See <https://www.casd.eu/en/source/dads-panel-with-matched-data-from-edp/>.

- destatis: Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). Statistisches Bundesamt. (2025). <https://www.destatis.de/DE/Methoden/Klassifikationen/Gueter-Wirtschaftsklassifikationen/Klassifikation-wz-2008.html>
- Dustmann, C., Ludsteck, J., Schönberg, U.: Revisiting the German Wage Structure. *Q. J. Econ.* **124**(2), 843–881 (2009). <https://doi.org/10.1162/qjec.2009.124.2.843>
- Eberle, J., Müller, D., Heining, J.: A Modern Job Submission Application to Access IAB's Confidential Administrative and Survey Research Data (Nos. 2017–01; FDZ Method Reports). Research Data Centre of IAB (FDZ-IAB). (2017). [https://doku.iab.de/fdz/reporte/2017/MR\\_01-17\\_EN.pdf](https://doku.iab.de/fdz/reporte/2017/MR_01-17_EN.pdf)
- Esping-Andersen, G.: The three worlds of welfare capitalism. Princeton University Press (1990)
- Eurostat: NACE - Statistical classification of economic activities in the European Community Rev. 2. European Union. (2008). <https://www.bundesbank.de/re-source/blob/619164/ccae9c8d5f91f785cb1693e60c77f3a9/mL/nace-rev-2-0-ata.pdf>
- Falcon, J.: SocialPosition: Social Position Indicators Construction Toolbox R package [Computer software]. (2015). <https://cran.r-project.org/src/contrib/Archive/SocialPosition/>
- FDZ-IAB: Overview on classifications. Key Working Tools of the FDZ. (2023). <https://fdz.iab.de/en/manuals-and-working-tools/key-working-tools/>
- Filser, A., Achard, P., Frodermann, C., Müller, D., Wagner, S.: Stratification of Post-Birth Labour Supply in a High- and Low- Maternal Employment Regime. *Res. Social Stratification Mobil.* 101133 (2026). <https://doi.org/10.1016/j.rssm.2026.101133>
- Fitzenberger, B., Seidlitz, A.: The 2011 break in the part-time indicator and the evolution of wage inequality in Germany. *J. Labour Market Res.* **54**(1), 1 (2020). <https://doi.org/10.1186/s12651-019-0265-0>
- Fitzenberger, B., Osikominu, A., Völter, R.: Imputation Rules to Improve the Education Variable in the IAB Employment Subsample. *Schmollers Jahrbuch: J. Appl. Social Sci. Stud. / Z. Für Wirtschafts- Und Sozialwissenschaften.* **126**(3), 405–436 (2006)
- Frodermann, C., Müller, D., Abraham, M.: Determinanten des Wiedereinstiegs von Müttern in den Arbeitsmarkt in Vollzeit oder Teilzeit. *KZfSS Kölner Z. für Soziologie und Sozialpsychologie.* **65**(4), 645–668 (2013). <https://doi.org/10.1007/s11577-013-0236-3>
- Frodermann, C., Ganzer, A., Schmucker, A., Vom Berge, P.: Sample of Integrated Labour Market Biographies Regional File (SIAB-R) 1975–2019. FDZ-Datenreport, 05–2021. (2021). [https://doku.iab.de/fdz/reporte/2021/DR\\_05-21\\_EN.pdf](https://doku.iab.de/fdz/reporte/2021/DR_05-21_EN.pdf)
- Frodermann, C., Wrohlich, K., Zucco, A.: Parental Leave Policy and Long-run Earnings of Mothers. *Labour Econ.* 102296 (2022). <https://doi.org/10.1016/j.labeco.2022.102296>
- Frodermann, C., Filser, A., Bächmann, A.-C.: Parental leave of married couples: Mothers tend to return to the labour market faster when their partners take parental leave [Elternteile von verheirateten Paaren: Mütter kehren meist schneller auf den Arbeitsmarkt zurück, wenn ihre Partner Elternteil nehmen]. IAB-Kurzbericht, 1(2023), 1–8. (2023). <https://doi.org/10.48720/IAB.KB.2301>
- Ganzer, A., Schmucker, A., Stegmaier, J., Wolter, S.: Establishment History Panel (BHP) – Version 7521 v1 [Dataset]. Forschungszentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). (2022). <https://doi.org/10.5164/IAB.BHP7521.DE.EN.V1>
- Gartner, H., IAB: (with) The imputation of wages above the contribution limit with the German IAB employment sample. FDZ Methodenreport 02/2005. (2005)
- Graf, T., Griebemer, S., Köhler, M., Lehnert, C., Moczał, A., Oertel, M., Schmucker, A., Schneider, A., Seth, S., Thomsen, U., Berge, V.: P. Weakly anonymous Version of the Sample of Integrated Labour Market Biographies (SIAB) – Version 7521 v1 [Dataset]. Forschungszentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). (2023). <https://doi.org/10.5164/IAB.SIAB7521.DE.EN.V1>
- Hook, J.L., Li, M.: National Work-Family Policies and Gender Earnings Inequality in 26 OECD Countries, 1999 to 2019. *Socius.* (Sage: Los Angeles, CA). (2025). <https://doi.org/10.1177/23780231251360042>
- INSEE: La PCS 2003. (2024). <https://www.insee.fr/fr/information/2497952>
- INSEE: Annual declaration of social data. (2025a). <https://www.insee.fr/en/metadonnees/source/serie/s1163>
- INSEE: Permanent Demographic Sample. (2025b). <https://www.insee.fr/en/metadonnees/source/serie/s1166>
- Kleven, H., Landais, C., Søgaard, J.E.: Children and Gender Inequality: Evidence from Denmark. *Am. Economic Journal: Appl. Econ.* **11**(4), 181–209 (2019). <https://doi.org/10.1257/app.20180010>
- Kleven, H., Landais, C., Leite-Mariante, G.: The Child Penalty Atlas. *Rev. Econ. Stud.* **92**(5), 3174–3207 (2025). <https://doi.org/10.1093/restud/rdae104>
- Kluve, J., Tamm, M.: Parental leave regulations, mothers' labor force attachment and fathers' childcare involvement: Evidence from a natural experiment. *J. Popul. Econ.* **26**(3), 983–1005 (2013). <https://doi.org/10.1007/s00148-012-0404-1>
- Lietzmann, T., Frodermann, C.: Gender Role Attitudes and Labour Market Behaviours: Do Attitudes Contribute to Gender Differences in Employment in Germany? *Work Employ. Soc.* 09500170211011318 (2021). <https://doi.org/10.1177/09500170211011318>
- Mazur, A.G.: Drawing Comparative Lessons from France and Germany. *Rev. Policy Res.* **20**(3), 493–524 (2003). <https://doi.org/10.1111/1541-1338.00033>
- Meurs, D., Pora, P.: Traditional Views, Egalitarian Views, and the Child Penalty: Insights from Immigrant Populations in France. (2025). <https://hal.science/hal-04947430>
- Müller, D., Filser, A., Frodermann, C.: Update: Identifying mothers in administrative data. In FDZ-Methodenreport (Nos. 2022–01; FDZ-Methodenreport, pp. 1–15). Institute for Employment Research (IAB). (2022). <https://doi.org/10.5164/IAB.FDZM.2201.en.v1>
- Robert-Bobée, I., Gualbert, N.: L'échantillon démographique permanent. *Courrier Des Statistiques.* **47**, 47–63 (2021)
- Rosenfeld, R.A., Trappe, H., Gornick, J.C.: Gender and Work in Germany: Before and After Reunification. *Ann. Rev. Sociol.* **30**(1), 103–124 (2004). <https://doi.org/10.1146/annurev.soc.30.012703.110531>
- Schmidtlein, L., Wolter, S., Ganzer, A., Stegmaier, J.: Establishment History Panel (BHP) – Version 7519 v2 (Version v2) [Dataset]. Forschungszentrum der Bundesagentur für Arbeit (BA) im Institut für Arbeitsmarkt- und Berufsforschung (IAB). (2020). <https://doi.org/10.5164/IAB.BHP7519.DE.EN.V2>
- Schmucker, A., Seth, S., Vom Berge, P.: Sample of Integrated Labour Market Biographies (SIAB) 1975–2021. FDZ-Datenreport, FDZ-Datenreport. (2023). <https://doi.org/10.5164/IAB.FDZD.2302.EN.V1>
- Statistik der Bundesagentur für Arbeit: Meldeverfahren zur Sozialversicherung. (2023), December
- Stüber, H., Dauth, W., Eppelsheimer, J.: A guide to preparing the sample of integrated labour market biographies (SIAB, version 7519 v1) for scientific analysis. *J. Labour Market Res.* **57**(1), 7 (2023). <https://doi.org/10.1186/s12651-023-00335-w>
- Thomsen, U., Ludsteck, J., Schmucker, A.: Skilled or unskilled—improving the information on qualification for employee data in the IAB Employee Biography. FDZ-Methodenreport No. 09/2018. FDZ-Methodenreport, FDZ-Methodenreport. (2018). <https://doi.org/10.5164/IAB.FDZM.1809.EN.V1>
- Tomaskovic-Devey, D., Rainey, A., Avent-Holt, D., Bandelj, N., Boza, I., Cort, D., Godechot, O., Hajdu, G., Hällsten, M., Henriksen, L.F., Hermansen, A.S., Hou, F., Jung, J., Kanjuo-Mrčela, A., King, J., Kodama, N., Kristal, T., Křížková, A., Lippényi, Z., Tufail, Z.: Rising between-workplace inequalities in high-income countries. *Proceedings of the National Academy of Sciences*, 117(17), 9277–9283. (2020). <https://doi.org/10.1073/pnas.1918249117>
- van der Vleuten, M., Evertsson, M., Moberg, Y.: Joint Utility or Sub-optimal Outcomes? Household Income Development of Same-Sex and Different-Sex Couples Transitioning to Parenthood in Denmark, Finland, Norway, and Sweden. *J. Fam. Issues.* 0192513X231194305 (2023). <https://doi.org/10.1177/0192513X231194305>
- Wagner, S., Filser, A., Achard, P., Amend, I.M.: Gender Inequalities and Motherhood Penalties across French and German Local Labor Markets. *Socius.* **11**, 23780231251374110 (2025). <https://doi.org/10.1177/23780231251374110>
- Zoch, G.: Thirty Years after the Fall of the Berlin Wall—Do East and West Germans Still Differ in Their Attitudes to Female Employment and the Division of Household Work? *Eur. Sociol. Rev.* **37**(5), 731–750 (2021). <https://doi.org/10.1093/esr/jcab002>
- Zoch, G., Hondralis, I.: The Expansion of Low-Cost, State-Subsidized Childcare Availability and Mothers' Return-to-Work Behaviour in East and West Germany. *Eur. Sociol. Rev.* **33**(5), 693–707 (2017). <https://doi.org/10.1093/esr/jcx068>

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.