Essays on the Relation Between Higher Education and the Labour Market

Shihang Hou

St Cross College

University of Oxford

A thesis presented for the degree of

Doctor of Philosophy

June 2023

Acknowledgements

I am extremely grateful to my supervisors, Professors Margaret Stevens and Abigail Adams-Prassl, for their invaluable guidance over the course of this thesis. I have learned much from their guidance both in conducting research and in navigating the ups and downs of academia. I am also very thankful to all the faculty members at the university, especially Professors Barbara Petrongolo, Romuald Meango, Simon Quinn, and Howard Smith, who have been generous with their time in discussing my work over the years.

I am also deeply indebted to my friends on the DPhil for contributing to a febrile intellectual environment and for providing invaluable camaraderie, good humour and advice during this long journey. I am especially grateful to Binta Zahra Diop, Luke Heath Milsom, Hannah Zillessen, Nick Ridpath, Vatsal Khandelwal, Sam Altmann, Yiliang Li, Tatjana Schulze, Manos Perdikakis, Maria Balgova, Verena Wiedermann, Sanghamitra Mukherjee and many others who have enriched my life during the writing of this thesis. Many thanks to Diana Beltekian, Robert Jagiello, Henian Su, Jingtian Chen, and Audrey Li, who have been long-standing pillars of support.

Finally, I want to thank my family for their unwavering support, and am extremely grateful for all they have sacrificed in allowing me to pursue this DPhil.

Declaration of Authorship

I certify that this thesis, submitted for the degree of Doctor of Philosophy in Economics at the University of Oxford, is my own work, carried out between October 2018 and May 2023, except where otherwise clearly stated.

I hereby declare the existence of joint work in this thesis. The first and third essays in this thesis are solely my own work, and the third essay is based on parts of my MPhil thesis. The second essay is co-written with Mr Luke Heath Milsom. We certify that the work on both the conceptualisation and execution of the project was split evenly between the two of us.

Abstract

An important theme in the economic analysis of education is the relation between students' decisions to invest in education and the implication of these investments for their wages and productivity on the labour market. Despite being an area of very active research, there is much that remains unknown about why and how students make education decisions, whether these decisions are optimal, and what the effect of these investments are. This thesis contributes to the literature by exploring three questions directly or indirectly relating to these open issues. In the first essay, "Individual and Aggregate Mismatch in Higher Education", I study the broad public concern that there are too many students who choose to go to university. I produce a model that combines endogenous education choice with heterogeneity and uncertainty in returns, with a labour market framework based on matching that creates supply and demand-like wage responses. In this setting, uncertainty about returns generates individual education mismatch, while endogenous education choice in a matching market generates aggregate inefficiency in education choice. In the second essay, "The Butcher, the Brewer, or the Baker: The Role of Occupations in Explaining Wage Inequality", we study the importance of the occupational structure in wage determination. We introduce a model that nests Abowd et al. (1999), and estimate it on British data. We find that even if workers and firms were identical, 25% of current log-wage variance would remain, due solely to heterogeneity between occupations. Thus, occupational pay-premia are important sources of wage heterogeneity in the economy. In the third essay, "Determinants of University Subject Choice in the UK", I study the elasticity of students' demand for university courses in particular subjects with respect to their expected earnings in that subject. I find that earnings, defined as earnings at age 25 or as lifetime earnings, are positively associated with choice probability, although the quantitative impact of expected earnings on subject choice is not large. I document substantial differences in choice probability between students of different sexes, ethnicities and to a lesser extent, socio-economic classes.

Contents

Introduction

One of the most robust empirical findings in economics is that there is a positive causal relation between education and earnings (Card (1999),Heckman et al. (2018)). This finding naturally has led researchers to think about the decision to undertake education as an investment, where the student trades time and money for an increase in their future earnings potential. Given that higher education has a positive impact of earnings (Hanushek & Woessmann (2010)), many governments have also turned to subsidise higher education as a pro-growth policy. The large numbers of graduates and the cost of higher education subsidies have drawn renewed interest about the optimality of these policies and the policy instruments governments have to affect individuals' higher education investments.

In this thesis, I produce three essays which are directly or indirectly related to the discussion about the effect of higher education on the labour market, the resulting incentives to invest in higher education, and implications for higher education policy. I first consider the normative question of the senses in which there might be too much higher education in the economy. I propose two ways in which one can argue that there is too much higher education in the economy, even when the *average* return to higher education is positive. I show how this can appear in the data as what is commonly called overeducation, that is, graduates working in occupations commonly labelled low-skilled. I then consider the quantification of the relative importance of the occupation structure in overall wage inequality, and show how being matched to high-productivity occupations is just as important as firm and individual heterogeneity as a source of wage inequality. Finally, I consider how students choose between field of study at the university level. I find a positive elasticity of choice probability to expected earnings, and document that gender, ethnicity and to a lesser extent, socio-economic status, are also factors that are correlated with subject choice.

The Optimal Level of Education

In the first essay, "Individual and Aggregate Mismatch in Higher Education", I start with the observation that in many countries, there is both a high share of students selecting into higher education (HE) and a large share of graduates who end up working in non-graduate occupations. This has driven politicians in some countries to argue that there are too many graduates, and that policy should look into redirecting some of these graduates into other forms of human capital investment.

There is relatively little work on the normative question of how much HE investment in an economy is too much. The causal literature typically finds positive local average treatment effects, or even policy-relevant treatment effects (Heckman & Vytlacil (2001)). However, there are two concerns about the relevance of these causal estimates in forecasting the effects of a change in the level of higher education investment. First, even if the average effect is positive under the current existing policy, it nevertheless could be the case that there are alternative policies that could do better than the status quo. Second, there may be supply and demand effects which are assumed away by the potential outcomes framework, which imposes the stable unit treatment value assumption. The graduate supply and demand literature (Katz & Murphy (1992); Katz & Goldin (2008); Acemoglu & Autor (2011)) for example interprets the relative wage premium for graduates as an indicator of the relative supply of graduates, and so predicts higher education expansion will reduce the college wage premium. However, this literature also does not offer implications about whether there should be policy action to affect the level of higher education choice in the economy. While it can tell us what the relative changes in supply and demand for graduates are, there is little sense in which we can say that the supply and demand is ‘too high’ or ‘too low’.

Does the existence of “over-educated” workers mean that the level of education is inefficiently high? In my essay, I propose a model that features uncertainty about the causal effect of HE and incorporates general equilibrium wage effects in a labour market matching model. My model highlights that there are two distinct senses in which one can understand the term “overeducation”. Workers could be individually mismatched in the sense that they would be better off if they made alternative education choices. There could also be aggregate mismatch if average welfare would be higher if the level of HE attendance changed. I find that uncertainty about returns generates the former, while endogenous education choice in a matching market generates the latter.

I contribute the empirical literature on overeducation by estimating the model’s structural parameters using UK data. I calculate that 32.9% of the population is individually mismatched. I find in policy simulations that policy-makers can improve aggregate welfare by reducing HE share by 1.6 percentage points, but workers would in fact prefer lower rates of education attendance to reduce the degree of competition they face in the matching market. Reducing uncertainty about returns is unambiguously good for workers but increases income inequality and worsens firms’ outcomes by making labour more expensive.

The Occupation Structure

An important contribution to the literature studying the effect of education on earnings has been the task-approach, pioneered by Acemoglu and Autor (Acemoglu & Autor (2011)). This approach suggests that there are a number of tasks which are the primary inputs to production. Tasks can be performed by different workers at different levels of productivity, leading to specialisation and segmentation in the labour market. They also argue that technological change affects the relative productivity in different tasks differently. For example, they argue that computer advancements may enhance productivity in cognitive, non-routine tasks but not necessarily manual, routine tasks.

One of the implications of this literature is that differences in the complexity of the tasks themselves give rise to differences in pay between workers. More generally, even within firms, there should be differences in pay based on the tasks that workers perform; Google does not pay its canteen chefs the same premium as it does its software engineers. How important are these pay differences in determining aggregate wage inequality? The standard Abowd et al. (1999) framework that has been a workhorse in the study of the sources of wage heterogeneity does not allow for wage differences by occupation.

In my second essay, "The Butcher, the Brewer, or the Baker: The Role of Occupations in Explaining Wage Inequality", written with Luke Heath Milsom, we propose an econometric framework that allows for wage heterogeneity by occupation even within a firm. We estimate a two-way, worker-job, fixed effect model that builds upon the canonical model by Abowd et al. (1999) by allowing for within-firm heterogeneity. Estimating our model on UK administrative data, we find that even if workers, and firms, were identical, 25% of current log-wage variance would remain, due solely to heterogeneity between occupations. We further document that worker heterogeneity accounts for a relatively small proportion of log-wage variance, and that between-firm pay heterogeneity is more important in higher-wage occupations and larger labour markets. Our paper shows that occupational differences have been vastly under-estimated as a cause of income inequality in the UK economy.

Higher Education Choice

Macroeconomic models of the relation between education and the labour market (see Deming (2023) for an overview) typically assume that changes in the wage differential between investment decisions drive shifts in the composition of education decisions in an economy. For example, when demand for graduates increases, the relative wages for graduates increases, inducing more non-graduates to acquire higher education. In a similar vein, Kirkeboen et al.

(2016) finds that there are substantial differences in the return to field of study at university as well. Do these differentials also drive choice of field of study in different subjects?

This question has been active line of research in recent years, driven by the availability of new administrative data and relatively new survey techniques. A major difficulty in answering the research question lies in the fact that students typically do not know how much they will earn if they study different subjects. Rather, they are driven by their expectations and beliefs about their returns in different fields. These expectations and beliefs are not typically observable and it is not generally known how actual shocks to earnings in particular fields relate to students' beliefs about those earnings. Some recent exciting research, such as Wiswall & Zafar (2014), has used information experiments or elicited beliefs to make progress on this fundamental issue.

In my third essay, "Determinants of University Subject Choice in the UK", I make a contribution to the revealed preference analysis of subject choice by analysing students' subject choices in university using a cohort panel dataset. I use differences in expected earnings conditional on student characteristics, such as sex, ethnicity, and ability in English and Math at age 16, that are plausibly exogenous to subject choice, made at age 18. I construct estimates of employment-weighted lifetime earnings for each student conditional on subject choice, and estimate choice models of subject choice using these constructed expectations. I find that earnings, defined as earnings at age 25 or as lifetime earnings, are positively associated with choice probability, although the quantitative impact of expected earnings on subject choice is small. I also find that there are substantial differences in choice probability between students of different sexes, ethnicities and to a lesser extent, socio-economic classes. I also find differences in results depending on whether we use self-described subject preferences at age 17 and observed subject choices from age 18 onwards, suggesting that there are some constraints to students' choice of subject in practice.

Next Steps

While my thesis makes contributions to a number of fields related to the economic analysis of higher education, it also highlights the limits of what we currently know about the subject, and provides guidance on plausible avenues for future developments in the field. In this concluding section to my introduction, I wish to reflect on what I think are some two priorities for the field moving forward.

A recurring underlying theme in this thesis is the tension between microeconomic and macroeconomic approaches to the economics of education. There has been much work done from a microeconometric tradition on higher education. Much of this tradition comes from the analysis of selection, pioneered by Roy (1951), as well as the literature on the analysis of treatment effects, which has devoted much time and energy to the consideration of heterogeneity in recent years. This work has resulted in more credible estimates of the causal effects of education. However, it remains unclear what the policy implications of these causal estimates are, given that they do not speak directly about counterfactuals[1] and omit consideration of general equilibrium effects.

On the other hand, the macroeconomic literature has focused on the characterisation of the relation between the education profile and earnings, but in a way that has typically assumed away heterogeneous treatment effects and self-selection, issues at the core of the identification challenge for recovering credible causal effects. The combination of these perspectives has been a focus of the first essay in my thesis. The development of further work that combines these perspectives is, in my opinion, a priority for the future. For example, it is unknown how selection[2] affects current conclusions about the size of the wage premium and the supply and demand of graduates. This would require a framework that synthesises heterogeneity of

[1] A notable exception is the interesting work on policy-relevant treatment effects (e.g. Heckman & Vytlacil (2001)) which uses additional structure on consider the likely causal effects of particular policies. This points to the possibility of using this framework to analyse optimal policy. In a related exercise, work on policy learning using observational data (e.g. Athey & Wager (2021)) are also considering the question about how to use causal effect estimates to form policy.

[2] A notable exception is Carneiro & Lee (2011).

returns, endogenous choice, and a labour market framework that creates supply and demand effects as well as accounts for quality concerns.

Another tension between the microeconomic and macroeconomic perspectives is the relation between actual shocks, e.g. to the return to particular fields of study, and students' expectations and beliefs, which are what actually drives student choices. For example, I noted the difficulty of analysing choices when expectations and beliefs are not generally observed. Even in the existing work on students' expectations about their earnings in particular fields of study, it is not known how these expectations would change if students' were subject to particular shocks. It is possible that this analysis could prove important to the resolution to some puzzles in the field. For example, some studies (e.g. Abramitzky et al. (2019)) find much larger estimates of the elasticity of choice probabilities to changes in pecuniary remuneration than others (e.g. Berger (1988); Beffy et al. (2012)). These differences may relate to the different salience of the identifying variation used by the studies. Understanding the relation between real wage changes and student expectations is likely to prove important for the understanding of the policy tools governments have at their disposal to affect the skill mix and education profile of the population.

Chapter 1

Individual and Aggregate Mismatch in Higher Education

Abstract

In many economies, many graduates are observed to work in non-graduate occupations. Does this mean that the level of education is inefficiently high? I propose a model that features uncertainty about the causal effect of HE and incorporates general equilibrium wage effects in a matching model of the labour market. My model highlights that there are two distinct notions of overeducation. Workers could be individually mismatched in the sense that they would be better off if they made alternative education choices. There could also be aggregate mismatch if average welfare would be higher if the average level of HE attendance changed. I find that uncertainty about returns generates the former, while endogenous education choice in a matching market generates the latter. Structurally estimating my model on UK data, I calculate that 32.9% of the population individually mismatched. I find in policy simulations that policy-makers can improve aggregate welfare by reducing HE share by 1.6 percentage points. Reducing uncertainty about returns is unambiguously good for workers but increases income inequality and worsens firms' outcomes by making labour more expensive.
JEL Codes: D81, I23, I26, J24
Keywords: Overeducation, efficient human capital investment, labour market matching

1.1 Introduction

In many developed countries, a large and increasing share of young people attend higher education (HE), completing undergraduate degrees. HE attendance is often subsidised by governments, either directly through tuition fee subsidies (e.g. Germany, France) or the provision of below market-rate loans (e.g. United Kingdom, Australia). The large numbers of HE students and the high cost of such HE subsidies to the state have often prompted concerns that there are too many students who select into higher education. These concerns are exacerbated by the observation that a significant share of graduates end up working in occupations which typically do not require a college degree, a phenomenon which is known as overeducation or underemployment. These underemployed graduates typically earn less than their peers, and are persistently employed in low-skilled employment (Barnichon & Zylberberg (2019)).

How can policy-makers evaluate whether there is in fact overinvestment in HE and whether policies to reduce HE attendance will have a positive effect? Over-education has typically been studied in the macroeconomics tradition; in these models, workers are either exogenously heterogeneous in their human capital (Albrecht & Vroman (2002); Dolado et al. (2009); Barnichon & Zylberberg (2019)) or can select irreversibly into human capital levels (Jackson (2021)). There are at least two types of jobs with varying skill requirements, and workers compete for jobs either in unsegmented or partially segmented labour markets. Skilled workers may accept low-skilled jobs either due to search frictions or because they wish to escape competition from other skilled workers. The normative implications depend on the specific model but arise from the common finding that the decentralised equilibrium in such models is typically inefficient, due to common externalities which occur in search models like the hold-up externality, the thick markets externality and the congestion externality. One weakness of these models is that they do not typically feature heterogeneity in returns to education[1].

[1] These models typically assume some other form of individual heterogeneity to account for workers making different education decisions. Chade & Lindenlaub (2021) assume heterogeneity in the cost of education while Jackson (2021) assumes heterogeneity in endowments.

This paper argues that popular discussions of over-education use the term in a different sense to that in the macroeconomic literature on underemployment. In those models, there is overeducation or undereducation in the sense that aggregate utility in the economy, defined in some specific way, would increase if the profile of HE attendance in the economy changed. This is the sense in which one may argue that there is credentialism in the labour market or inefficient competition in HE attendance, and that average worker utility may increase if fewer students chose to engage in HE. I describe this definition of overeducation as aggregate overeducation. Another way in which people use the term is to argue that an individual is over-educated a graduate is working in an occupation which does not 'require graduate skills'. A possible definition based on this usage is as follows: individual education mismatch occurs when a worker selects into HE despite not benefiting from it, or if a worker does not select into HE despite benefiting from it. This usage clearly is similar to the notion of potential outcomes in causal analysis; a worker is individually mismatched if they make a choice that does not lead to the highest potential outcome. Such individual mismatch could occur ex-post when workers make ex-ante rational decision based on imperfect information (Cunha et al. (2005), Athey & Wager (2021)).

I propose a framework that rationalises both individual mismatch, and possible aggregate mismatch that arises from general equilibrium effects. To this end, I develop a model of education investment under uncertainty set in a labour market characterised by matching between workers and firms. In the model, the return to HE in terms of human capital is heterogeneous across workers, varying by an index which I call ability. This return is imperfectly observed at the point of choice; workers can only observe their grades in school, which I interpret as noisy signals of their underlying ability. As such, workers make an education decision based on an ex-ante expectation of their likely returns to education, but ex-post may experience regret if they are 'unlucky', i.e. they received a signal which convinced them that they were higher ability than they actually are. Post HE, workers match to jobs in a frictionless matching market based on their post-education human capital. This matching labour market generates general equilibrium price effects; when the share of workers in HE increases, workers have less bargaining power as there are more skilled workers in the economy and their wages conditional on their human capital fall as a result.

In my framework, the main determinant of the extent of individual mismatch is the amount of uncertainty that workers face about their underlying return to HE. This suggests that an important function of testing is to enable workers to make more accurate choices about their human capital investments. I also find that there are two forces that generate aggregate mismatch in my model. First, there is a hold-up externality where workers do not internalise the effect of their education on firms' profits because they cannot appropriate this surplus. Thus, they will under-invest in higher education. Second, because workers' wages depend on their position on the overall skill distribution in the economy, they exert positional externalities on others when they invest in education. Workers do not consider the impact of their education choice on the bargaining power of other workers in the economy in equilibrium. This leads to over-investment in higher education relative to the social optimum. Because these externalities go in opposite directions, it is ambiguous whether the level of education is too high in the economy. In general, workers and firms have differing policy preferences, so policy-makers have to decide how to weight the preferences of workers and firms in deciding aggregate policy.

I contribute empirical estimates of the extent and welfare costs of education mismatch by estimating a parametric version of my model on data from the UK using the Simulated Method of Moments (SMM). I find that students face substantial uncertainty about their eventual returns to university; the estimated correlation between unobserved ability and the signal is only 0.324. I estimate that 32.9% of workers would have been better off if they made a different education decision to the one they actually did. 18.2% of workers or 39.1% of graduates did not benefit from university despite going, and 14.7% of all workers or 27.6% of non-graduates would have benefited from university despite not going. The costs in utility terms are small on average but non-negligible. The average utility loss to going to university for overeducated workers is equivalent to earnings increases of £0.183 per hour on average, while the average utility loss to missing out on university for undereducated workers is equivalent to £0.264 per hour on average. These results suggest that while over-education is the focus of headlines in the education debate, under-education is a substantially more costly problem when it occurs, which is why workers tend to err on the side of over-investment.

I conduct a number of counterfactual simulation exercises to shed light on optimal policy in this framework. First, I consider the implementation of a compensated graduate tax or subsidy aimed at affecting the share of students who choose to go to university. I show that workers and firms are not aligned in their policy preferences. A compensated graduate tax to reduce the level of education would increase average worker utility by increasing wages when the supply of skilled workers falls. Firms on the other hand would prefer a maximal subsidy as they do not pay for workers' education and would like to free-ride on workers' human capital investments both in terms of increased productivity of matches and a lower wage because of labour supply. In practice, it seems to be welfare-improving to reduce the level of education by 1.6pp by implementing a small revenue-neutral graduate tax.

Second, I consider the effects of targeting the quality of the information students have about their return to education. I find that, as expected, increasing the informativeness of students' signals about their return to education increases their welfare, but surprisingly, it decreases firms' profits. Increasing informativeness leads to increased inequality, because it allows high return workers to more perfectly sort into education and low return workers to sort out of education. The increased inequality also leads to a more polarised distribution of skilled workers, leading to higher wage costs for firms and a reduced worker quality for less productive firms.

One of the main contributions of this paper is to provide a bridge between the microeconometric literature that focuses on uncertainty and heterogeneity in the returns to education (Altonji (1993), Cunha et al. (2005), Cunha & Heckman (2016), Arcidiacono et al. (2016),Heckman et al. (2018)), and the macroeconomic analysis of the relation between education and the wage structure. Many macroeconomic treatments of the supply of graduates do not model the education decision (Albrecht & Vroman (2002), Dolado et al. (2009), Barnichon & Zylberberg (2019)), and those that do have typically abstracted away from Roy-type selection on returns to higher education, opting for simpler mechanisms with constant returns and heterogeneous costs (Chade et al. (2017), Jackson (2021)). My model expands on these macroeconomic models by adding an endogenous education choice featuring both Roy-type selection based on heterogeneous returns and uncertainty about one's return to education.

Using this new model, I contribute to the analysis of overeducation in the economy by (1) distinguishing between two senses (individual and aggregate) in which over-education may exist, (2) providing a mechanism for individual overeducation, and (3) showing that workers' and firms' incentives are not necessarily aligned in the conduct of optimal policy. Previous work has focused on over-education has been to explain why it must exist in equilibrium with rational agents. Proposed explanations include labour market frictions leading to educated workers choosing to accept 'non-graduate' jobs (Albrecht & Vroman (2002), Dolado et al. (2009), Jackson (2021)), skilled workers choosing to enter low-skilled jobs in order to escape competition from other skilled workers (Barnichon & Zylberberg (2019)), and workers having preferences for particular low-skilled occupations (Gottschalk & Hansen (2003)). Besides Jackson (2021), these models do not analyse the normative implications of their models for the level of education in the economy.

My model introduces the possibility of an 'involuntary' kind of underemployment due to mismatches between ex-ante expectations and ex-post realisations of the return to education. It also highlights a new channel that is important to determining the level of individual mismatch in higher education, the quality of the signals that students receive about their ability in college. This point relates to work by Athey & Wager (2021) about how best to assign treatments (e.g. university attendance) based on information available to colleges and the government. It also relates to a line of research based on research discontinuity designs (e.g. Tan (2022)) suggesting that testing results and grades may have an informative purpose in informing students about their underlying ability and helping them make more allocatively efficient life decisions. I also show that while increasing workers' information is beneficial to workers, it may also exacerbate wage inequality in leading to more perfect Roy-type sorting in equilibrium and reduce firm profits by increasing workers' wages.

This paper proceeds as follows. I conclude the introduction by discussing the literature that my paper relates to. I begin by describing some stylised facts about UK higher education in section 1.2 that my model seeks to explain. I then describe my model in section 1.3. In section 1.4, I describe the empirical parameterisation of my model, identification and the data sources which I use. Section 1.5 describes the results of my estimation, and the model fit. Section 1.6 analyses worker welfare in this model given the estimated parameters. I consider whether policy makers can increase welfare using a compensated graduate tax or subsidy in section 1.7, relating the exercise to current policy proposals being considered by the UK government. Section 1.8 concludes.

1.2 Stylised Facts

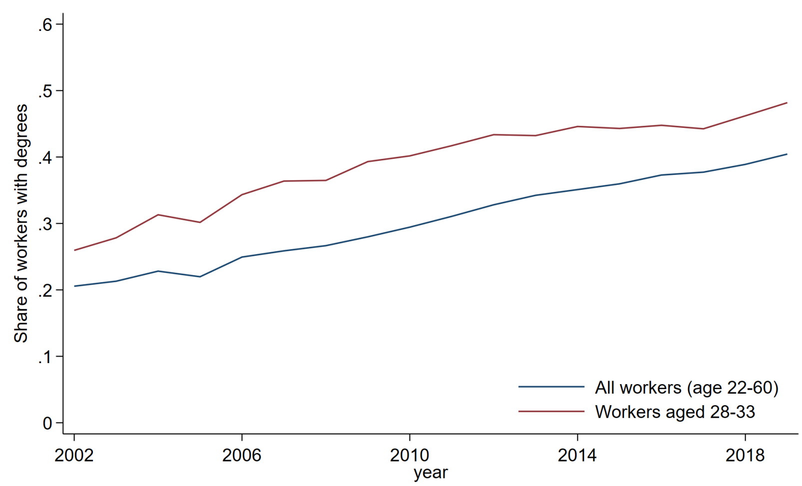
I motivated the paper by noting that in many developed country contexts, there is increasing concern about there being too many students selecting into higher education. In this section, I present some stylised facts which prompt such concerns in the context of the UK, which is the country from which the data for my empirical application comes. A recent paper, Jackson (2021), has documented similar patterns in the American context. The data used in this section comes from the UK Labour Force Survey, a nationally representative quarterly survey providing self-reported information on a cross-section of workers.

In summary, I find that the share of students who get university degrees has increased substantially over time. The simple college wage premium has fallen over time, and a large reason for this seems to be that the premium from matching to higher paying occupations has fallen. I show that a substantial part of this decline seems to be driven by the lowest earning graduates. Finally, I note that a large share of graduates are observed to work in occupations that are in some sense low-skilled. These facts motivate a model, presented in the following section, emphasising labour market matching to particular occupations, where ex-post some graduates end up with low labour market performance.

1.2.1 The higher education wage premium

There has been a large and sustained increase in higher education in the UK over the last two decades. Figure 1.1 plots the share of workers with a degree or equivalent in each year from 2002-2019 for the group of 22-60 year old workers and the subset of 28-33 year old workers. The share of workers with degrees increased from 20.6% in 2002 to 40.4% in 2019, an increase of 19.8 percentage points. This fact has been previously documented in Walker & Zhu (2008), Devereux & Fan (2011), and Blundell et al. (2022). Devereux & Fan (2011) argues that this increase in the number of graduates is plausibly exogenous, and Blundell et al. (2022) argues that the increase is due to policy choices by the British government.

Figure 1.1: Changes in the share of workers with degrees from 2002-19



Data from the UK Labour Force Survey 2002-19. I consider a worker to have a degree if their highest achieved qualification is recorded as a first degree or a higher degree.

Is there any indication that there are now too many graduates in higher education? In the supply and demand framework used in Katz & Murphy (1992), when the supply of graduates increases faster than demand for graduates, the college wage premium should fall[2]. A falling wage premium may thus be indicative of an excess of graduates. Carneiro & Lee (2011) further argue that when more students select into education, the marginal student is likely to be of lower quality and thus should further lower the wage premium by changing the quality composition of the groups of graduates and non-graduates.

To compute the wage premium, I regress log hourly earnings on a degree indicator, sex and age fixed effects with the LFS sample in each year from 2002-19. Like in Lemieux (2014), I compare this base regression to a second specification with occupation fixed effects. Lemieux (2014) argues that two possible mechanisms for the causal effect of education on earnings are that workers are more productive within the jobs they work at, and that workers match to higher paying occupations due to their training[3]. Controlling for occupation disables the occupation channel, leaving only the human capital channel, as explanations for why graduates earn more than non-graduates. Thus, comparing the coefficient of the base regression to the coefficient of the regression with occupation controls gives a (non-causal) indication of the relative importance of the occupational upgrading story as explanations of why graduates earn more than non-graduates.

Figure 1.2 plots the coefficient on degree in each year for the base regression in blue and the regression with occupation fixed effects in red. I find that the college wage premium has declined, especially in the latest years from 2013 to 2019, from 0.481 to 0.429[4]. The wage premium conditional on the occupation match has stayed largely stable, falling from 0.184 to 0.171. This seems to suggest that a large part of the fall in the college wage premium comes from a fall in the difference between the occupations that graduates and non-graduates

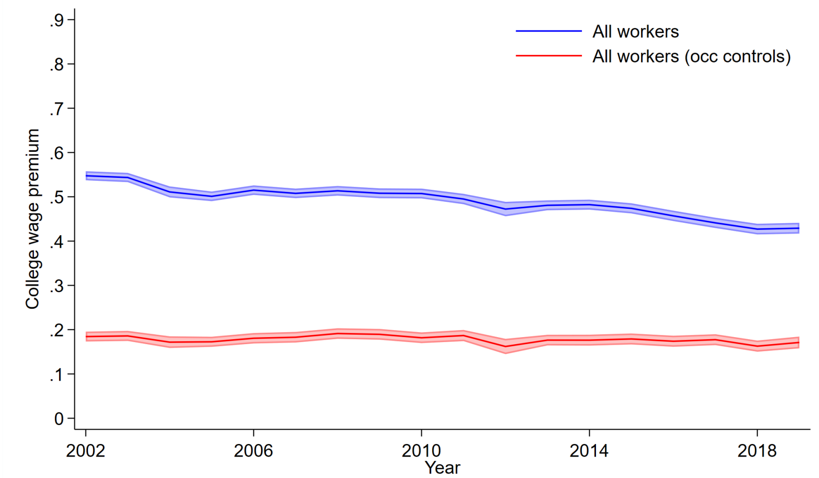
[2]The calculated wage premium is also not an estimate of the causal wage return to education, since it does not take into account selection into education and other forms of endogeneity. There have been three main papers (Blundell (2000); Belfield et al. (2018a); Waltmann et al. (2020)) estimating the return to higher education in a British context. All these papers estimate Mincer equations, accounting for heterogeneity using a matching approach; they take advantage of detailed observations of schooling and family variables to argue that any residual selection effects do not substantially affect the estimate.

[3]Lemieux also considers a match channel, where workers may earn more if the occupation they work in 'matches' in some sense the education that they achieve. I abstract from this channel in this analysis.

[4]Much of the older literature on the UK higher education premium has tended to conclude that the higher education expansion that started in the 90s did not substantially reduce the observed college wage premium. This seemed to be true from 2002 to 2010, when wage premium stayed at about 0.50.

work in. Finally, the absolute size of the fall in the premium, by roughly 10-15 percentage points, is small relative to massive expansion in both the share and numbers of students now in education.

Figure 1.2: The college wage premium from 2002-19



Data from the UK Labour Force Survey 2002-19. I consider a worker to have a degree if their highest achieved qualification is recorded as a first degree or a higher degree. The mean college wage premium for each year is calculated as the coefficient on the indicator for whether the individual has a degree of a regression of log hourly earnings on sex, age fixed effects and the degree indicator. The occupational-conditional premium for each year is calculated as the coefficient on the indicator for whether the individual has a degree of a regression of log hourly earnings on sex, age fixed effects, occupation fixed effects and the degree indicator.

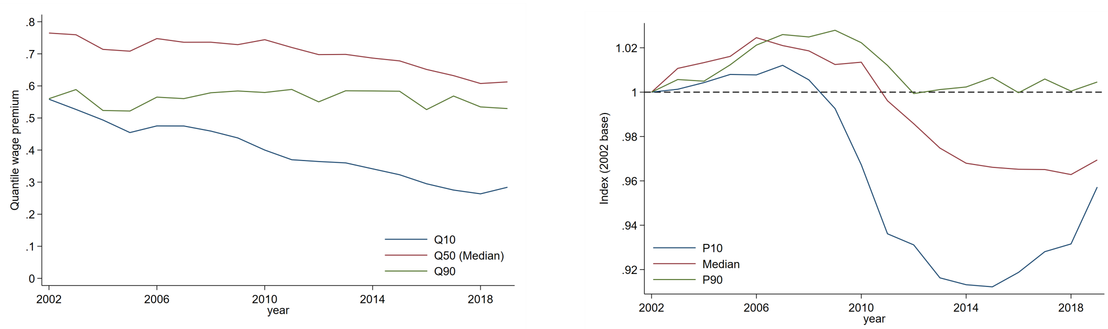
The percentiles of the log wage distributions conditional on education

Many recent studies on the causal impact of higher education have highlighted the heterogeneity of the effect higher education has on students (in the UK, Belfield et al. (2018a); Waltmann et al. (2020); in the US, Heckman et al. (2018); Andrews et al. (2022)).

It is possible that the mean or median wage premia may mask different and divergent trends for graduates at the tails of the wage distribution. To check this, I run a number of quantile regressions at the 10th percentile, median and 90th percentile to study how the earnings gap between graduates and non-graduates have changed at these points. I run quantile regressions of log hourly earnings on degree status, age fixed effects and sex in each year from 2002-19. Figure 1.3a plots the coefficients on degree status by year for the 10th, 50th and 90th percentiles.

As the blue line (for the 10th percentile) and the red line (for the median) shows, the difference of the conditional 10th percentile and median log wage between graduates and non-graduates has fallen substantially over the years, suggesting that the fall in the wage premium is largely due to the lower quantiles of the graduate wage distribution. Figure 1.3b plots the indices of the log-wage of graduates at the 10th, 50th, and 90th percentiles, with 2002 as the base year. The real log wage has fallen substantially for the median and 10th percentile graduate. Any theoretical model should ideally explain why the wages at the 10th percentile and median have fallen over the course of substantial expansion of higher education.

Figure 1.3: Changes in the distribution of earnings conditional on degree status



(a) Coefficients on quantile regressions at p10, p50 and p90

(b) Log wages for graduates at p10, p50 and p90

Data from the UK Labour Force Survey 2002-19. Controls for the quantile regressions in figure 1.3a include sex and age fixed effects. The wage measure used in figure 1.3b are real log hourly earnings.

1.2.2 The share of graduates in low-skill occupations

Another stylised fact that is often found in the UK context is that there is a significant and increasing share of workers with graduate degrees working in so-called ‘non-graduate’ jobs[5].

[5]As discussed in the literature review, the demarcation between graduate and non-graduate jobs is controversial, and a main concern of the large overeducation literature has been the clarification of the demarcation between the two categories. There are two main approaches to this issue. First, some researchers, including Green & Zhu (2010), and Green and Henseke (2016), argue that a classification should be based on some notion of an occupation requiring certain skills. Approaches which make use of existing skills classifications, such as the Office for National Statistics SOC classifications or the Home Office Appendix J classifications

Given the policy interest in the subject, especially since the government implicitly subsidises graduates who cannot fully repay their student loans, the UK Office of National Statistics has published a number of technical reports attempting to quantify overeducation using a variety of methods[6].

In this section, I first show that there are more graduates in occupations classified by the government statistical agencies as non-skilled occupations. The narrative that emerges from this exercise seems consistent; there are more graduates today working in occupations that would traditionally be considered 'non-graduate'. Figure 1.4a plots the share of graduates working in occupations considered to require a NVQ 1 or 2 level education (equivalent to having a good high school qualification), as classified by the ONS occupational classification. Over the studied time period, the share of workers working in low skill occupations increased from 11.4% in 2002 to a high of 19.4% in 2018. Figure 1.4b plots a similar graph, but instead uses a classification used by the UK Home Office to classify occupations into different skill levels[7]. Occupations were classified into low skill (requiring less than an A-level qualification), requiring an A-level qualification (RQF 3), requiring one year of a bachelor's degree (RQF 4), requiring an undergraduate degree (RQF 6), and requiring a postgraduate degree. At the start of the period, 9.8% of graduate workers worked in an occupation classified as low skill, but at the end, 14.2% of graduate workers do. 22.1% of workers were in an occupation requiring a high-school qualification or below at the start of the period, and at the end, 27.4% did. Thus, on the face of it, it seems that more graduate workers are now employed in so-called low-skill qualifications requiring a high-school degree or lower. Appendix A shows that while the share of graduates in so-called low-skill occupations do vary between different subjects, non-negligible shares of graduates in most subjects, both in STEM and non-STEM, work in

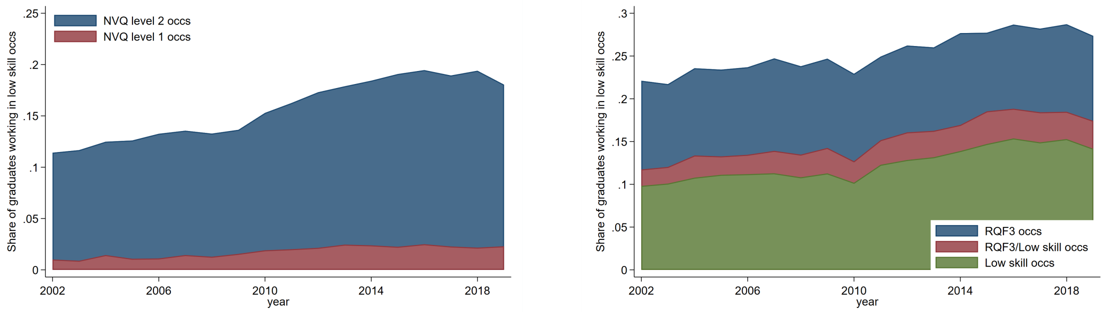
for occupation skill levels, would fall under this approach, as do approaches which turn to the task content of occupations to stipulate whether a job is 'graduate'. Second, other researchers have tried to classify occupations into 'graduate' and 'non-graduate' categories by examining outcomes, either the college-noncollege wage premium within an occupation (Gottschalk & Hansen (2003), O'Leary & Sloane (2016)) or the share of workers with degrees within that occupation (Elias and Purcell (2003)).

[6]See Clegg (2017) and Saric (2019). Academic work in a UK context attempting to quantify overeducation include Elias & Purcell (2004), Green & Zhu (2010), and O'Leary & Sloane (2016).

[7]This classification has also been used in Aghion et al. (2019)

low-skilled occupations[8].

Figure 1.4: Share of graduates in low-skill jobs by two different classifications



(a) Share of graduates working in occupations requiring NVQ1/2 qualifications

(b) Share of graduates in low skill occupations by Home Office classification

Data from the UK Labour Force Survey 2002-19. The occupational classifications produced by the Office for National Statistics are assigned skill levels relating the qualification level (based on the old NVQ classification) required for the job. The appendix J classifications are produced by the Home Office for the then-purpose of classifying prospective migrants by their skill level according to the Regulated Qualifications Framework.

The interpretation of this fact is however not straightforward. As the number of graduates increase, the market readjusts[9], and jobs which were previously seen as non-graduate may be increasing taken up by graduates. This is in fact a feature of the model proposed in the following section, that as the supply of graduates changes, the composition of graduates within particular occupations similarly change. Appendix A documents that the increase in the number of graduates between 2011-19 led to these graduates finding employment in occupations which previously had low shares of workers with university degrees, further showing that the occupations graduates work in change when the supply of graduates changes.

1.3 A Model of Education Investment and the Labour Market

In this section, I study a model in which workers first decide in an initial period whether or not to invest in education, and subsequently receive a wage depending on their heterogeneous

[8]This does not apply to degrees for vocational subjects, such as medicine, subjects allied to medicine, architecture, and to a lesser extent, education.

[9]Models in which the composition of jobs changes as a result of technological change or an increase in the number of graduates include Acemoglu (1999), Albrecht & Vroman (2002), and Gottschalk & Hansen (2003).

characteristics and their education choices. The model of education choice in the first period is a Willis-Rosen type model of self-selection into education, but with incomplete information and uncertainty about returns, like in Cunha et al. (2005). The labour market is modelled with a frictionless matching model with transferable utility, like in Sattinger (1993)[10]. The central aim of the model is to rationalise the patterns in the wages and occupational outcomes of graduate and non-graduate workers with their choices to invest in higher education, as well as to show how both individual and aggregate overeducation can arise in this setting.

This section proceeds as follows. I describe the model and maintained assumptions in section 1.3.1, including the characterisation of equilibrium in this setting. I consider optimal private investment and optimal social investment in education in this model in section 1.3.2, and discuss how individual mismatch arises in this setting. Section 1.3.3 argues that the selection into higher education that occurs based on individual choice is not generally socially optimal, and leads to aggregate mismatch.

1.3.1 Description of the model

Setting

There are two masses of workers and jobs of equal length. As a primitive, workers are heterogeneous in two respects. First, they differ in their labour market ability, denoted by a and distributed with density function $f_A(a)$. Second, they differ in their preference for higher education, denoted by two random variables η_1 and η_0 representing preference for HE and not going to HE respectively. Their net preference for HE is denoted by $\Delta\eta$ which is a random variable with distribution $f_{\Delta\eta}(\Delta\eta)$.

The random variable a determines a worker's return to education. In the model, workers

[10]This is, in some ways, the simplest characterisation of a matching market with continuous types on both sides of the market. The main alternatives include two-sided matching with search frictions (Shimer & Smith (2000)), or reducing the number of types of either workers or firms (Acemoglu (1999), Shephard & Sidibe (2019)).

can choose to select into higher education, denoted by $e \in \{0, 1\}$. Attending HE increases their skill on the labour market; in the model, it is the skill level s, instead of initial ability a, that determines the worker's return on the labour market. The relation between labour market ability and education, and skill is summarised by the skill function $S(a, e)$. I assume that the skill function $S(a, e)$ is continuous and differentiable in a for both values of e. Furthermore, I assume that $\frac{\partial S(a,1)}{\partial a} > \frac{\partial S(a,0)}{\partial a}$ for all values of a, such that the difference of $S(a, 1) - S(a, 0)$ is increasing. Intuitively, this implies that the skill return to education, denoted by $\Delta S(a) \equiv S(a, 1) - S(a, 0)$, is increasing in ability.

Workers are unable to fully observe their labour market ability a, and instead receive a noisy signal θ, which imperfectly reveal their underlying ability to the worker. I assume that the noisy signal is additive in ability and a white noise term ε, as follows.

$$\begin{aligned} &\theta = a + \varepsilon \\ &\varepsilon \sim N(0, \sigma_\varepsilon^2) \end{aligned} \tag{1.1}$$

The signal is imperfectly correlated with labour market ability a because (i) the current testing technology available to us cannot perfectly capture worker talent and ability, and (ii) in the time between when the signal was drawn and labour market performance is realised, workers could experience a number of shocks to their labour market performance, e.g. adverse health shocks, or luck in meeting the right opportunities. More generally, the variance of ε governs the correlation between the information workers have at the time they have to make an education decision and their actual labour market ability.

Jobs have varying levels of productivity, denoted by y. Suppose that there are K discrete occupations, indexed by k, where the distribution of productivity of jobs within that occupation is denoted by $f_y^k(y)$, and the share of jobs in that occupation in the economy is denoted by p_k, where $\sum_{k \in \{1,\cdots,K\}} p_k = 1$. The distribution of productivity in the entire economy is given by $f_y(y) = \sum_{k=1}^{K} p_k f_y^k(y)$. In this model, occupations do not serve any direct role in the equilibrium match (i.e. there are no preferences for occupations), but they are important for two reasons. First, they help the estimation by providing information on the productivity of the job the worker is matched to, especially if information on the matched job is not available. Second, they lead to testable implications about the resulting share of workers with degrees and the college wage premium within occupations. I do not make any assumptions about the ordering of occupation-specific productivity distributions. I discuss implications of the model for the occupation structure in the economy further in appendix B.

There are two periods in the model. In the first period, workers decide whether to invest in education based on their expected returns, and their preferences for education. The decision to invest in education is described in detail in section 1.3.1. This determines the distribution of skill among workers in the economy. In the second period, workers with skill s and jobs with productivity y can match pairwise with each other to produce joint output $g(s, y)$. I assume that $g(s, y)$ is twice continuously differentiable, increasing and supermodular[11]. A model of this kind is known in the matching literature as a frictionless matching model with transferable utility and has been analysed thoroughly in other contexts. The resulting wage function for workers with skill s is the wage that individuals receive in the second period and is analysed in the following subsection 1.3.1.

In summary, the following assumptions are maintained throughout.

Assumption 1. *In the model, I maintain the following assumptions.*

1. *(Information structure of workers) Workers do not observe their initial ability a, and observe a noisy signal θ, related to a by the following expression. θ is the only information that the worker has about a and their return to education. ε has a mean zero, log-concave distribution.*

$$\theta = a + \varepsilon$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2)$$

[11] In this case, the supermodularity condition is equivalent to $\frac{\partial^2 g}{\partial s \partial y} \geq 0$ (Topkis (1998)).

2. *(Assumptions on $S(a, e)$) The skill function $S(a, e)$ is continuous and differentiable in a for both values of e. Furthermore, I assume that $\frac{\partial S(a,1)}{\partial a} > \frac{\partial S(a,0)}{\partial a}$ for all values of a, such that the difference of $\Delta S(a) \equiv S(a, 1) - S(a, 0)$ is increasing in a.*

3. *(Assumptions on joint output $g(s, y)$) The joint output function $g(s, y)$ is increasing in both s and y, and twice continuously differentiable. The function is assumed to be supermodular, which is equivalent to the following condition since it is twice-continuously differentiable: $\frac{\partial^2 g}{\partial s \partial y} \geq 0$.*

The labour market and the wage function

Post-education, workers come to the labour market where they seek to find a single job. After completion of higher education, firms are able to observe workers' skill types; this assumption coheres with work by Lange (2007) which shows that employers typically learn quickly about their employees productivity types. For simplicity, I assume that once matched, workers stay matched forever in the job; there are no dynamics in this model that can lead workers to transition between jobs[12]. I also assume for simplicity that there is no unemployment[13]. To fix ideas, I start by describing this labour market as resolving a discrete assignment problem, where N workers match to N jobs. The results that I need however are expressed in continuous terms with continuums of jobs and workers, which can be thought of as corresponding to the situation when N becomes very large.

On the labour market, workers are heterogeneous to job-operators only insofar as that they have different levels of skill. The distribution of worker skill can be described by distribution $f_s(s)$, which is exogenous in the labour market stage but endogenous to workers' earlier education choices[14]. Jobs differ in their productivity (as described in the previous section),

[12]Because there are no frictions in this model, the matching that obtains is stable, which implies that there is no coalition of workers and jobs which would block it because it is profitable to do so. In this sense, the outcome specified here is closer to an end state for a dynamic matching process, and one could interpret the adoption of this matching process as abstracting from the frictional process of arriving at this end state.

[13]It is possible to relax this assumption by assuming that there are more workers than jobs. See Chade et al. (2017) for an overview.

[14]The determination of $f_s(s)$ is discussed in sections 1.3.1 and 1.3.1.

which governs the total joint output that a worker with skill s and a job with productivity y produces. The primitives of the matching problem that characterises the labour market are thus the distribution of skills $f_s(s)$, the overall distribution of productivity $f_y(y)$, and the joint output function $g(s, y)$.

Workers and firms simultaneously play what Shapley & Shubik (1971) refer to as an assignment game, that is, workers and jobs have to find a counterparty to match to, and to divide the joint output that results from the match. Because there are no restrictions on how the joint surplus is divided besides feasibility and rationality, workers and firms alike can attempt to induce more productive parties to match with them by offering to accept less of the surplus themselves. The outcome of this game is a matching, which is a specification of which workers match with which jobs, denoted μ, and a set of wages paid to workers w and a set of profits π paid to job-operators which come from the division of the joint output. In the continuous case, the matching function is a continuous function, denoted μ, mapping the domain of s to the domain of y, while the wage ($w(s)$) and profit ($\pi(y)$) functions become continuous functions in s and y respectively; i.e. the outcome of the assignment game in the continuous case is a triple of functions $\{\mu, w, \pi\}$.

This is a well-studied question in the field of matching theory[15], and in this section, I only present the results and intuition relevant to my application, omitting the proofs. In particular, I am interested in what the outcome will be in the assignment game in the labour market. An important notion here is that of stability; a stable outcome is one which no pair of workers and jobs can profitably block. The technical condition related to the no-blocking requirement is that for any values of s, y:

$$w(s) + \pi(y) \geq g(s, y), \quad \forall s, y \tag{1.2}$$

If this condition did not hold, then worker s and firm y could increase their payoffs by matching

[15] The exposition of the problem, its solution and attendant proofs have been discussed in many other papers and books, including Galichon (2016), and Chade et al. (2017).

with each other and increasing their output by splitting the excess $g(s, y) - w(s) - \pi(y)$. A stable outcome can be defined as a feasible outcome where the no-blocking condition given by equation 1.2 is satisfied, where feasibility means that all agents are matched and that the sum of the surpluses received by all workers and jobs does not exceed the total output of the economy.

Combining results from Shapley & Shubik (1971) and Becker (1973), it turns out given the supermodularity of the joint output function $g(s, y)$, the unique stable outcome is positive assortative matching, i.e. when the most skilled workers match to the most productive firms. Then the assignment that results in the labour market is given as follows, where $F_s(s)$ and $F_y(y)$ denote the cumulative distribution functions of the skill distribution and the productivity distribution respectively.

$$\mu(s) = F_y^{-1}(F_s(s)) \tag{1.3}$$

What are the payoffs that can sustain this positive assortative matching? One can pin down the pay-offs $\{w, \pi\}$ by demonstrating the properties it has to satisfy if such pay-offs exist, and then in turn show that the pay-offs are feasible[16]. This exercise is often interpreted as showing that the optimal assignment can be sustained in a competitive (Walrasian) market in which agents take the prices they have to pay to match as given, and optimise over who they want to match with. Let $w(s)$ and $\pi(y)$ denote a pay-off that can sustain the unique stable assignment. Then, workers' and job-operators' individual problems are to choose an agent of the opposite kind to match with so as to maximise their individual pay-off, as follows:

$$\text{Worker's problem for worker with skill } s^*: \quad \mu(s^*) = \arg\max_y g(s^*, y) - \pi(y) \tag{1.4}$$

$$\text{Firm's problem for firm with prod } y^*: \quad \mu^{-1}(y^*) = \arg\max_s g(s, y^*) - w(s) \tag{1.5}$$

In equilibrium, the first-order condition of the maximisation problem faced by firms needs

[16] As this is not relevant to my application, I refer those interested in the second part of this proof to chapter 4 of Galichon (2016).

to be satisfied by the worker with skill s that it matches to. That is, for all values of s, the following condition should hold, pinning down the shape of the resulting wage function. Analogously, the first-order condition of the maximisation problem faced by the worker fixes the shape of the profit function as follows, where $\mu^{-1}(y)$ specifies the skill of the worker that a firm with productivity y matches to. This also suggests that the wage and profit schedules are convex functions[17].

$$\frac{\partial g(s, \mu(s))}{\partial s} = w'(s) \tag{1.6}$$

$$\frac{\partial g(\mu^{-1}(y), y)}{\partial y} = \pi'(y) \tag{1.7}$$

We can then solve for the wage and profit functions by integration.

$$w(s) = w_0 + \int_{-\infty}^{s} \frac{\partial g(x, \mu(x))}{\partial x} dx \tag{1.8}$$

$$\pi(y) = \pi_0 + \int_{-\infty}^{y} \frac{\partial g(\mu^{-1}(x), x)}{\partial x} dx \tag{1.9}$$

The wage function is thus fixed up to location, given by w_0 which is a constant of integration[18]. The outcome of the labour market is thus the matching $\mu(s)$, which predicts positive assortative matching between workers and jobs (given by equation 1.3), and the wage and profit functions, given by equations 1.8 and 1.9. These wage and profit functions depend on the primitives of the matching problem, in particular the distribution of skill in the economy, which is in turn dependent on workers' education choices. In this way, there is a feedback channel from the education decisions of workers to the wages that they obtain in the model.

[17] To see this, note that for any $s' > s''$, $\mu(s') \geq \mu(s'')$, with non-equality for most continuous distributions of y. Then, by the assumption of supermodularity, $\frac{\partial g(s,\mu(s))}{\partial s}(s') > \frac{\partial g(s,\mu(s))}{\partial s}(s'')$, and so $w'(s') > w'(s'')$.

[18] This indeterminacy is discussed in Sattinger (1993); one intuition is that since the incentives for matching with any type depends on the relative incentives, any wage function that preserves the shape of the function leads to the same assignment in Walrasian equilibrium. The values of w_0 and π_0 must be such that $w_0 + \pi_0 = g(s_0, y_0)$, where s_0 and y_0 denote the minimum (or infimum) value of s and y in the domain of the distribution of both variables. The precise value of w_0 relative to π_0 can be interpreted as the bargaining outcome between the least skilled worker and the least productive firm, possibly reflecting the relative outside options of each agent.

The decision to invest in education

As described before, workers have heterogeneous preferences for higher education, denoted by η_1, and for non education, denoted by η_0, with $\Delta\eta$ denoting their net preference for HE. I assume that the heterogeneous preference shocks for education are independent of the other shocks in the model, i.e. of the information shock ε, as well as the initial distribution of ability a. They also receive utility from their wages in the subsequent period, with time preference factor β. Students also face a fixed cost for investing in education, which I denote by $-\kappa$.

Workers have beliefs about the wage they would receive conditional on the skill that they have. Let $w^B(\cdot)$ denote a wage function which maps a level of skill to a wage on the labour market. $w^B(\cdot)$ defines the wage environment that workers believe will prevail in the labour market at the time of workers' educational choice. For simplicity, I assume that all workers believe that the same wage environment prevails.

Denote the difference between the value functions under HE and no HE by ΔV, where the arguments are θ, $\Delta\eta$ and w^B.

$$\Delta V(\theta, \Delta\eta, w^B(\cdot)) = \kappa + \Delta\eta + \beta \left\{ E\{w^B(S(a,1)) - w^B(S(a,0))|\theta\} \right\} \tag{1.10}$$

Workers face a discrete choice problem. Their optimal education choice e^* given their beliefs $w^B(\cdot)$ therefore is given as follows.

$$e^*(\theta, \Delta\eta, w^B(\cdot)) = \underset{e \in \{0,1\}}{\arg\max}\, e \times \Delta V(\theta, \Delta\eta, w^B(\cdot)) \tag{1.11}$$

Intuitively, the addition of endogenous education choice under uncertainty leads to education mismatch in two senses. First, the uncertainty that workers face leads to mismatch in the sense that their ex-ante optimal education choices may not line up with ex-post optimal ones. Second, as I will show subsequently, the endogeneity of education choice in a matching model also results in the aggregate level of education under free choice not being socially optimal in general. These two kinds of mismatch are conceptually separate; mismatch in the first sense occurs with choice under uncertainty even without matching labour markets and mismatch in the second occurs with matching markets even without uncertain education choice. However, I show in subsequent simulations that with higher choice uncertainty, the inefficiency of education choice under equilibrium is more severe.

The distribution of skill in the economy

The distribution of skill in the economy is an important determinant of wages in the economy as it enters the wage equation through the matching function, equation 1.3. In this model, the distribution of skill in the economy is dependent on who chooses to go into education. When more workers are educated, there is a greater supply of skilled workers in the economy, which affects the wages that workers receive. Let $p(\theta, \Delta\eta)$ denote a function that maps $\mathbb{R} \times \mathbb{R} \to \{0, 1\}$. Intuitively, this function represents an education profile, which maps workers by their observed grade to whether they attend higher education. Using the law of total probability, we can compute the distribution of skill in the economy using the Law of Total Probability.

$$\begin{aligned} f_S(s) = \int_{\varepsilon \in \mathbb{R}} \int & p(S^{-1}(s,1) + \varepsilon, \Delta\eta) f_A(S^{-1}(s,1)) \left| \frac{dS^{-1}(s,1)}{ds} \right| + \\ & (1 - p(S^{-1}(s,0) + \varepsilon, \Delta\eta)) f_A(S^{-1}(s,0)) \left| \frac{dS^{-1}(s,0)}{ds} \right| dF_{\Delta\eta}(\Delta\eta) dF_\varepsilon(\varepsilon) \end{aligned} \tag{1.12}$$

$$F_S(s) = \int_{-\infty}^{s} f_S(x) dx \tag{1.13}$$

Equilibrium under Rational Expectations

To complete the model and define equilibrium, I first introduce two notions. An education profile p is consistent with wage beliefs w^B if:

$$\forall\theta \in \mathbb{R}, \forall\Delta\eta \in \mathbb{R} : p(\theta, \Delta\eta) = e^*(\theta, \Delta\eta, w^B(\cdot)) \tag{1.14}$$

Second, a wage function w is generated by an education profile p if the education profile implies a skill distribution (by equation 1.12),and thus a matching function (by equation 1.3) which imply w (by equation 1.8). That is to say, the skill distribution $F_S(\cdot)$ is a functional with p as an argument, the matching function $\mu(\cdot)$ is a functional that takes F_S as an argument, and $w(\cdot)$ is a functional that takes $\mu(\cdot)$ as an argument. Thus, this series of functionals give a mapping between each education profile p to a wage function w. For convenience, I denote this mapping of functions by R in the following text.

To complete the model, I need to specify how students generate beliefs about the wage functions they will face on the labour market. This is an area where much research still need to be done. Many authors now realise that it is the perceived return to higher education that matters for studying students' investment decisions (Jensen (2010), Wiswall & Zafar (2014)), but there is still relatively little work on how students form such beliefs about returns. In the absence of more sophisticated theories, I resort to a simple assumption of rational expectations in this paper.

In this context, rational expectations imply that workers are able to anticipate the wages that actually obtain when they are on the labour market and set their beliefs to the predicted wages. If workers have rational expectations, then workers' beliefs w^B, the education profile generated by their optimal education choices p, and the resulting wage structure w have to agree.

Definition 1. *(Equilibrium) Rational expectation equilibrium obtains when workers have beliefs w^{B*}, an education profile p^* obtains, and wages w^* are generated such that:*

1. p^* is consistent with wage beliefs w^{B*},

2. w^* is generated by p^*,

3. and $w^{B*}(s) = w^*(s), \forall s$.

Note that the definition of "generated by" specifies a mapping of education profiles to wage functions, denoted by R. Similarly, the notion of consistency with a set of beliefs maps a function representing wage beliefs w^B to an education profile p. Denote this mapping T. The final condition trivially maps wages to wage beliefs; denote this by S. Then the conditions for equilibrium imply the following mappings:

$$R(p) = w$$

$$T(w^B) = p$$

$$S(w) = w^B$$

Putting these mappings together, the equilibrium condition implies a self-mapping of the wage belief function to itself, and an equilibrium wage belief is a Banach fixed point in this self-map:

$$w^B = S(R(T(w^B))) \tag{1.15}$$

Another way to define equilibrium is thus as follows. It is also possible to analogously define the self-mapping in terms of education profiles p or wages w.

Definition 2. *(An equilibrium wage belief) Let Q denote the set of functions that map $\mathbb{R}_+$ to $\mathbb{R}_+$. A function $w^B \in Q$ is an equilibrium wage belief if it is a fixed point in the self-mapping $w^B = S(R(T(w^B)))$.*

1.3.2 Optimal education investment

Under the assumptions maintained, it turns out that the solution follows a cut-off structure; workers should invest in higher education if their signal and preference for education is 'high enough'. I first discuss the characterisation of the optimal education decision under both the full-information and incomplete-information settings considering only private utility returns. Finally, I describe ex-post regret in this model, for individuals who made the ex-post welfare-reducing choice despite making the ex-ante rational decision, and show that this is entirely characterised by the relevant investment thresholds.

Full-Information Benchmark

Consider the case when ability is fully observable. Define by $e^P(a, \Delta\eta)$ the optimal education choice under full information for a worker with ability a and net preferences $\Delta\eta$. I assume for the subsequent analysis that workers are in rational expectations equilibrium where their beliefs about wages w^B are equal to actual wages.

$$\begin{aligned} e^P(a, \Delta\eta) &= \underset{e\in\{0,1\}}{\arg\max}\, e \times \{\kappa + \Delta\eta + \beta\{w(s(a,1)) - w(s(a,0))\}\} \\ &= \begin{cases} 1 \text{ if } \kappa + \Delta\eta + \beta\{w(s(a,1)) - w(s(a,0))\} \geq 0 \\ 0 \text{ otherwise} \end{cases} \end{aligned} \tag{1.16}$$

Denote for convenience $\Delta w(a) \equiv w(s(a,1)) - w(s(a,0))$. The equation defining optimality also implicitly defines a cut-off value of ability $a^P(\Delta\eta)$ for each level of education preference. We can visualise the optimality condition as a line in a-$\Delta\eta$ space. By implicitly differentiating the condition by $\Delta\eta$, we can further show that this line must be downward-sloping, thus defining a boundary, beyond which workers invest in education and before which they do not. This is visualised in figure 1.5 in the red line.

Proposition 1. *(Optimal private education choice under full-information)*

1. *The optimal education decision under full information is characterised by a cut-off point $a^P(\Delta\eta)$ for each value of $\Delta\eta$, which is the unique point that satisfies the optimality condition $w(s(a,1)) - w(s(a,1)) = -\left\{\frac{\kappa+\Delta\eta}{\beta}\right\}$. This cut-off is decreasing in $\Delta\eta$.*

2. *The optimal education decision under full information $e^P(a, \Delta\eta)$ is increasing in a, and $\Delta\eta$.*

Proof: *See appendix C.*

Imperfect information

Now, I consider my main setting where ability is not observable by the worker. Workers face an optimisation decision where they choose their education based on a signal θ, which is related to initial ability a through the structure specified by equation 1.1. The standard deviation of signal noise, σ_ε, determines how informative the signal is about the worker's underlying ability.

Denote by $e^I(\theta, \Delta\eta)$ the optimal education decision conditional on the signal and education preferences.

$$e^I(\theta, \Delta\eta) = \underset{e\in\{0,1\}}{\arg\max}\, e \times \{\kappa + \Delta\eta + \beta E\{w(s(a,1)) - w(s(a,0))|\theta\}\} \tag{1.17}$$

We can show that the optimal decision $e^I(\theta)$ follows an optimal cut-off point structure under the maintained assumptions.

Proposition 2. *(Partial information case under signals with additive, normally distributed noise)*

1. *Denote by $e^I(\theta, \Delta\eta)$ a function that maps θ and $\Delta\eta$ to the optimal education decision of a worker with signal θ and preferences $\Delta\eta$ who maximises their expected utility under*

imperfect information. Then, conditional on Δeta*,* $e^I(\theta, \Delta\eta)$ *is weakly increasing in* θ *under the maintained assumptions.*

2. *The optimal education decision under imperfect information is characterised by a cut-off signal* $\theta^I(\Delta\eta)$*, which is the point that satisfies equation 1.18 below. This cut-off is decreasing in* $\Delta\eta$*, such that workers with greater preference for higher education require a lower signal to induce them to choose to invest in education.*

$$E\left\{w(s(a,1)) - w(s(a,0))|\theta^P(\Delta\eta)\right\} = -\left\{\frac{\kappa + \Delta\eta}{\beta}\right\} \tag{1.18}$$

Proof: *Direct application of Athey (2002), theorem 2, p. 200. See appendix C for details.*

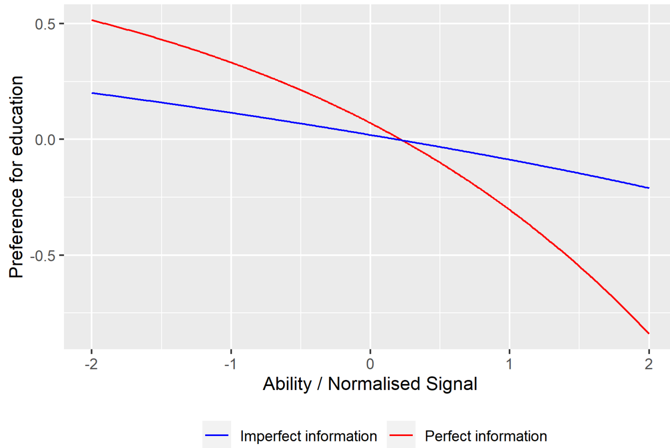
Propositions 1 and 2 can be interpreted as saying that in a two-dimensional representation of worker heterogeneity, we can draw a continuous downward-sloping line that divides the space of workers into those that select into HE and those who do not. Figure 1.5 plots a parameterised version the optimality condition with equality under perfect and imperfect information; as discussed, this condition implicitly defines a decreasing function in $\Delta\eta$ and a space which serves as a boundary for whether students will invest in HE. Students will invest in HE in imperfect information if they are to the right of the blue line, and will not invest in education if they are to the left of the line.

Individual mismatch

In the motivation of this paper, I discussed two notions of education mismatch, individual mismatch and aggregate mismatch. It turns out that the cut-off point structure of the solution to individuals' education decision problems offer a natural definition of individual education mismatch in a way that relates to the potential outcomes framework used in causal analysis. It also relates to the notion of ex-post regret as described in Cunha et al. (2005).

A worker experiences individual education mismatch when they choose a level of education

Figure 1.5: Boundary Lines Describing Optimal Education Investment in Perfect and Imperfect Info



This graph plots a parameterised example optimality condition under perfect information and imperfect information. The parameterisation comes from the empirical implementation, described in subsequent sections.

ex-ante which leads than a worse outcome ex-post.

Definition 3. *(Education mismatch, over-education, under-education) A worker with ability a, signal θ, and education preferences $\Delta\eta$...*

1. *...experiences individual education mismatch if $e^P(a, \Delta\eta) \neq e^I(\theta, \Delta\eta)$.*
2. *...is over-educated if $e^P(a, \Delta\eta) = 0$ and $e^I(\theta, \Delta\eta) = 1$.*
3. *...is under-educated if $e^P(a, \Delta\eta) = 1$ and $e^I(\theta, \Delta\eta) = 0$.*

There are two kinds of workers who are ex-ante rational but ex-post worse off in my model. There are workers who do not select into education when they would have received a benefit ex-post; in this sense, they under-invest in education. There are then the corresponding type of worker who do invest in education when they do not receive a benefit ex-post; these type of workers over-invest in education. This notion of mismatch is consistent with the traditional phenomenon of overeducation. Workers with high signals but low actual ability are more likely to match with lower productivity jobs, more of which are likely to be traditionally considered low-skill. Instead of being the definition of over-education, being in a non-graduate occupation is instead an indicator of being over-educated.

Propositions 1 and 2 show that conditional on a worker's education preferences, their education choice under perfect and imperfect information depends on whether their ability or signal exceeds cutoffs, denoted by $a^P(\Delta\eta)$ and $\theta^I(\Delta\eta)$. Thus, we can redefine education mismatch in terms of these cut-offs. We can also use the cut-offs to computed the shares of workers who are mismatched given the joint distribution of θ and a.

$$\begin{aligned} Pr(\text{overinvest}) &= Pr(a < a^P(\Delta\eta), \theta > \theta^I(\Delta\eta)) \\ &= \int_{-\infty}^{a^P(\Delta\eta)} (1 - F_\varepsilon(\theta^I(\Delta\eta) - a)) f_A(a) da \end{aligned} \tag{1.19}$$

$$\begin{aligned} Pr(\text{underinvest}) &= Pr(a > a^P(\Delta\eta), \theta < \theta^I(\Delta\eta)) \\ &= \int_{a^P(\Delta\eta)}^{\infty} F_\varepsilon(\theta^I(\Delta\eta) - a) f_A(a) da \end{aligned} \tag{1.20}$$

Outside of rational expectations equilibrium, mismatch can also occur if workers' beliefs about the wage schedule that they will face does not align with the actual wage schedule. I abstract from this issue in my empirical application, and leave the quantification of mismatch due to mistaken beliefs to future work.

1.3.3 Socially optimal education investment

Is the education profile that prevails under rational expectations equilibrium efficient? To investigate this question, I formulate a simple version of the planner's problem where the planner chooses an education profile to maximise the sum of worker and firm utility in the economy. It turns out that the decentralised equilibrium is not efficient in general in this sense, and there is aggregate education mismatch in the economy.

Suppose there were a social planner who could choose an education profile, that is, a mapping $p : \mathbb{R}\times\mathbb{R} \to \{0, 1\}$ that specifies a level of education investment for each level of grades and education preference. Denote by $\mathcal{P}$ the set of education profiles p. This planner aims to choose an education profile to maximise the sum of joint output and students' preferences in the economy[19]. The social planner's objective function is as follows, denoted by W. The social planner's problem is to find the education profile $p \in \mathcal{P}$ that maximises $W[p]$. I denote this optimal profile by p^*.

$$\begin{aligned} W[p] = \int_{\Delta\eta}\int_{\varepsilon}\int_{a} \{ & p(a+\varepsilon, \Delta\eta)\left[g(s(a,1), \mu(s(a,1))) + \kappa + \Delta\eta\right] + \\ & (1 - p(a+\varepsilon, \Delta\eta))\left[g(s(a,0), \mu(s(a,0)))\right]\} \, dF(a) dF(\varepsilon) dF(\Delta\eta) \end{aligned} \tag{1.21}$$

Note that besides directly appearing in the equation, $\mu(\cdot)$ is also a functional in p, as p determines the supply of skill in the economy, which affects the match conditional on s. Denote by $\mu(s,p)$ the function with the function p as a secondary argument. By extension, this implies that $g(s, \mu(s))$ is a functional of p.

It is in general difficult to solve the social planner's problem; since the object optimised over is a function, it is not straightforward to use conventional derivative based methods to solve the problem. It is simpler to show however, that the education profile that prevails in equilibrium may not be socially optimal.

Proposition 3. *Suppose the economy is in rational expectations equilibrium, where the education profile in equilibrium is $\bar{p}$. Then, $\bar{p}$ is not the socially optimal education profile, i.e. $\bar{p} \neq p^*$, except in a specific knife-edge case.*

Proof: *See appendix C.*

The basic strategy behind the proof is to redefine the problem so that one could take functional derivatives of W, and show that in general, this derivative at $\bar{p}$ is not equal to 0. Inspecting the first-order condition of the derivative of $\tilde{W}$ at $\bar{\psi}$[20] also explains why equilibrium is not efficient. There are two externalities (corresponding to the two terms in the equation) which lead to inefficiency at equilibrium. First, there is a hold-up externality at equilibrium because the equilibrium in the assignment game is not perfectly competitive. This is because

[19]In my notation, workers' and firms' utility are equally weighted. It is straightforward to account for different weighting of workers and firms, and does not affect the qualitative conclusions below.

[20]See equation 1.32 in appendix C. This text describes the intuition derived from the proof without going into the mathematical detail.

workers and firms on either side of the labour market are not fully substitutable, and thus have a degree of market power. Workers are not able to fully appropriate the surplus that they generate when they invest in education, and as such, underinvest in it relative to the social optimum[21].

Second, an indirect effect of choosing to invest in education is that it makes other workers' relatively lower on the overall skill distribution, which leads them to match to less productive firms. This is apparent in a discrete model; if a worker's education choice leads them to leapfrog another worker on the skill ranking, they match with the firm that those workers matched with and reduce those workers' wages by pushing them to match with less productive firms. I call this a congestion externality[22] or a positional externality, using these terms interchangeably. In making their education choice, workers do not factor in this indirect effect on other workers' wages. Because of this, they are likely to over-invest in education.

In general, the rational expectations equilibrium is thus not socially efficient. I verify this result by showing that it is possible to improve aggregate utility by implementing flat subsidies or taxes in policy simulations. Because there are two counteracting externalities, it is generally ambiguous whether education is too high in my model in aggregate.

1.4 Empirical Implementation

I structurally estimate a parametric version of my model and use it to perform welfare analysis and counterfactual analysis. In this section, I will describe my empirical approach. Section

[21]Note that equation 1.6 does not imply that workers appropriate their whole marginal surplus, which corresponds to the total derivative with respect to s, not the partial derivative. Note that:

$$\frac{dg(s,\mu(s))}{ds} = \frac{\partial g(s,\mu(s))}{\partial s} + \frac{\partial g(s,\mu(s))}{\partial y}\mu'(s)$$

The latter term represents the job-operator's increase in their profits from the workers' increased productivity, which the worker does not fully appropriate.

[22]The interpretation of the congestion externality is different from that in search and matching models. In those models, the congestion externality is the effect of workers' search or firms' vacancy posting on the value of search or vacancy posting. In my model, the congestion effect comes from education choices of some workers squeezing other workers down the skill ranking.

1.4.1 describes a parametric specification of the model described above. Then, I discuss the identification of my parameters and estimation in subsection 1.4.2. Finally, section 1.4.3 describes the data sources I use for my empirical application.

1.4.1 Parameterization

To bring the model to data, I impose a number of parametric assumptions on the exogenous functions and distributions in my model, summarised below in table 1.1. Most of the parameterisations are chosen for simplicity and tractability; I assume that ability is normally distributed, with the mean and variance normalised to 0 and 1[23]. The signal noise is Gaussian, with variance σ_ε^2 which is interpreted as the relative degree of uncertainty that workers face about their true underlying ability. This parameterisation of the prior and the noise is commonly used in the economics of information, and a well-known result is that the resulting posterior distribution, that is the distribution of $a|\theta$, is also normal with a mean and variance that depend on the signal and the signal noise, i.e. $a|\theta \sim N(\frac{\theta}{1+\sigma_\varepsilon^2}, \frac{\sigma_\varepsilon^2}{1+\sigma_\varepsilon^2})$. The normality of the posterior distributions is convenient as it allows the use of quadrature methods to compute objects of interest, particularly the expected return to education conditional on the received signal.

To parameterise the joint output function, I follow the Cobb-Douglas parameterisation chosen by Chade & Lindenlaub (2021), where the joint output is given by $qs^{\gamma_1}y^{\gamma_2}$. This implies that the slope of wages as a function of skill is parameterised by $q(\gamma_1)s^{\gamma_1-1}\mu(s)^{\gamma_2}$. This parameterisation satisfies supermodularity if the parameters q, γ_1 and γ_2 are greater than 0, and if the domain of s and y is restricted to $\mathbb{R}_+$. Because only the shape of the wage function is determined, and not the location, I assume a minimum wage that will be estimated from the data, denoted by w_0.

To parameterise the skill function, which relates underlying ability and education to skill,

[23]This assumption is commonly made, and gave rise to Pigou's paradox, which asks why wages are lognormal when ability is assumed to be symmetric. The seminal contribution of Roy (1951) can be interpreted as a response to this paradox. An interesting discussion of this early work is in Heckman & Sattinger (2015).

I use a simple exponential form, where a worker with skill a has skill $\exp(a)$ if uneducated and $(1+\delta)\exp(a)$ if educated. This satisfies the assumptions made on the skill function in section 1.3, particularly that the marginal increase in skill to ability is greater for educated than non-educated workers. At this point, I note that the parameterisation that was chosen assumes that there are no effects on the occupational match and wages of higher education besides through its effects on skill. This model does not contain signalling (since firms observe worker skill, there is no asymmetric information) and sheep-skin effects.

I also parameterise the distribution of job productivities as a mixture distribution with K components, where K is the number of occupations. The full specification of this distribution involves specifying the weights, a $K \times 1$ vector denoted by $p = \begin{pmatrix} p_1, \cdots, p_K \end{pmatrix}$, and the distribution for each component distribution. I assume that the job productivity within each distribution is log-normal with mean μ_y^k and variance σ_y^k. These moments are collected in the two $K \times 1$ vectors μ_y and σ_y respectively. While this parametric assumption may still be restrictive, I argue that it is an improvement on current approaches in the literature, which simply assumes that each occupation is associated with a particular productivity level[24].

Finally, I assume that preferences for higher education are consist of two components, an aggregate net preference for higher education, κ, and non-systematic preferences for higher education are parameterised by $\eta(1)$ and $\eta(0)$, which I assume to be mean 0 type I extreme value distributions with scale parameter $\frac{1}{\xi}$. The difference $\eta(1)-\eta(0)$ is logistically distributed with location parameter 0 and scale parameter $\frac{1}{\xi}$. For simplicity, I assume β in equation 1.10 is equal to 1. It is not separately identified from the scale of the net preference for education ξ.

There are two aspects of the parameterisation which may be seen as extreme. First, the model starkly predicts that workers with the same ability should receive the same wage, and thus any variation in wages is attributed to the worker's observed signal deviating from their actual ability. A reasonable objection is that there may be sources of variation in observed

[24]This is the case in Chade & Lindenlaub (2021), as well as the occupational mismatch literature (e.g. Guvenen et al. (2020); Lise & Postel-Vinay (2020a)).

Table 1.1: Parametric Specification

No	Object	Notation	Parametric Form	Parameters
	Exogenous functions			
1	Skill function	$s(a,e)$	$\exp(a)(1+\delta e)$	δ
2	Joint output function	$g(s,y)$	$qs^{\gamma 1}y^{\gamma 2}$	$q, \gamma 1, \gamma 2$
	Exogenous distributions			
3	Ability distribution	$f_A(\cdot)$	$N(0,1)$	-
4	Dist. for signal noise	$f_\varepsilon(\cdot)$	$N(0,\sigma_\varepsilon^2)$	σ_ε
5	Demeaned dist. for net heterogeneous educ pref	$f_{\Delta\eta}(\cdot)$	Logistic with loc 0 and scale $\frac{1}{\xi}$	ξ
6	Aggregate preference for HE	κ	-	κ
	Other parameters			
7	Minimum wage	-	-	w_0
8	Job productivity distribution	$f_Y(\cdot), f_Y^k(\cdot)$	$logN(\mu_y^k, (\sigma_y^k)^2)$	μ_y^k, σ_y^k

wages that is not due to this mechanism, particularly data-driven sources of variance like observation error. Unfortunately, adding observation error to the model leads to problems with optimisation, which suggests that the variance of the observation error and the variance of the signal noise is not well identified separately. Thus, my estimates are likely to overstate the true uncertainty in the system if there is significant observation error in the data.

Second, my mechanism for assigning observed occupations to workers assumes that workers are indifferent to the occupation choice beyond the productivity of the job offered. This excludes concerns like preference for particular occupations, e.g. for prestige reasons. Implementation of these elements in my model is unfortunately not straightforward; incorporating these elements requires a more complex matching problem than has been described thus far. As I shall show in the discussion of model fit, this simple assumption seems to fit outcomes for non-graduate workers well but workers seem to have a preference for skilled preferences even when the productivity of the job they match to is the same. Thus, my model captures the dependence of the share of graduate workers in skilled occupations on the observed signal, but misses the level.

1.4.2 Identification and estimation

In this section, I provide a heuristic discussion of the features of the data which help to identify my model. I start by noting that it is possible to construct the distribution of skill in the economy under a particular set of parameters with an estimate of the share of workers who get degrees, conditional on θ, according to equation 1.12, conditional on δ and σ_ε. I assume that I can observe the components of the productivity distribution by observing the occupation shares, and occupation-specific means and variances, p_k, μ_y^k, σ_y^k for all $k \in \{1, \cdots, K\}$. This implies that conditional on parameters δ and σ_ε, I know the distributions F_S and F_Y in the economy. In that case, the identification of the joint output parameters q, γ_1, γ_2 and the minimum wage w_0 is the same as that in Chade & Lindenlaub (2021), and I can use their identification argument in my case. They note that in this case, Ekeland et al. (2004) argues that the shape of the wage function identifies the parameters. Thus, it is possible to identify parameters q, γ_1, γ_2 and the minimum wage w_0 conditional on parameters ε and δ.

Given the joint output parameters, the education technology parameter δ is governed by the wage return to education conditional on a; since a is not observed, the natural counterpart is the wage return conditional on θ. In a straightforward way, if δ is large, there should be a larger gap between the observed wages conditional on θ for $E[w|e = 1, \theta]$ and $E[w|e = 0, \theta]$. The signal noise standard deviation σ_ε is identified by the variance of wages conditional on θ. In the model, wage dispersion conditional on θ is due solely to the variation of a conditional on θ. Thus, if workers with the same signal have significantly varying wages, this must be because the observed signal is fairly weakly correlated with true underlying ability. Conversely, if σ_ε was small, then we should observe that the signal θ is highly predictive of wages, conditional on education. This intuition is illustrated in figure 1.6, which plots scatterplots of log wages against the signal for three simulations using different values of σ_ε, overlaid with each other. Decreasing the value of σ_ε reduces the dispersion of the points around a common relationship. As the value of σ_ε goes towards 0, the R^2 of a non-parametric regression of log wages on the signal and degree status increases. Thus, estimates of the joint output parameters and

w_0 help pin down the education technology parameter δ and the signal noise parameter σ_ε. These parameters in turn determine the empirical skill distribution, which pins down the joint output parameters and w_0.

Figure 1.6: Intuition for the relevance of R^2 in identifying σ_ε

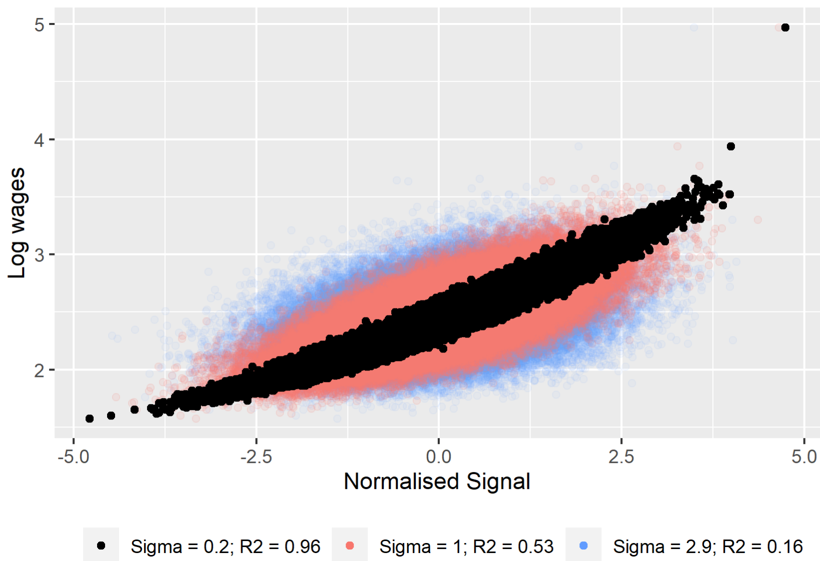


Figure 1.6 plots the scatterplot of log wages against the normalised signal from three simulations of the model, each with a different value of σ_ε. The R^2 is from regressing log wages on a polynomial of degree 10 on the signal, interacted with degree status.

Finally, the parameters κ and ξ cannot be identified without taking a stand on workers' beliefs about their returns to education. To estimate these parameters, I assume that workers have rational expectations, that is, they have knowledge of the parameters of the model and can accurately assess their likely returns in expectation prior to education investment, and that the model is in equilibrium, i.e. that the observed distribution of workers choosing education is consistent with their incentives to pursue education[25]. In this case, since the parameters $\delta, q, \gamma_1, \gamma_2, \sigma_\varepsilon$ and w_0 are sufficient to construct workers' incentives conditional on their signal θ, I can identify ξ and κ using a binomial logit model, especially since I assume that the heterogeneous preference variables $\eta(1), \eta(0)$ are distributed according to extreme value type I.

[25]It is possible that the model is not in equilibrium and thus that a different set of workers would choose education than is actually observed under rational expectations. This would make identification of the preference parameters κ and ξ impossible. However, this cannot be ruled out conclusively ex ante.

This discussion suggests the following procedure for identifying the model parameters:

1. Non-parametrically[26] construct the share of workers who choose higher education in the data conditional on the observed signal θ at a set of points, and denote this $\hat{P}(\theta)$.

2. Estimate the parameters governing wages, the occupational match and the returns to education conditional on $\hat{P}(\theta)$ using simulated method of moments. This recovers the estimated partial parameter vector, $(\hat{\delta}, \hat{q}, \hat{\gamma}_1, \hat{\gamma}_2, \hat{w}_0, \hat{\sigma}_\varepsilon)$.

3. Using $(\hat{\delta}, \hat{q}, \hat{\gamma}_1, \hat{\gamma}_2, \hat{w}_0, \hat{\sigma}_\varepsilon)$ and $\hat{P}(\theta)$, I can calculate the return to education $\Delta(\theta) \equiv E[w(s(a,1)) - w(s(a,0))|\theta]$. Under the assumptions that workers have rational expectations and that the model is in equilibrium, I can identify κ, ξ using a binary logit model, with $\Delta(\theta)$ as the RHS variable.

Estimation of the joint output function, education technology function, and signal correlation

To jointly estimate $\delta, q, \gamma_1, \gamma_2, \sigma_\varepsilon$ and w_0, I use the simulated method of moments (Gourieroux et al. (1996); Adda & Cooper (2003)). The basis of this estimation method is to simulate datasets based on the parameters, and to find a set of parameters that produce moments that best approximates moments computed from the data. Let φ denote the vector of parameters to be estimated, and let $\hat{m}$ denote a vector of n_m targeted moments computed from the data. For a given set of parameter vectors, I simulate the model R times with N observations in each simulation, and compute a vector of equivalent moments from the simulated data. I then average the R vectors over the simulations to derive the average simulated counterparts of the moments in $\hat{m}$, which I denoted by $\tilde{m}(\varphi)$[27].

Define by $q(\varphi) \equiv \hat{m} - \tilde{m}(\varphi)$, which is a $n_m \times 1$ vector. Then the estimator is given as

[26] I do this by essentially constructing a histogram, and taking the mean share of workers with degrees at 19 points across the signal distribution.

[27] I use 50 simulations, each with 2000 simulated observations.

follows:

$$\hat{\varphi} = \arg\min_{\varphi} q(\varphi)' W q(\varphi) \tag{1.22}$$

, where W is a $n_m \times n_m$ positive semi-definite weight matrix. While any choice of W yields consistent estimates of the parameters φ, the optimal weighting matrix is the inverse of the covariance matrix of the moments used, which I compute using the bootstrap. Additional details about computation are available in appendix D.

Using this estimation strategy, it is crucial to choose the moments that can identify the parameters specified. Table 1.2 summarises the moments that I use for my estimation. There are 126 moments in all, in seven categories. The mean log wage within log-wage deciles (categories 1 and 2) help to identify the joint output function by summarising the shape of the wage distribution. The means, especially conditional on education and signal quintile, help to identify δ, which governs the return to education. The variances as well as the quartiles of log wages conditional on education and signal quintile help to identify σ_ε, which governs how good the signal is of underlying ability. Another important moment is the R^2 of regressing log earnings on a high degree polynomial conditional on education status. In a world without noise, the signal is as good as observing ability and thus the R^2 should be 1. As the degree of noise increases (i.e. σ_ε increases), the R^2 declines as the signal no longer perfectly tracks wages conditional on education status. Thus, this moment is important for identifying the degree of noise in the grade signal of ability.

Table 1.2: Moments

No.	Moments category	Number of moments
1	Mean log wage within income deciles	10
2	Mean log wage within income deciles cond. on education status	20
3	Log wage quartiles cond. on signal quintile	20
4	Log wage quartiles cond. on signal quintile and education	40
5	Mean and variance of log wages	2
6	Mean and variance of log wages cond. on education	4
7	Mean and variance of wages cond. on signal quintile	10
8	Mean of wages cond. on signal quintile and education	10
9	R^2 of regressing log earnings on a polynomial of grades conditional on degree	2

Estimation of education preference parameters

Conditional on estimates of δ, σ_ε, the parameters of the joint output function, q, γ_1, γ_2 and w_0 from the estimation procedure described in the previous section, as well as the probability of higher education conditional on the signal, I can construct the expected return to education conditional on the signal, $\Delta(\theta) \equiv E[w(s(a,1)) - w(s(a,0))|\theta]$.

Suppose that workers have rational expectations, and the system is in equilibrium (i.e. the incentives to invest in education align with the observed choices to undergo education). Under this assumption, I then estimate the preference parameters $\psi = \{\kappa, \xi\}$, the location and scale of the extreme value type I distribution of $\eta_1 - \eta_0$, by the Generalised Method of Moments (GMM). The natural moments are the shares of workers with degrees conditional on the worker's signal. Denote by $\hat{\psi}$ the estimated values of ψ. Suppose that $\hat{P}_\theta$ denotes a vector containing the probability of having a degree conditional on $\tilde{\theta}$, where $\tilde{\theta}$ denotes a set of values of θ corresponding to the 5th quantile to the 95th quantile of the distribution of θ. Let $\tilde{P}_\theta(\psi)$ denote the probability of investing in education implied by $E[w(s(a,1)) - w(s(a,0))|\theta]$ and ψ at the points $\tilde{\theta}$. Then, the minimisation problem underlying the estimation is as follows. The standard errors are computed based on the Jacobian of the moments to the parameters (Greene (2003)).

$$\hat{\psi} = \arg\min_{\psi} \left(\tilde{P}_\theta(\psi) - \hat{P}_\theta\right)' W \left(\tilde{P}_\theta(\psi) - \hat{P}_\theta\right) \tag{1.23}$$

There are two reservations with my education choice set-up as it stands. First, this estimation approach does not account for the possibility that preferences for higher education may differ substantially across demographic groups. Suppose that there are G distinct demographic sub-groups, with the generic sub-group indexed by g. Then, a straightforward strategy is to use data on the observed higher education choice shares for each sub-group g, $\hat{P}_g(\theta)$, to estimate sub-group specific choice parameters. While I do not pursue this strategy in this paper for sample size reasons, it is conceptually straightforward to account for this concern in future research with a bigger dataset.

Second, this method does not allow for psychic costs of education that are correlated with workers' grades. This is, for example, in contrast to Chade & Lindenlaub (2021) which attribute all the correlation between ability scores and education choice to such costs. Without additional variation of earnings not due to differences in grades, such psychic costs are not identified in my model[28]. To the extent that this is a big concern, my estimation method would overstate the degree to which workers are responsive to pecuniary incentives in education choice, and understate the role of non-pecuniary factors in HE choice.

1.4.3 Data

To estimate the model, I need two sources of data. First, I need the set of moments that are used to identify the parameters in the model, described in table 1.2, which requires observations on labour market wages, whether the worker has a degree, a measure of their observed ability, and an observation of the occupation they ended up matching to (for analysis of untargeted moments). I use a sample of individuals from the Understanding Society longitudinal study[29]. This dataset allows administrative linkage to data on high school grades in a national exam, the GCSE, which are the grades that students rely on when applying to higher education. I pool observations from individuals born across five years from 1988-1993, who would have been 18 and considering higher education around the years 2006 to 2011[30], and use this as my main sample. For this sample of students, I have data on their signal, whether they attended higher education, a set of controls, and an unbalanced panel of earnings between

[28] A promising method to address this concern may be to use variation across cohorts due to wage variations across the business cycle or changes in tuition fees due to policy. I do not have sufficient observations in this dataset to pursue this strategy and leave this for future research.

[29] Understanding Society (University Of Essex (2022)) is a longitudinal study which follows households over a long period of time, collecting data on a wide variety of topics for people within the household of all ages. It is a successor of the British Household Panel Survey (BHPS). The linkage to administrative education data is provided by Education & University Of Essex (2022).

[30] The government's policy towards higher education funding changed often in the late 1990s and 2000s, and thus I deliberately choose to pool cohorts which faced similar conditions for higher education funding. In this period (from the 2005/06 academic year to 2012/13), students typically borrowed about £6000 per year, under student loan plan 1. These debt details are taken from a government research report, accessed at https://researchbriefings.files.parliament.uk/documents/SN01079/SN01079.pdf.

the ages of 25 to 32.

For wages, I use hourly labour income, net of taxes and transfers. I face two issues which commonly affect empirical analyses of the relation between education choice and the labour market. First, the 'correct' wage measure to consider in calculating a worker's return to education is their total life-time earnings, which is unfortunately not observed in most datasets. Second, wages are typically the influence of many observable factors, a large number of which is not considered in my model. To address both issues, I run a Mincer regression of log earnings on age, grades, and a set of controls[31], on a panel of earnings for my sample. I adjust earnings using the coefficients from the regression to control for the effects of sex, ethnicity and wave fixed effects. I then take the predicted earnings at age 30 as the relevant earnings that I consider in the structural estimation[32]. Considering earnings at age 30 is common in the literature on higher education given the data typically available and practical constraints on collecting data on the whole life-cycle; see for example similar measures in papers including Blundell (2000); Beffy et al. (2012); Delavande & Zafar (2019); Belfield et al. (2018a). It is helpful to consider these age 30 wages as a proxy for lifetime earnings.

As a measure of the signal that individuals consider, I use a variable recording students' performance at their Key Stage 4 exams[33] (Total GCSE/GNVQ new style point score). I normalise this score within the student's academic year; due to attrition, the distribution of this normalised score in my sample is not standard normal. There is a growing literature documenting that grades play a significant role in providing individuals with information on their ability (Tan (2022)) and their comparative advantage in subject choice (Avery et al. (2018); Li & Xia (2022)). Similarly, my set-up proposes that students learn about their relative aptitude for university and their heterogeneous return from higher education from their grades

[31] The controls I include are wave fixed effects, sex, and ethnicity.

[32] For each individual, I compute the mean residual of the regression summarised in table 1.3 over the periods that they are observed for in the panel. I add this mean residual to predicted earnings conditional on their grades predicted at age 30.

[33] In the UK, the education system is based on five key stages. Key Stage 4 refers to two years typically called Years 10 and 11, when students are between 14-16. This stage is capped by National Examinations, most typically the GCSE (General Certificate of Secondary Education), although other vocational qualifications are available. The point score used in this paper is a system used to encode students' performance in the range of exams into a numerical point score.

Table 1.3: Results of a Mincer regression on sample

	(1)
	Log net hourly labour income
Degree=1	0.0721***
	(3.43)
Normalised KS4 score	0.0752***
	(6.02)
Degree x Normalised KS4 score	0.0462
	(1.86)
Female	-0.0781***
	(-3.95)
Age	0.0363***
	(5.96)
Observations	3122
Individuals	1128

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

This table presents results of a Mincer-style regression of log earnings on degree status interacted with KS4 scores, age, gender, ethnicity and year fixed effects, estimated on an unbalanced panel dataset of workers born between 1988-93. The standard errors were clustered at the individual level.

at GCSE. An important issue is that the researcher does not observe the individual's true information set and thus may be liable to overestimate the degree of uncertainty individuals face[34]. Workers may have further private information about their returns than is revealed through grades. I leave a satisfactory resolution of this important issue to future work and note that if private information is indeed significant, my results represent an overstatement of the degree of individual uncertainty.

I thus construct a dataset comprising 1113 observations containing three variables: whether workers have a degree at age 32, their adjusted log wage at age 30, and their normalised KS4 grade, which I interpret as the worker's signal of their underlying return. The summary statistics for the dataset from which the moments are computed are tabulated in table 2.1.

Table 1.4: Summary statistics

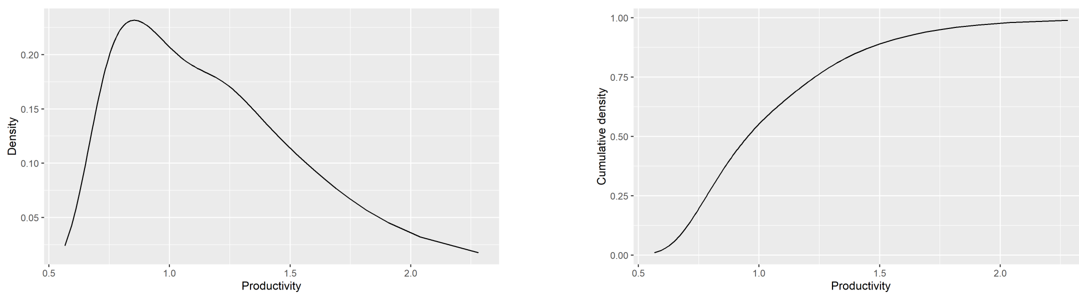
Variable	N	Mean	Sd
Log hourly labour earnings net of taxes and transfers	1113	2.43	0.26
New style KS4 point score, normalised within student's academic year	1113	0.22	0.91
Whether worker has a degree by age 32	1113	0.49	0.50
Female	1113	0.53	0.50
Non-white ethnicity	1113	0.27	0.44
In 1988 birth cohort	1113	0.17	0.38
In 1989 birth cohort	1113	0.20	0.40
In 1990 birth cohort	1113	0.18	0.39
In 1991 birth cohort	1113	0.16	0.36
In 1992 birth cohort	1113	0.16	0.36
In 1993 birth cohort	1113	0.13	0.34

Second, I need data on the distribution of the productivity of jobs that workers match to, that is, the distribution of $f_Y(\cdot)$. As a proxy for productivity, I use the distribution of firm fixed effects within an occupation, calculated from a two-way fixed effect regression of log wages on individual worker and firm-occupation pair fixed effects. I assume that within each occupation (at the SOC00 3 digit level), the distribution of firm productivity is log-normal, with the mean and variance of the log productivity given by the mean and variance

[34]In Cunha et al. (2005); Cunha & Heckman (2016), Cunha and Heckman discuss an empirical strategy for estimating workers' information sets using theoretical restrictions from the permanent income hypothesis and data on consumption. Another approach which is theoretically less onerous is to ask respondents in surveys about their expectations over outcomes, and to compare it against realised outcomes in the future. Unfortunately, these strategies are not available with the data I have.

of occupation-firm pair fixed effects within the firm. More details about the computation of the fixed effects can be found in Hou & Milsom (2021). The effective distribution of firm productivity is plotted in figure 1.7. There is precedence for interpreting these fixed effects as informative about productivity (e.g. in Card et al. (2018a)); Hou & Milsom (2021) document that these fixed effects are correlated with firm revenue per worker.

Figure 1.7: Productivity distribution in empirical exercise



(a) Productivity density

(b) Productivity cumulative density

Panel 1.7a plots the density of productivity used in the empirical exercise, while panel 1.7b plots the cumulative density function of productivity used in the empirical exercise.

1.5 Empirical Results

Table 1.5 summarises the headline parameter estimates that I obtained from my estimation procedure. These results are conditional on a non-parametrically estimated probability of investment in education function conditional on θ, $\hat{P}(\theta)$.

Table 1.5: Parameter Estimates

No.	Parameter	Notation	Value	SE
Stage 1				
1	Signal noise	σ_ε	2.92	0.0000846
2	Skill return to education	δ	0.364	0.00269
3	Joint output function scale	q	7.83	0.172
4	Joint output function - exponent on s	γ_1	0.344	0.00713
5	Joint output function - exponent on y	γ_2	0.0318	0.0226
6	Minimum wage	w_0	4.46	0.0225
Stage 2				
7	Location parameter of het pref for educ relative to no educ	κ	-11.0	0.00447
8	Scale parameter of het pref for educ/no educ	ξ	11.6	0.00417

In subsection 1.5.1, I discuss the estimation of the variance of signal noise, the parameter governing the education technology function, the parameters of the joint output function, and the constant of the wage function $(\hat{\delta}, \hat{q}, \hat{\gamma}_1, \hat{\gamma}_2, \hat{w}_0, \hat{\sigma}_\varepsilon)$. I loosely refer to this group of parameters as labour market parameters. In subsection 1.5.2, I discuss the empirical education investment function in θ, $\hat{P}(\theta)$, as well as the estimation of κ and ξ, which I refer to as the choice parameters.

1.5.1 Estimated labour market parameters and goodness-of-fit

The main feature of these estimates is that the estimate of σ_ε is large, implying that the correlation between actual ability and observed grades is 0.324. As I shall show in the subsequent welfare analysis, this implies that there will be substantial scope for actual returns to education to differ from students' expected returns to education.

The other parameters are not as naturally interpretable. Instead, to evaluate how well the model performs in fitting the data, I analyse how well the simulated moments match the actual moments of the data. Table 1.6 presents the means and variances of log wages, overall as well as conditional on whether the worker has a degree, for both the actual data and for the simulated outcomes from my model. The simulated means match the actual moments very well. The simulated variances are still within the confidence intervals of the estimated data variances, although my model implies that the log wage variance of graduates is higher than the log wage variance of non-graduates. This is not what I find in this dataset, but is in line with evidence from other papers like Lemieux (2006).

Finally, I plot the R^2 of a regression of log earnings on grades for the subsets of graduates and non-graduates. In my model, if grades are perfect signals of ability, the share of variation captured by grades should be 1, and if grades are non-informative about ability at all, the share of variation captured should be 0. The data suggests that the R^2 of this regression is 0.125 for graduates and 0.0659 for non-graduates, while my model predicts a value of around

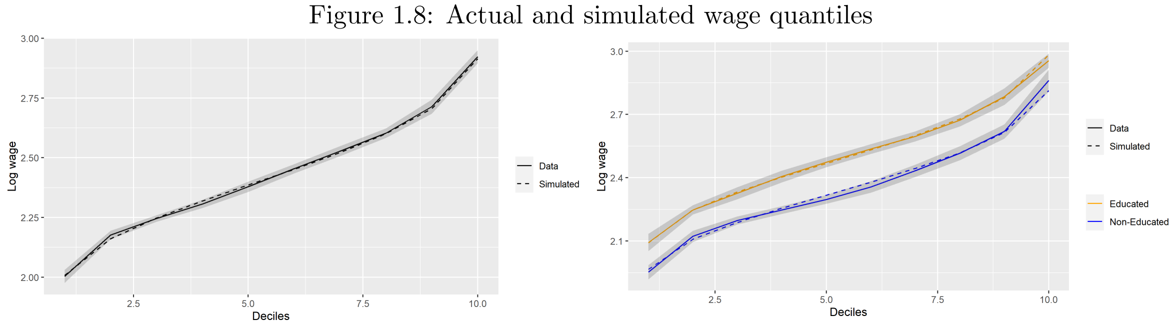
0.09 for both groups. While these simulations lie within the 95% confidence intervals, it seems likely that the lower R^2 for non-graduates than graduates is a feature that is not captured by my model (e.g. through multi-dimensional skills).

Table 1.6: Targeted log wage moments, overall and conditional on degree

Statistic	Data	Conf Interval	Simulated
Mean log wage	2.43	[2.42,2.45]	2.43
Mean log wage (e=1)	2.51	[2.49,2.53]	2.51
Mean log wage (e=0)	2.36	[2.34,2.38]	2.36
Variance log wage	0.0688	[0.0634,0.0743]	0.0675
Variance log wage (e=1)	0.062302	[0.0553,0.0693]	0.0652
Variance log wage (e=0)	0.064314	[0.0566,0.0720]	0.0588
R^2 of regressing log wages on grades (e=1)	0.125	[0.0737,0.176]	0.0917
R^2 of regressing log wages on grades (e=0)	0.0659	[0.0251,0.107]	0.0964

Panel 1.8 plots the actual and simulated log-wage quantiles, while panel 1.8b plots the quantiles conditional on education. My model is able to achieve a good match to the wage quantiles, both overall and conditional on education status.

Figure 1.8: Actual and simulated wage quantiles

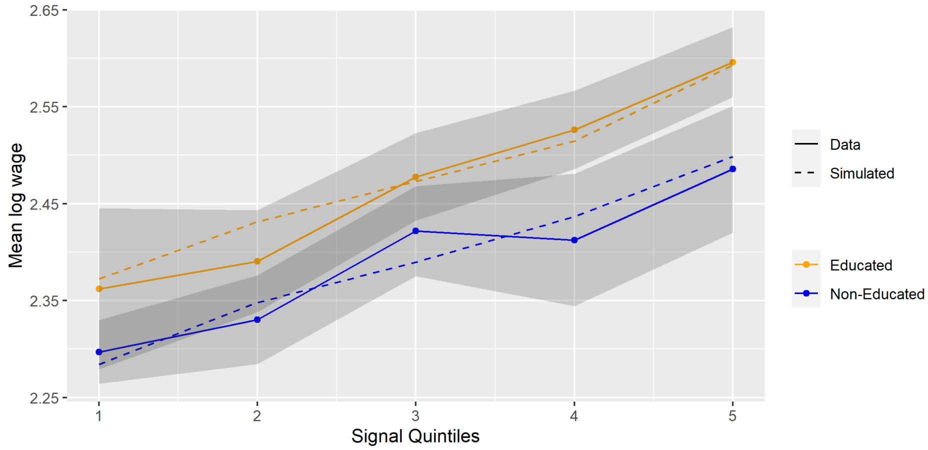


(a) Simulated and actual wage quantiles (b) Simulated and actual wage quantiles cond. on educ

Panel 1.8a plots the mean log wage within each wage decile in the data in the solid line and the simulated equivalents in the dashed line. Panel 1.8b plots the actual and simulated wages for educated workers (in orange) and non-educated workers (in blue). The shaded ribbon plots the confidence intervals of the estimates from the data.

Furthermore, my model is able to capture the interdependence between grades, education status and the individual's wage. Figure 1.9 plots the mean wage conditional on a worker's signal quintile and whether they are educated. Again, my model fits the data relatively well. In appendix E, I plot the fit of the simulated model to means of log earnings within four quartiles conditional on signal quintile and degree status, and show that the model is able to fit even those more detailed moments well.

Figure 1.9: Actual and simulated mean wages conditional on education and signal quintile



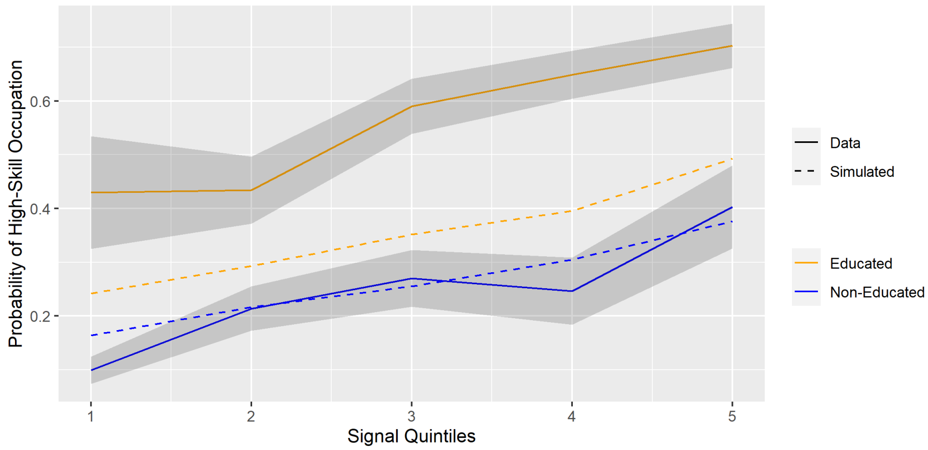
This figure plots the mean log hourly earnings conditional on education (non-graduates in blue and graduates in orange) and five grade quintiles for both the actual data moments (in the solid line) and the simulated data moments (in the dashed line). The shaded ribbon represents the 95% confidence interval for the data moments.

Untargeted moments

Finally, my model makes predictions about the probability of a worker being in particular occupations, conditional on their signal quintile and education. Given a worker's rank in the skill distribution post-education, it specifies a job productivity that the worker matches to. Conditional on this job productivity, a worker randomly matches to an occupation, with the probability of being in occupation k conditional on job productivity s given by equation 1.28. I then classify these SOC00 3 digit occupations into high skill/low skill categories based on a classification used by the Home Office for immigration guidance purposes.

To analyse how well my model fits the data on the matching of workers to occupations, I plot the share of workers in high-skill occupations, conditional on degree and five signal quintiles, in figure 1.10. While my model is imperfect in capturing the levels of graduates in high-skill occupations, it captures the qualitative facts that there are substantial shares of workers in both education groups in high-skill jobs and that the share of workers in high-skill jobs increases with the observed signal for both groups. This lends support to the modelling approach that there is an underlying skill index that determines both wages and the occupa-

Figure 1.10: Actual and simulated probabilities of matching to high-skill occupations

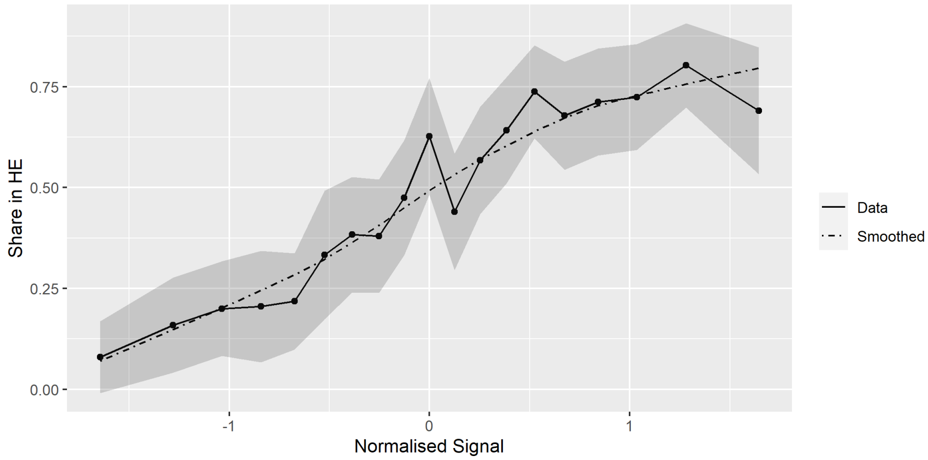


This figure plots the probability of being in a high-skill occupation conditional on education (non-graduates in blue and graduates in orange) and signal quintiles for both the actual data moments (in the solid line) and the simulated data moments (in the dashed line). The ribbons give the 95% confidence intervals for the data moments.

tional match, of which the degree and grades are only indications. However, the significant mismatch in the level of graduates in high-skill jobs suggests that my characterisation does not fully capture sorting into occupations.

An interesting element to this failure to match the high-skill probability of graduates is that while my model underestimates the probability of working in a high-skill occupation for graduates and overestimates the probability for non-graduates, it captures the wages of both graduates and non-graduates quite well. This suggests that the failure in modelling the match to occupations does not affect the model's specification of wages. This could imply that while the model specifies the right productivity of the job that both graduates and non-graduates match to, it does not correctly specify how graduates choose the occupation conditional on the matched job productivity. This could be indicative that graduates may have non-pecuniary preferences for occupations classified as high-skill (e.g. if they are white collar occupations), even when they are remunerated at the same rate as low-skilled occupations.

Figure 1.11: Estimated share of workers with degrees conditional on θ



This figure plots the share of workers with degrees conditional on θ at 19 points across the signal distribution. These shares were calculated by averaging the share in a 0.05 interval around each point, corresponding to the 5th to 95th quantiles of θ. The confidence intervals use standard errors calculated by the bootstrap. The blue mixed line represents a smoothed approximation of the points, calculated by locally weighted smoothing, which is used to compute the distribution of skill in the estimation.

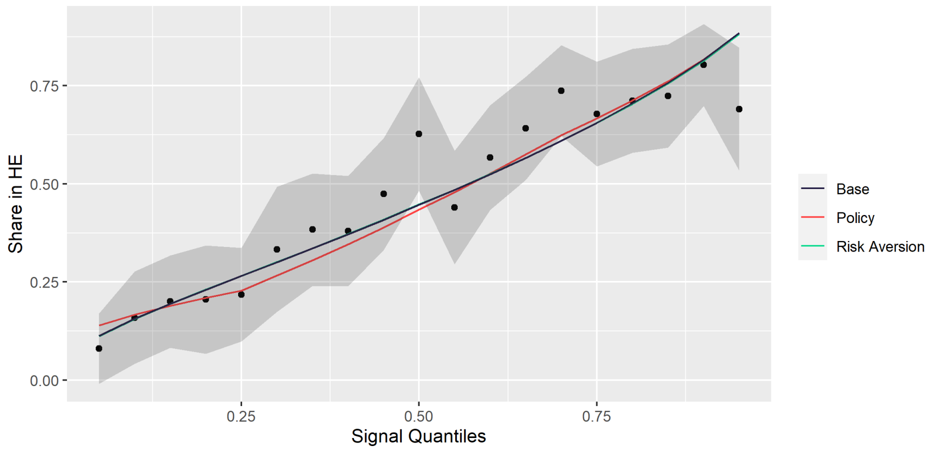
1.5.2 Estimated choice parameters

Figure 1.11 plots the shares of students who chose to go to higher education within a local region around 19 points across the signal distribution[35]. I compute standard errors by the bootstrap. To summarise this relation between probability of choosing higher education and θ as a function, I smooth the observed points using local polynomial smoothing. As expected, workers are more likely to invest in education if they receive a higher grade, with choice probabilities ranging from 8.0% to 80.3% over the grade distribution.

I follow the procedure set out in section 1.4.2 to estimate the parameters κ and ξ. The parameters are summarised in table 1.5.The fit of the conditional choice probabilities to the observed data is shown in figure 1.12 below, in the black line. The simple base choice model achieves a good fit to most points of the data, besides the point at the median signal and the point at the 95 signal percentile.

[35]For example, the first point summarises the share in higher education from the 2.5th to 7.5th quantile of the signal.

Figure 1.12: Predicted probabilities of investing education conditional on θ



This figure plots the share of workers with degrees conditional on θ at 19 points across the signal distribution. The confidence intervals (in dashed black lines) use standard errors calculated by the bootstrap. The mixed line represents a smoothed approximation of the points, calculated by locally weighted smoothing, which is used to compute the distribution of skill in the estimation. The blue solid line plots the predicted probability of investment in higher education conditional on θ considering the income-contingent loan policy, and the red solid line plots the predicted probability of higher education choice conditional on θ neglecting the income-contingent loan policy, using estimates in table 1.7.

I now turn to considering two extensions to this simple model of education choice. First, in this paper, I have abstracted from risk aversion in my discussion. As a robustness check, I then estimate a version of the choice set-up when workers have CRRA utility as follows, where w denotes their wage:

$$u(w) = \begin{cases} \frac{w^{1-\zeta}-1}{1-\zeta} \text{ if } \zeta \in [0,1) \cup (1,\infty) \\ \log(w) \text{ if } \zeta = 1 \end{cases} \tag{1.24}$$

The risk aversion parameter is identified by the relative concavity of the log-odds function in θ relative to the expected returns function to θ. If the log-odds function is highly concave when the expected returns function is convex, then, this suggests that workers' choices are driven by a degree of risk aversion. The coefficient of risk aversion parameter is estimated as part of the GMM estimation described before. The results are presented in column 3 of table 1.7. I estimate a very low degree of risk aversion, consistent with the fairly non-concave shape of the conditional choice probabilities conditional on grades. Although the minimised value is lower than it was for the base model, the improvements are fairly minor. I plot the predicted probabilities of the model with CRRA utility in green in figure 1.12. The line coincides basically perfectly with the black line until the end, where it does better in fitting the points at the top end of the signal distribution.

Table 1.7: Choice Parameters under Various Specifications

Parameter	Base	Policy	Risk Aversion
ξ	11.5	24.2	17.4
	(0.00447)	(0.00948)	(0.440)
κ	-11.0	-16.2	-12.0
	(0.00417)	(0.00626)	(0.0717)
ζ			0.127
			(0.00765)
Minimised Value	600.	719.	582.

Second, students typically pay for higher education in the UK with an income-contingent loan, where they are not required to repay the loan in full if their income is not sufficiently high. This offers a significant level of insurance to workers and is very generous for workers with lower grades, who have to pay a much lower expected tuition. This changes the incentives that workers face; in appendix F, I describe further how the implementation of the tuition policy changes workers' incentives under the estimated parameters. I estimate a set of parameters where workers face expected returns that would prevail under the tuition policy. These estimates are presented in column 2 of table 1.7 above. The implied choice probabilities are also plotted on figure 1.12 in a red line. The minimised value suggests that this model is a substantially worse fit to the moments, driven by a bad fit for the middle of the signal distribution. This may suggest that students do not fully internalise the generosity of the income-contingent loan, especially for workers with lower grades.

1.6 Welfare costs of individual education mismatch

As section 1.3.2 describes, my model implies that imperfect information about ability can lead workers to make ex-ante optimal decisions which nevertheless lead to outcomes that are regretted ex-post. This gives rise to individual mismatch in education choices. In this section, I use the estimates from my structural estimation to analyse (i) whether individual mismatch is a prevalent problem, and (ii) how significant the utility losses from individual mismatch are.

As a measure of welfare, I use the difference between the value a worker places on investing in education and not investing in education measured in terms of foregone pounds per hour earnings (at age 30). Intuitively, this measure can be thought of as the willingness to pay (WTP) for an investment in education; that is the amount in pounds per hour at age 30 equivalent that an uneducated worker is willing to give up to be in higher education (on top of any other tuition or psychic costs already captured by the choice parameters). Conversely, a worker with higher education has to be compensated up to their WTP to be induced out of higher education (on top of being refunded any tuition or psychic costs). A worker's willingness to pay would depend on their net preferences for education $\Delta\eta$ and either their ability a under perfect information or their signal θ under imperfect information. A worker with a positive willingness to pay would select into higher education, and vice versa[36].

With the estimated parameters of my model described in the previous section, I can simulate the joint distribution of WTP under both perfect information and imperfect information. In table 1.8, I summarise the shares of the simulated population in each of the four quadrants: workers who invest and would benefit, those who invest but would not benefit (over-education), those who do not invest but would have benefited (under-education), and those who do not invest and would not have benefited. The noise around individuals' ability implies significant levels of individual mismatch; 39.1% of graduates are individually overeducated in the sense that they go to university despite a net negative return and 27.6% of

[36]The WTP under perfect information is the difference between the wages under education and no education, and the difference between heterogeneous preferences for education and no education. The WTP under imperfect information is the difference between the value functions.

$$\text{WTP}_{\text{perf}}(a, \Delta\eta) = \beta\left(w(s(a,1)) - w(s(a,0))\right) + \kappa + \Delta(\eta) \tag{1.25}$$

$$\text{WTP}_{\text{imperf}}(\theta, \Delta\eta) = V(\theta, 1) - V(\theta, 0) \tag{1.26}$$

Table 1.8: Shares of workers by education choice and whether they receive a net benefit

	Would invest	Would not invest	Total
Would benefit	0.286	0.147	0.433
Would not benefit	0.182	0.385	0.567
Total	0.468	0.532	1.00

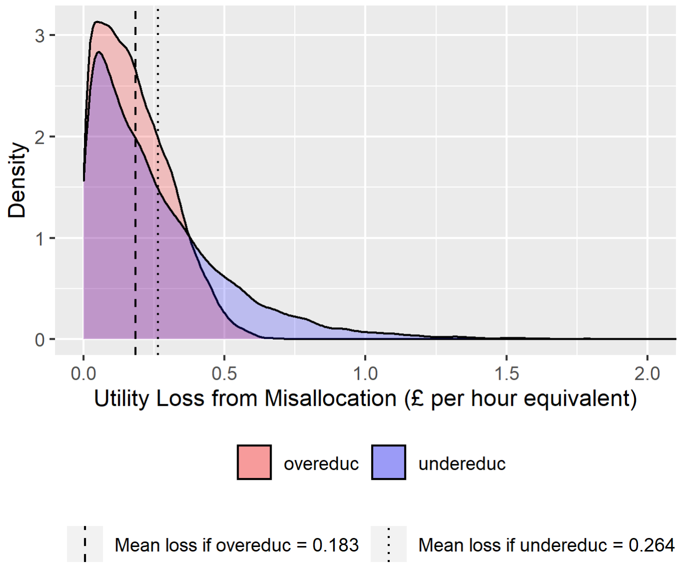
This table summarises the simulated population by whether they would invest in higher education and whether they would experience a net positive utility from doing so ex-post. The bottom-left quadrant shows the share of workers who are ex-post overeducated, while the top-right quadrant shows the share of workers who are ex-post undereducated.

non-graduates are individually undereducated in the sense that they could have benefited from university if they had gone. In total, 32.9% of the population would have been better off if they made a different education choice.

How costly is the misallocation of higher education? In figure 1.13, I plot the distribution of welfare losses in terms of willingness to pay, conditional on being over-educated and being under-educated. In general, the losses are small and right-skewed, in large part because they are censored to the left. Small deviations between willingness to pay under the perfect information and imperfect information scenarios may not be sufficient to induce the worker to change their education choice, and thus do not lead to misallocation. The mean utility loss in expected return terms comes out to be £0.183 per hour (£323.35 per year) conditional on being over-educated, and £0.264 per hour (£465.22 per year) conditional on being under-educated.

The mean loss from being over-educated is smaller than the mean loss from being under-educated, which suggests that while over-education is a more salient problem because one can observe graduates matched to low-skill jobs in the data, under-education is on average more costly to those who experience it. This is driven by long right tail of utility loss conditional on being under-educated; table 1.9 below summarises the quantiles of the utility loss in £ per hour money-metric terms. Intuitively, this is driven by the convexity of returns to education with regards to ability. An under-educated worker is a high ability worker who mistakenly thinks that they are low ability, and the counterfactual is that they may miss out on a large return higher education as a result. On the other hand, a over-educated worker is a low ability

Figure 1.13: Distribution of welfare loss due to misallocation



This figure plots the distribution of the utility loss from misallocation, conditional on being over- and under-educated, from a simulated dataset with the estimated parameters from the structural estimation exercise. The results are presented in terms of willingness to pay, in pounds per hour terms.

worker who mistakenly thinks that he is high ability; their counterfactual earnings without education investment is also relatively low, and the degree of loss from mismatch is capped by the low counterfactual earnings. This asymmetry of potential gains and losses lead workers to err in favour of investing in education; thus, the share of over-educated workers exceeds the share of under-educated workers.

Table 1.9: Quantiles of WTP loss due to misallocation (£ per hour equivalent)

	25%	50%	75%	90%
Overeducated	0.0797	0.163	0.269	0.365
Undereducated	0.0829	0.194	0.368	0.586

This table summarises four quantiles of utility loss for over- and under-educated workers from a simulated dataset with the estimated parameters from the structural estimation exercise. The results are presented in terms of willingness to pay, in pounds per hour terms.

There are no directly analogous results to mine in the literature, but this result falls within a plausible range implied by other studies on the UK higher education system. Waltmann et al. (2020) uses an administrative dataset including tax data and education records to estimate lifetime returns to education and find that about 20% of students would experience negative lifetime pecuniary returns to education. They also find that the government is expected to make a loss on the degrees of 40% of men and 50% of women due to the UK's income contingent loan scheme. My finding that 39.1% of graduates do not benefit from university in aggregate falls within the ballpark of their results. In appendix G, I consider the share of workers who receive a 'sufficiently' high monetary return from higher education; this metric is popular in policy circles interested in ensuring that students' education decisions pass a cost-benefit analysis. I find that particularly in this period (when tuition fees were low at roughly 3000 pounds per year), the cost of education is very low relative to the returns and most workers pass a basic value for money check ex-post (23.2% do not ex-post) and almost all workers do ex-ante (only 0.3% do not ex-ante).

1.7 Policy counterfactuals

In the previous sections, I have showed how individual and aggregate mismatch can arise in a model with education decisions under uncertainty in a labour market characterised by matching. In this section, I perform a number of counterfactual policy experiments to numerically compute the welfare effects of a number of possible policies to affect the education profile. Are there any policy instruments which can reduce the welfare costs of mismatch and increase aggregate welfare? First, I consider the effects of flat graduate subsidies/taxes which alter workers' incentives to engage in HE, and document how that affects aggregate welfare and the incidence of individual mismatch. Second, I consider an abstract policy to decrease the relative noise in workers' signal of their prospective return to HE, and show that unexpectedly, the policy is not Pareto-improving; while it makes workers unambiguously better off, it can make firms worse off by making skilled workers more expensive.

1.7.1 Graduate taxes and subsidies

A simple policy instrument to affect the share of students selecting into university is a graduate subsidy or tax, which can be made revenue-neutral by a corresponding lump-sum tax or subsidy on the whole population. Because the lump-sum compensating tax or subsidy would apply regardless of education status, the compensation does not affect the incentives to invest in education. Denote the flat tax by τ, where a negative value implies a subsidy. The policy then modifies the value function for a worker with signal θ and education e as follows:

$$V^*(\theta, e) = \left((\kappa - \underbrace{\tau}_{\text{Graduate tax/subsidy}}) \times e \right) + \eta(e) + \underbrace{\left[\int P(\theta) d\theta \right] \tau}_{\text{Compensation}} + \beta E\{w(s(a, e))|\theta\} \quad (1.27)$$

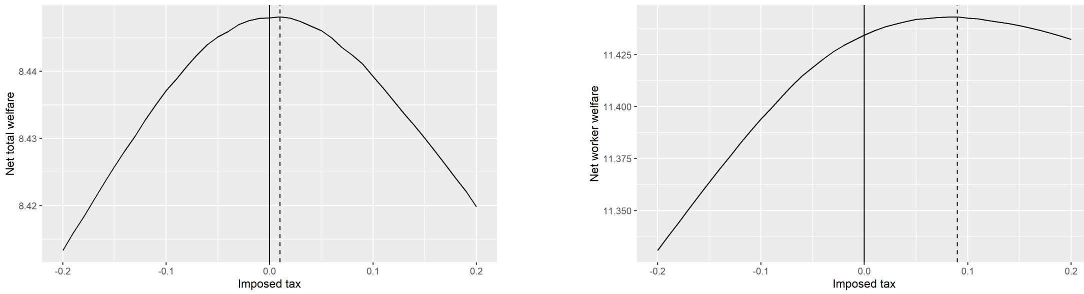
The tax/subsidy is thus a revenue-neutral method of shifting the aggregate preference for higher education, and can be equivalently thought of as a tool to achieve an optimal level of higher education in the economy. Each possible tax level shifts the economy into a different rational expectations equilibrium, which I compute by iteration. In each resulting equilibrium, I document the incidence of individual mismatch and the aggregate utility for workers and firms.

Figure 1.14 plots the total net welfare and the net worker welfare for a range of possible tax and subsidy values. I find that quantitatively, the optimal policy for total welfare is to instate a small tax which will slightly reduce the level of education from 46.2% to 44.6%. The optimal policy from a worker welfare perspective is to instate a relatively larger tax[37] that will reduce higher education substantially, from 46.2% to 28.6%.

Qualitatively, the striking result in this case is that what is best in aggregate is not what is best for workers. In the model, over-education is bad for workers but good for firms; they are able to take advantage of the greater stock of skill in the economy to reduce their wages for all skilled workers in equilibrium. Workers on the other hand are better off on average if fewer workers chose to invest in education precisely because skilled work would be more scarce and they would be able to receive higher wages. Thus, the graduate tax or subsidy has to trade off between the welfare of workers and firms.

[37] This is accomplished by imposing a tax equivalent to approx 9 p per hour at age 30, or adding 3315.36 pounds to the initial cost of education (amortised over 30 years at a 2.5% interest rate).

Figure 1.14: Net welfare under different compensated tax levels

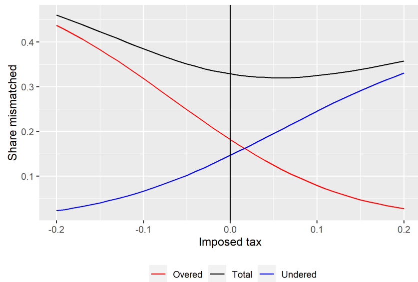


(a) Net total welfare (b) Net worker welfare

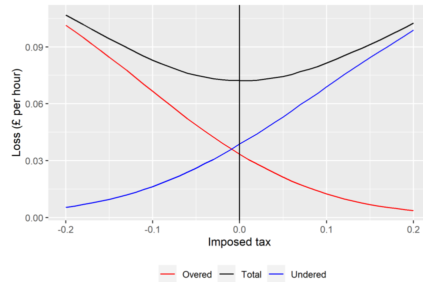
This figure plots the simulated counterfactual net welfare and net worker welfare under various levels of compensated tax. The welfare loss is expressed in willingness to pay. Figure 1.14a considers wages and utility from higher education, while figure 1.14b also includes average profits. The dashed vertical lines denote the utility-maximising value of the tax. The optimal tax is 0.00475 in figure 1.14a and 0.0864 in figure 1.14b.

Is it possible to reduce the incidence and welfare cost of mismatch through the proposed compensated tax scheme? I simulate the welfare effects of a tax for a range of values between -0.2 and 0.2, and consider minimising the number of workers affected by mismatch, or the total loss in WTP terms of mismatch. Figure 1.15a plots the share of mismatched workers under various compensated tax levels, while figure 1.15b plots the total welfare loss in WTP terms under various such levels. The optimal policy if we are aiming to minimise those affected is a tax of 0.06 pounds per hour, reducing the share mismatched from 32.9% to 31.9% under the no-policy baseline. However, this increases the average loss of the mismatch by 4% from 0.0723 pounds per hour to 0.0755. Intuitively, this follows from the earlier observation that under-education is more costly on average than over-education, and thus there are more over-educated than under-educated workers under ex-ante utility maximisation. Thus, the optimal tax in this situation promotes costly under-education (from 14.7% to 20.5%) to reduce the share of over-educated workers (from 18.2% to 11.4%). The optimal tax to minimise the degree of welfare loss in WTP is 0.01 pounds per hour, which is close to the baseline.

Figure 1.15: Degree of mismatch under different compensated tax levels



(a) Share affected



(b) WTP Loss

This figure plots the share affected by mismatch and the welfare loss due to the information friction under various levels of compensated tax. The welfare loss is expressed in willingness to pay. The black line plots total loss, and the blue and red lines respectively plot the loss due to undereducation and overeducation.

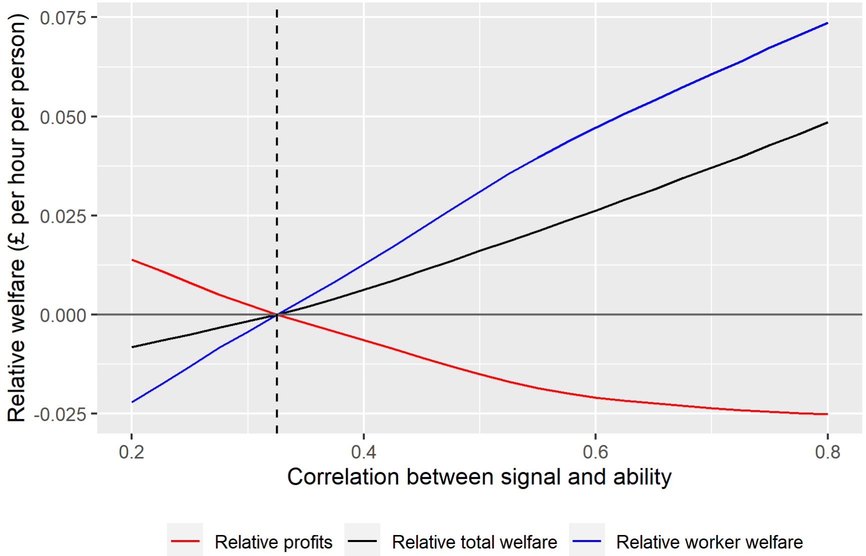
1.7.2 Reducing worker uncertainty

An implication of my theoretical model is that the quality of information that workers have about their return to education is important in determining the extent of individual ex-post mismatch in education choices. If the signal workers receive about their labour market returns was more correlated to their actual labour market return, there would be less individual mismatch in the model. Policy makers may be able to increase the information available to workers about their likely return to education. For example, a policy maker might be able to achieve this by offering more standardised tests to students for them to determine their relative ranking in a cohort, or changing the timing of tests so that workers have access to information before they make important education decisions[38]. More radically, governments may have extensive data on the relation between background and grades, and labour market performance, and can offer data-based recommendations about whether students should go to university (e.g. Athey & Wager (2021)). To my knowledge, there have been few causal studies on policy actions that governments can take to increase the quality of information that workers have about their education decisions.

At present, I consider in an abstract way the value of information about their returns for

[38] In the UK context, students learn the outcome of their age-18 examinations only after they apply and receive offers for university study.

Figure 1.16: Relative changes in welfare as workers' information improves

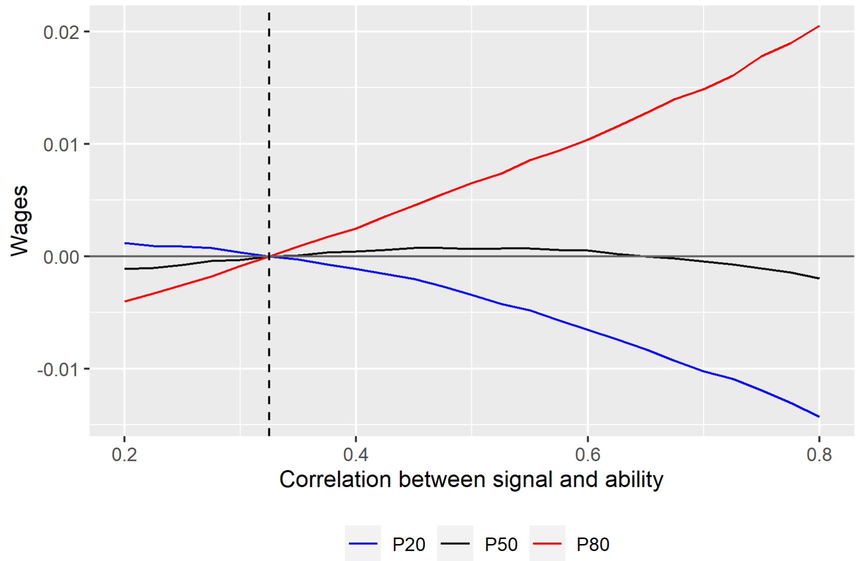


This figure plots the relative change in average worker utility, average firm profits, and the sum of the two for 25 values of the implied correlation between labour market ability and the observed signal. For each value, I iterate the economy towards the rational expectations equilibrium, and simulate the resulting welfare with 50 repetitions with 2000 observations per draw.

workers and in aggregate. I consider the resulting rational expectations equilibrium for 25 economies, where all the parameters were as estimated but the degree of noise in the signal σ_ε implies correlations with unobserved ability of 0.2 to 0.8. For each economy, I simulate the equilibrium economies, and plot the average worker utility, firm profit and total utility, normalising the value at baseline to 0. The results from this exercise are presented in figure 1.16.

I find that worker welfare increases as the degree of correlation increases. This is to be expected in some ways as workers in these economies are able to make better decisions and experience less ex-post regret. However, this does not mean that all workers receive higher wages. Rather, wage inequality increases in this economy, in part because the distribution of skill becomes more polarised. With weak signals of ability, the ability distribution of workers who select into higher education is close to the ability distribution of workers who do not. However, if signals were very informative about ability, the former distribution is more right-skewed than the latter one, increasing the dispersion of skill in the economy. Figure 1.17 plots

Figure 1.17: Relative changes in workers' wages as workers' information improves



This figure plots the relative percent change in average wages for the 20th, 50th, and 80th percentile workers. For each value, I iterate the economy towards the rational expectations equilibrium, and simulate the resulting welfare with 50 repetitions with 2000 observations per draw.

the wages for the 20th, 50th and 80th percentile workers as the information that workers have becomes more accurate. 20th percentile wages fall, and 80th wages increase as information for workers improve.

Interestingly, overall welfare increases at a slower rate than worker welfare, because firm profits actually decrease as uncertainty decreases. This is for two possible reasons. First, because the wage function is convex in skill, the average share of workers who select into education actually decreases as information improves. The decrease in the share of workers leads to firms, particularly on the lower end of the distribution, facing a less skilled labour market. Second, firms have to pay more to recruit skilled workers with more bargaining power since there are fewer medium skilled workers in the economy as the distribution of skill polarises. These factors lead to profits declining with workers having more information about their returns to education.

1.8 Conclusion

There is often public concern about the level of investment in higher education, and that too many students go to university when they would not benefit from going. These concerns are frequently based on observations that many graduates end up working in occupations which typically do not require higher education to perform. However, causal studies of the wage returns to higher education typically find a substantial positive return. In this paper, I propose that both observations are consistent with a model in which returns to higher education are heterogeneous and uncertain at the point of university choice, and in which workers match to jobs in a Sattinger-style matching labour market after investing in education. Workers who are observed to work in non-graduate occupations despite getting a degree are interpreted in this framework as experiencing ex-post regret; they correctly invested in higher education with an ex ante expectation of a positive return but ended up with a negative return ex-post. There are two kinds of education mismatch that arises in this model: individual mismatch as outlined above, and aggregate mismatch, when the decentralised equilibrium outcome is not efficient and aggregate utility can be increased by changing the education profile. Aggregate mismatch arises in the model because workers' do not internalise the externalities of their human capital decisions on firms and other workers in the economy.

I estimate a parametric version of my proposed model on a pooled dataset of five cohorts born between 1988 and 1993, using data from the Understanding Society panel survey. I find that the degree of noise inherent in the signal is large; workers only have a signal with a 0.324 correlation to their underlying ability. I find that this leads 32.9% of workers in the economy to end up with ex-post regret. 18.2% of workers end up over-educated, that is, they invest in education but end up with a net negative utility return, and 14.7% are under-educated, missing out on a positive ex-post utility return. On average, the utility loss is in money-metric terms is £323.35 a year for over-educated workers, and £465.22 a year for under-educated workers. Being under-educated is on average more costly than being over-educated, which explains why more workers prefer to over-invest in education rather than vice versa.

I consider the welfare impacts of government policies to (1) impose compensated graduate taxes or subsidies to manipulate the returns to higher education, and (2) reduce worker uncertainty about their individual return to higher education. I find that the optimal policy differs depending on the policy objectives of the government. I find that if the policy is to minimise the incidence or cost of mismatch in the economy, the government should instate a compensated tax to reduce the level of higher education attendance in the economy. For a welfare-maximising social planner, the equilibrium is in general efficient and trades off two kinds of externalities that work in opposite directions. I find that empirically, my estimates imply that the optimal graduate incentive policy would be to reduce the incidence of higher education by 1.6 percentage points by imposing a small graduate tax equivalent to 1 p per hour at age 30 (roughly equivalent to increasing the cost of education by 368 pounds). Reducing worker uncertainty about returns unambiguously benefits workers but surprisingly harms firms by increasing the wage cost. Lower uncertainty about returns also leads to higher inequality by allowing workers to engage in more perfect Roy-type sorting.

There are many unanswered questions arising from my proposed model that I have not been able to address in this paper, which I leave for subsequent research. First, my empirical analysis has been based on observational outcomes data and an assumption of rational expectations. This rules out potentially interesting causes of education relating over-confidence and mistaken beliefs. Another promising avenue of research is to further consider how uncertainty about multiple dimensions of skill could affect the analysis in this model. My framework also leads to the possibility that graduate taxes or subsidies conditional on school grades may be more efficient than flat graduate taxes or subsidies. I hope to eventually tractably extend the welfare analysis to consider a greater policy set than has been considered here.

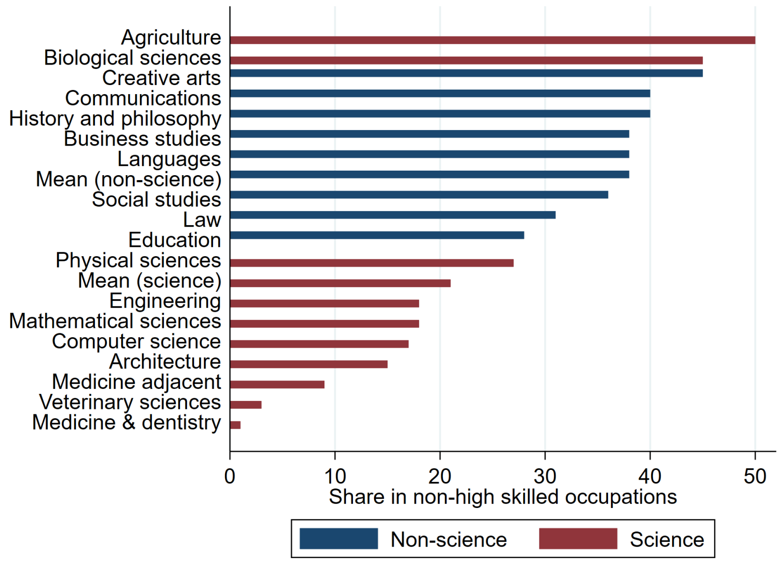
1.9 Appendix

A Additional Stylised Facts

Employment in low-skilled occupations by subject

Figure 1.18 plots the share of graduates in non-high-skill occupations by their degree subject. In general, it is true that this rate is higher for workers in non-STEM subjects than in STEM subjects, with biological sciences being an important main exception. However, with the exception of a few vocational subjects like medicine and dentistry, veterinary science and medicine-adjacent subjects like pharmacy, most graduates even in STEM subjects experience at least 10% of workers in non skilled occupations, with the within-science mean being 20%. This suggests that the issues discussed in this paper also apply to STEM fields, even if it is not to a similar extent. Explicitly modelling a multinomial education choice with choice of subject substantially complicates the matching problem considered and is not feasible with the current data used (the Understanding Society which forms the source for the core of the data work does not have information on university subjects). I hope to incorporate it in future work.

Figure 1.18: Workers by share in non high-skill occupation by degree subject (2018-19)



Data from HESA outcomes survey for 2018-19 for graduates one-year from graduation. The red bars represent STEM subjects while the blue pairs represent non-STEM subjects. A high-skilled occupation is defined as an occupation which requires at least a degree, according to the ONS occupation classification, and consists of occupations with occupation codes beginning 1 to 3.

New graduates are increasingly working in occupations with previously low shares of degree holders

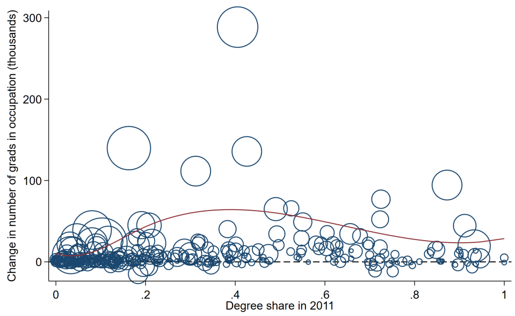
A natural question given the patterns in figure **??** showing a substantial increase in the number of graduates is what occupations these new graduates found employment in. To this end, I use information on the moments of hourly earnings within each occupation from the Annual Survey of Hours and Earnings (ASHE), matched to information about the share of workers with degrees within the occupation. Figure 1.19a plots a scatter diagram of the occupations, with the share of workers within the occupation with degrees in 2011 (the start of the available period[39]) on the x-axis, and the absolute change in the number of graduates (in thousands) on the y-axis[40]. A non-parametric line of best fit, weighted by the number of workers in that

[39]To maintain consistent occupational definitions, I used a period, 2011-19, for which SOC10 classifications were available.

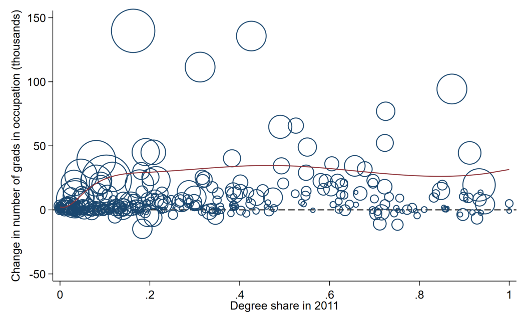
[40]The absolute number of graduates was calculated by multiplying the number of workers in the occupation from the ASHE by the share of workers with degrees estimated from the LFS.

occupation in 2011, is plotted over the points. Figure 1.19b plots a similar scatterplot, but excludes the occupation with the largest change in figure 1.19a, nursing[41]. The size of the circles in both diagrams shows the size of the occupation in 2011.

Figure 1.19: Change in number of graduates within each occupation (2011-19)



(a) Change in number of graduates working in occ



(b) Figure 1.19a (excluding nurses)

Data on the share of workers with degrees from the UK Labour Force Survey 2011-19. Data on the number of workers in the occupation from the Annual Survey of Hours and Earnings. A non-parametric line of best fit is drawn in red in both diagrams. The size of the bubbles reflects the size of the occupations.

In each figure, I plot a non-parametric line of best fit for the relation between the initial share of workers in the occupation with degrees in 2011 and the change in the number of graduates in the occupation between 2011 and 2019. A general pattern that arises from these figures is that the occupations which received the greatest influx of workers were typically occupations with a fairly medium levels of workers with degrees at the start of the period[42].

The patterns shown in figures 1.19a and 1.19b may potentially reflect only the relative sizes of occupations, not that their compositions have changed. Thus to complement those diagrams, I plot a scatterplot of the change in the share of workers with degrees within an occupation from 2011 to 2019 against the initial degree share in 2011 in figure 1.20. The figure shows that occupations with fairly low shares of workers with degrees experienced a

[41]From 2008, the minimum award for completing a nursing qualification became a degree, and so, new nursing graduates in subsequent cohorts have substantially increased the share of workers with degrees in the nursing profession. Since this change in classification arguably does not reflect any real economic change yet drives the roughly quadratic shape in figure 1.19a, it is removed as a robustness check in figure 1.19b.

[42]The occupation seeing the largest change is nursing, with 288 thousand new graduates entering the profession; this is likely to be a mechanical affair driven by a change in the classification of the nursing qualification. The occupations absorbing the next most graduates are respective "other administrative occupations" (+140 thousand), "sales accounts and business development managers" (+135 thousand), and "production managers and directors in manufacturing" (+111 thousand).

Figure 1.20: Change in the share of workers with degrees within occupations from 2011-19

Data from the UK Labour Force Survey 2011-19. A non-parametric line of best fit is drawn in red. The size of the bubbles reflects the size of the occupations.

substantial increase in the share of workers in the occupation with degrees, consistent with figure 1.19. Thus, any theory of the impact of higher education expansion on the graduate labour market has to account for the fact that most of these new graduates end up working in occupations with fairly low shares of workers with degrees.

B Additional Model Description Details

Implications of the model for the occupational structure

While the solution of the individual's problem and the determination of equilibrium in the model does not involve analysis of occupations, the mixed distribution set-up of the distribution of job productivity allows the model to analyse how graduates match to occupations in this model. Conditional on a worker having skill s, the probability of the worker having a job in occupation k is as follows.

$$Pr(occ = k|S = s) = \frac{p_k f_Y^k(\mu(s))}{\sum_{l\in\{1,\cdots,K\}} p_l f_Y^l(\mu(s))} \tag{1.28}$$

There are two main advantages to this set-up. First, this allows occupations to be used in the identification of the model in a more natural way. Many past papers have used information on occupations, typically indices constructed from the O*NET dictionary of occupation titles (DOT), to identify structural matching models between workers and jobs. This assumes job heterogeneity within occupations. By allowing for a distribution of job productivities within occupations, I allow both the mean and variance of job productivities to affect the wages and matching to graduate workers in the model.

Second, the model's structure allows the causal effect of education be interpreted as the combination of two effects: the causal effect of education on a worker's productivity in any job, and the causal effect of education on a worker's probability of placing at a better occupation. This coheres with previous research on the channels of the education wage premium, such as Lemieux (2014), which finds that matching to better occupations accounts for at least half of the estimated college wage premium.

It is worth noting that some implicit assumptions made in the analysis. First, an important implicit assumption is that workers have no heterogeneous preferences for occupations. This, of course, is unrealistic; deviations in the predictions of the model for the wages and share of graduate workers in an occupation could plausibly be due to unobserved compensating differentials. Second, I assume that there are no frictions in the labour market part of the model, which implies that there are no occupational mismatch in the labour market. Thus, the entire source of mismatch in this model comes from informational frictions in the educational decision stage. The inclusion of either of these concerns will substantially complicate the matching problem that workers face, and increase the computational costs of the model. I leave the extension of this model to encompass those concerns for future work.

C Proofs

Proof of proposition 1

The optimality condition is as follows:

$$\kappa + \Delta\eta + \beta\Delta w(a) \geq 0$$

First, note that the LHS of the equation is strictly increasing in both $\Delta\eta$ and a. To show that $\Delta w(a)$ is increasing in a, note that:

$$\begin{aligned}\frac{d\Delta w(a)}{da} &= \frac{dw(s(a,1))}{da} - \frac{dw(s(a,0))}{da} \\ &= w'(s(a,1))\frac{\partial s(a,1)}{\partial a} - w'(s(a,0))\frac{\partial s(a,0)}{\partial a}\end{aligned}$$

By assumption, $\frac{\partial s(a,1)}{\partial a} > \frac{\partial s(a,0)}{\partial a}$. Furthermore, we know that the wage function $w(\cdot)$ is convex, so $w'(s(a,1)) > w'(s(a,0))$. Thus, $\frac{d\Delta w(a)}{da}$ has to be greater than 0; the wage return is increasing in ability. Since the LHS is increasing in a and $\Delta\eta$, it follows that for each value of $\Delta\eta$, there must be a unique value of a which satisfies the optimality condition with equality.

Since the LHS is increasing in a and $\Delta\eta$, it follows that the LHS will only cross the zero threshold once, implying that e^P is increasing in a and $\Delta\eta$.

Let $a^P(\Delta\eta)$ denote the value of a which satisfies the optimality condition with equality, for any given value of $\Delta\eta$. Implicitly differentiate by $\Delta\eta$.

$$\kappa + \Delta\eta + \beta\Delta w(a^P(\Delta\eta)) = 0 \tag{1.29}$$

$$1 + \beta\frac{d\Delta w(a^P)}{da^P}\frac{da^P}{d\Delta\eta} = 0 \tag{1.30}$$

$$\frac{da^P(\Delta\eta)}{d\Delta\eta} = -\frac{1}{\beta\frac{d\Delta w(a^P)}{da^P}} \tag{1.31}$$

The RHS is negative since $\frac{d\Delta w(a^P)}{da^P} > 0$. Thus, we can show that the optimal cut-off $a^P(\Delta\eta)$

is decreasing in $\Delta\eta$.

The solution to each individual's education problem is completely characterised by $a^P(\Delta\eta)$, the cut-off level of ability for each preference state. We know that for any value of a smaller than it, the optimal decision would be to not invest in education, and similarly for any value of a greater than it, the optimal decision would be to invest in education. Also, $\frac{da^P(\Delta\eta)}{d(\Delta\eta)} < 0$, such that when a worker has a greater preference for university, it takes a lower ability threshold to induce them into education. Thus, $e^P(a, \Delta\eta)$ is increasing in a, and $\Delta\eta$.

Proof of proposition 2

The proof uses results about comparative statics under uncertainty provided in Athey (2002), particularly theorem 2 in her paper. She considers problems where workers aim to maximise a stochastic function $U(x,\theta)$ with respect to $x \in B \subset \mathbb{R}$, where $U(x,\theta)$ can be expressed as follows.

$$U(x,\theta) \equiv \int u(x,s)f(s;\theta)d\mu(s)$$

She provides conditions on the primitive functions $u(x,s)$ and $f(s;\theta)$ such that the argument maximising $U(x,\theta)$ is non-decreasing in θ and B. Theorem 2 states that two conditions are jointly sufficient:

1. The density function f is log-supermodular, which is satisfied when it obeys the monotone likelihood ratio (MLR) order.

2. The payoff function $u(x,s)$ satisfies single-crossing in (x,s); this means that for all $x' > x$, then $u(x',s) - u(x,s)$ crosses 0 at most once and from below.

The first condition is satisfied when $f(s;\theta)$ obeys the monotone-likelihood ratio order. Consider a family of distribution functions $\{\lambda(\cdot,\theta)\}_{\theta\in\mathbb{R}}$; this family of functions obey the MLR order if $\frac{\lambda(a,\theta'')}{\lambda(a,\theta')}$ is increasing in a whenever $\theta'' > \theta'$. It turns out that under the additive error structure assumed in 1.1, the conditional distribution satisfies MLR order in θ regardless of the initial distribution of a. Thus, given additive noise, the conditional distribution of the underlying random variable conditional on signal θ obeys MLR order with respect to the size of the signal θ as long as the error ε has a log-concave distribution (see a proof of this, see appendix C). This is assumed in assumption 1.3.1.

For the second condition to be satisfied, we have to show that $\kappa + \Delta\eta + \beta\{\Delta w(a)\}$ crosses 0 at most once and from below. This is satisfied when the term is strictly increasing, which was proved in appendix C as part of the proof of proposition 1. Thus, we can prove that the optimal investment is non-decreasing in θ, conditional on $\Delta\eta$, by directly applying Athey's results.

The corollary of this conclusion is that since the condition for the worker choosing $e = 1$ over $e = 0$ is $V(\theta, 1) - V(\theta, 0) > 0$, then if $e^I(\theta)$ is non-decreasing in θ, then $V(\theta, 1) - V(\theta, 0)$ is also weakly increasing in θ. This implies that $E\{w(s(a, 1)) - w(s(a, 0))|\theta\}$ is also weakly increasing in θ.

Denote by $\theta^P(\Delta\eta)$ the cut-off value of θ such that given heterogeneous preferences $\Delta\eta$, a worker with signal $\theta^P(\Delta\eta)$ would be indifferent between investing and not investing in education. Since $E\{w(s(a, 1)) - w(s(a, 0))|\theta\}$ is weakly increasing in θ, $\theta^P(\Delta\eta)$ should be weakly decreasing in $\Delta\eta$, and $e^I(\theta)$ would be weakly increasing in $\Delta\eta$. Workers with a greater preference for higher education would be more likely to invest in education across all values of θ. As is the case in the perfect information scenario, the cut-off function differs a boundary in two-dimensional θ-$\Delta\eta$ space which separates students who will invest in HE and those who will not.

Under additive noise, log-concavity of the density function of the error term is sufficient for MLR order

Under an additive structure as specified in equation 1.1, the joint distribution of θ and a can be derived by convolution as the product of the density function of a and of ε. The conditional densities can then be computed as follows:

$$
\begin{aligned}
f_{\Theta,A}(\theta, a) &= f_A(a) \cdot f_\varepsilon(\theta - a) \\
f_\Theta(\theta) &= \int_{-\infty}^{\infty} f_A(a) f_\varepsilon(\theta - a) da \\
f_{A|\Theta}(a|\theta = \theta_1) &= \frac{f_{\Theta,A}(\theta_1, a)}{f_\Theta(\theta_1)} \\
f_{\Theta|A}(\theta|A = a) &= \frac{f_{\Theta,A}(\theta, a)}{f_A(a)} \\
&= f_\varepsilon(s - a)
\end{aligned}
$$

Consider the ratio $\frac{f(a|\theta=\theta'')}{f(a|\theta=\theta')}$ where θ'' and θ' are arbitrary values such that $\theta'' > \theta'$. We begin by substituting the expression for the conditional density into the expression.

$$
\begin{aligned}
\frac{f(a|\theta = \theta'')}{f(a|\theta = \theta')} &= \frac{f_{\theta,a}(\theta'', a)}{f_{\theta,a}(\theta', a)} \cdot \frac{f_\Theta(\theta')}{f_\Theta(\theta'')} \\
&= \frac{f_A(a) f_\varepsilon(\theta'' - a)}{f_A(a) f_\varepsilon(\theta' - a)} \cdot \frac{f_\Theta(\theta')}{f_\Theta(\theta'')} \\
&= \frac{f_\varepsilon(\theta'' - a)}{f_\varepsilon(\theta' - a)} \cdot \frac{f_\Theta(\theta')}{f_\Theta(\theta'')}
\end{aligned}
$$

Since $\frac{f_\Theta(\theta')}{f_\Theta(\theta'')}$ is not a function in a, to establish MLR order in θ, we have to show that $\frac{f_\varepsilon(\theta''-a)}{f_\varepsilon(\theta'-a)}$ is increasing in a. Differentiating the expression and signing it, we get the following inequality for $\theta'' > \theta'$:

$$
\frac{f'_\varepsilon(\theta' - a) f_\varepsilon(\theta'' - a) - f'_\varepsilon(\theta'' - a) f_\varepsilon(\theta' - a)}{[f_\varepsilon(\theta' - a)]^2} > 0
$$

Since the denominator must be positive, we can examine the following identity:

$$f'_\varepsilon(\theta'' - a)f_\varepsilon(\theta' - a) < f'_\varepsilon(\theta' - a)f_\varepsilon(\theta'' - a)$$

Given that distribution functions have a positive range, we can rearrange terms as follows:

$$\frac{f'_\varepsilon(\theta' - a)}{f_\varepsilon(\theta' - a)} > \frac{f'_\varepsilon(\theta'' - a)}{f_\varepsilon(\theta'' - a)}$$

Thus, the condition for the conditional distributions obeying the monotone likelihood ratio order in the signal θ reduces to the following condition on the distribution of the error f_ε.

$$\theta' < \theta'' \rightarrow \frac{f'_\varepsilon(\theta' - a)}{f_\varepsilon(\theta' - a)} > \frac{f'_\varepsilon(\theta'' - a)}{f_\varepsilon(\theta'' - a)}$$

This condition is equivalent to the condition that $\frac{f'(x)}{f(x)}$ is monotone decreasing in x, and this is sufficient for the distribution f_ε to be log-concave (Bagnoli & Bergstrom (2005)).

The condition for the MLR order also holds when the errors are normally distributed; this can be computed directly as follows.

$$\begin{aligned}
\frac{f_\varepsilon(\theta'' - a)}{f_\varepsilon(\theta' - a)} &= \frac{\frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp(-0.5\cdot(\frac{\theta''-a}{\sigma_\varepsilon})^2)}{\frac{1}{\sigma_\varepsilon\sqrt{2\pi}}\exp(-0.5\cdot(\frac{\theta'-a}{\sigma_\varepsilon})^2)} \\
&= \exp\left(-0.5\cdot\left(\left(\frac{\theta''-a}{\sigma_\varepsilon}\right)^2 - \left(\frac{\theta'-a}{\sigma_\varepsilon}\right)^2\right)\right) \\
\frac{\partial\frac{f_\varepsilon(\theta''-a)}{f_\varepsilon(\theta'-a)}}{\partial a} &= \left(\frac{\theta''-\theta'}{\sigma_\varepsilon^2}\right)\cdot\exp\left(-0.5\cdot\left(\left(\frac{\theta''-a}{\sigma_\varepsilon}\right)^2 - \left(\frac{\theta'-a}{\sigma_\varepsilon}\right)^2\right)\right) > 0
\end{aligned}$$

Proof of proposition 3

The broad strategy of the proof is to show that the functional derivative of W with respective to p at the equilibrium education profile $\bar{p}$ is non-zero, and thus the necessary condition for optimality does not hold. However, it is not possible define the derivative in terms of p, because the set of education profiles $\mathcal{P}$ does not constitute a vector space as the education profiles cannot be summed to produce education profiles. Thus, I first define a subset of education profiles which can be represented by a function the set of which is a metric space. If $\bar{\psi}$ is not the maximiser of the counterpart function $\tilde{W}$ within the set Ψ, which corresponds to a subset of $\mathcal{P}$, then $\bar{p}$ also cannot be the maximiser of W within $\mathcal{P}$.

Proposition 2 shows that any equilibrium education profile p can be defined alternatively with a cut-off signal function, which decreases with $\Delta\eta$. A worker selects into education then if their signal exceeds the cut-off signal given their preference level. This means that such education profiles correspond one-to-one to continuous, non-increasing functions which maps the real line to the real line. Then, note that we can re-express the domain of θ and $\Delta\eta$ as $[0, 1]$ intervals, by re-expressing any value of θ and $\Delta\eta$ as a quantile. I denote this re-scaled values of θ and $\Delta\eta$ as $\tilde{\theta}$ and $\tilde{\Delta\eta}$, where:

$$\tilde{\theta} \equiv F_\theta(\theta)$$
$$\tilde{\Delta\eta} \equiv F_{\Delta\eta}(\Delta\eta)$$

Denote by $\psi : [0, 1] \to \mathbb{R}$ a function which maps a normalised net preference in the interval $[0, 1]$ to an un-normalised signal in the real line, and is continuous and non-increasing. Intuitively, this denotes a line in θ-$\Delta\eta$ space, which divides workers who select into education and those who don't. This relates to an education profile p as follows:

$$p(\theta, \Delta\eta) = 1[\theta > \psi(F_{\Delta\eta}(\Delta\eta))]$$

Denote the set of ψ functions by Ψ. Unlike the set $\mathcal{P}$, the set Ψ is a vector space, and we can define a norm for the vector space as follows: $||\psi|| = \int_0^1 |\psi(\tilde{\Delta\eta})|d\theta$. Denote the counterpart of $\bar{p}$ in Ψ by $\bar{\psi}$, where $\bar{p}(\theta, \Delta\eta) = 1[\theta > \bar{\psi}(F_{\Delta\eta}(\Delta\eta))]$ for all values of θ and $\Delta\eta$.

We are now ready to define the derivative. Consider a slightly modified function W, denoted by $\tilde{W} : \Psi \to \mathbb{R}$, which takes ψ as an argument instead of p with the replacement defined above. Denote the Gateaux derivative of $\tilde{W}$ at $\bar{\psi}$ in the direction $\phi \in \Psi$ as follows:

$$d\tilde{W}(\bar{\psi}, \phi) = \lim_{\tau \to 0} \frac{\tilde{W}[\bar{\psi} + \tau\phi] - \tilde{W}[\bar{\psi}]}{\tau}$$

Consider the numerator of the fraction in this limit. For notational clarity, I omit the argument of the function ψ and write $\psi' \equiv \bar{\psi} + \tau\phi$. Noting that the joint output function $g(s, \mu(s))$ is also a functional of the matching function μ, which depends on ψ through F_S, I write $g(s, \mu(s; \psi)) \equiv g(s, \psi)$. I similarly write the wage and profit schedules for a worker with skill s and the firm that matches to a worker with skill s under the education profile corresponding to the function ψ as $w(s, \psi)$ and $\pi(s, \psi)$ respectively. Finally, instead of writing the full expression for the correspondence between p and ψ, I write the function $1[\theta > \psi F_{\Delta\eta}(\Delta\eta)] \equiv \chi(\theta, \Delta\eta, \psi)$, occasionally suppressing the arguments $\theta, \Delta\eta$ to save space to write $\chi(\psi)$. Denote the difference between joint output functions under ψ by $\Delta g(a, \psi) \equiv g(s(a, 1), \mu(s(a, 1), \psi)) - g(s(a, 0), \mu(s(a, 0), \psi))$.

We can rearrange the $\tilde{W}[\bar{\psi} + \tau\phi] - \tilde{W}[\bar{\psi}]$ into the following three terms:

$$\begin{aligned}
\tilde{W}[\psi'] - \tilde{W}[\psi] &= \int\int\int (\chi(\psi') - \chi(\bar{\psi}))\{w(s(a,1), p) - w(s(a,0), p) + \kappa + \Delta\eta\} dF(a)dF(\varepsilon)dF(\Delta\eta) \\
&+ \int\int\int (\chi(\psi') - \chi(\bar{\psi}))\{\pi(s(a,1), p) - \pi(s(a,0), p)\} dF(a)dF(\varepsilon)dF(\Delta\eta) \\
&+ \int\int\int \chi(\psi')[\Delta g(a, \psi') - \Delta g(a, \bar{\psi})] + [g(s(a,0), \psi') - g(s(a,0), \bar{\psi})] dF(a)dF(\varepsilon)dF(\Delta\eta)
\end{aligned}$$

Note that at $\bar{p}$, $\bar{p}(\theta, \Delta\eta) = 1$ if and only if $E[w(s(a,1), p) - w(s(a,0), p) + \kappa + \Delta\eta | \theta] \geq 0$. Thus, integrating over θ and $\Delta\eta$, it follows that $\bar{p}$ must maximise $E_{\theta,\Delta\eta}[p\{w(s(a,1), p) -$

$w(s(a,0),p)+\kappa+\Delta\eta\}]$. This means that $\bar{\psi}$ must also be the maximiser of $\int\int\int(\chi(\bar{\psi}))\{w(s(a,1),p)-w(s(a,0),p)+\kappa+\Delta\eta\}dF(a)dF(\varepsilon)dF(\Delta\eta)$, as if a function is the maximiser in a set, it is also the maximiser in any subset of that set.

Furthermore, a necessary condition of $\bar{\psi}$ being the maximiser of the objective in the set Ψ is that the Gateaux derivative of $\int\int\int(\chi(\bar{\psi}))\{w(s(a,1),p)-w(s(a,0),p)+\kappa+\Delta\eta\}dF(a)dF(\varepsilon)dF(\Delta\eta)$ with respect to ψ at $\bar{\psi}$ is equal to 0. Denote:

$$k[\psi] = \int\int\int(\chi(\bar{\psi}))\{w(s(a,1),p) - w(s(a,0),p) + \kappa + \Delta\eta\}dF(a)dF(\varepsilon)dF(\Delta\eta)$$

Then, for all $\phi \in \Psi$,

$$dk(\bar{\psi},\phi) = \lim_{\tau\to 0}\frac{k[\bar{\psi}+\tau\phi] - k[\bar{\psi}]}{\tau} = 0$$

Thus for any direction ϕ, the Gateaux derivative of $\tilde{W}$ at $\bar{\psi}$ consists of the following two remaining terms:

$$d\tilde{W}(\bar{\psi},\phi) = \underbrace{\lim_{\tau\to 0}\left\{\frac{\int\int\int(\chi(\psi') - \chi(\bar{\psi}))\{\pi(s(a,1),\bar{\psi}) - \pi(s(a,0),\bar{\psi})\}dF(a)dF(\varepsilon)dF(\Delta\eta)}{\tau}\right\}}_{\text{Hold-up externality}} +$$
$$\underbrace{\lim_{\tau\to 0}\left\{\frac{\int\int\int\chi(\psi')[\Delta g(a,\psi') - \Delta g(a,\bar{\psi})] + [g(s(a,0),\psi') - g(s(a,0),\bar{\psi})]dF(a)dF(\varepsilon)dF(\Delta\eta)}{\tau}\right\}}_{\text{Positional externality}} \tag{1.32}$$

In general, the last two terms are not in general non-zero. The surplus profits are positive, and thus for $d\tilde{W}[\bar{\psi}] = 0$, the last term must coincidentally exactly offset the profits term. When this is not the case, the necessary condition for optimality does not hold, implying that the equilibrium education profile is not the optimal education profile from the social planner's point of view.

D Computation and Estimation Details

The estimation procedure requires the simulation of earnings and grades conditional on the distribution of education choices in the economy. To derive the wage function, we require the evaluation of the cumulative density of each value of skill s, which requires numerical integration of a complex object $f_S(s)$, given by equation 1.12. The probability density function given by equation 1.12 itself requires integration over the signal noise term ε. I compute the integration in equation 1.12 using Gauss-Hermite quadrature, since the distribution of ε is assumed to be normal with variance σ_ε^2. I then compute the integration in equation 1.13 using standard numerical integration.

To compute the wage function, I first compute the derivate of the wage function, which has a simple functional form under my parameterisation: $w'(s) = q\gamma_1 s^{\gamma_1 - 1}\mu(s)^{\gamma_2}$. I approximate this function using a monotone cubic spline using Hyman filtering. To derive the wage condition on s, I integrate this spline approximation according to equation 1.8. Since each set of trial parameter values requires the computation of only a single wage function, I then approximate this wage function using a monotone cubic spline in s to save on computation when simulating the dataset.

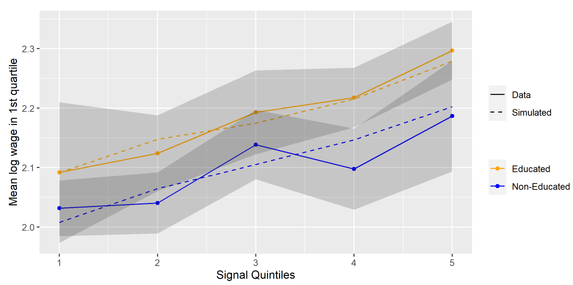
Finally, to compute the difference between the value functions under higher education and not under higher education, it is necessary to take expectations over the wage function conditional on the signal θ. Under the parametric assumption of normality of both the distribution of a and the distribution of ε, the posterior distribution of $a|\theta$ is also normal. I use Gauss-Hermite quadrature to reduce the computational burden of computing these large expectation terms.

E Additional Targeted Moment Fits

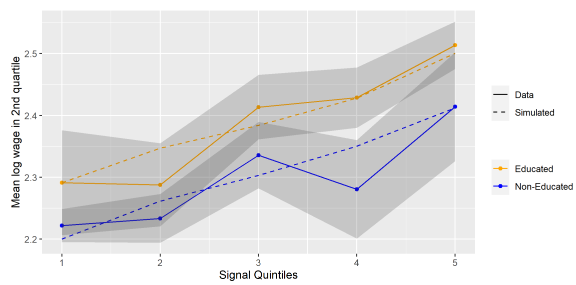
The main figure in this appendix, figure 1.21 plots the mean log wage within four quartiles of log hourly wages, conditional on degree status and signal quintiles. These quartiles were used as moments in the structural estimation and are used to pin down both the degree of uncertainty about the signal, and the shape of the wage function. The simulated data seems to fit all the moments well, although the some moments are estimated with significant variance.

Figure 1.21: Simulated versus actual means of log wage within wage quartiles conditional on degree and grade quintile

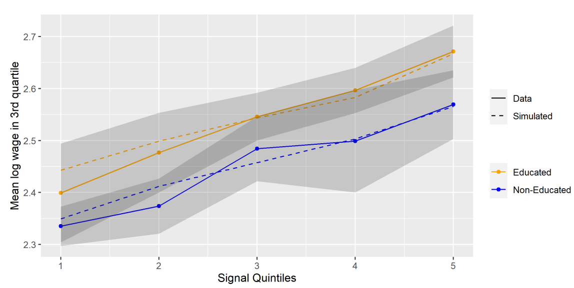
(a) First Quartile



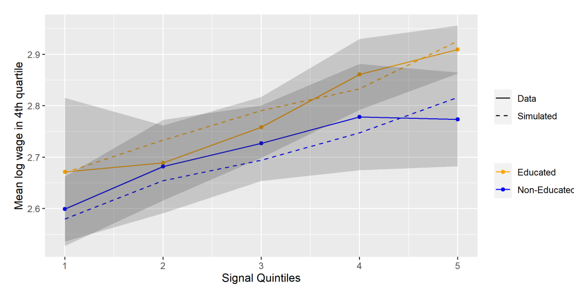
(b) Second Quartile



(c) Third Quartile



(d) Fourth Quartile



Notes: This figure plots the mean log-wage within each of four log wage quartiles, conditional on degree status and five signal quintiles.

F Details about implementation of the income-contingent loan policy

In the UK, students typically pay for higher education in the UK with an income-contingent loan, first introduced under the Labour government in 1998. Under the terms of the loan, students borrow the sum required to pay for tuition fees. They then are obliged to make repayments, where the minimum repayment is a certain percentage of their income above some income threshold. They do not have to make any repayments if their income is below the specified threshold. Repayments stop after the initial loan plus interest is fully repaid, or a certain time period has elapsed[43].

The actual loan policy is implemented over a lifetime, which creates scope for dynamic complexities around the optimal repayment of the loan given a certain wage path. This is not captured in my framework as I focus on effectively a single period; in my implementation, I simplify it for a single period as follows. The parameters of the policy are the repayment threshold ι_1, the repayment rate above the threshold ι_2, and the initial loan sum ι_3. Then, the post-policy wage, given an initial wage of w, is given by the following expression.

$$\text{post policy wage}(w) = \begin{cases} w \text{ if } w < \iota_1 \\ \iota_1 + (1 - \iota_2)(w - \iota_1) \text{ if } w \in [\iota_1, \iota_1 + \frac{\iota_3}{\iota_2}] \\ w - \iota_3 \text{ if } w > \iota_1 + \frac{\iota_3}{\iota_2} \end{cases} \tag{1.33}$$

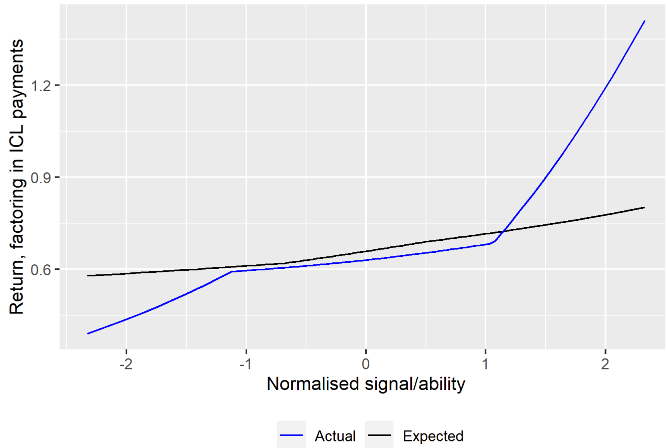
The expression says that a worker pays nothing under the scheme if their wage is below the income threshold ι_1. Past the threshold, they pay a share ι_2 of their income in excess of the threshold. Once this payment exceeds the initial loan sum ι_3, then the total payment is simply ι_3. I abstract away from the possibility of strategic repayment of the debt to minimise total repayment for high income workers, in this application. I calibrate the parameters as follows. I set the repayment threshold, ι_1, to be 15711 pounds per year, which is the post-tax average of repayment thresholds from 2016-19[44]. I convert this to pounds per hour terms by assuming that workers work for 44 weeks per year and 40 hours per week. The repayment rate is 9% as has been the policy rate for the loan plan taken by workers in my sample. Finally, I set the initial loan sum to be 22,229 pounds, amortised over 30 years at a 2.5% interest rate and converted to pounds per hour terms. The initial sum assumes on average 3.5 years in university, adding together the sum of the mean maintenance and tuition payments for 2006-2009[45].

[43] More details can be found on government websites such as https://www.gov.uk/government/statistics/student-loans-in-england-2021-to-2022/income-contingent-student-loan-repayment-plans-interest-rates-and-calculations-england. The UK income contingent loan policy has also been analysed in Britton et al. (2019).

[44] See https://www.gov.uk/guidance/previous-annual-repayment-thresholds

[45] The loan sizes are taken from student loan statistics produced for the British parliament by

Figure 1.22: Post-income contingent loan return to higher education, conditional on grades



This figure plots the relation between ability and the return to higher education net of payments due because of the ICL policy in the blue line, under the parameters obtained in the estimation procedure (described in subsection 1.5). The black line plots the relation between ability and the expected return to higher education, post the implementation of the ICL policy.

Figure 1.22 plots the return to higher education after factoring in the payment of the income contingent loan under perfect information conditional on ability (in the blue line), and in expectation conditional on grades (in the black line). The blue line shows that the details of the policy creates two kinks, which create non-linearity. The black line is substantially flatter but also kinked at around -0.8. The increase in the expected return is lower up to that point because the relative upset is depressed by the upside of being higher ability than expected is depressed by the flat portion of the blue line between approximately -1.2 and 1.2. Past that point however, the larger upside of the part of the blue line to the right of approximately 1.2 becomes more relevant, and increases the rate at which the expected return increases.

The policy effectively offers a significant level of insurance to workers, and workers who end up with sufficiently low earnings pay almost nothing for higher education. Importantly, the degree of insurance is greater for workers with low grades than those with high grades, as they are more likely than the latter to have low earnings later in life. Furthermore, workers are not liable to pay more than the original sum that they borrowed. Thus, the amount that

Paul Bolton for the Commons Library. See Bolton (2019). The report was accessed online at https://researchbriefings.files.parliament.uk/documents/SN01079/SN01079.pdf.

a worker would pay stops rising after a certain level of income. In this sense, the greatest relative contribution comes from workers in an intermediate range of income.

G Welfare in terms of value for money

Although willingness to pay is a natural concept of welfare to focus on, the policy debate, especially in the UK, has focused on the concept of 'value for money' or 'earnings potential'[46]. Intuitively, this seems to refer to the exclusion of non-pecuniary motivations for attending university, roughly implying that students should attend higher education if their decision passes a cost-benefit analysis considering only pecuniary returns to university net of pecuniary costs.

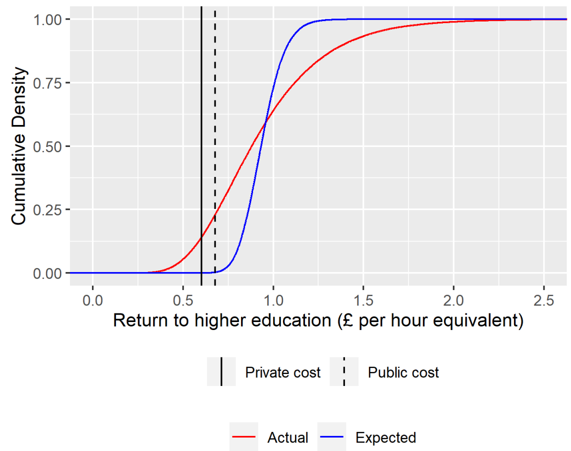
In this paper, I interpret value for money roughly as the pecuniary wage return to attending university exceeding the pecuniary costs of doing so, excluding the consumption value that workers may receive from higher education. In this sense, a worker may attend university in my model despite not receiving value for money for doing so because of (1) they have a positive non-pecuniary preference for attending university, or (2) because the information friction leads them to believe that they will receive a positive pecuniary return despite not actually doing so.

Figure 1.23 plots the distribution of actual and expected wage returns to higher education in a simulated dataset with 20000 workers as cumulative density functions. A striking feature of these distributions is that expected returns (conditional on the grade signal) are much less variable than average returns (conditional on ability). This is to be expected, especially given the substantial noise with which grades correlates to actual underlying ability.

To conduct a full cost-benefit analysis, I need to specify the pecuniary cost of higher

[46] Earnings potential was discussed by the former Chancellor of the UK, and Conservative party leader hopeful Rishi Sunak, who pledged in the 2022 leadership contest, that he would crack down on university courses, assessing them "through their drop-out rates, numbers in graduate jobs and salary thresholds, with exceptions for nursing and other courses with high social value". See https://www.theguardian.com/politics/2022/aug/07/rishi-sunak-vows-to-end-low-earning-degrees-in-post-16-education-shake-up.

Figure 1.23: Cumulative density of simulated actual and expected returns to higher ed



This figure plots the distribution of actual (conditional on unobserved ability) and expected returns (conditional on the signal) in a simulated dataset generated from the model and the estimated parameters. The red line plots the cumulative density of actual returns and the blue line plots the cumulative density of expected returns. The two vertical lines represent two cost benchmarks.

education. I consider two alternatives. First, I take an estimate of the cost of providing higher education from Belfield et al. (2018b), £25,000, and amortise this over 30 years at a 2.5% interest rate. This calculation implies the per-year payment to be £1194.441 per year, or £0.679 per hour (assuming a 40 hour work week for 44 weeks). Second, I take the total student loan that students take out on average, similar amortised over 30 years at a 2.5% interest rate. The average loan is £22,230, and is taken from Bolton (2019). The implied per year per hour equivalent cost is 0.599. Two vertical lines are plotted on the figure, representing two cost benchmarks. The dashed vertical line presents the cost of providing a higher education course during the period, including the private cost borne by the student and the public cost borne by the Treasury in terms of teaching grants given to universities. The solid vertical line represents the cost borne directly by workers, including tuition costs and living costs covered under the maintenance loan.

The figure implies that most workers in this cohort would receive a benefit net of either cost, in part because this cost, especially spread over a lifetime, is not particularly high. Only 14.3% of workers would not have received a return exceeding their private cost of education, and only 23.2% of workers would not have received a return exceeding the total cost of provision. Only 0.3% workers would not have had an expected income above the total cost of provision, implying that if workers considered only the pecuniary accounting of costs and benefits, more workers would have selected into higher education[47]. From this point of view, higher education was an extremely valuable investment for workers relative to its cost for the studied cohorts[48].

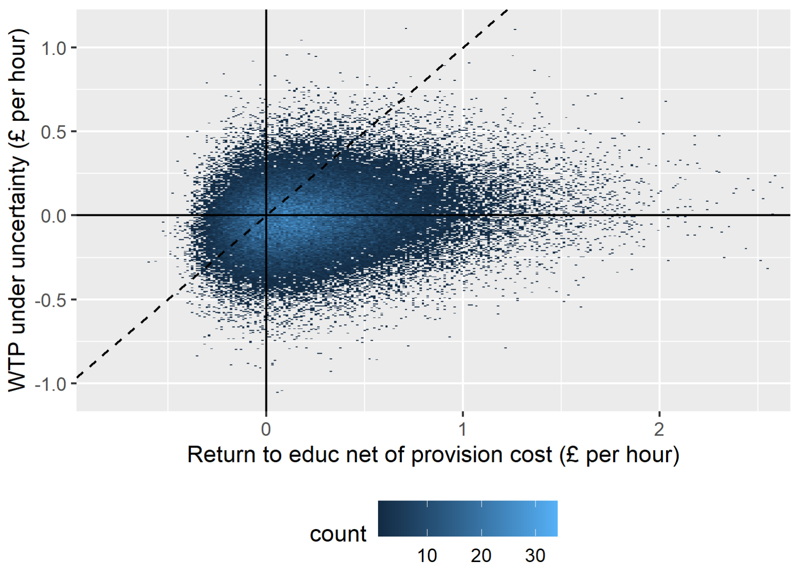
In practice, only 46.8% of workers do invest in education. Figure 1.24 plots the joint density of the return to education net of the cost of provision and the expected willingness to pay for higher education. The figure shows that going by the returns net of the cost of provision, most graduates receive a positive return to higher education, and only 18.3% of graduates would not have a positive return. On the other hand, most non-graduates are under-investing in education, and 72.5% of non-graduates could have benefited. In my model, this misallocation can be generated by two channels. First, the noisiness with which grades measures the actual labour market performance leads to an information imperfection which impedes choice. Second, heterogeneous net preferences, such as a high consumption value for university or particularly low costs due to family endowments, may lead some workers to find higher education value for money even with a lower labour market return. In table 1.10, I first summarise the share of workers who would invest in higher education, and then divide the population into workers who invest and receive a return greater than the mean preference benchmark, those who invest but receive a return lower than the benchmark, those who do not invest but would have received a return greater than the benchmark, and those who do not invest and would not have received a return higher than the benchmark. Then, I perform a counterfactual exercise, where I isolate each of these two channels in turn to analyse which factor contributes more to the mismatch between returns and choice (keeping the skill environment constant).

First, column 1 suggests that there is significant mismatch between wage returns in the

[47]Note however that if all workers selected into higher education, the prospective return under the resulting skill distribution might not sustain all of them selecting into higher education after wages fall.

[48]Note that the per-year tuition cap was 3000 pounds for the studied cohorts. Since then, this cap has tripled to 9250 pounds per year in 2021.

Figure 1.24: Simulated joint density of actual return net of cost of provision and expected WTP for higher ed



This figure plots the distribution of actual (conditional on unobserved ability) return to higher education minus the cost of provision and expected (conditional on the signal) willingness to pay for higher education in a simulated dataset generated from the model and the estimated parameters. Each square represents a set of binned values for WTP under perfect info and imperfect info, and the colour of the square denotes the count of points within the bin (with N = 100,000). The dashed line denotes the 45 degree line.

Table 1.10: Breakdown of investment choices and whether return exceeds mean costs benchmark, under four counterfactual scenarios

	Het + Noise channels	Het channel only	Noise channel only	No channels
Share choosing degree	0.468	0.433	0.997	0.768
Would invest, VfM	0.382	0.429	0.767	0.768
Would invest, not VfM	0.0856	0.004	0.230	0
No invest, VfM	0.386	0.340	0.00108	0
No invest/Not VfM	0.146	0.228	0.00187	0.232

This table summarises the share of workers in higher education, and a breakdown by whether they would experience "sufficiently" high returns to higher education. I use as a benchmark for "sufficiently high" the level of returns that would cover the cost of providing higher education, and is computed to be £0.679 in this setting. This is denoted as VfM, short for "value for money". These shares are computed for four scenario, the base scenario with both uncertainty of returns and heterogeneous preferences, and three counterfactual scenarios alternative with heterogeneous preferences only, uncertain returns, and one with neither.

system; of the 46.8% of workers who would choose university in the model, 38.2 percentage points, or 81.6% would generate a return greater than the cost of provision benchmark. Furthermore, 38.6 percentage points, or 72.5% of non-graduates, would have received a return higher than the mean preference benchmark but do not end up in higher education; this suggests that too few workers invest in education from a value for money point of view. Heterogeneous non-pecuniary preferences for higher education seems to drive under-investment in higher education, since even knowing the actual returns to higher education, 44.3% who would have received a net positive return would not have attended higher education. On the other hand, the uncertainty about true returns seems to drive over-education, leading workers who might not have benefited to invest in higher education. Thus, both channels are important in different ways in driving the mismatch between the net wage returns to education and workers' choices in the cohort studied.

Chapter 2

The Butcher, the Brewer, or the Baker: The Role of Occupations in Explaining Wage Inequality

Abstract

Google does not pay its canteen chefs the same premium as it does its software engineers; there is heterogeneity in pay premia within a firm. To study the implications of this observation, we estimate a two-way, worker-job, fixed effect model that builds upon the canonical model by Abowd et al. (1999) by allowing for within-firm heterogeneity. Estimating our model on UK administrative data, we find that even if workers, and firms, were identical, 25% of current log-wage variance would remain, due solely to heterogeneity between occupations. We further document that worker heterogeneity accounts for a relatively small proportion of log-wage variance, and that between-firm pay heterogeneity is more important in higher-wage occupations and larger labour markets.
JEL Codes: J01, J30, J31, J40, J42
Keywords: Wage inequality, occupations.

2.1 Introduction

Google does not pay its canteen chefs the same premium as it does its software engineers. As a result, the best engineers go to Google, but the best chefs work at Michelin star restaurants: workers match to jobs, not firms. This simple observation implies that the canonical model of wage decomposition pioneered by Abowd et al. (1999) (hereafter AKM) – which proposes a statistical model that explains log-wages as the sum of a set of worker fixed effects, firm fixed effects, and idiosyncratic errors – is misspecified because it implies that firms pay the same wage premium to all their workers[1].

In this paper, we argue that AKM models lead to the omission of an important dimension of inequality – inequality of wages across occupations. This omission leads to the overstatement of the degree of wage heterogeneity explained by individual differences, and the understatement of significant heterogeneity of earnings across occupations within firms. This implies that even if we could eliminate all wage dispersion between firms or all wage dispersion between individuals, there would still be substantial wage dispersion due to differences in occupations.

We extend the canonical framework to allow different firms to pay occupations different wage premia, by estimating a model with individual and *job* fixed effects, where jobs are defined as occupation-firm pairs. Our model embeds the AKM model as a particular case but is more flexible by allowing firms to pay different wage premia to workers in different occupations. To estimate such a model we use a two-stage approach suggested by Bonhomme et al. (2019) to overcome limited mobility bias (Andrews et al. (2012), Jochmans & Weidner (2019)), first clustering jobs, and secondly employing worker job-cluster fixed effect regressions.

Estimating worker-job wage variance decompositions using matched employee-employer

[1]The line of research pioneered by Card et al. (2013) has largely emphasised the role played by firms in explaining wage inequality (Song et al., 2019; Card et al., 2018b; Bonhomme et al., 2019; Alvarez et al., 2018). The point that the model used is misspecified is taken up by recent theoretical work by Haanwinckel (2020) and Bloesch et al. (2021), which both argue that workers who are in some sense more important to a firm are paid more within it.

data from the UK[2], we show that the share of overall log-wage variance due to individual heterogeneity is about 18%. This is significantly lower than what is typically found in AKM regressions. Card et al. (2013, 2016); Song et al. (2019); Alvarez et al. (2018); Bonhomme et al. (2020) all find that between 50% and 60% of log wage variance is explained by worker differences. We show this reduction relative to worker-firm decompositions is to be expected theoretically if more productive workers match to more productive occupations within firms. If the true model has varying job wage premia within a firm, then estimating an AKM model would cause this varying occupation premia to become an omitted variable, which will be loaded onto the estimated worker FE and firm FE. This results in AKM estimates of log-wage variance decompositions to attribute job premia heterogeneity to worker heterogeneity.

Consequently, our estimates of the role of job heterogeneity are larger than is typically found in the AKM context for firm heterogeneity. 38% of the log-wage variance can be explained by job heterogeneity, and 29% can be explained by positive assortative matching between workers and jobs. Both of these components are considerably larger than those found in firm-worker decompositions. In bias-corrected specifications, Bonhomme et al. (2020) find that firm heterogeneity accounts for between 5% and 13% across various European countries and the US, and that sorting between workers and firms accounts for at most 20%. Thus, previous work somewhat puzzlingly finds that sorting is strong even when the associated differentiation between firms is relatively less significant. Our results resolve this puzzle by showing both considerable job heterogeneity and considerable positive assortative matching.

We attempt to reconcile our results with those of the AKM literature by decomposing the variance in job fixed effects and in sorting into that which can be attributed to between-occupation variance and within-occupation variance. Variation within occupation across firms captures heterogeneity in firm-level pay premia, and a weighted average of this variation is

[2] We use the Annual Survey of Hours and Earnings (ASHE) data in this paper. The UK offers a particularly interesting setting for several reasons. First, although significant prior work has studied UK wage inequality (Faggio et al., 2010; Mueller et al., 2017; Machin, 2011; Lee et al., 2016; Rice et al., 2006; Schaefer & Singleton, 2020) to the authors' knowledge this is the first time large-scale employee-employer data has been brought to bear on the question using AKM style wage decompositions. Second, the UK typifies many debates surrounding wage inequality in western economies with spatial and occupational wage inequality squarely positioned in everyday political and social discourse.

conceptually comparable to the variance of firm fixed effects estimated in a standard AKM decomposition. When considering the variation in job fixed effects, we find that the between-occupation component is twice as important as the within-occupation component and that they account for 25% and 13% of overall log wage inequality respectively. Therefore, most variation is a between- rather than within-occupation phenomenon. We find similar effects when decomposing sorting with 20% explained by sorting between occupations and 8% by sorting within occupations. These results highlight the importance of occupations in explaining log wage inequality but do not completely discount the role of firms.

Armed with worker and job fixed effects, we test and reject the assumption in the canonical model that firms pay the same wage premia over occupations. We reject the null hypothesis at traditional levels of statistical significant, but also show that differences are of a magnitude that is economically significant. We then study the distribution of match effects between occupations and firms by computing the differences between our framework and a naive framework that uses linear worker, firm, and occupation fixed effects. We find that there are substantial match effects between occupations and firms; low-skill occupations have a negative match effects at most firms, while high-skill occupations have a positive match effect at the most productive firms. This result supports the results by Haanwinckel (2020) and Bloesch et al. (2021) that firms pay disproportionate large wage premia to certain, typically high-skill, workers, and disproportionately small wage premia to typically low-skill workers.

This result naturally points to the hypothesis that since wage premia are especially large for high-wage occupations, it is likely that occupational labour markets in those occupations ought to be characterised by more heterogeneity and sorting. To investigate this, we analyse the sources of wage variance in each occupation. Looking at differences across occupations we find that sorting between workers and firms is only a substantial component of wage variance within higher-wage occupations. In relatively lower-wage occupations, the sorting component is negligible. This evidence suggests that models which emphasise positive matching between workers and firms are more appropriate for characterising high-wage occupations. In lower-wage occupations where it appears workers and jobs are not so heterogeneous, search is likely to be less important.

Turning to heterogeneity across local labour markets and using a methodology first developed in Dauth et al. (2018), we show that high-paying jobs and workers co-locate in larger labour markets. This finding lends evidence to a large literature in spatial and urban economics which highlights agglomeration economies and/or sorting across space leading to a city wage premium (see D'Costa & Overman (2014), Glaeser & Mare (2001), and Rice et al. (2006) for example). We also find that the variance of wages is increasing in local labour market size and sorting, especially within-occupation, explains an increasingly large proportion of this variation with size. In the smallest labour markets, almost none of the wage variance is explained by within-occupation sorting (between individuals and firms within an occupation) whereas in the largest labour markets over 10% is. Differences in pay between firms conditional on occupation and worker are only a feature of large labour markets. To our knowledge, this result is new to the literature.

Our work relates to the large and active literature exploring the factors underlying wage inequality using the two-way fixed effects approach. This literature was initiated by Abowd et al. (1999) and Abowd et al. (2002), using French data. More recently, Card et al. (2013) used this framework to argue that increasing wage inequality in recent years was driven by increasing dispersion in firm wage premia. Their result was replicated in the American context by Song et al. (2019). Alvarez et al. (2018) find that the implementation of the minimum wage drove reductions in wage inequality in the Brazilian context. Recent econometric developments in addressing the limited mobility bias faced by two-way fixed effect regressions in relatively sparsely connected networks (Andrews et al. (2012), Jochmans & Weidner (2019), Kline et al. (2020), Bonhomme et al. (2019) and Bonhomme et al. (2020)) have suggested that the true share of log-wage variance explained by firm effects is potentially much smaller than the 20-25% found in the earlier literature. Dauth et al. (2018) use the AKM framework to study geographical income inequality in the German context.

We contribute to this literature by emphasising that if we allow firms to pay workers in different occupations varying premia, as suggested by Haanwinckel (2020) and Bloesch et al. (2021), we obtain qualitatively very different results from the AKM approach. In particular, while results from most AKM regressions find that wage inequality is largely due to individual heterogeneity, we find that this proportion is substantially smaller when we consider between-occupation moves. Instead, a lot of wage inequality is due to disparate pay between occupations, and this is particularly important in explaining wage inequality in small geographical labour markets. While there is less within-occupation between-firm variation in pay for smaller labour markets and low-wage occupations, there are large match effects between high-wage occupations and high-premium firms, and thus within-occupation variation is more important for higher-paying occupations. Partly due to these occupations typically being located in larger labour markets, and partly due to a market size effect, we also find that within-occupation heterogeneity between firms only matters substantially in larger geographical labour markets.

The rest of our paper proceeds as follows. Section 2.2 describes the specification and identifying assumptions underlying our model. Section 2.3 discusses the UK data that we use. We present our main results relating to the log wage variance decomposition in section 2.4. Section 2.5 considers differences across occupations. We discuss the implications of our model for wage inequality across space in section 2.6. Section 2.7 concludes.

2.2 Econometric model

In our econometric framework, we aim to disentangle log-wage variation into the components due to worker specific heterogeneity, job specific heterogeneity, and that due to the sorting of workers into jobs. For each individual $i \in \mathcal{I}$ in each period $t \in \mathcal{T}$ we observe the firm $F(i,t) \in \mathcal{F}$ and occupation $O(i,t) \in \mathcal{O}$ they work in as well as the wage w_{it} received. We denote by $J(i,t)$ the *job* an individual i is employed doing, in period t where j is a firm-occupation pair i.e. $J(i,t) \in \mathcal{J} = \mathcal{F} \times \mathcal{O}$. We assume that log wages $\ln(w_{it})$ can be described by the sum of worker, α_i, and job, $\psi_{J(i,t)}$, components[3] as well as an idiosyncratic error as shown in equation 2.1.

$$\ln(w_{it}) = \alpha_i + \psi_{J(i,t)} + \varepsilon_{it} \tag{2.1}$$

This setup follows the literature started by Abowd et al. (1999) and continued by others such as Card et al. (2013) and Song et al. (2019). In this model person fixed effects α_i can be interpreted as a combination of skills and other factors that are awarded equally across jobs j. Similarly, ψ_j can be interpreted as a pay premium offered to all workers in job j irrespective of unobserved worker-specific characteristics. Note that the usual worker-firm models (as opposed to worker-job) use ψ_f instead of ψ_j and interpret it as a pay premium offered to all workers at a specific firm. This means that, for example, software developers, administrators, managers, cleaners, and chefs at Google all receive the same log-additive pay premium.

Our model nests this as a special case where all occupations within a firm have the same ψ_j, but in general is significantly more flexible by allowing this premium to vary within a firm by occupation. For example, this would allow Google to pay its software developers a higher premium than its chefs and the Ritz to pay a higher premium to its chefs than its software developers. Given the nested structure we can test formally whether firms do indeed offer the same wage premium across all occupations and do so in section 2.4.3 finding evidence of statistically and economically significant varying premia.

The fixed effects in the AKM model are identified by *movers* between firms. In the same way, the job fixed effects are identified by the same individual moving from one job to another and noting how their wages changed due to the move, and individual fixed effects were identified by observing workers' wages between jobs relative to other workers in those jobs. Denote by y the stacked vector of individual wages, $A = [a_1, \cdots]$ the design matrix of individual indicators, $J = [b_1, \cdots]$ the design matrix of job indicators and ε the stacked vector

[3] Throughout we additionally condition on a set of time varying worker covariates X_{it} such as age, these have been omitted in this discussion for notational brevity.

of error terms. Then we can write equation 2.1 as $y = A\alpha + J\psi + \varepsilon$ or more succinctly as $y = Z\beta + \varepsilon$ where $Z = [A, J]$ and $\beta = [\alpha, \psi]$. We will estimate equation 2.1 by OLS to find $\hat{\beta} = \beta + (Z'Z)^{-1}Z'\varepsilon$. This result highlights the two main hurdles facing identification. First note that implicit in the above derivation is the invertibility of $Z'Z$. This is not the usual innocuous assumption; in general $Z'Z$ is not invertible in this set up if there are *unconnected* individuals or firms. That is, $Z'Z$ is only invertible in the subset of the data representing the largest connected set. In our data the largest connected set covers over 90% of observations and will be the sub-sample that we use for estimation (this is summarised in table 2.4).

The specification in equation 2.1 allows a simple decomposition of the observed variance of log wages as given in equation 2.2 below.

$$\mathbb{V}[\ln(w_{ijt})] = \underbrace{\mathbb{V}[\alpha_i]}_{\substack{\text{Variance due to individual} \\ \text{heterogeneity}}} + \underbrace{\mathbb{V}[\psi_j]}_{\substack{\text{Variance due to job} \\ \text{heterogeneity}}} + \underbrace{2 \cdot Cov[\alpha_i, \psi_j]}_{\substack{\text{Variance due to workers} \\ \text{sorting into jobs}}} + v \tag{2.2}$$

This shows how, for each period, wage variance can be decomposed into four components. $\mathbb{V}[\alpha_i]$ captures the variance due to individual characteristics, and $\mathbb{V}[\psi_j]$ similarly captures the variance due to job characteristics. Finally, $2Cov(\alpha_i, \psi_j)$ captures the component of the variance which is due to sorting or matching of individuals to jobs, and v captures unexplained variance.

2.2.1 Limited mobility bias

AKM type decompositions suffer from limited mobility bias which was recently found to be qualitatively and quantitatively of importance (Bonhomme et al. (2020)). In our dataset for example, although we have almost 1.4 million observations, equation 2.1 requires estimating 303,000 worker fixed effects and 282,000 job fixed effects. However, the problem is significantly worse than it at first appears. This is because intuitively job fixed effects are only identified from *movers* between jobs — of which there are even fewer. While this doesn't lead to biased estimates of worker or job fixed effects, it does lead to biased second order (or higher) moments calculated from the resulting estimated fixed effects[4].

In the main body of this paper, we follow the two-step approach suggested in Bonhomme et al. (2019) and Bonhomme et al. (2020) to overcome the limited mobility bias, where we first reduce the dimensions of the problem by classifying jobs into L job classes using a k-means algorithm[5]. This gives us a very connected network of workers and job classes, on which we can estimate a two-way linear regression on workers and job classes and related variance terms with little bias. In the first step, we cluster jobs into L classes using an unsupervised machine learning algorithm, k-means clustering based on 20 points (percentiles) of the job wage distribution. We find that these thirty clusters capture some 85% of the underlying variation in job wages; this is encouraging evidence of the success of the dimensionality reduction with minimal information loss[6]. In the second step we then estimate equations of the form given in 2.1 using job-class rather than job fixed effects.

2.2.2 Endogeneity and specification tests

To uncover unbiased estimates of the fixed effects, we require that $\mathbb{E}[Z'\varepsilon|Z] = 0$, the classic exogeneity assumption in OLS. The validity of this assumption for the classic AKM framework was discussed in detail in Card et al. (2013), and in our paper, we focus on what changes in our context relative to the standard AKM model. As in Card et al. (2013), we focus on whether the assignment of workers to jobs satisfies exogeneity, i.e. $\mathbb{E}[J'\varepsilon|Z] = 0$, and consider the sufficient condition that $\forall j, Pr(J(i,t) = j|\varepsilon) = Pr(J(i,t) = j)$.

The assumption made is not violated if the labour market is characterised by sorting on individual fixed effects and job fixed effects, i.e. if workers move to higher paying jobs or firms sort to more productive workers. Rather, there are two ways in which the assumption

[4]The econometrics of this bias are reviewed in the appendix in section D.

[5]K-means is an unsupervised machine learning algorithm that classifies data into K groups by minimising each within-group sum of square differences from the corresponding group mean.

[6]All results presented here are not sensitive to the number of clusters used. We have replicated very similar findings using 10, 20, 50, and 100 clusters.

could be violated. First, the assumption is violated if workers sort to jobs on the basis of a worker-job match-specific characteristic not captured by α_i and $\psi_{J(i,t)}$. Relative to the AKM specification, this is more of a concern, as the skills necessary for different occupations are likely to be non-unidimensional (as would be implied in our model) (see e.g. Lindenlaub (2017) and Lise & Postel-Vinay (2020b) for work arguing for the importance of multi-dimensional skills). Thus, there may be dimensions of individual skill not captured by the fixed effect that generate match-specific surplus at particular occupations which are pushed to the error term.

A second possible problem is that temporary variation in wages may be correlated to the job that workers perform. A concern of this type in Card et al. (2013) is that the statistical model is incompatible with models of the labour market where workers move to jobs due to high transitory wage offers not related to the firm fixed effect[7]. In our model, an additional problem of this kind is if temporary occupation-specific productivity shocks lead to substantial movement between occupations in the period.

An upshot of both concerns is that wage changes due to moves in one direction should be larger than wage changes due to moves in the other. If workers sort to jobs on the basis of elements not captured in either the worker fixed effect nor the firm fixed effect, we should expect the relative wage change of a worker moving from a low fixed effect job to a higher fixed effect job to exceed in magnitude the relative wage change of workers moving the other way. On the other hand, if the exogeneity assumptions holds, the average wage change of movers from a less productive job to a more productive job should be equal to the negative of movers from the more productive job to the less productive job.

We check for this endogeneity in appendix B, using methods developed in Card et al. (2018) in the context of the AKM model. The main idea behind these tests is to check if wage changes in moves of workers between job clusters is symmetric. We find that wage changes between estimated job clusters are approximately symmetric.

Finally, another possible criticism of this approach is that the two-way fixed effect re-

[7]A prominent model with such a mechanism is Postel-Vinay & Robin (2002).

gression imposes an inappropriate log-additive functional form on firm and job fixed effects. The log-additive structure we impose on the data could mask important heterogeneity, mechanisms or disallow potentially relevant theoretical channels. However, we feel relatively assured that, even when taking a coarse job clustering approach[8], most of the variation of interest is captured by our approach. This is due to two reasons. First, a relatively small part is played by unexplained variation in wage decompositions (roughly 8%). Second, we follow the approach in Card et al. (2013) and show that the mean of the residuals of the two-way FE regression are near zero on average, and by job and worker cells. This exercise and its results are summarised in appendix C. We find that the specification performs well, except for very low-skill workers at very low wage premia jobs. This could be due to institutional factors like minimum wage that are imperfectly captured by this model. This finding is qualitatively similar to what other authors find when analysing the suitability of linearity in the context of more standard AKM models (e.g. Card et al. (2013); Bonhomme et al. (2019)).

2.3 Data and empirical setting

We apply the empirical specification discussed in section 2.2 to matched employee-employer data from the United Kingdom. We study the UK for two main reasons. First, to the best of our knowledge, this is one of the first studies (other than Schaefer & Singleton (2020)) to bring employee-employer data and AKM-like decompositions to bear on the question of UK wage inequality. Second, the UK typifies contemporary debates surrounding wage inequality in aggregate as well as across space, common in developed economies.

Our main data source for this paper is the UK Annual Survey of Hours and Earnings (hereafter ASHE) (Statistics (2020)). The ASHE is a matched employee-employer data set collected annually by surveying employers and comprises of a 1% sample of the government tax agency's job records. The survey, which is administered in April each year, records

[8]We check the robustness of the results to various numbers of clusters ($K = 30, 50, 100$) and clustering algorithms and find no difference.

information about the identity of the establishment, its location and the industry classification of its primary business, the worker's identity, sex, and home address, and details about the job itself, including earnings, hours, and occupational classifications. The survey is administered by the UK Office for National Statistics (ONS) and response to the ASHE is compulsory. The survey provides a panel data set as the same individuals are tracked over time, and as filling in the ASHE survey is a legal requirement of the employer attrition is very low. The ASHE can be compared to other data sets used in the AKM literature, such as the Integrated Employment Biographies data in Germany (used by Card et al. (2013)), the Master Earnings File in the USA used in Song et al. (2019), and a mix of microdata sets including the ANST in Sweden used in Bonhomme et al. (2019). We supplement this data with information on firm size, revenue and firm industry from the Business Structure Database (Statistics (2022)), an administrative dataset which tracks almost all businesses in the UK.

Details of our data cleaning process can be found in appendix A. Table 2.1 below provides summary statistics on the entire data set and the largest-connected set used in our main clustered worker-job fixed effect specification. We lose data on about 50,000 workers and 15,000 firms in the largest connected set, or about 16% of all unique workers and 10% of all unique firms in the data set. This amounts to 3.5% of the sample in the second period in total. Reassuringly, the hourly wage moments in the two periods do not change significantly between the connected set and the full sample. Table 2.4 in the appendix summarises how much of our data is retained in the largest connected set for various specifications.

Table 2.1: Summary statistics

	02-10	11-19
All observations		
Total observations	1398098	1598919
No. unique workers	303,122	339,697
No. unique firms	127,014	156,178
No. unique firm-occ pairs	282,251	335,298
Average log hourly wage	2.21	2.18
SD log hourly wage	0.53	0.51
Largest connected set of main specification		
Total observations	1346163	1542383
No. unique workers	251,187	283,161
No. unique firms	112,817	139,533
No. unique firm-occ pairs	258,676	308,095
Average log hourly wage	2.22	2.19
SD log hourly wage	0.53	0.51

Notes: This table shows summary statistics from the ASHE data for two time periods: 2002-2010 and 2011-2019 as well as for the full data and the largest connected set.

2.4 Log wage decomposition

2.4.1 Overall wage variance decomposition

We present the main results of our decomposition in table 2.2 below, as well as several alternative specifications that have been produced in the literature[9]. In column (1), we produce estimates from the classic AKM two-way regression with worker and firm fixed effects. We find results that are qualitatively similar to those reported in other countries: a small and negative covariance between firm and worker fixed effects and 24% of log-wage variance due to firms' fixed effects, with most of the variance in log-wages being due to worker fixed effects. Column two shows the two-way fixed effect regression with worker and firm-cluster fixed effects, where firms are clustered into 30 firm clusters using a k-means algorithm on the basis of their wage distributions. In this specification, sorting becomes positive and increases at the expense of the contribution of firms and to a lesser extent workers. This again serves to highlight the importance of accounting for limited mobility bias.

[9]We compare these estimates against estimates in the literature in table 2.5 in appendix section F

Table 2.2: Comparison of different specifications (2011-2019)

Proportion of overall log-wage variance	Basic-firm	Cluster-firm	Cluster-Occ	Cluster-Job
Worker component	0.64	0.54	0.56	0.18
Firm component	0.24	0.10		
Occupation component			0.06	
Job component				0.38
Sorting component	-0.02	0.19	0.20	0.29
Residual component	0.08	0.10	0.11	0.08
Adjusted R2	0.89	0.88	0.87	0.90

Notes: This table presents log-wage variance decompositions for four specifications: (i) simple AKM, (ii) bias-corrected AKM, (iii) worker-occ twfe, and (iv) worker-job twfe (main specification). We use the ASHE data over the period 2011-19 for these results.

In column (3)[10], we consider a specification where log earnings are decomposed into worker and occupation-specific effects. This regression serves two purposes. First, it is of interest in and of itself and shows a remarkable similarity to the cluster firm regression in column (2). A large share of the log-wage variance appears to be due to worker heterogeneity, with between-occupation heterogeneity accounting for only 6% of the total variance. Second, it serves to highlight that the differences between (low mobility adjusted) traditional decompositions in column (2) and those we suggest in column (4) are not just driven by the inclusion of occupation fixed effects but rather by specifically the inclusion of firm-occupation fixed effects.

Our preferred job-level analysis on the other hand in column (4) produces different results in several ways[11]. There are three qualitative differences in these results relative to the results using the traditional AKM approach. First, the share of log-wage variance attributed to differences between workers falls to only about 18% of overall wage variance, from 64% in the standard AKM decomposition. As discussed in section 2.2, this result arises due to the misspecification of the wage process in traditional AKM models. Intuitively, where the job fixed effect is not included in the regression in column (2), it shows up as part of the omitted variable bias in the estimate of α_i in the misspecified model if there is residual

[10]As there are only roughly 367 occupations, there is no limited mobility issue in individual-occupation fixed effect regressions. The results do not differ when clustering is used.

[11]We check that the job-level estimates are related to firms' revenue in appendix H, suggesting that these estimates can plausibly be interpreted as job-level wage premia.

covariance between α_i and $\psi_{J(i,t)}$ after accounting for the firm workers work at. This bias component is counted as part of the individual fixed effect in the variance decomposition of the misspecified model and must be larger than the actual variance of α_i if the covariance between the individual fixed effect and the residualised, weighted sum of job fixed effects across an individual's career is positive. This explanation implies that, in practice, workers who work at high-wage occupations tend to work at high-wage occupations for most of their career, and vice versa[12]. We discuss this analysis of the differences between our results and the canonical AKM results in appendix E, and show that in general, if there is positive sorting of workers to jobs conditional on firms, then in the misspecified model, the variance of worker fixed effects will be over estimated and that additionally the variance of firm fixed effects in worker-firm specifications will be incorrectly estimated (although the bias is of ambiguous sign).

These results paint a fundamentally different picture of wage inequality than those from the AKM regressions. In the AKM story, the substantial amounts of wage inequality are primarily because workers are very different, due to different degrees of ability or human capital accumulation. In our case, we find that when we make use of information from moves by workers across occupations, we see that differences between workers are much smaller than previously suggested. As a caveat, there is substantial heterogeneity in the importance of the individual worker differences in explaining wage dispersion between different occupations. As we discuss in section 2.4.3, wage dispersion in high-wage occupations is more readily explained by individual differences than in low-wage occupations. But, our results suggest that policies which reduce job-level differences in pay can be effective in reducing overall income inequality.

Second, we find that job heterogeneity and sorting both contribute significantly to wage inequality, specifically 38% and 29% respectively. Therefore, whereas previous specifications find low firm-level complementarities (and indeed we find low occupation level complementarities in column (3)) we find large job level complementaries in wages. We also find a larger impact of sorting of workers into jobs, which explains 29% of the variation, validating the

[12] We provide some direct evidence of this in appendix G.

original intuition that workers match to jobs, not firms.

A strange qualitative feature of results from AKM regressions, corrected for low-mobility bias, is that they have typically found low firm heterogeneity (10% in our sample) but a relatively larger share due to sorting between firms and workers, implying high degrees of sorting to firms when there is much smaller heterogeneity between firms. This feature is not present in our findings; we find that there is substantial job-level heterogeneity, and consequently, substantial degrees of sorting between jobs and workers. In the next section, we present evidence that suggests that a possible explanation of the differences between our specification and the standard AKM specification is that the latter approach only captures within-occupation job heterogeneity while we also capture heterogeneity between occupations.

2.4.2 Between and within-occupation log-wage decompositions

A natural question in our framework is whether heterogeneity in the job fixed effect is due primarily to differences between firms or between occupations. To answer this, we can further decompose the variation of job fixed effects, and the covariance between job and worker fixed effects, into between- and within- occupation components, using the law of total variance. The index of occupations partitions the data, and let $\mathbb{E}_o(\cdot)$, $\mathbb{V}_o(\cdot)$, and $Cov_o(\alpha_i, \cdot)$ denote the mean, variance and covariance with α_i of the random variable within the partition indexed by o. Equation 2.3 presents the decomposition of job variance into between and within occupation components, while equation 2.4 does the same for the decomposition of covariance.

$$\mathbb{V}[\psi_j] = \underbrace{\mathbb{V}[\mathbb{E}_o[\psi_j]]}_{\text{Between } o} + \underbrace{\mathbb{E}[\mathbb{V}_o[\psi_j]]}_{\text{Within } o} \tag{2.3}$$

$$Cov[\alpha_i, \psi_j] = \underbrace{Cov[\mathbb{E}_o[\alpha_i], \mathbb{E}_o[\psi_j]]}_{\text{Between } o} + \underbrace{\mathbb{E}[Cov_o[\alpha_i, \psi_j]]}_{\text{Within } o} \tag{2.4}$$

Results of this decomposition are reported in table 2.3. Table 2.3 also splits the sample into two periods and compares the decomposition across both. We find no significant differences in the drivers of inequality in the UK over time. This is in contrast to Card et al. (2013) who find, in Germany, that increases in firm heterogeneity and sorting explain changes in wage inequality.

We find that about two-thirds of the variation in job effects can be explained by between-occupation variation and a third can be explained by within-occupation variation. The latter part, an occupation-weighted summary of wage variation between firms that is not explained by the occupation title of the worker, is about 13% in our data, which is close to the 10% estimated for the variance of firm fixed effect in the main AKM specification. This implies that the standard AKM approach misses most of the variability in job effects, which is due to differences in occupations across jobs rather than firm heterogeneity within occupations.

Similarly sorting between workers and jobs can be decomposed into within and between occupation components. The between occupation component is given by the covariance between mean worker type within an occupation and mean job fixed effect within that occupation. Whereas the within occupation component is given by the expected covariance between workers and jobs in a given occupation, across all occupations. We find that over two-thirds of the variation due to sorting can be explained by sorting between occupations, which reflects the sorting of higher-paid workers into higher-paying occupations, whereas only 9% of the log-wage variation can be explained by sorting of workers within occupations to higher-paying jobs.

Table 2.3: Variance decomposition from jobs model over time

Component of total log-wage variance	Period 2002-2010	Period 2011-2019
Variance of worker FE	0.20	0.18
Variance of job FE	0.34	0.38
Between-occ variation	0.23	0.25
Within-occ variation	0.11	0.13
Correlation between worker and job FE	0.32	0.29
Sorting between occupations	0.23	0.20
Sorting within occupations	0.09	0.08
Adjusted R2	0.90	0.90
RMSE	0.17	0.16

Notes: This table decomposes the log wage variance according to our job-worker two-way fixed effects model, for two period, 2002-10, and 2011-19. We further decompose the job variance and covariance components into between and within-occupation parts using the law of total variance.

Our results thus offer a different perspective on log wage inequality than most of the literature. Results from standard AKM regressions have tended to primarily emphasise the role of individual heterogeneity and secondarily the role of firms. In contrast, our results show that the occupational structure plays a large role in the story of wage inequality, and most of the variance due to job heterogeneity and sorting of workers to jobs can be attributed to between-occupation heterogeneity and sorting of workers to occupations. Wage heterogeneity between firms within occupations is quantitatively less important, and we show in a subsequent section that it is relatively more important in high-wage occupations than in low-wage occupations.

2.4.3 Evidence on varying wage premia across occupations within firms

Armed with fixed effects estimates, we can turn to test the idea that motivated this paper: that firms offer different wage premia to workers in different occupations[13], One reason for this may be that workers in some occupations are in some sense more central or important

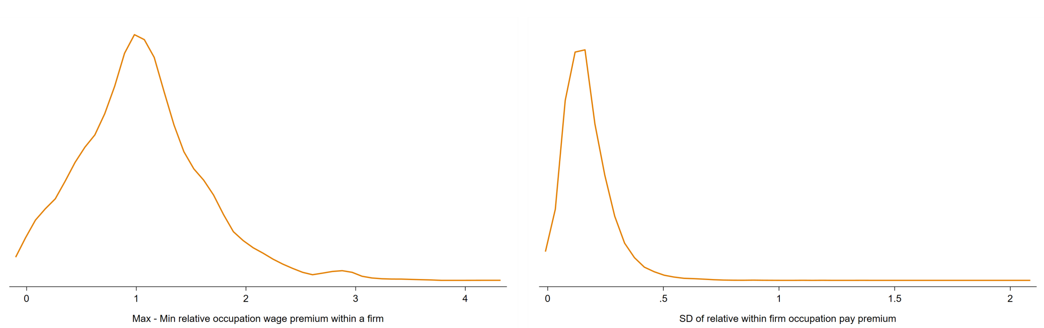
[13]In appendix H, we regress the estimated job fixed effects on firm characteristics to check whether the rent-sharing interpretation is appropriate and conclude that it is. We further check that the relation between the job FE and firm revenue is weaker for public firms than for private firms without collective bargaining agreements, and document heterogeneity across occupations. We leave a fuller analysis with firm profit data, which is sadly not available in the Business Structure Dataset (BSD) data set, for future work.

to a firm's profit[14]. In this section, we present empirical evidence to suggest that firms offer varying wage premia across the occupations they employ. We take two approaches, first we analyse the estimated job fixed effects ψ_J and show that these vary within-firm across occupations. Second, we consider to what extent a naive additively linear specification in firm and occupation effects would be misspecified relative to our more general specification, which is equivalent to testing for occupation-firm specific match effects.

First, we plot the within-firm cross-occupation job fixed effect variance and range having removed aggregate occupation variation, in figure 2.1. This figure shows significant dispersion in wage premia across occupations within firms, even after accounting for aggregate occupation variation. In the left-hand side panel, the within-firm cross-occupation range is shown, and in the right-hand panel the within-firm, cross occupation variance. These statistics are calculated from the residuals of a regression of estimated job fixed effects on occupation and firm effects and therefore reflect varying wage premia within a firm accounting for aggregate occupational differences.

Figure 2.1: Distribution of the variance of job fixed effects across occupations within firms

(a) Cross-occupation range (b) Cross-occupation variance



Notes: This figure shows the within firm, cross occupation dispersion in wage premia removing aggregate cross-occupation effects. Both figures show density plots over firms, in the left-hand panel the cross occupation range (max-min) is displayed and in the right-hand panel, the cross occupation variance is displayed. The sample is restricted to firms with 10 or more employees. To calculate the statistics graphed estimated job fixed effects are residualised on firm and occupation effects.

Figure 2.1 shows significant variation, and therefore gives evidence to suggest that job wage

[14]In the model of Haanwinckel (2020), firms require tasks in different intensities for production, and workers in different occupations produce units of different tasks. In Bloesch et al. (2021), workers in different occupations can receive different wage premia because they have different hold-up power in the production of the firm.

premia do vary across occupations within firms. However, it is unclear from the figure alone whether this variation is statistically significant. Therefore, we formally test the differences graphically observed, by performing a test of the hypothesis that the variance of job fixed effects within the firm is no greater than five percent of the overall variance in job effects[15]. Performing this test we find that we can reject the null hypothesis for 49% of firms and therefore conclude that there is significant within-firm, cross-occupation, wage variance.

An alternative hypothesis is that while firms may offer an occupational premium to workers in different occupations, they also impose a constant firm premium on all workers within a firm. That is, we can approximate the job fixed effect estimated in our model by an additive specification given by equation 2.5, where $occ(i,t)$ denotes the occupation worker i is in at time t and $F(i,t)$ denotes the firm the worker is in at time t.

$$\psi_{J(i,t)} = \phi_{occ(i,t)} + \chi_{f(i,t)} \tag{2.5}$$

To test this interpretation, we estimate regressions of the form $\ln(w_{it}) = \alpha_i + \tau_{O(i,t)} + \xi_{F(i,t)} + \varepsilon_{it}$, where τ_O are occupation fixed effects and ξ_F are firm fixed effects[16]. Then we can study the difference between these estimated fixed effects and our previously estimated job fixed effects $\psi_{J(i,t)}$. We interpret this difference as an occupation-job specific match effects on the offered wage premium. If firms do not pay differential wage premia across occupations, $\Delta = \psi_J - (\tau_O + \xi_F)$ will be small and not systematically related to any firm or occupation characteristics. A positive match effect of e.g. 0.1 implies that the premium offered by the particular job is roughly 10% greater than the premium implied by an additive occupation and firm fixed effect (and vice versa).

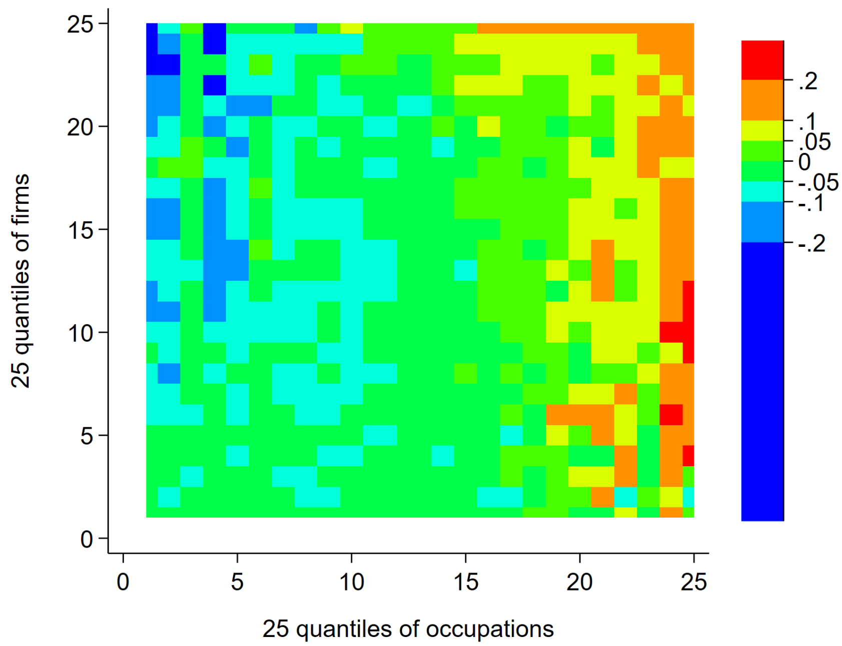
In figure 2.2, we produce a heat map of the difference between the job fixed effect in our

[15]To be precise we test whether within-firm, cross-occupation variation in job fixed effects is statistically significantly greater than 5% of the overall variance in job fixed effects across all firms and occupations using a Chi-Squared test (Snedecor & Cochran, 1989). We find that we cannot reject the null hypothesis at the 5% level in 49% of firms with more than 10 employees in our data.

[16]This specification is similar to the one examined in Torres et al. (2018).

main model and the sum of the occupation and firm fixed effects in the auxiliary model, that is of match effects. We plot this for 25 quantiles of occupations ranked by the average log hourly pay in the occupation and 25 quantiles of firms ranked by the average log hourly pay in the firm. Thus, a positive number (e.g. 0.05) can be interpreted as follows: if we naively replace the job fixed effect by the sum of the two fixed effects in our predicted values in our main model, we would underestimate the hourly wage by 5%. Green tiles represent quantile combinations where the error is less than 5% in either direction.

Figure 2.2: Contribution of worker-job match effects by firm and occupation mean pay quantiles



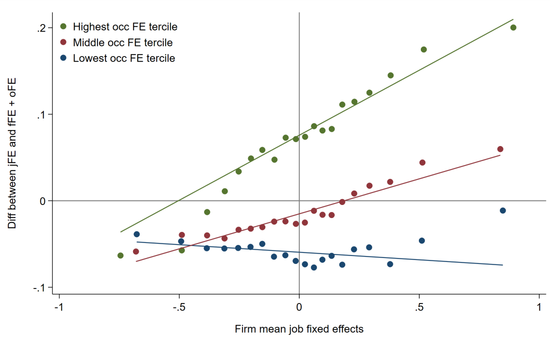
Notes: The figure shows a heatmap of the difference between the estimated job fixed effect in our main specification, and the sum of a firm fixed effect and an occupation fixed effect in an auxiliary specification, which we interpret as match effects between occupation and firms. Green squares imply that if the job fixed effect were replaced by the additive firm and occ fixed effects, the resulting predicted log wage would be within 5% of the true predicted log wage with job FE. Red implies that using the additive specification would underpredict the true wage by over 10% and blue implies that overprediction by over 10%. The x-axis indexes 25 quantiles of occupations, sorted by the mean wage within the occupation, and the y-axis indexes 25 quantiles of firms, sorted by the mean wage within the firm.

This figure shows that the linear form performs quite badly, especially at the extremes for each occupation. The linear combination would severely overestimate the wages of workers in low type occupations at all firms (as these have a negative match effect), and severely underestimate the wage of high type occupations at all types of firms. The problem seems especially acute for high wage occupations at low wage firms and low wage occupations at high wage firms, leading to mistaken predictions by over 20%. This leads us to conclude that the proportionate firm fixed effect is inappropriate for capturing the relation of pay premia to occupation and firm characteristics.

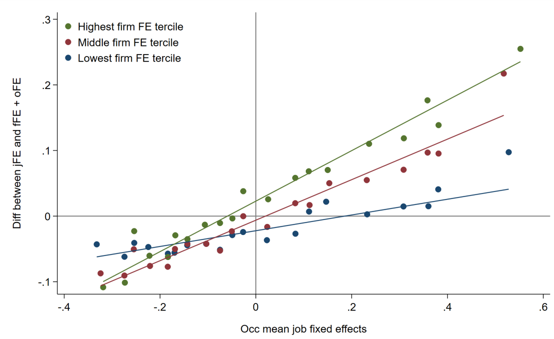
To analyse this dynamic further, we produce two complementary figures, first plotting the relation between the estimated residual against the mean log hourly pay at the firm for three terciles of occupations sorted by mean occupation wage, and the relation between the estimated residual against the mean log hourly pay at the occupation for three terciles of firms sorted by mean firm wage.

Figure 2.3: Plotting estimated match effects

(a) Match effect by firm average job fixed effects for three occupation mean wage terciles



(b) Match effect by occupation job fixed effects for three firm mean wage terciles



Notes: This figure plots the distribution of the difference between the estimated job fixed effect in our main model, and the sum of estimated firm and occupation fixed effects in an auxiliary regression, that is, match effects. Panel 2.3a plots the differences by firms, sorted by the mean job fixed effect within the firm for three terciles of occupations, while panel 2.3b plots the differences by occupation, sorted by mean occupation job fixed effect, for three terciles of firms. par

Figure 2.3 shows binned scatterplots of match effects which emphasise the importance of match effects for job-level pay premia[17]. Panel 2.3a shows that match effects increase from negative to positive as mean job fixed effects within firms increase. This relationship, however, is mediated by occupation terciles. Low-wage occupations have a negative match effect at all

[17]The scatterplots can be interpreted as similar to the data in figure 2.2, 'sliced' into a 2-D graph by reducing the dimensions of one of the axes.

firms, regardless of the firm-level pay premium. On the other hand, high-wage occupations have a positive match effects at most firms, even those who offer on average below-average wages. Panel 2.3b shows the analogous results when plotting match effects over occupation average job fixed effects and cutting by firm occupation fixed effect terciles. We find that low occupation fixed effects occupations consistently have negative match effects and that although these increase with occupation average job fixed effects it does so faster for higher firm fixed effect firms.

Finally, we can formally test for the difference between the naive model: $\ln(w_{it}) = \alpha_i + \tau_{O(i,t)} + \xi_{F(i,t)} + \varepsilon_{it}$ and our more flexible specification $\ln(w_{it}) = \alpha_i + \psi_{J(i,t)} + \varepsilon_{it}$ using an F-test. This gives a F statistic of 3.93 and 3.24 when tested against the naive model with and without including occupation fixed effects, in both cases the critical value at the 5% level is very close to 1, so we resoundingly reject equality between models.

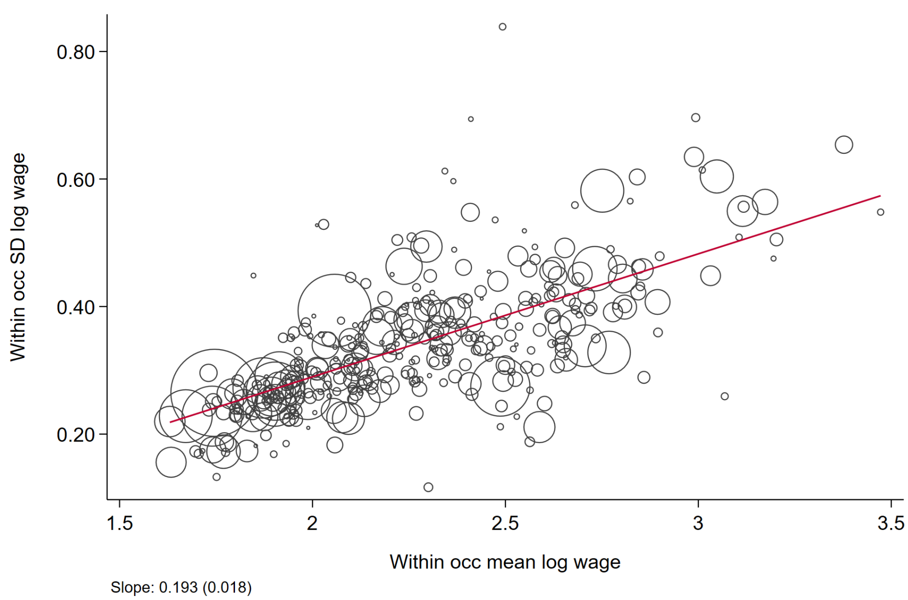
Our results support papers that argue that the standard AKM approach is misspecified, including Haanwinckel (2020), and Bloesch et al. (2021). Haanwinckel (2020) argues that firms should pay larger premia to workers who perform tasks which the firm use more intensively, and in a similar vein, Bloesch et al. (2021) show that workers which have more hold-up power within a firm should be paid more. Panel 2.3a shows that these jobs are likely to be more highly paid and skilled occupations. Assuming a proportional firm wage premium leads to the systematic overestimation of the pay of low-wage occupations and the underestimation of the pay of high-wage occupations at high productivity firms.

2.5 Differences in the log-wage decomposition across occupations

An implication of the variation of wage premia across occupations within firms is that the sources of wage inequality within occupations may be quite heterogeneous. An advantage of our more flexible specification is that we can investigate the extent to which the overall wage decomposition presented is representative of the wage variance in individual occupations. Within an occupation, variation in the job fixed effect reflects differences in wage premia offered by firms, and we can more straightforwardly compute decompositions according to equation 2.2 within occupations. We find that the overall wage decomposition however disguises substantial heterogeneity in the wage decomposition in each occupation.

We start with the stylised fact that in the UK labour market, the variance of log hourly wages is co-increasing in the mean log hourly wage within an occupation. This is shown in figure 2.4, which plots a scatter plot of occupations with the mean log hourly wage on the x-axis and the standard deviation of log hourly wage within that occupation on the y-axis. The size of the bubble represents the number of workers in that occupation in the data. There is a clear approximately log-linear relationship between the mean and variance of log earnings, and dramatic outliers to this trend tend to be fairly small occupations.

Figure 2.4: Within occupation mean vs SD of wages

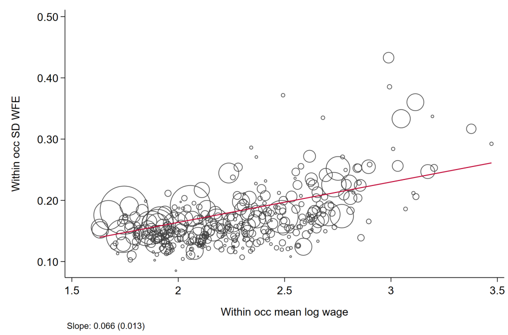


Notes: This figure plots the mean log-wage (x-axis) and the standard deviation of log-wages (y-axis) within each occupation for all occupations in the data. Each bubble represents a particular occupation and the larger the bubble is, the more workers there are in the occupation. A red linear trend line is drawn, with slope 0.193.

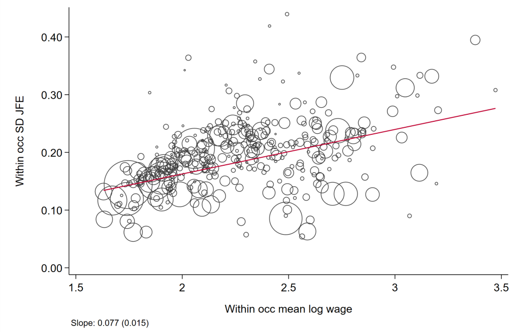
What accounts for the greater wage variance in high-wage occupations relative to lower wage occupations? To understand this, we decompose the variance in each occupation into an individual component, a firm component and a sorting component. We plot the log-variance due to each component for each occupation in figure 2.5, panels (a) to (c). In panel 2.5d, we plot the variance of the components as a share of total log-wage variance within the occupation.

Figure 2.5: Variance of worker FE, job FE and correlation between the two components

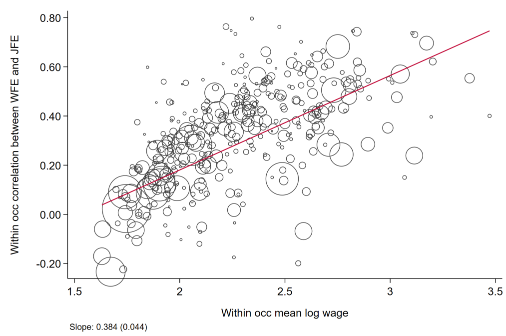
(a) Worker component



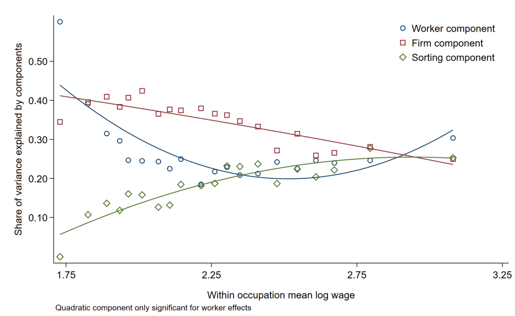
(b) Job component



(c) Sorting component



(d) All together



Notes: Panel 2.5a plots the standard deviation of worker fixed effects in each of the 368 occupations in period 2 of the data (2011-19). Panel 2.5b plots the standard deviation of within-occupation job fixed effects within each occupation; since a job is identified as an occupation-firm pair, the variation is from firms offering different wages for the same occupation. Panel 2.5c plots the correlation between worker and firm fixed effects for each occupation. The occupations are ordered by the mean log-wage within the occupation on the x-axis.

The variances of all three components are increasing in the within-occupation mean wage, but at different rates. In particular, sorting is zero or even negative for many low-wage occupations and does not contribute to within-occupation wage variance for those occupations. However, it becomes as important as individual and firm heterogeneity for high-wage occupations. The variance of individual worker fixed effects also increases exponentially; it is relatively low for low and medium wage occupations but is quite high for the highest wage occupations. Thus, the drivers of wage variance appear to be quite different for low, medium and high wage occupations.

Our results suggest that different occupational labour markets require different modelling approaches based on the degree to which firms and individuals are heterogeneous. In particular, the approaches emphasising firm-level wage heterogeneity and the role of matching (summarised in Mortensen (2003)) seem more applicable to higher-wage occupations, in which matching to a less productive firm or worker leads a larger loss of wages or profits. Lower-wage occupations exhibit almost no sorting and much smaller degrees of both worker and firm heterogeneity, and thus may be better approximated by classical or monopsonistic models of the labour market which do not emphasise matching. This conclusion resembles that reached by Lise & Postel-Vinay (2020b), who arrive at it using a structural model of occupational mismatch with heterogeneous skills.

2.6 Understanding spatial wage variation through worker-job decompositions

An important dimension of wage inequality, in the UK[18] but also more broadly in the world, is geographical inequality. It is well-documented that wage inequality is higher in larger labour markets (e.g. Nord (1980), Baum-Snow & Pavan (2013), and specifically in the UK context,

[18] Agrawal & Phillips (2020) show that nominal earnings gaps across space are large with incomes in Wales on average 40% below those in London. However, Agrawal & Phillips (2020) suggest that real differences are smaller and stress that within-city inequalities can be large, especially so in London. Spatial differences in income and opportunity are also increasingly dominating the political discourse in the UK, in September 2021 the department for Housing, Communities and Local Government became the Department for Levelling Up, and the new Secretary of State for Levelling Up, Michael Gove MP, called the Leveling up Agenda "the defining mission of this government".

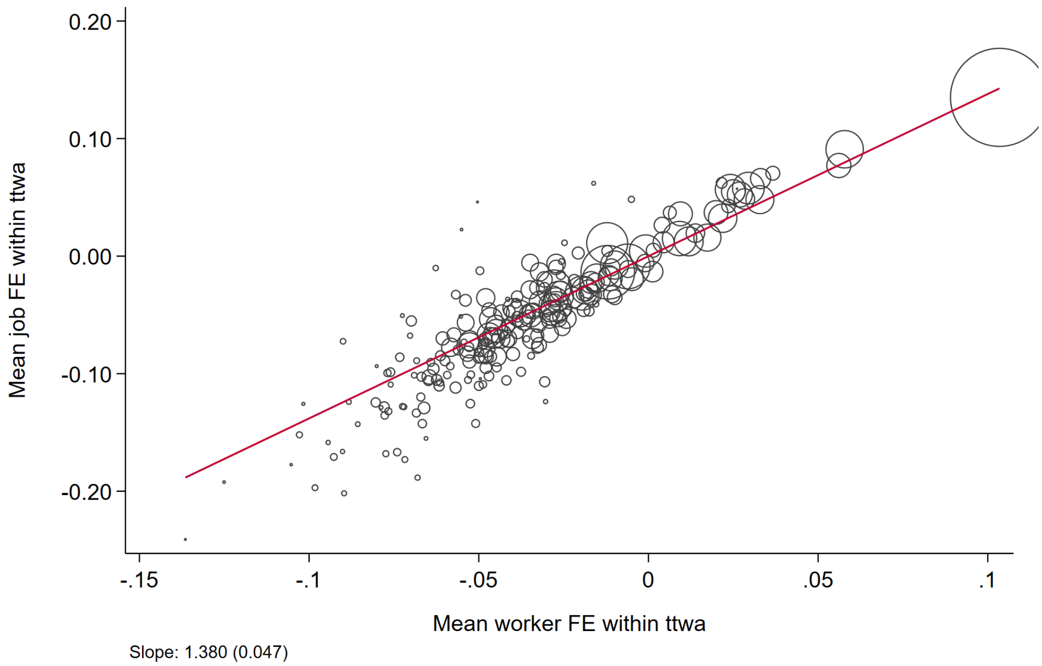
Lee et al. (2016))[19]. Our novel[20] wage variance decomposition allows us to study inequality within and across geographies in the UK with a new lens, and better understand the salient features of inequality across space.

First, we find that higher type workers tend to co-locate with higher type jobs in larger labour markets. Figure 2.6 plots the mean worker fixed effect and mean job fixed effect within a travel-to-work area, where the size of the circle reflects the size of the local labour market. This finding supports a similar finding by Dauth et al. (2018) using a worker-firm two-way fixed effect regression, and papers based on the observation that higher productivity firms sort to larger labour markets, like Gaubert (2018) and Bilal (2021). Analogously, it coheres with a large body of literature, including Venables (2011), Davis & Dingel (2019), and Behrens et al. (2014), which has found evidence to suggest and propose explanations for why more productive individuals sort to larger cities.

[19] We replicate this fact in the UK in appendix I, and show that the mean and variance of earnings is increasing in log travel-to-work area population. Our spatial unit is the UK travel-to-work area, as defined using data from the 2011 census.

[20] For an analogous study that we build on based on using the canonical AKM model to study geographic income inequality, see Dauth et al. (2018).

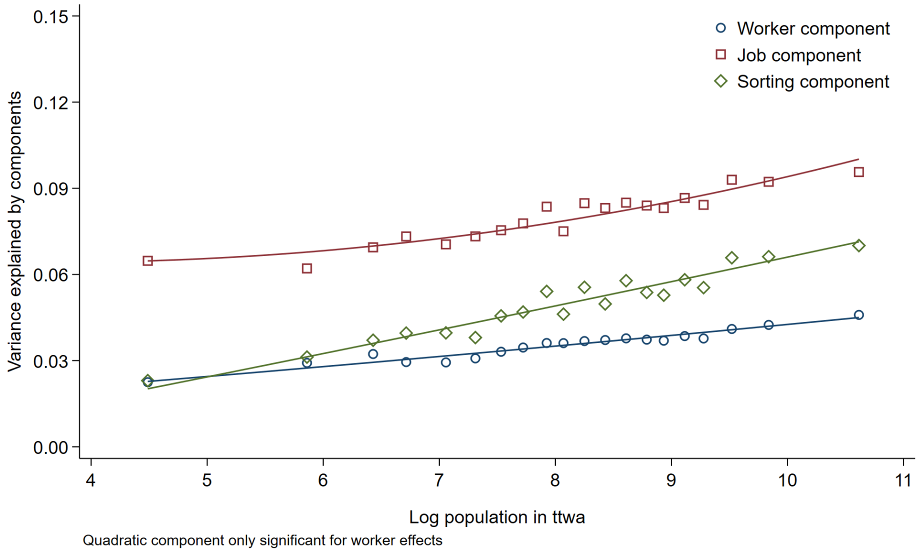
Figure 2.6: Within TTWA mean of worker FE and job FE (2011-19)



Notes: This figure plots the mean worker fixed effect (x-axis) and the mean job class fixed-effect (y-axis) within each travel to work area for each of 218 travel to work areas for 2011-19. Each bubble represents a particular travel-to-work area and the larger the bubble is, the more workers there are in the travel-to-work area observed in the data. A red linear trend line is drawn.

To study differences across labour markets in the constituent parts of within-labour market wage variation we apply our standard decomposition. Splitting log wage variance into a component due to the variance of individual fixed effects, a component due to the variance of job fixed effects, and a component due to the covariance between the two. We relate each of these components to (log) travel-to-work area size in figure 2.7 to explain the increasing within labour-market variation. All three constituent parts are increasing in local labour market size, but at different rates. It's clear from figure 2.7 that sorting becomes relatively more important in larger labour markets. This result coheres with a previous result in Dauth et al. (2018) which finds that sorting to firms accounts for more variance in larger labour markets in Germany. This finding could be explained by the value of search being greater in larger cities due to scale effects, as was found in Petrongolo & Pissarides (2006).

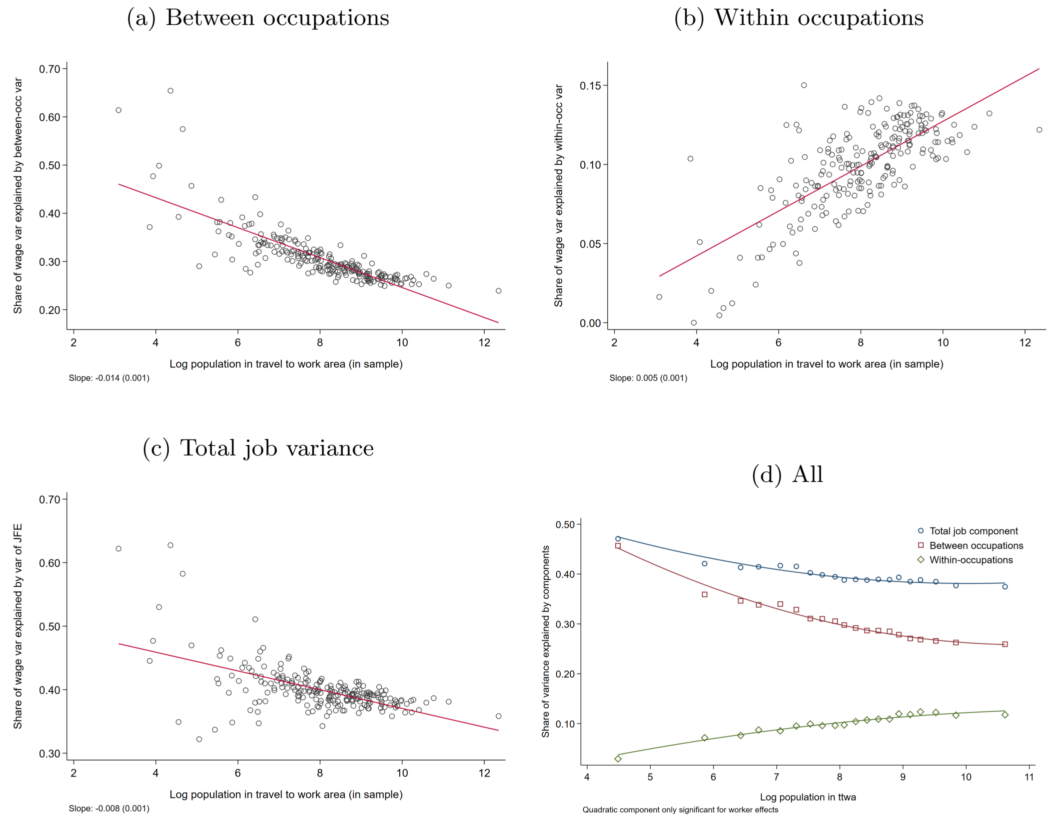
Figure 2.7: Within TTWA variance components vs local LM size



Notes: We compute the variance of individual fixed effects (blue line), the variance of job fixed effects (red line), and the covariance between the two (green) line for each travel-to-work area. We then plot a binned scatterplot against log-population within the ttwa and draw quadratic lines of best fit for all three components.

To delve deeper into what explains the patterns in figure 2.7, we decompose the local labour market job-class variance and covariance components into a between-occupations component and a within-occupations component using equations 2.3 and 2.4. First, we examine whether the decline in the relative importance of job heterogeneity is due to within or between occupations variance. We plot the relative importance of each factor in each TTWA in figure 2.8 below. Panels (a) and (b) plot the between- and within- occupation part of total log-wage variance within the TTWA as a function of log population in the TTWA. Panel (c) plots the total job variance as a share of total log-wage variance, and panel (d) combines all the other three panels in a binned scatterplot for clarity of comparison. The fall in the relative importance of job heterogeneity is primarily due to the falling relevance of between-occupation job variance as the local labour market gets larger. In larger local labour markets, wage dispersion between firms for the same occupation becomes relatively more important from a low base, eventually accounting for about 10% of the job component of log-wage variance. Absolute, rather than relative, trends are plotted in figure 2.17 in appendix I.

Figure 2.8: Decomposing the within TTWA job variance component into within and between occupations



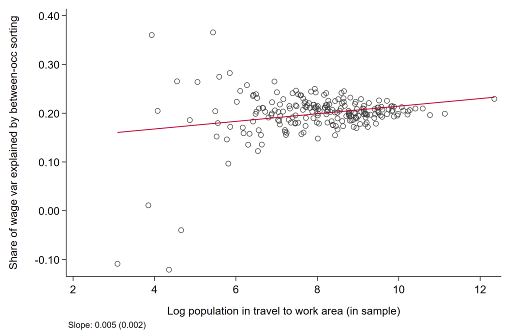
Notes: This figure presents scatterplots with log-population within the TTWA (x-axis) against between-occupation job-class variance in panel (a), within-occupation job-class variance in panel (b), total job-class variance in panel (c), and all components in bin-scatter plots in panel (d).

Similarly, figure 2.9 decomposes the share of within local labour market wage variance due to sorting between workers and jobs into the component between occupations and the component within occupations. (The between- and within- occupation covariances as raw values are included in figure 2.18 in appendix I.) What is immediately clear from this figure is that the majority of the previously documented increased importance of sorting in larger labour markets can be attributed to increased occupation sorting. That is, in larger labour markets better workers sort into better firms within occupation classes to a much greater degree than they do in smaller labour markets[21].

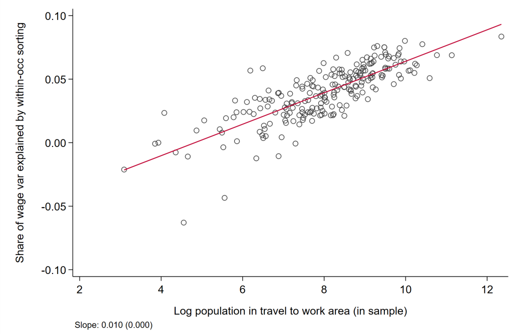
[21] The increase in within-occupation sorting could reflect the tendency for jobs in high-wage occupations to locate in larger labour markets, given that we know that the degree of sorting is higher than in higher wage occupations. In appendix J, we show that while a share of this relation is explained by occupational

Figure 2.9: Decomposing the within TTWA wage variance sorting component into within and between occupations

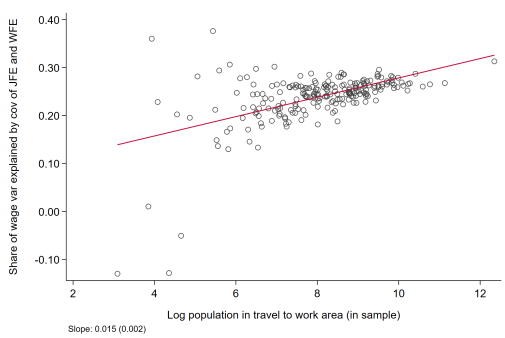
(a) Between occupations



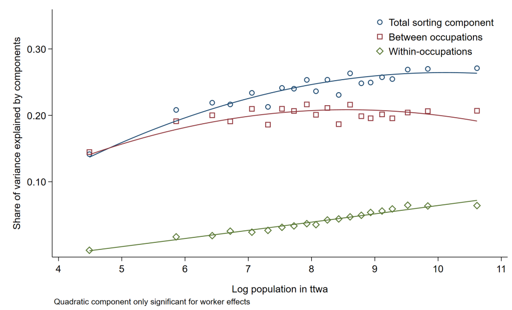
(b) Within occupations



(c) Total sorting



(d) All



Notes: This figure presents scatterplots with log-population within the TTWA (x-axis) against between-occupation sorting between worker and job-class in panel (a), within-occupation sorting between worker and job-class in panel (b), total worker and job-class sorting in panel (c), and all components in bin-scatter plots in panel (d).

In sum, although overall wage variation in the UK can mainly be described by a between-occupation story, within-occupation variation becomes is a relevant feature of larger labour markets. Colloquially, wage inequality in villages and small towns is largely due to occupations, because of the differences in pay of the plumber and the shopkeeper rather than to differences between pay of plumbers at particular firms. As the local labour market gets larger, wage dispersion between firms within an occupation becomes a quantitatively more important element of wage variance. The fact that better software engineers work at Google instead of smaller software firms is more relevant to wage inequality in London than in a

composition effects, not all of it is explainable by composition.

provincial town.

2.7 Conclusion

In this paper, we propose that the classic two-way fixed effect model with worker and firm fixed effects is misspecified and that this misspecification produces a misleading decomposition of the components of log-wage variance. Instead of matching to firms, workers match to jobs which may differ within firms — e.g. a software engineer and a chef at Google are different positions. Further, it is likely that the firm would want to offer differing wage premia to workers in different occupations at the firm.

We estimate a two-way fixed effects model with individual and job fixed effects. We find that the component of variance due to individual heterogeneity, usually found to be 50-60%, accounts for only 18% of log wage inequality in our model. Conversely, we find that job heterogeneity and the sorting of workers into jobs account for 38% and 29% of the log-wage variance respectively. The majority of variation in these components can be explained by cross rather than within-occupation differences, highlighting the importance of occupations and jobs rather than firms, in explaining UK wage inequality.

Our approach also allows us to analyse how wage inequality within occupations differs across occupations. We find that wage variance in higher mean wage occupations is explained by sorting to a greater degree. We thus conclude that occupational labour markets should be analysed using different approaches, depending on the degree to which individual and firm heterogeneity affect productivity. Occupations in which heterogeneity is more important are better explained by search and matching models which produce more wage dispersion, while occupations, where workers and firms are more interchangeable, are better modelled by spot models of the labour market.

Finally, we consider the implications of analysis for understanding wage inequality in local labour markets. We find that sorting is increasingly important in larger labour markets, which could be due to scale effects in search markets. We also find that sorting to firms within occupations are relatively much more important in larger labour markets than in smaller ones; while wage inequality in small towns is likely to be due to differences in occupation, wage inequality in large cities are as much due to workers being at different firms within the same occupation as being in differing occupations.

2.8 Appendix

A Details on data cleaning

We focus our decomposition analysis on the most recent period from 2011-19, during which the UK standard occupation classification 2010 is used. We observe the most fine-grained 4-digit occupational classifications, giving us 368 occupations in total[22]. Our preferred wage variable, used throughout as the focus of our analysis, is the hourly wage, which is computed as the gross pay divided by the total number of hours. We trim the top and bottom 1% of the sample to remove outliers. Our main unit of geographical analysis is the travel to work area (hereafter TTWA), a non-overlapping geographical unit provided by the Office for National Statistics. The TTWA is defined as an area where at least 75% of the resident workforce work in the area and 75% of the people who work in the area also live in the area and is thought of as a local labour market. TTWAs are the UK equivalent of commuter zones in the US which have become the standard unit for measuring local labour markets. TTWAs differ in size, and urban areas such as London can be very large. We use the 2011 classification of the TTWA in our analysis, giving us 218 travel-to-work areas in the data in total. Observations with incomplete information about their home address are used in the estimation of the fixed effects model but are not included in the descriptive analysis on geographical dimensions.

[22] One of these, [2321], has fewer than 9 observations in the entire period, and is excluded for statistical disclosure purposes.

In the main decomposition analysis we can only use the largest connected set of workers and jobs to estimate fixed effects. Table 2.4 shows the overall size of the data set and the proportion thereof that is retained when considering the largest connected set over various decompositions reported in the main text. As one can see, the majority of the data is retained in most specifications with over 96% of observations used in our main decomposition reported in the final row.

Table 2.4: Largest connected set of common specifications (2011-19)

	No. obs	**No. workers**	**No. firms**	**No. occs**	**No. jobs**
All	1598919	339697	156178	368	335298
LCS for firm FE	93.57	81.01	63.57	100.00	79.30
LCS for clustered firm FE	96.46	83.36	89.34	100.00	91.89
LCS for job FE	88.50	77.98	59.14	100.00	56.93
LCS for clustered job FE	96.46	83.36	89.34	100.00	91.89

Notes: This table documents the proportion of the data retained in the largest connected set over various decompositions performed in the main text.

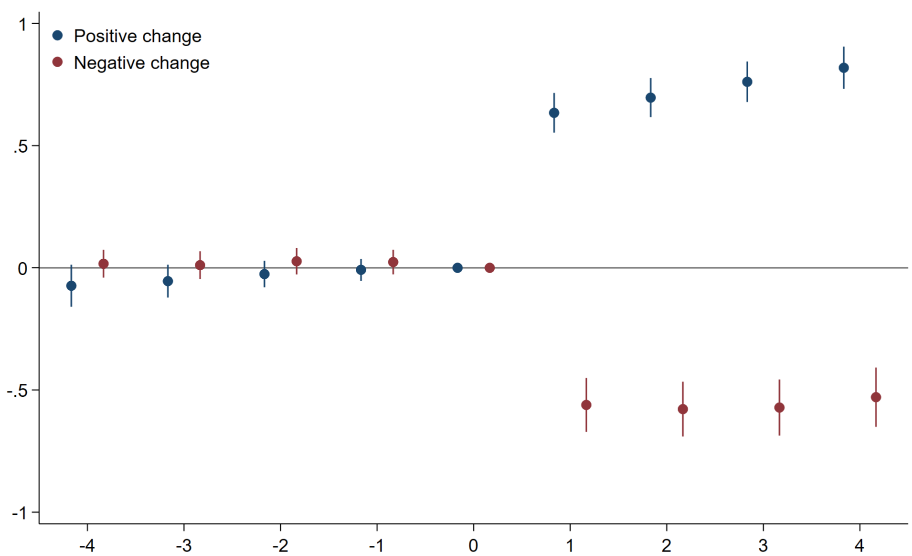
B Testing the identifying assumption

In section 2.2, we described the exogeneity assumption we need for our estimated fixed effects to not be biased, and listed some examples under which exogeneity would fail. In this section, we perform a number of tests to show that this assumption is likely to hold. We follow analogous strategies in our context to the approaches that Card et al. (2018b) takes to establish the suitability of the base AKM model.

The two-way fixed effect framework implies that when workers change jobs, they would experience a log wage change equal to the difference in the job fixed effect at the old job and at the new job. One implication of this is that for similar workers, the log wage change of moving from a former job to a latter job should be equal to the negative of log wage change of moving from that latter job to the former job. This may not be the case if for example job transitions occur due to workers matching to more suitable jobs over time. In that case, the mean wage changes associated with moves from less to more productive jobs should exceed that of moves in the other direction.

This idea is the basis of both tests presented in this section. First, we adopt the event study approach in Card et al. (2013) when we consider the wage effects of a worker moving from a less productive job to a more productive job and vice versa. Figure 2.10 considers the wage rate in years relative to the year before a job switch occurs of workers who will (or have) switched jobs from the 1st to 4th (or 4th to 1st) quartile of jobs in terms of job fixed effects. This figure estimates wage trajectories using only within job switch pair variation and taking into account individual fixed effects. Wage trajectories are normalised to the year before the move (0'th year).

Figure 2.10: 1 to 4 (and 4 to 1) job switches

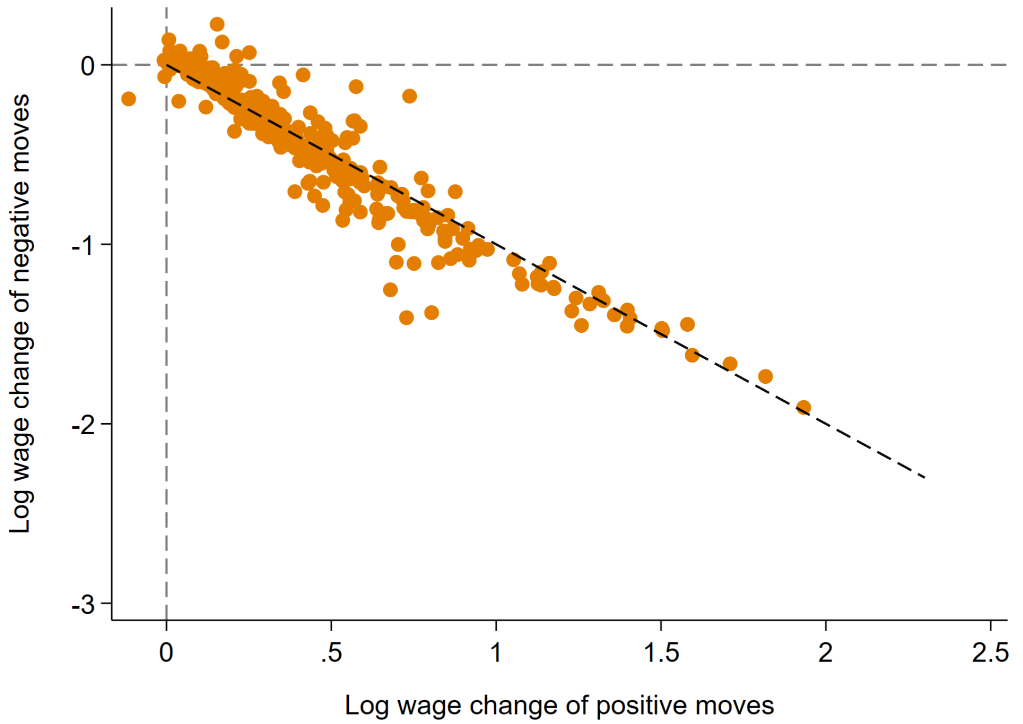


Notes: This figure shows the estimated log wage over time for workers who switch between jobs in the first and fourth quartile of the job fixed effect distribution. By including job pair fixed effects and individual fixed effects this figure shows the relative wages of workers who make a positive vs negative switch where wages are normalised to the year before the switch (0th year).

Two things are apparent from figure 2.10. First, there are no clear increases or decreases in pay in the year prior to, or proceeding, job changes, supporting what Card et al. (2018b) describe as the step-like pattern of wage changes following job changes. Second, the degree to which wages change is symmetric; positive switches are associated with a similar increase in wages as negative switches are associated with a decrease. This gives evidence to suggest that there is not a large worker-job match-specific component of wages.

Second, the implied symmetry of wage changes in between job pair moves can be tested more directly by considering the average log wage change when workers move from a job to another. Because the set of moves between jobs is fairly sparse, we consider the average wage changes of moves between the wage clusters which we use to correct for limited mobility bias. We plot the changes in figure 2.11, restricting our analysis to firm-cluster pairs where we have observed at least 10 moves. If the exogeneity assumption is satisfied, then all the points should lie on the 45 degree line. We find that wage changes are roughly asymmetric as the model predicts, although there are a few outliers. The severity of the problem seems more acute for small moves, while larger moves between clusters seem largely symmetric as the model predicts. Many of the points also show that the wage loss associated with a move to a lower productivity cluster exceeds the wage gain of moving to a more productive cluster, which similarly contradicts prominent narratives of between job transitions that do not cohere with the implications of the two-way fixed effect framework.

Figure 2.11: Average log wage changes of moves between firm clusters

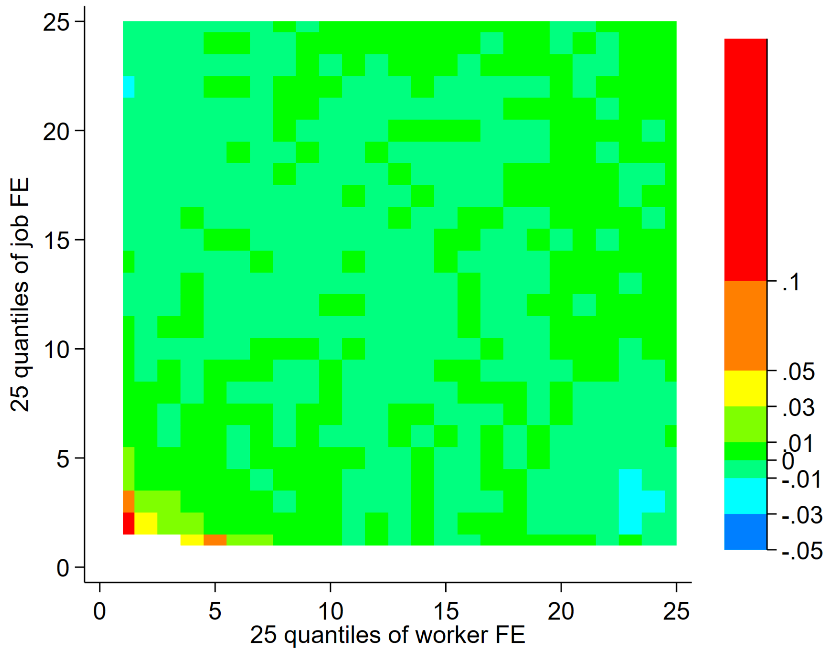


Notes: This figure plots the within-cluster average wage changes of moves between firm clusters. We use 30 firm-clusters which you use to correct for limited mobility bias. On the x-axis, we plot the wage change associated with moves from less productive to more productive clusters, while on the y-axis we plot the average wage change of moves from a more productive cluster to a less productive cluster. Each point represents a firm cluster pair, and there are 435 such pairs. We drop 94 such pairs where there are fewer than 10 moves in either direction. The dashed line is the 45 degree line which represents perfect symmetry.

C Is the log-additive functional form a good specification for wages?

In this section, we check whether the log additive functional form in fixed effects is a good approximation to the data. To do this, we follow the test in Card et al. (2013), and aggregate our estimated fixed effects into 50 quantile groups. For each worker quantile and job quantile, we take the mean of the residual of the two-way fixed effect regression, reasoning that if the specification is a good fit to the data, then the mean of the residuals should be 0 or close to zero. We plot a heat map in figure 2.12 with the quantiles of worker FE on the x-axis, the quantiles of job FE on the x-axis, and the size of the mean residuals as the colour within the cell. Warmer colours (yellow, orange, red) mean that the model underpredicts the true wage in the data while cooler colours (cyan, blue) means that the model overpredicts the actual wage data. Green implies that the mean residual in the cell is within 1% on each side of 0.

Figure 2.12: Mean of residuals by worker FE and job FE quantiles (25 groups each)



Notes: This figure plots the mean of the two-way fixed effect residuals for 25 worker FE quantiles by 25 job FE quantile cells for the two-way fixed effect regression with worker and job fixed effects without clustering.

The figure shows that for the vast majority of cells the difference between the predicted value and the true data is within 1% of the actual log wage. The model is particularly bad for a small number of low-skill workers at low-wage premium jobs, severely underestimating the pay of such workers by more than 10% and up to 30% at worst. This could be because of institutional factors, like minimum wage which are not explicitly considered in the model specification.

D Review of limited mobility bias

Limited mobility bias occurs when second (or higher) order moments are estimated for fixed effects, which are estimated off few observations. In this case, the estimated fixed effects themselves are unbiased, but the high-order moments are biased. Andrews et al. (2012)) pointed out this potential threat to identification and Jochmans & Weidner (2019) show how to understand the likely size of the bias as a function of the number of moves between workers and firms, or of the connectedness of the network of workers and firms. To see this more formally consider the variance of the fixed effects β, denoted as $V_Q = \beta' Q \beta$ for some matrix Q. Then our estimate of the variance would be given by $\hat{V}_Q = \hat{\beta}' Q \hat{\beta}$ the bias of which can be characterized as follows:

$$\mathbb{E}[\hat{V}_Q] = V_Q + \mathbb{E}[\varepsilon' Z(Z'Z)^{-1}Q(Z'Z)^{-1}Z'\varepsilon] \tag{2.6}$$

This bias arises from an insufficient number of movers between jobs (or firms in the classic AKM literature) and is therefore termed the *limited mobility bias*. As a result, the variance of job (or firm) fixed effects is overstated at the expense of understating the co-variance between workers and jobs. This is a problem the whole literature attempting AKM decompositions faces but the scale of the issue has only recently become apparent with Bonhomme et al. (2020) for example showing that the limited mobility bias is severe and that bias correction is very important. In our case, relative to more standard worker-firm decompositions, by introducing jobs we increase mobility intuitively by capturing within-firm, cross-occupation moves, but simultaneously increase the number of parameters to be estimated. Therefore it's unclear ex-ante whether limited mobility bias represents a smaller or larger problem in our framework.

We solve the limited mobility bias by following Bonhomme et al. (2019) in first clustering job using a k-means algorithm, into K clusters. By performing this first step we greatly increase movers between clusters whilst retaining much of the underlying variation (30 clusters capture over 85% of the underlying job level wage variation).

E Analysing the result of misspecifying the model with worker-firm FE when the true DGP is with worker-job FE

In the paper, we noted that when the true data-generating process underlying the data involves worker-job (firm-occupation pair) fixed effects but a worker-firm FE two-way fixed effect regression is specified, the model leads to the overestimation of variance attributed to individuals under certain conditions. We show this in more detail in this section. This result is conceptually similar to the work in Abowd et al. (1999) to show the effects of misspecifying the model with worker-industry FE when the true model is worker-firm.

We can rewrite the true data-generating process as follows.

$$\begin{aligned}\ln(w_{it}) &= \alpha_i + \phi_{F(i,t)} + (\psi_{J(i,t)} - \phi_{F(i,t)}) + \varepsilon_{it} \\ \phi_f &= \frac{1}{n_f}\sum_{j\in \bar{f}} \psi_j\end{aligned}$$

It seems from this rewriting that what misspecification does is to add a term to the part that varies across individuals and periods which is now an unobserved omitted variable, and which will therefore cause the regression to estimate the individual fixed effects and the mean firm fixed effects with bias. Let W denote a $NT \times 1$ vector of wages, L a $NT \times N$ matrix of indicators for individuals $i = 1, \cdots, N$, and J a $NT \times H$ matrix of indicators for jobs $j = 1, \cdots, H$. Let F denote a $NT \times G$ matrix of indicators for the firm the job is at (for $f = 1, \cdots, G$). Let ε denote a $NT \times 1$ vector of ε_{it} errors. Finally, let K be a $H \times G$ matrix where each row $j = 1, \cdots, H$ indexes which firm the job is at. Let α denote a $N \times 1$ vector of personal fixed effects, Ψ a $H \times 1$ vector of job fixed effects and Φ a $G \times 1$ vector of firm-specific mean job fixed effects.

$$
\begin{aligned}
W &= L\alpha + J\Psi + \varepsilon \\
&= L\alpha + F\Phi + J(\Psi - K\Phi) + \varepsilon
\end{aligned}
$$

When J is not observed, estimating an OLS regression on L and F leads to omitted variable bias. Suppose we estimate the misspecified model. Then the parameters estimated have omitted variable bias as follows, where $M_A = I - A(A'A)^{-1}A'$ for an arbitrary matrix A.

$$
\begin{aligned}
\alpha^* &= \alpha + (L'M_F L)^{-1} L'M_F J(\Psi - K\Phi) && (2.7)\\
&= \alpha + (L'M_F L)^{-1} L'M_F J\Psi && (2.8)\\
\Phi^* &= \Phi + (F'M_L F)^{-1} F'M_L J(\Psi - K\Phi) && (2.9)\\
&= (F'M_L F)^{-1} F'M_L J\Psi && (2.10)
\end{aligned}
$$

The M_F and M_L matrices are residualising matrices, where $M_F L$ refers to the residual of a linear projection of F on L as in the Frisch-Waugh-Lovell theorem. It is instructive to re-write the expressions above as follows.

$$
\begin{aligned}
\alpha^* &= \alpha + \beta && (2.11)\\
\Phi^* &= \iota && (2.12)
\end{aligned}
$$

Where β is the regression coefficient found from estimating the following by OLS: $J\Psi = (M_F L)\beta + \xi$ and ι is similarly the regression coefficient from estimating $J\Psi = (M_L F)\iota + \xi$.

Thus, the bias in variance estimation is easily characterisable.

$$\mathbb{V}\left[\alpha^*\right] - \mathbb{V}\left[\alpha\right] = \mathbb{V}\left[\beta\right] + 2Cov\left(\alpha, \beta\right) \tag{2.13}$$

$$\mathbb{V}\left[\Phi^*\right] - \mathbb{V}\left[\Phi\right] = \mathbb{V}\left[\iota\right] - \mathbb{V}\left[\Phi\right] \tag{2.14}$$

From this formula we can see that the individual fixed effects will be over estimated if $Cov(\alpha, \beta)$ is not sufficiently negative. To gain intuition as to the signs and magnitudes of $\mathbb{V}[\beta]$ and $2Cov(\alpha, \beta)$, we must understand β. β is an $N \times 1$ vector of coefficients, an individual element of β is large if for that individual the correlation between the firm-residualised individual and job effect is large. If the strength of this relationship varies over individuals $\mathbb{V}[\beta]$ will be large. If this relationship is stronger for higher-FE individuals $Cov(\alpha, \beta)$ will be positive, this is the most likely scenario.

It's harder to sign the difference in firm fixed effect variance. ι is the firm-analogous version of β, it is large if the firm conditional on worker is correlated with job FE. However, it's unclear whether the strength of variation in this coefficient will be less than or greater than the variation in firm effects itself and so signing the difference is difficult.

In general, from the above we can say that we expect the variance of the estimated worker and firm fixed effects in the mis-specified model to be wrong, and for it to be increasingly wrong the more residual variation there is in the relationship over workers between workers and jobs (conditional on firms) and over firms between firms and jobs (conditional on workers).

F Comparing log-wage decomposition results against other results in the literature

In this section we compare results from estimating the usual worker-firm AKM model decomposition on our UK data to that from previous studies using data in other countries. Table 2.5 displays the main results. Panel [A] compares results using the basic AKM decomposition which does not correct for limited mobility bias whereas panel [B] shows results correcting for limited mobility bias. The main take away from this table is that in these specifications the UK data appears to behaviour similarly to that from other contexts. One possible exception is in the estimated negative effect found due to sorting in panel [A]. However, this is not as unusual a finding as it would appear from this table; in a recent survey, Bonhomme et al. (2020) show that in 10 of the 37 AKM estimates from previous studies, a negative sorting effect is found.

Table 2.5: Comparison of basic AKM and clustered correlated fixed effect decomposition results with select estimates from previous studies.

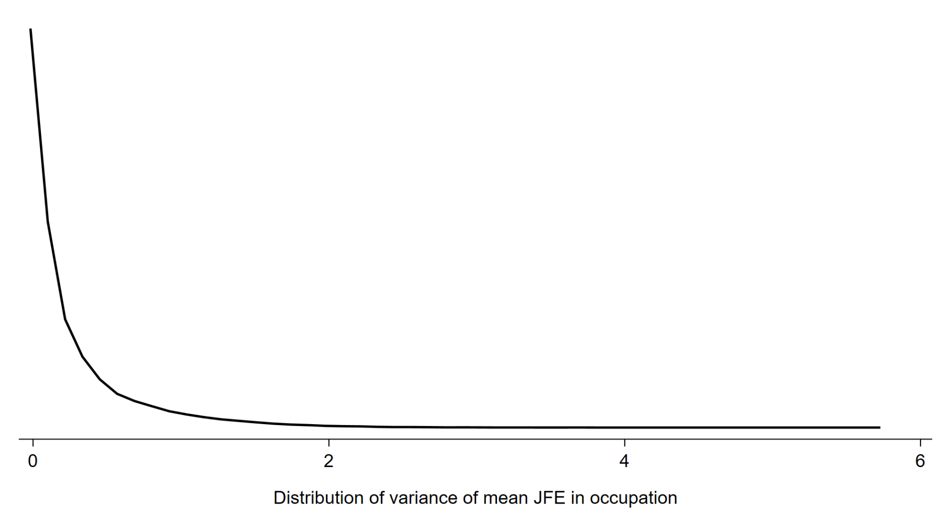
[A] Basic AKM Decomposition					
Paper	**Period**	**Country**	**WFE**	**FFE**	**Sort**
Card et al. (2013)	2002-09	Germany	51%	21%	16%
Card et al. (2016)	2002-09	Portugal	58%	20%	11%
Song et al. (2019)	2007-13	USA	52%	9%	12%
Alvarez et al. (2018)	2008-12	Brazil	58%	16%	18%
Our data	2011-19	UK	64%	24%	-2%
[B] Clustered Random Effects Bias Correction					
Paper	**Period**	**Country**	**WFE**	**FFE**	**Sort**
Bonhomme et al. (2019)	2002-04	Sweden	60%	3%	12%
Our data	2011-19	UK	54%	10%	19%

Notes: This table compares results using the UK data used in this paper when estimating classic and bias-corrected AKM models, to those found in selected papers from the previous literature.

G Direct evidence that occupational changes are relatively unlikely

In section 2.4.1, we argued that the misspecification of the AKM leads to over estimation of the variance of the individual fixed effect especially if an individual's occupation is persistent over time. In this section we provide some direct evidence of the persistence of occupation choice. We use mean job fixed effect within an occupation as a proxy for the relative productivity of the job. We then calculate the variance of the mean job fixed effect across occupations for each individual, and normalise it by the value of $\mathbb{V}[\mathbb{E}_o[\psi_j]]$ for the whole population. This measures an individual's relative variance in occupational quality for their time in the panel relative to the overall population between-occupation job fixed effect variance. We plot the values for workers who have been in the panel for at least 5 years in figure 2.13 below. We find that this variance is very low for most individuals and the distribution is skewed towards 0. This suggests that most workers do not substantially change the quality of the occupation they are employed at over time.

Figure 2.13: Distribution of the variance of mean JFE within occupations over individuals



Notes: This figure plots the variance of the occupation mean fixed effect for the occupations an individual was observed in over the length of the panel, normalised over the overall variance in occupational mean job fixed effects.

H Additional results on the relation between job FE as wage premia related to rent-sharing

As discussed in the introduction, estimates of the firm fixed effects in implementations of the AKM model have been interpreted as wage premia arising from firms rent-sharing with their workers (see Card et al. (2018b) in particular for an attempt to synthesise the rent-sharing and AKM literatures). We check this story in our setting by investigating whether the job fixed effect $\psi_{J(i,t)}$ correlate with the firm's average productivity. Specifically, we consider a model where the job fixed effects can be modelled as follows, where $\phi_{F(i,t)}$ denotes industry fixed effects and $\chi_{Occ(i,t)}$ denotes occupation fixed effects.

$$\psi_{J(i,t)} = \alpha_0 + \beta \cdot \ln(\text{prod}_{F(i,t)}) + \phi_{F(i,t)} + \chi_{Occ(i,t)} + u_i \tag{2.15}$$

Since $\psi_{J(i,t)}$ is an estimated coefficient, we have to take into account its covariance matrix when estimating the parameters of equation 2.15 including β. We do this by substituting equation 2.15 into equation 2.1, and estimating the whole model[23]. We use data on firm characteristics from the Business Structure Database (BSD), which contains detailed panel information on firm formation, as well as some basic characteristics like industry, employment, tax status and revenue. A major disadvantage of the BSD is that it does not contain information on firm profits, and we are unable as a result to test directly whether the estimated coefficients relate to firm productivity. We proxy for productivity using revenue per worker.

We run our analysis on three samples. Besides the full sample, we also consider splitting the sample into two groups (1) private workers not covered by a collective bargaining agreement and (2) private workers who are covered by a collective bargaining agreement and public/non-profit sector workers[24]. If the estimated fixed effects were driven by rent-sharing, then we would expect that workers in private firms not covered by a collective bargaining agreement to have a higher value of β_o over all occupation groups. We thus estimate equation 2.15 on sub-samples of the two groups. Results for private, non-CA workers are presented in column (2) while results for public sector or workers under CA are presented in column (3)[25].

We present the results in table 2.6. We find that the job fixed effect is increasing in the revenue per worker of the firm, and that this correlation is greater for the sub-sample of

[23] An alternative approach is to use Generalised Least Squares after estimating the covariance matrix of $\psi_{J(i,t)}$.

[24] We use indicators in the ASHE dataset on whether the worker works at a public, private or non-profit institution. We also use indicators in the dataset on whether the worker is covered by a collective agreement. We lump workers under all kinds of collective agreements in the same category.

[25] We also calculated the regression on public/private, and CA/non-CA sub-samples separately. Results for public sector workers are similar to those for CA workers and results for private workers are similar to those for workers not under CAs. Thus, we present only results for the two combined groups, noting that the results are not driven in particular by either the public/private or the CA/no-CA dimension.

jobs at private firms that are not covered by collective agreements. This coheres with the interpretation of these job fixed effects as wage premia paid to workers.

Table 2.6: Estimates of β in equation 2.15 (2011-19)

Variable	**All workers (1)**	**Private, no CA workers (2)**	**Public or CA workers (3)**
Firm revenue per worker	0.018	0.027	0.005
	(0.002)	(0.002)	(0.002)
Obs	1537793	740,090	715,920
R^2	0.897	0.911	0.906

Notes: This table shows the results from estimating $\psi_{J(i,t)} = \alpha_0 + \beta \cdot \ln(\text{prod}_{F(i,t)}) + \phi_{F(i,t)} + \chi_{Occ(i,t)} + u_i$ on the full sample in column one, a sample of private workers not covered by a collective bargaining agreement in column two, and on public workers or those covered by a collective bargaining agreement in column three.

We run a similar regression with different specifications to check if the relation to firm characteristics differs across occupations. Table 2.7 summarises the coefficients of a similar regression to equation 2.15, but where we interact the log worker productivity variable with the 1 digit occupation group; we also plot the coefficients in figure 2.14. We find that for all occupation groups, there is a positive relation between revenue per worker and the job fixed effect. We find statistically significant relationships for all occupation categories for the sample of jobs at private firms and jobs not covered by collective agreements. On the other hand, we find that the relations between the wage premium for managers, professionals, customer service and machine operators and log revenue per worker are not statistically significant. This again seems to be evidence in support of the wage premium interpretation of the job fixed effects.

An interesting aspect of the results in figure 2.14 is how the coefficient varies across occupations. The coefficient is broadly higher for more skilled workers than for less skilled workers for jobs in private firms, with workers in care and leisure having a particularly low correlation and customer service having a high correlation. In the public sector however, the premium paid to managers and professionals have no relation to firm revenue per worker, but the premium paid to mid and low-skill occupations seem to still exhibit some relation. This could reflect workers in high-skill occupations in these public institutions being better able to insure for wage variation.

For robustness, we check whether this pattern persists if we use a measure of the skill requirement of the occupation[26]. The results for these regressions are summarised in table 2.8 and figure 2.15. We similarly find that overall the relation between revenue per worker and the job fixed effect is greater for lower-skilled workers than for higher-skilled workers, but this is primarily driven by a composition effect. The relation is statistically insignificant and near zero for high skill workers at public firms and at jobs covered by collective agreements, but very large for similarly skilled workers in jobs at private firms not covered by a collective agreement. In contrast, the correlation of job fixed effect with revenue per worker is positive and statistically significant in all samples for lower skilled workers. We leave a full explication of these puzzles to future work.

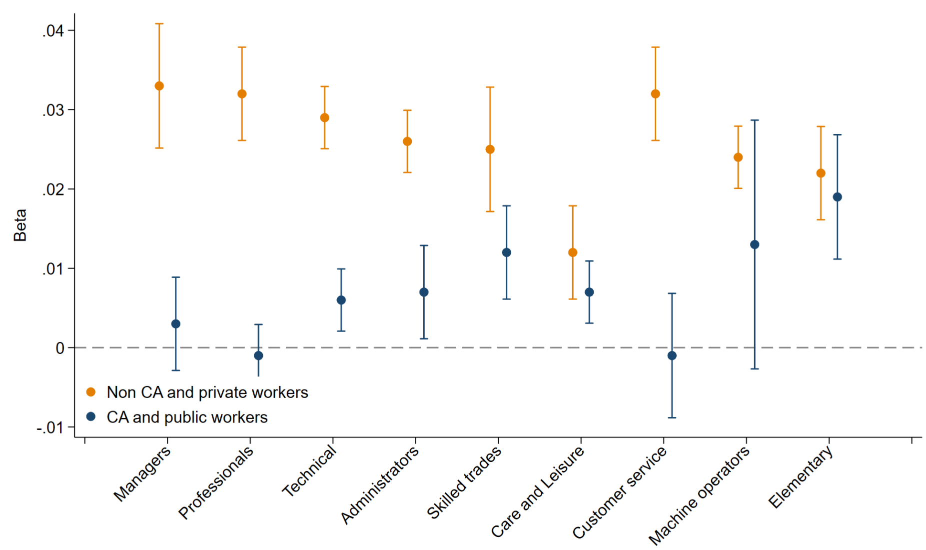
[26]We use the skill requirements for occupations used by the UK Home Office, and set out in their Appendix J tables. RQF 3 corresponds to a high school degree, while RQF 4 corresponds to 1 year of a bachelors degree. RQF 6 corresponds to a bachelor degree.

Table 2.7: Coefficients on log revenue per worker interacted with 1 digit occupation group

Occupation Group	**All workers** (1)	**Private, no CA workers** (2)	**Public or CA workers** (3)
Managers, directors and senior officials	0.023	0.033	0.003
	(0.003)	(0.004)	(0.003)
Professional occupations	0.010	0.032	-0.001
	(0.004)	(0.003)	(0.002)
Associate professional and technical occupations	0.019	0.029	0.006
	(0.002)	(0.002)	(0.002)
Administrative and secretarial occupations	0.021	0.026	0.007
	(0.002)	(0.002)	(0.003)
Skilled trades occupations	0.024	0.025	0.012
	(0.003)	(0.004)	(0.003)
Caring, leisure and other service occupations	0.010	0.012	0.007
	(0.002)	(0.003)	(0.002)
Sales and customer service occupations	0.022	0.032	-0.001
	(0.004)	(0.003)	(0.004)
Process, plant and machine operatives	0.025	0.024	0.013
	(0.003)	(0.002)	(0.008)
Elementary occupations	0.022	0.022	0.019
	(0.002)	(0.003)	(0.004)
Obs	1537793	740,090	715,920
R^2	0.897	0.911	0.906

Notes: This table shows the coefficients and corresponding standard errors found when estimating equation 2.15, interacting log revenue per worker with occupation group.

Figure 2.14: Relation between job fixed effect and firm revenue per worker, by occupation group (2011-19)



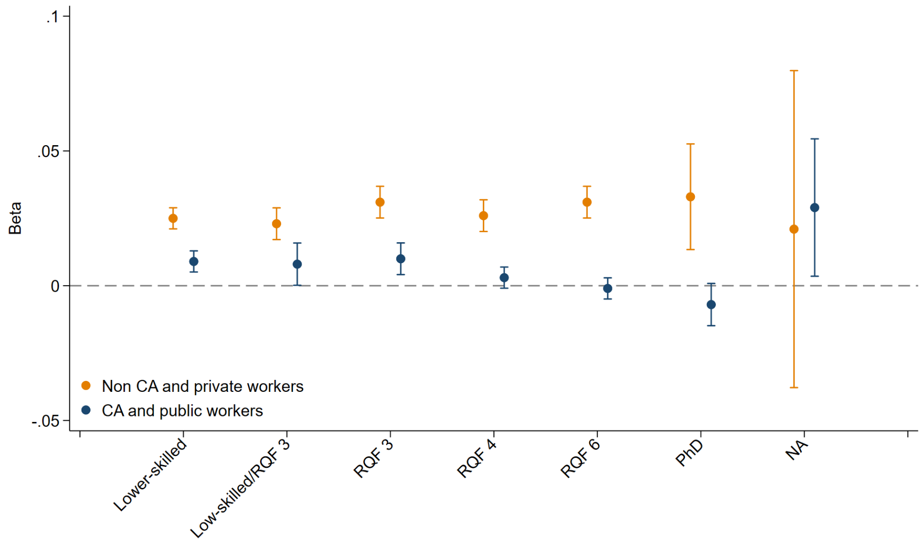
Notes: This figure shows the coefficients and corresponding 95% confidence intervals found when estimating equation 2.15, interacting log revenue per worker with occupation group.

Table 2.8: Coefficients on log revenue per worker interacted with skill groups

Occupation Group	**All workers (1)**	**Private, no CA workers (2)**	**Public or CA workers (3)**
Lower-skilled	0.019	0.025	0.009
	(0.002)	(0.002)	(0.002)
RQF 3/Lower-skilled	0.016	0.023	0.008
	(0.003)	(0.003)	(0.004)
RQF 3	0.023	0.031	0.010
	(0.002)	(0.003)	(0.003)
RQF 4	0.015	0.026	0.003
	(0.002)	(0.003)	(0.002)
RQF 6	0.014	0.031	-0.001
	(0.004)	(0.003)	(0.002)
PhD	0.014	0.033	-0.007
	(0.006)	(0.010)	(0.004)
Unclassified	0.016	0.021	0.029
	(0.010)	(0.030)	(0.013)
Obs	1534631	738,048	714,953
R^2	0.897	0.911	0.906

Notes: This table shows the coefficients and corresponding standard errors found when estimating equation 2.15, interacting log revenue per worker with the skill level of the occupation, as determined by the UK Home Office.

Figure 2.15: Relation between job fixed effect and firm revenue per worker, by skill group (2011-19)



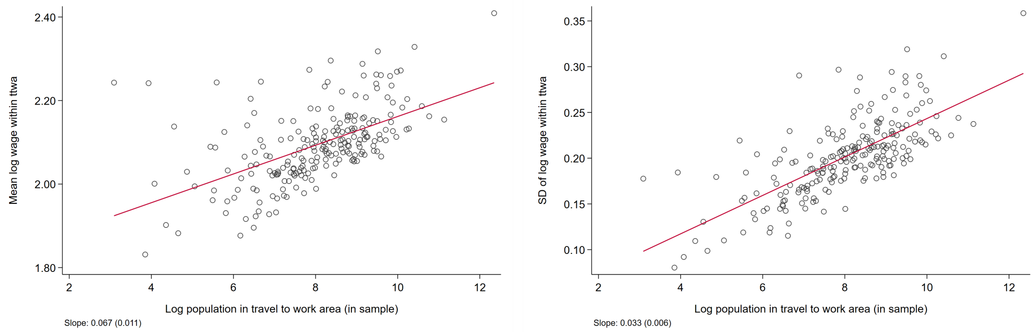
Notes: This figure shows the coefficients and corresponding 95% confidence intervals found when estimating equation 2.15, interacting log revenue per worker with the skill level of the occupation, as determined by the UK Home Office.

I Appendix figures and tables relating to wage inequality across geography

Panel 2.16a reproduces the well-known fact that mean wages are higher in larger labour markets, while panel 2.16b shows that the variance of log-wages is also increasing log-linearly with log-population within a TTWA.

Figure 2.16: Within TTWA log population vs moments of log pay (2011-19)

(a) Mean (b) Variance

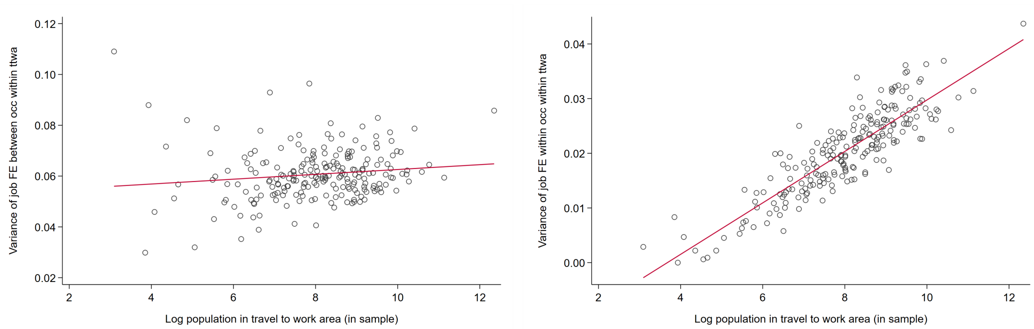


Notes: Panel 2.16a plots the mean of log-hourly wage (y-axis) against log-population (x-axis) in the travel-to-work area. Panel 2.16b plots the variance of log hourly wage (y-axis) against log-population (x-axis) in the travel-to-work area.

In figure 2.17, we plot scatterplots of the absolute values of the part of log-wage variance explainable by between-occupation job fixed effect variance (panel 2.17a), and the part explainable by within-occupation job fixed effect variance (panel 2.17b). Analogously, figure 2.18 plots the absolute values of the covariance of job fixed effects and individual fixed effects attributable to between-occupation sorting (panel 2.18a) and within-occupation sorting (panel 2.18b).

Figure 2.17: Within TTWA log population vs variance of job FE between and within occupation

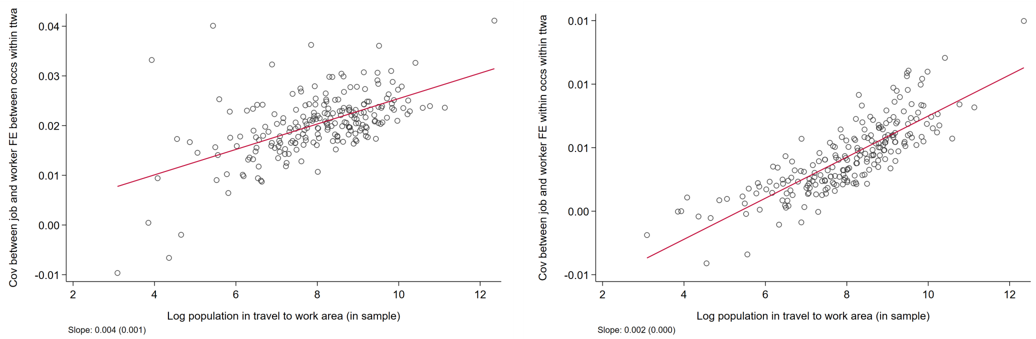
(a) Between occupation total variance (b) Within occupation total variance



Notes: Panel (a) plots a scatterplot of TTWAs with log-population on the x-axis and the between-occupation com ponent of job-class fixed effect variance on the y-axis. Panel (b) plots a scatterplot of TTWAs with log-population on the x-axis and the within-occupation component of job-class fixed effect variance on the y-axis.

Figure 2.18: Within TTWA log population vs covariance components decomposed into between and within occupation parts

(a) Between occupation total variance (b) Within occupation total variance



Notes: Panel (a) plots a scatterplot of TTWAs with log-population on the x-axis and the between-occupation component of worker-job-class covariance on the y-axis. Panel (b) plots a scatterplot of TTWAs with log-population on the x-axis and the within-occupation component of worker-job-class covariance on the y-axis.

J Is the relation between size of travel to work area and variance due to occupational composition?

The increase in within-occupation sorting documented in figure 2.9d could reflect the tendency for jobs in high-wage occupations to locate in larger labour markets, given that we know that the degree of sorting is higher in higher wage occupations. Thus, the relation documented here between TTWA size and positive assortative matching could reflect a composition effect rather than a true correlation with size. To obtain direct evidence on this, we calculate the standard deviation of job fixed effects within each occupation for each travel-to-work area in each period ($\sigma_{kl\tau}$) as well as the correlation between worker fixed effects and job fixed effects within each travel-to-work area, occupation and period ($\rho_{kl\tau}$) (where τ indexes the period). We then perform simple regressions of these moments on log population within the travel-to-work area, controlling for occupation and period fixed effects (denoted by γ_k, and γ_τ respectively for the standard deviation regression, and by δ_k and δ_τ for the correlation regression).

$$\sigma_{kl\tau} = \gamma_0 + \gamma_1 \cdot \ln(pop_l) + \gamma_k + \gamma_\tau + u_{kl\tau} \tag{2.16}$$

$$\rho_{kl\tau} = \delta_0 + \delta_1 \cdot \ln(pop_l) + \delta_k + \delta_\tau + v_{kl\tau} \tag{2.17}$$

The results of this regression for γ_1 and δ_1 are reported in table 2.9 for both the full set of TTWA-occupation-period cells and for the set of cells with at least 30 observations. We find that even after controlling for occupational differences, there is still a residual correlation between travel-to-work area size and both the standard deviation of job fixed effects and the correlation between worker and job fixed effects. This qualitative result is robust to dropping smaller cells for which these moments might not be as reliable. Thus, we conclude that being in a larger local labour market is associated with larger variance of job fixed effects, larger variance of worker fixed effects, and a greater correlation between worker and job fixed effects, even conditional on the occupational composition of the travel-to-work area.

Table 2.9: Coefficients on log population from regression of occ-ttwa moments

Coefficient on log pop	All > 10	All > 30
Standard deviation of job FE	0.018	0.013
	(0.001)	(0.001)
Standard deviation of worker FE	0.015	0.008
	(0.001)	(0.001)
Correlation of worker and job fixed effects	0.002	0.001
	(0.000)	(0.000)
N	42,328	18,749

Notes: This table shows the results from estimating equations of the form given in 2.16 of the relationship between the TTWA-occupation-period standard deviation of job and worker FE as well as the correlation between the two and TTWA population. Column one gives results for all cells with more than 10 observations and column two for all cells with more than thirty observations. Standard errors are clustered at the TTWA level.

Chapter 3

Determinants of University Subject Choice in the UK

Abstract

Students applying to British universities have to choose to specialise in a particular subject at the point of application. What factors play into their choice of field of study? I analyse students' subject choices in university using a cohort panel dataset, and in particular, whether those choices respond to their expected wages conditional on studying the subject. I find that expected earnings are positively associated with choice probability, although the quantitative impact of changes in expected earnings on subject choice is small. I also find that there are substantial differences in choice probability between students of different sexes, ethnicities and to a lesser extent, socio-economic classes.
JEL Codes: I21, I23, J10
Keywords: Subject choice, college major, higher education

3.1 Introduction

Students applying to British universities have to choose to specialise in a particular subject at the point of application. How do they choose between different subjects of study? Traditional models of education such as Becker's human capital model or Spence's signalling model alike predict that demand for higher education ultimately derives from the desire to supply labour at better wage levels, in particular industries or professions[1]. Both theories predict that the relative demand for different subjects should be responsive to differentials in the expected lifetime wages conditional on achieving a degree in each subject. Thus, if policy makers were interested in affecting the mix of the subjects students study at university, e.g. if they wanted to promote STEM education as part of industrial policy, they could in theory do so by subsidising STEM degrees and taxing humanities degrees. However, the level of such subsidies would depend on the elasticity of subject choice to changes in the wage structure. It is not clear what the quantitative effect of exogenous changes in the net pecuniary return to education would be on subject choice probabilities.

There has been substantial empirical research trying to measure the elasticity of subject choice to earnings, mainly in the US and French context[2]. I provide a short literature review in subsection 3.1.1. In this paper, I contribute to this existing literature in a few ways. First, I provide a set of estimates of multinomial discrete choice models of university subject choice comparable to those in the literature from a British context. Since the UK uses a different system of university financing, it is of interest whether British students behave differently to those in the US and France. Another major difference is that in the UK, students choose a university and major simultaneously, while in other systems, students can enrol in a college and choose a major after. This may plausibly affect students' responsiveness to pecuniary factors in making college major choices. Second, since I have information on both students'

[1]According to the human capital theory, agents invest in education in the current period to raise their productivity, resulting in higher wages in future periods. According to the signalling theory, agents consume education to signal to future employers that they have high productivity or ability and hence earn a higher wage.

[2]For a survey of this literature, see Altonji et al. (2012), Altonji et al. (2015) and Patnaik et al. (2020b).

self-stated subject choice and the subject they were actually observed to have chosen at university, I can study the effect of using observed subject choice instead of stated subject choice on parameter estimates. Finally, since I use a cohort study with rich background information, I provide extensive descriptive evidence on the correlations between unobserved preference for subjects and background characteristics. I can also investigate whether concerns for expected earnings when making the subject choice decision varies between subgroups in the population.

My main result is that as with other studies using a variety of methods, I find that expected earnings is not a quantitatively large driver of differences in subject choice. A 1% increase in expected earnings for a given subject changes the estimated probability of choosing that subject by between 0.28% (for not going to university) to 0.93% (for STEM). I find some suggestive evidence that higher ability students are more responsive to changes in the pecuniary incentive for studying particular subjects.

I find that a student's level of achievement at school is an important driver of subject choice. There is some evidence that students choose subjects that align with their academic strengths even after accounting for earnings; higher ability in English language, interpreted as language skill, is associated with higher choice probability of humanities and languages and social science and lower choice probability of STEM, and vice versa. Higher overall ability, as proxied by mean KS4 score excluding English and Mathematics, is associated with higher choice of academic subjects like STEM and the humanities and lower choice probability of more instrumental degrees like business and administration.

I also document differences in subject preference by sex and ethnicity, and minor differences between social economic classification groups. I find no differences in subject choice probability between groups with different eligibility for free school meals and groups living in areas with different degrees of income deprivation. Instead, I find that ethnicity seems to be an important correlate to subject choice in the British context, a finding that seems unique to the British setting. I document that minority Asian, Black and 'Other' groups (including mostly Chinese and Arab students) are far more likely to go to university than white students, controlling for relevant correlates. They are also more likely to study social sciences, and professional degrees, and less likely to study humanities and arts.

Finally, I find that there are some differences in the estimated choice probabilities between models that use stated preference and observed choice as the relevant measure of subject choice. I argue that the former is closer to the notion of student preference and that the latter involves an unmodelled supply element. I find that estimated choice elasticities are slightly lower in models using observed choice than in models using stated preference as the choice variable. My results are suggestive of the importance of an unmodelled supply side in the subject choice decision; this would suggest that relying only on student preference to model policy counterfactuals may lead to misleading inference about what would happen in practice.

This paper is structured as follows. I start by providing a short literature review in the following subsection 3.1.1. In section 3.2, I describe the econometric specification of the choice model and the model of individual-specific expected lifetime earnings. Section 3.3 describes the details of the estimation and the data used. Section 3.4 presents the results for estimated predicted life-time earnings while section 3.5 presents the results of the choice models regarding expected earnings while section 3.6 describes the results of the correlates of unobserved preference for different subjects. Section 3.7 concludes.

3.1.1 Literature Review

The question of whether students respond to market incentives when choosing subjects of study is a long-standing one in the economics of education literature. The difficulty of reaching a satisfactory conclusion is due to the difficulty of determining what the relevant notion of expected income is, measuring it accurately, and finding plausibly exogenous variation in earnings in different subjects. The main problem for the researcher is that theoretically, students make their subject choice decision based on their subjective expectations about the relative returns to different subjects conditional on their heterogeneous information sets and self-perceptions of their own ability over their lifetime. This is typically not observed by the researcher, who instead observes wages of those who had selected into the subject of study. One solution is to make assumptions about how students form expectations about conditional wages given observed wages of workers who had selected into that field of study. My work falls within this approach. Examples of work that use observed wages include Berger (1988), Montmarquette et al. (2002), Beffy et al. (2012), and Long et al. (2015). While Berger (1988) and Beffy et al. (2012) both use rational expectations, Long et al. (2015) assume that students observe the labour market returns of past cohorts. Berger finds a positive and statistically significant parameter on expected lifetime earnings, but does not provide choice elasticities. The later papers similarly find statistically significant positive effects, but quantitatively small elasticities of subject choice to expected earnings.

Most of the aforementioned papers, as well as other papers that study the decision to select into university attendance more generally such as Blundell et al. (2016), do not have unambiguously exogenous variation in expected wages. Blundell et al. (2016) argues that heterogeneous individuals have different expectations of their likely earnings and thus would respond to exogenous tax changes differently; i.e. people of different sexes, ethnicities and family background have different expectations of how much they will earn given how much actually observed wages do vary with those variables. This is the approach taken by this paper. Beffy et al. (2012) uses variation of earnings in different subjects across the business cycle as the primary identifying variation. There are few studies using observed earnings where expected earnings vary quasi-experimentally; a notable exception is Abramitzky et al. (2019), which studies how major choice in Israel was affected by changes in earnings for different occupations due to movements within individual kibbutzim from equal sharing to productivity-based earnings. In contrast to the other papers discussed in this review, the authors find that after the reform, students moved into BA subjects which had high wage returns. However, this study was done in a very unique context and it is unknown whether these results would generalise to large changes in expected income for different subjects in other settings.

Another solution to the unobserved expectations problem is to collect subjective expectations data in a controlled survey setting. This approach derives from the criticisms of Manski (1993), and pioneering work in Dominitz & Manski (1994). This approach obviates the need for many assumptions relating to the connection between observed earnings data and subjective expectations about earnings, but introduces the problems of possible measurement error and, due to the costly nature of data collection, unrepresentative samples. Papers in this vein (e.g. Wiswall & Zafar (2014)) have tended to use a smaller sample of students already studying in universities; thus, the estimates derived from these studies may not be representative of population parameters. Examples of such work include Zafar (2013) using a sample of Northwestern students, Arcidiacono et al. (2012) using a sample of Duke students, and Wiswall & Zafar (2014) using a sample of NYU students. Arcidiacono et al. (2012) finds that equalising the abilities of students would increase the share of students doing economics and reduce the share of students doing the humanities while equalising the average earnings across majors would lead to the opposite. Wiswall & Zafar (2014) finds elasticities of major choice to changes in future earnings in each period of around 0.03-0.07 for most majors.

Regardless of method, most papers find that a substantial part of the subject choice decision is not explained by expected income. There is an active and on-going literature exploring what other factors matter for subject choice. The literature has found that various other factors that affect subject choice, such as differences in competitiveness (Buser et al. (2014)), differences in risk preferences and patience (Patnaik et al. (2020a)), and differences in preferences for job characteristics (Wiswall & Zafar (2017)). My work documents ethnic patterns to subject choice that do not show up in the US context that have strong correlations to subject choice.

3.2 Econometric Specification

In the UK, students apply to universities to study specific courses in different subjects roughly around the age of 18. Unlike in other countries in the US where they apply first to be admitted to a university and then select a subject to major in, students make a simultaneous decision whether or not to apply to university, and what to study at university. Since university education is heterogeneous across subjects, students expect that their labour market experience will differ based on their field of study as well as their own characteristics and ability. I first describe a specification of student expectations about the present value of their stream of lifetime earnings conditional on subject studied at university. I then describe a choice model that describes their preference of university subject between available options.

3.2.1 Labour market performance

In this model, students graduate into the labour market at the age of 21, having either no university education or a university education in a subject, indexed by j, and retire at age T (set to 60). In each period, they have a probability R_{ija} of being employed which varies by age, indexed by a. If they are employed, they receive annual earnings W_{ija}. As a simplification, selection into occupations and economic activity is not modelled. Let X_i^r and X_i^w be vectors of variables, that vary over individuals but not over time, that affect employment status and wages respectively.

Denote an individual's present value of lifetime earnings at age 22 conditional on studying subject j by W_{ij}^l. Let T be the age of retirement (set to 60) and δ be the discount rate (set to 0.96). Then W_{ij}^l is the time discounted sum of the product of the probability of being unemployed at age a (R_{ija}) and the wage received at age a (W_{ija}), both conditional on having studied subject j, as follows.

$$W_{ij}^l = \sum_{a=22}^{T} \delta^{a-22} R_{ija} W_{ija} \tag{3.1}$$

The object that matters for subject preferences is the individual's expectation of their lifetime income conditional on studying each subject j, which I denote by $E_i(W_{ij}^l)$. A key problem however is that this term, which is a psychological object, is generally unobserved in representative cohort or panel datasets, as discussed in the literature review. I follow an older literature that aims to specify lifetime expectations with a set of structural assumptions to resolve the issues that arise from not observing expected lifetime earnings[3].

A well-known problem with inferring causal effects of education on wages from observed wage data is that there might be unobserved selection effects, as first observed by Roy (1951) and discussed by Willis & Rosen (1979). An individual might choose subject j over an alternative j' because of factors which affect wages and are observed by the individual but unobserved by the researcher. In this paper, I use the matching approach (e.g. as described in Heckman et al. (2005)), which assumes that there are variables Q, such that $F(W_j|X,Q) = F(W_j|X,Q,S=j)$, that is, conditional on observable variables Q, there is no selection problem. In this case, I use two grades from GCSE English and GCSE Mathematics exams, which all students in England have to take, and which control for their respective verbal and mathematical ability respectively. That is, my model for the data-generating process of earnings at each age is given as follows, where X_i^w denotes a set of individually varying covariates:

$$W_{ija} = X_i^w \beta_{aj} + Q_i \kappa_{aj} + \varepsilon_{ija} \tag{3.2}$$

[3] Manski (1993) argues that imposing such structure implies that students make decisions on the basis of a similar econometric exercise and that is an unrealistic model of student behaviour. Although students are unlikely to recover rational expectations from data on the basis of regressions, it is not inconceivable that they respond to wage changes due to market conditions that are picked up in aggregate wage data. Furthermore, the alternative method in the literature, the elicitation of beliefs about hypothetical choices is relatively more expensive and is typically performed on specific populations with smaller data sets.

The vector of variables Q_i is parameterised in a linear form by vector κ_{aj}. ε_{ija} is a mean zero i.i.d error. This matching approach has also been used in prior research on British returns to higher education (Blundell (2000), Belfield et al. (2018a), Waltmann et al. (2020)) to identify causal effects of university attendance and specific returns to studying particular subjects.

Second, I follow a similar approach to specify a data-generating process for the student's expected probability of unemployment. The probability of being employed conditional on having graduated in subject j at age a is denoted by R_{ija}. I assume that the probability of being employed follows a logit model, such that the probability of being employed is as given in equation 3.3.

$$R_{ij} = \frac{exp(X_i^r \gamma_j)}{1 + exp(X_i^r \gamma_j)} \tag{3.3}$$

Third, to estimate a discrete choice model with both alternative-specific variables (expected earnings conditional on subject) and individual-specific variables (grades, personal characteristics), it is necessary that alternative-specific variables vary across individuals. Moreover, for credible identification of the preference for expected earnings, we require plausibly exogenous individual variation in alternative-specific expectations of lifetime earnings in each of the choices. Previous studies (e.g. Beffy et al. (2012)) use intertemporal variation over the business cycle which was argued to be plausibly exogenous. I follow the approach in Blundell et al. (2016) and Blundell et al. (2019) which argues that students with different family backgrounds have differing expectations over expected earnings in each of the alternatives conditional on their English and Maths ability. These family background parameters are argued to be plausibly exogenously assigned to students at birth, and hence causality on wage flows from these variables to wages, not vice versa. In this application, I use variables on students' family socio-economic status, their family's income (proxied by whether the student qualifies for free school meals the income deprivation score of their home at age 14), whether they are of white ethnicity, quintiles of the KS4 English and Maths scores, and their normalised KS4 point scores for their best 8 subjects. These variables make up the vector

X_i^w specified in equation 3.2.

There is some evidence that students' pecuniary beliefs do indeed vary systematically by observable characteristics. For example, Boneva & Rauh (2020) presents some evidence that students in British high schools believe that being female and having a low socio-economic status (denoted by having parents who did not go to university) would negatively affect their earnings and having good grades would positively affect their earnings. Tan (2022) shows that students seem to respond to high letter grades by adopting more confident education choices.

Finally, at the point of application, I assume that students have rational expectations about their wages conditional on their information set being limited to Q_i and X_i^w. The individual does not observe the realisation of the error term u_{ij} affecting his wages and so takes its expectation which is assumed to be mean 0. Denote the estimated parameters of equations 3.2 and 3.3 from a representative sample by $\hat{\beta}_j$, $\hat{\kappa}_j$ and $\hat{\gamma}_j$. Then $\hat{W}_{ija} = X_i^w \hat{\beta}_j + Q_i \hat{\kappa}_j$ evaluated at age a, and $\hat{R}_{ija} = \frac{exp(X_i^r \hat{\gamma}_j)}{1+exp(X_i^r \hat{\gamma}_j)}$ evaluated at age a. For simplicity, I assume that the expected wage at age a after factoring in unemployment is simply the product of $\hat{W}_{ija}$ and $\hat{R}_{ija}$.

Expected present value of lifetime earnings is denoted by $\hat{W}_{ij}^l$ and is given by the following expression:

$$\hat{W}_{ij}^l = \sum_{a=22}^{T} \delta^{a-22} \hat{R}_{ija} \hat{W}_{ija} \tag{3.4}$$

3.2.2 Choice of subject

Students considering university application choose between J different possible subjects of study plus the outside option of not going to university. They make their decision at least partly based on their expectations about the present value of the stream of earnings over their lifetime conditional on subject choice $\hat{W}_{ij}^l$ as described in the previous section. In the data, we observe their preferred subject $S_i \in \{0, 1, \cdots, J\}$ and also their actually observed education outcome $T_i \in \{0, 1, \cdots, J\}$.

The utility an individual derives from option j is denoted by U_{ij}, and is given by the following expression. The individual chooses S_i to be the option $j \in \{0, 1, \cdots, J\}$ that maximises utility U_{ij}. η_j captures the residual preference or taste for each subject, allowed to vary over characteristics X_i^c. The error term ϵ_{ij} captures aspects of the decision observed by the individual but unobserved by the researcher. The parameter λ_i captures the influence of the expected present value of lifetime earnings for the individual i and plays a role in calculating elasticities of subject choice to exogenous changes in earnings expectations.

$$U_{ij} = \eta_j + X_i^c \alpha_j + \lambda_i \hat{W}_{ij}^l + \epsilon_{ij} \tag{3.5}$$

$$S_i = \underset{j \in \{0,1,\cdots,J\}}{\arg\max} \; U_{ij} \tag{3.6}$$

The terms $\eta_j + X_i^c \alpha_j$ capture a black box of influences sometimes simplistically described as tastes. The parameters in α_j can be interpreted as describing the correlation of personal characteristics like sex with the residual preference for different subjects not captured by the expected present value of lifetime earnings variable. This does not necessarily describe pure preference for different subjects, and could reflect unmodelled deep parameters of behaviour like risk preference, preference for likely job characteristics, etc.

Often, the distribution of λ_i is of interest for policy purposes, for example, if governments wish to increase enrolment in certain subjects. Because nobody makes the subject choice decision multiple times with the same information set, it is impossible to analyse the question with panel data and identify individual-specific coefficients outside of a hypothetical choice setting. Instead of imposing a distribution on the responsiveness to earnings parameter λ_i, I account for possible heterogeneity in the parameter by interacting it with a rich set of personal characteristics variables.

The parameter λ_i is assumed to be of a linear form relating to a vector of personal characteristics variables X_i^λ, which contains indicator variables for sex, ethnicity, socio-economic classification, and quartiles of GCSE scores as a proxy for ability. The coefficient $\bar{\lambda}$ should be interpreted as the responsiveness to earnings of white, male, low-SES students in the bottom GCSE quintile. The vector λ_x captures the correlations between personal characteristics in X_i^λ and responsiveness to expected lifetime earnings.

$$\lambda_i = \bar{\lambda} + X_i^\lambda \lambda_x \tag{3.7}$$

This specification does not allow for heterogeneity in earnings preference within the groups demarcated by the personal characteristics variables X_i, but allows for a more data-driven understanding of heterogeneity between groups. This specification also breaks the implication that substitution patterns are identical in the population. Although individuals in the same group will have identical substitution patterns, substitution patterns will differ between the groups demarcated in X_i^λ.

I assume that the error term ε_{ij} is type I extreme value, and estimate the choice model described by the random utility function in equation 3.5 as a multinomial logit model. The multinomial logit imposes an independence of irrelevant alternatives assumption that implies that the ratio of the choice probabilities of any two options j and k does not depend (in any aspect) on a third option m. Importantly, it imposes a restriction that the cross elasticity of the choice probability P_{ij} with regards to the expected wage of option k is the same for all $j \neq k$ for each individual, although these elasticities vary across individuals.

Nevertheless, the multinomial logit model imposes constant cross-elasticities across alternatives with regards to changes in the associated earnings of any particular alternative. I hope to ameliorate this inflexibility by choosing appropriate choice categories; the IIA property becomes more reasonable if the choice set contains alternatives that are less similar, in an intuitive sense. I aggregate the number of alternatives to 6, chosen to be in relatively distinct fields. Table 3.4 in the summary statistics in the next section summarises the distribution of subjects in my chosen categories in the data.

3.3 Data and Estimation

3.3.1 Estimation

The ideal data for estimating the model above is a large panel of earnings at many points in the life-cycle that records subject choice and a rich set of personal characteristics variables. A simple way to estimate the model would then be to first estimate equations 3.2 and 3.3 by OLS and maximum likelihood to derive parameters $\hat{\beta}_j$, $\hat{\kappa}_j$ and $\hat{\gamma}_j$, and use these parameters to calculate $\hat{W}^l_{ij}$ for each individual and each subject (making an assumption about δ).

Unfortunately, such a dataset was not available during the execution of the project and instead I have merged two separate datasets. First, I use the LSYPE, a cohort study of a representative sample of the cohort born in 1989-90, as my main dataset for data on higher education choices, achievement variables in secondary school and high school, personal characteristics, and earnings at age 25. The results should thus be interpreted as representative of a particular cohort of students born in 1989/90 and who first started attending higher education in 2008. Unfortunately, the wages of the cohort used are only observed once at age 25 in the LSYPE dataset. Thus, I am unable to estimate life-cycle profiles using the LSYPE data. Alternative datasets with a longer panel of wages[4] however do not contain the same breadth of information about students' personal characteristics.

I deal with this issue by approximating lifetime earnings by estimating an age profile of earnings estimated from a less detailed specification of equation 3.2 from an alternative data source, the UK Labour Force Survey. I then apply these estimated wage profiles to point estimates of expected annual earnings at age 25 in the main dataset; in practice, this amounts to multiplying the predicted earnings at age 25 by a multiplier that differs by the student's

[4] In recent years, the LEO dataset from the UK Department of Education contains many of these features, minus perhaps a rich set of personal characteristics variables.

individual characteristics. First, I use the age 25 earnings from the LSYPE to calculate the regression suggested by equation 3.2. This gives me individually varying values of expected earnings at age 25, which I denoted $\hat{W}_{ij,25}$. As a robustness check, I compute an analogous expected median wage at age 25 for each subject j using quantile regression, denoted $\hat{W}^{med}_{ij,25}$.

I then use the LFS to perform a binary logit regression of being unemployed on a quadratic polynomial of age, and sex to estimate equation 3.3. I use these regressions to generate estimated probabilities of being unemployed conditional on studying subject j at each age from 25-60. I denote these estimated probabilities of unemployment at age a $\hat{R}_{ija}$. In practice, the vector X^r_i includes a quadratic polynomial in age and a dummy variable for sex and R_{ija} is equal to R_{ij} evaluated at age a.

Finally, I regress wages conditional on the subject studied on a smaller set of explanatory variables, and a quartic polynomial on age, together with various interactions, as in equation 3.8. I refer to the X variables in the estimation of earnings at age 25 by X^w_i and the X variables in estimating the age dynamics X^{w1}_i. Under the presented specification, individuals with different sex, GCSE and A-level attainment are allowed to have different age profiles of earnings conditional on the same subject studied at university.

$$W_{ij} = X^{w1}_i \beta_{0j} + \varepsilon_{ija} \tag{3.8}$$

I use these estimates to calculate an expected wage at each age from age 25 to 60 for each student in the sample. I then divide these wages by the predicted wage at age 25 (for the specification in equation 3.8) to generate a list of income multipliers for each student at each age and for each subject. I denote these estimated multipliers by $\hat{\mu}_{ija}$[5].

[5] Note that this differs slightly from the specification in the econometric section, which considers earnings from age 22. This is due to data limitations. This omission has the effect of making choices with front-loaded earnings less attractive.

My expected lifetime earnings variable is then calculated as follows, with δ assumed to be 0.97.

$$\hat{W}_{ij}^{l} = \sum_{a=22}^{60} [\delta^{a-25} \cdot \hat{R}_{ija} \cdot \hat{\mu}_{ija} \cdot \hat{W}_{ij,25}] \tag{3.9}$$

I also calculate an alternative measure applying the estimated multipliers to the predicted median wage (for each subject) instead of the predicted mean wage as follows:

$$\hat{W}_{ij}^{l,med} = \sum_{a=25}^{60} [\delta^{a-22} \cdot \hat{R}_{ija} \cdot \hat{\mu}_{ija} \cdot \hat{W}_{ij,25}^{med}] \tag{3.10}$$

I then estimate the choice model using maximum likelihood estimation. Given the assumptions used to approximate the age-wage trend to calculate life-time earnings, I also report results for the multinomial logit using estimated wages at age 25 as the earnings variable as a robustness check. I estimate alternative-specific coefficients for each of the variables in X_i^c. It is well-known that the individual choice probabilities have an analytical form:

$$P_{ij} = \frac{\exp(U_{ij})}{\sum_{k \in J} \exp(U_{ik})} \tag{3.11}$$

The objective function of the maximum likelihood problem is the log of the probability that the individuals make choose as they did in the sample. Let d_{ij} denote the probability that individual i chooses subject j.

$$LL(\beta, \gamma_j, \eta_j) = \sum_{i=1}^{N} \sum_{j \in J} d_{ij} P_{ij} \tag{3.12}$$

$$\{\hat{\beta}, \hat{\gamma}_j, \hat{\eta}_j, \forall j \in \{0, 1, \cdots, J\}\} = \underset{\beta, \gamma_j, \eta_j}{\arg\max}\, LL(\beta, \gamma_j, \eta_j a) \tag{3.13}$$

McFadden (1974) shows that if U_{ij} is linear in parameters, the log-likelihood function is globally concave in the parameters.

Table 3.1: Definition of variables in empirical analyses

Variable vector	Description	Variables
X_i^w	Covariates of earnings at age 25 (equation 3.2)	Sex, ethnicity (5 categories), socio-economic status, whether student has free school meals, index of income deprivation score of area student lived at when aged 14
X_i^{w1}	Covariates in estimation of wage-age profile (equation 3.8)	Quartic polynomial on age, interaction of quadratic polynomial terms of age on sex, whether the individual has more than 7 GCSEs, and whether non-white. Controls for year fixed effects, sex, ethnicity (5 categories), and whether more than 7 GCSEs.
X_i^r	Covariates in the employment probability equation (equation 3.3)	Quadratic polynomial of age, interacted with sex, whether non-white, whether has more than 8 GCSE passes; year fixed effects
X_i^c	Covariates in the choice equation (besides ability variables) (equation 3.5)	Sex, ethnicity (5 categories), socio-economic status, whether student has free school meals, index of income deprivation score of area student lived at when aged 14
X_i^λ	Covariates in heterogeneity of λ_i equation (equation 3.7)	Sex, socio-economic classification, whether student has free school meals, ethnicity, normalised point score for 8 best KS4 subjects
Q_i	Ability variables in equation 3.2	Normalised point score in GCSE English Language and Literature, normalised point score in GCSE Mathematics, normalised average KS4 point score (excluding English and Maths))

The details of the many different sets of explanatory variables are summarised in table 3.1.

To study the dependence of my results on various modelling assumptions, I estimate a number of versions of my model. For clarity, I number the permutations of my estimated models according to the numbering in table 3.2. There are three elements that are varied: the definition of the choice variable, the definition of the wage variable, and whether interactions are included on the preference coefficient for earnings. My preferred specification in this paper is model 1, which uses expected mean lifetime earnings, no interactions on the wage variable, and stated preference as the choice variable. This is the model consistent with expected utility theory, although it is possible that in practice students care about other moments of the distribution of outcomes when assessing lotteries.

Table 3.2: List of Estimated Model Variations

Model No.	Choice var	Wage var	Interactions?	Sample
1	Stated pref	Mean employment-weighted lifetime earnings	No	All
2	Stated pref	Median employment-weighted lifetime earnings	No	All
3	Observed choice	Mean employment-weighted lifetime earnings	No	All
4	Stated pref	Mean employment-weighted age 25 earnings	No	All
5	Stated pref	Mean employment-weighted lifetime earnings	Yes	All
6	Observed choice	Mean employment-weighted lifetime earnings	Yes	All

3.3.2 LSYPE

I use data from the Longitudinal Study of Young People in England (LSYPE), also known as Next Steps[6][7]. The study is a longitudinal study that followed the lives of 16,000 people born in 1989-90 in England. The survey began in 2004 and followed a panel of young people born between 1st September 1989 and 31st August 1990 for seven years until they were 20-21. The survey data is also linked with some administrative records from the National Pupil Database. All young people in the initial sample were contacted again in 2015-16 (when they were 25).

Table 3.3 summarises the timing of the LSYPE waves. Fieldwork is typically carried out in June to October of the listed year. As the table shows, waves 5 to 7 contain the most relevant information to the university application decision.

Table 3.3: LSYPE Waves and Academic Milestones

LSYPE Wave	Year of Fieldwork	Age (at 31 August)	Schooling
1	2004	14	Year 9 to 10
2	2005	15	Year 10 to 11
3	2006	16	Year 11 (GCSEs) to 12
4	2007	17	Year 12 to 13
5	2008	18	Year 13 (A-Levels) to HE
6	2009	19	Completes first year of HE
7	2010	20	Completes second year of HE
8	2015-16	25-26	-

The LSYPE observes the entire period over which young people typically apply for univer-

[6]University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2020). Next Steps: Linked Education Administrative Datasets (National Pupil Database), England, 2005-2009: Secure Access. [data collection]. 6th Edition. UK Data Service. SN: 7104, http://doi.org/10.5255/UKDA-SN-7104-6

[7]University College London. UCL Institute of Education. Centre for Longitudinal Studies, Next Steps: Sweeps 1 and 8, 2001 and 2016: Geographical Identifiers, 2001 Census Boundaries: Secure Access [computer file]. Colchester, Essex: UK Data Archive [distributor], June 2017. SN: 8189.

sity and records survey data about their actions and attitudes in this time. Furthermore, it is linked to the student record, allowing the researcher to understand in part the information students had about their ability, as inferred from their performance in Key Stages 3-5. As I describe later, I use information mainly from Key Stage 4, which corresponds to a compulsory national examination, called the GCSEs, administered to all school-leavers in the UK.

Sample

I form my sample as follows. Since I intend to include not attending university as the outside option, the sample to the group of people who might plausibly have gone to university but for whatever reason did not. I use the population of students with at least 5 GCSE results with grade C or above [8]. I further drop any observations with missing data and missing weights.

Subject choice

As mentioned in the econometric section, I consider two definitions of subject choice. First, in wave 5 of the LSYPE, young people are asked what subject they would like to study at university. This is the subject they had applied to university for (if they had applied that year) or what subject they are likely to apply to university for (if they state that they are likely to apply to university). I adopt this as my first definition of subject choice, and refer to it as “stated preferred subject”. Second, waves 5-7 record which subject the individual actually studied at university, if they did attend university. This is referred to as “observed choice”. In both cases, if the individual was recorded as having a not applicable value, I assigned their choice to be “no university”.

I aggregate the wide array of plausible subjects of study in university to 6 major categories, including an outside option for not attending university at all. STEM aggregates all science subjects, mathematics, computer science and engineering. Professional subjects aggregates

[8] This is the sample selection criteria used by Belfield et al. (2018a). Some other papers, like Blundell (2000), use the population of students with at least 1 A level qualification.

university subjects which are precursors to specific professions, like architecture, teaching, medicine and subjects related to medicine, such as pharmacy. This category also includes administration and business degrees. The classifications are based on the JACS 3 classification used in the UK by the Higher Education Statistics Agency (HESA) for administration purposes.

Table 3.4 below summarises the subject distributions in the estimation sample by whether observed choice is used or stated choice is used. Previous work on subject choice had been done in settings where students were not constrained in their subject choice by whether they would be admitted into the program of choice[9]. In the UK context however, students can be rejected from their university choices if their grades are insufficient, leading to their actually observed higher education courses not being the outcome of free choice. Thus, the stated preference should be viewed as the choice variable closest to those used in previous work.

As table 3.4 shows, there are notable differences between these notions of choice. While 48% of students express a preference for a university major over not attending university, only 37% do end up attending university. More students say that they want to pursue a degree in business administration, social science and medicine-related fields than students actually do. While I use stated preference as my main measure of choice in the following analysis, I also include estimates using observed choice as the dependant variable.

[9] For example, in the US, major choice occurs after students had been admitted to university.

Table 3.4: Distribution of subjects in HE for young people with more than 5 GCSE passes in estimation sample

Subject	Observed Choice	Stated Choice
No university	0.481	0.373
STEM	0.161	0.167
Humanities and languages	0.082	0.087
Social sciences	0.082	0.098
Professional subjects	0.177	0.239
Arts	0.017	0.035
N	4265	4265

1 This column presents the weighted distribution of actual subject choices by percent in the sample of students with at least 5 GCSEs observed in waves 5 to 7. The estimation sample used for choice analysis is a subset of all observed young people with 5 GCSE passes excluding those with missing personal characteristics data. The unweighted N is 4839.
2 This column presents the weighted distribution of stated preferred subject choices by percent in the estimation sample. The estimation sample used for choice analysis is a subset of all observed young people with 5 GCSE passes excluding those with missing personal characteristics data.
Source: Analysis on Next Steps data.

High school achievement variables

In the UK, students undergo five stages of education, known as Key Stages, as part of the National Curriculum. I use information mainly from Key Stage 4, which corresponds to education in years 10-11 or ages 15-16 and is compulsory for all Britons. In particular, I use three variables: the normalised point score the student achieved in GCSE English Language and Literature, the normalised point score the student achieved in GCSE Mathematics, the normalised total point score the student achieved, capped at 8, and the average KS4 point score, excluding English and Maths. The GCSE English and Literature, and GCSE Mathematics papers are compulsory for all Britons, and are used in this study as measurements of language ability and maths ability.

The KS4 point scores are the main measure for language and mathematical ability and importantly, they are recorded for all students, not just those who select into university. Summary statistics for these variables are included below in table 3.5 below. On average, students in the sample achieved a KS4 grade in English and Maths of about 44, equivalent to just below a B grade. Similarly, the mean capped GCSE point score is equivalent to 8 grades that are slightly below B. The average GCSE outside of the two compulsory subjects is 47.1, just above a B grade. I normalise the point score in each case over the entire cohort.

Table 3.5: Summary statistics of ability variables in the sample

Variable	Mean	Standard Deviation
KS4 maths GCSE points (best attempt)	44.045	7.522
KS4 English GCSE points (best attempt)	44.633	6.469
Capped GCSE and equivalents new style point score.	366.813	44.271
Average GCSE point score, excluding English and Maths	47.149	8.483

Personal characteristics

I use a relatively restricted set of personal characteristics variables: sex, ethnicity, socio-economic status, whether the student qualifies for free school meals, and the relative deprivation of the area the student was living at age 14. Table 3.6 summarises the distribution of ethnicities in the sample. The sample is comprised mainly of white students. The next largest ethnic group is Asian at 6% of the sample. Each of the other ethnic groups comprises about 2% of the sample.

Table 3.6: Distribution of personal characteristics in the sample

Ethnicity	Mean	Standard Deviation
Ethnicity		
White	0.868	
Asian	0.059	
Black	0.028	
Mixed	0.023	
Other	0.023	
Whether female	0.544	
Family SEC, coded to be high (professional occs),or low	0.484	
Free school meal	0.069	
IDACI score	0.170	0.157

Table 3.6 also summarises the distribution of personal characteristics in the data. Note that since the sample is comprised of students with at least 5 GCSE passes, it will not look like an average cross-section of the 18 year old population in that cohort. The socio-economic

classification variable is constructed from survey questions to the students' family asking them to list their socio-economic classification out of 8 choices. I aggregated these choices into two categories, where a student is in the higher SES if the family is from "managerial and professional occupations", and in the lower SES if the student is from a family with any other classification (including "intermediate occupations", "small employers and own account workers", "lower supervisory and technical occupations", "semi-routine occupations", "routine occupations", or "never worked/long-term unemployed"). The IDACI score represents the share of children aged 0-15 in the students' postcode considered to be income-deprived. The free-school meals variable records whether the student was eligible for free school meals, which occurs when parents are eligible for benefits like the Universal Credit, Income Support etc. and are enroled in a state school. As such, it is commonly used as one of the proxies of deprivation in the UK education literature.

3.3.3 Labour Force Survey

One weakness of the LSYPE is that it has relatively sparse information about individuals' labour market outcomes in their 20s. The individuals studied by the LSYPE were only surveyed once in their twenties at age 25 and there is no information about their earnings at ages above 25. I thus supplement the information drawn from the LSYPE with the Labour Force Survey. The LFS is a quarterly survey of a wide range of topics related to individuals' and households' economic activity, including occupation, training, hours of work and personal characteristics. Although it contains less detail than the LSYPE, the LFS nevertheless contains numerous variables about students' high school and university outcomes.

I use the LSYPE for two purposes. First, I check the earnings of individuals born in 1989 and 1990 at age 25 in the eight quarters in 2014 and 2015 as a sense check for my estimates of predicted earnings at age 25. These results are reported in table 3.8 in section 3.4 below.

Second, I use years 2007 to 2010 of the Labour Force Survey to estimate equation 3.2

for the expected age-wage profile, as discussed in section 3.2 below. I restrict the sample to individuals making a positive wage who had at least 5 GCSEs. The trend of the age-wage profile depends on the individual's sex, ethnicity, and whether or not they have more than 7 GCSE passes. I also use the table to estimate the probability of unemployment conditional on degree subject.

There are two issues with using the LFS. First, it often only records subject details of the highest qualification that an individual has achieved. They do not systematically record the subject of an individual's first degree until 2012. Second, it is possible that wage growth might be substantially higher for high-achieving students in different subjects. Since the available high school achievement variables in the LFS cannot identify such individuals, we might underestimate their own expectations of their lifetime earnings.

3.4 Labour Market Performance of Different Majors

3.4.1 Predicted annual earnings at age 25

My argument for expected earnings being exogenous with respects to subject choice is based on the idea that students with different schooling performance and socio-demographic characteristics will have different expectations of the wage they will achieve in different subjects. For example, students who have stronger language skills (as measured by a high KS4 english score) may expect to earn more after studying a humanities degree than one with a lower KS4 english score. I then argue that I can observe proxies of the relevant ability variables and thus can recover these expectations through standard regression techniques. To find the effects of these correlates to earnings, I regress wages on variables X_i^w and Q_i as described in table 3.1. I present the results for the 6 regressions that form the basis of the predicted mean age 25 earnings used in the choice model in table 3.7, with standard errors in parentheses.

Although not the primary focus of this study, it is interesting to examine the results of these Mincer regressions as they form part of the basis of the variation used to identify the elasticities of subject choice to earnings. I find that as expected, having higher maths and English ability is associated with higher earnings. There also appears to be some evidence of complementarities of language or maths abilities with labour market success conditional on degree; for example, having higher maths ability increases earnings conditional on studying STEM more than for some subjects like humanities.

I find gender wage gaps for all subjects except the arts, and this wage gap is significant at the 5% level for three of the five other subjects. The wage gap is smallest in the arts, humanities and languages, and the social sciences in that order, and largest for students without degrees, students in STEM and students in professional subjects. I find little evidence ethnicity-based disadvantages in earnings at age 25, although it is possible that ethnicity-based wage gaps may open up later in life. I find that being from disadvantaged areas is strongly correlated with lower labour market outcomes for all subjects besides the arts, and that this effect is strongest for STEM, professional subjects and humanities and the languages.

Table 3.8 below summarises the predicted average wage in each subject for the estimation sample in column 1, and its standard deviation in column 2. Columns 3 and 4 present the mean and standard deviation of the predicted median wage as a robustness check. These numbers were calculated by running quantile regressions of log gross annual earnings conditional on subject earnings on the same set of covariates as used in computing expected mean earnings at age 25. Columns 5 and 6 present the mean earnings for the cohort at age 25 from the LFS and the number of observations the mean is taken from respectively. I conclude that the predicted earnings using data from the LSYPE is broadly in line with the labour force survey data.

Table 3.8: Predicted age 25 earnings by subject

	Expected Mean Earnings		Expected Mean Earnings for Selected		LFS Observed Earnings	
Subject	Mean	SD	Mean	SD	Mean	N
No university	20440.28	3873.32	19492.72	3598.83	19095.58	496
STEM	22132.35	3860.08	24457.04	3729.40	23384.27	121
Humanities and languages	19501.31	3134.10	21504.06	2869.84	22964.72	55
Social sciences	21442.08	4668.11	22974.52	4721.19	22259.77	76
Professional subjects	23797.41	4862.50	25182.24	4965.78	23659.85	141
Arts	17686.28	2505.43	18327.89	2572.75	15624.16	45

Table 3.7: Regression coefficients for log annual gross pay at age 25

	(1) No uni	(2) STEM	(3) Humanities	(4) Social science	(5) Prof subjects	(6) Arts
Normalised KS4 maths point score	0.079*	0.116*	0.070	0.178**	0.133*	0.191*
	(0.033)	(0.054)	(0.069)	(0.060)	(0.054)	(0.086)
Normalised KS4 English point score	0.116**	0.096	0.139	0.127*	0.100*	-0.113
	(0.034)	(0.050)	(0.073)	(0.060)	(0.047)	(0.143)
Average GCSE point score (excluding English and Maths)	-0.029	0.009	0.063	0.023	0.058	-0.016
	(0.027)	(0.063)	(0.061)	(0.054)	(0.046)	(0.097)
Female	-0.307**	-0.189**	-0.088	-0.101	-0.158**	0.115
	(0.032)	(0.048)	(0.063)	(0.057)	(0.058)	(0.107)
Upper SEC	-0.046	-0.080	-0.023	0.137**	0.026	0.136
	(0.035)	(0.051)	(0.072)	(0.057)	(0.057)	(0.102)
Whether eligible for Free School Meal	-0.047	-0.084	-0.040	0.057	-0.132	-0.215
	(0.063)	(0.082)	(0.198)	(0.090)	(0.093)	(0.152)
Asian	0.120	0.069	0.059	-0.056	0.300*	0.057
	(0.075)	(0.070)	(0.221)	(0.089)	(0.131)	(0.183)
Black	0.164*	0.125	0.098	0.069	0.192*	-0.184
	(0.081)	(0.098)	(0.397)	(0.095)	(0.078)	(0.335)
Mixed	0.097	0.048	0.039	0.083	0.182	-0.101
	(0.116)	(0.074)	(0.124)	(0.134)	(0.114)	(0.134)
Other	-0.009	0.099	0.221**	0.199*	0.184*	-0.017
	(0.093)	(0.132)	(0.081)	(0.099)	(0.082)	(0.123)
IDACI score	-0.249*	-0.510**	-0.404	-0.113	-0.422*	0.285
	(0.103)	(0.181)	(0.292)	(0.226)	(0.183)	(0.262)
Constant	10.030	0.053	-0.201*	-0.271**	-0.016	-0.457**
	(0.035)	(0.081)	(0.102)	(0.086)	(0.085)	(0.104)

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$

1 This table presents regression coefficients of a Mincer regression of log annual earnings at age 25 on the variables listed. The regression was performed on the entire dataset by interacting the subject of study with the specified variables as follows. Robust standard errors are calculated.
Source: Analysis on Next Steps data.

Age 25 earnings are generally not representative of full lifetime earnings, and in particular, graduates tend to have faster wage growth in the middle of their career than non-graduates. High ability students who choose not to go to university tend to do well relative to their peers with similar ability who go to university early in their careers, but there is evidence in Waltmann et al. (2020) that over the lifetime, those who attend university end up doing better. This is also what I find in my analysis.

The following table presents the population average predicted mean and median lifetime, employment-weighted earnings weighted with an annual discount rate of 0.97. On average, I find that discounted lifetime earnings range between 580,000 pounds (for not attending university) to 821,000 pounds (for studying professional subjects) on average. Thus, while the gaps between subjects appear small in table 3.8, they seem to open up later in life, as seen in table 3.9.

Table 3.9: Mean Predicted Lifetime Earnings by Subject and Statistic

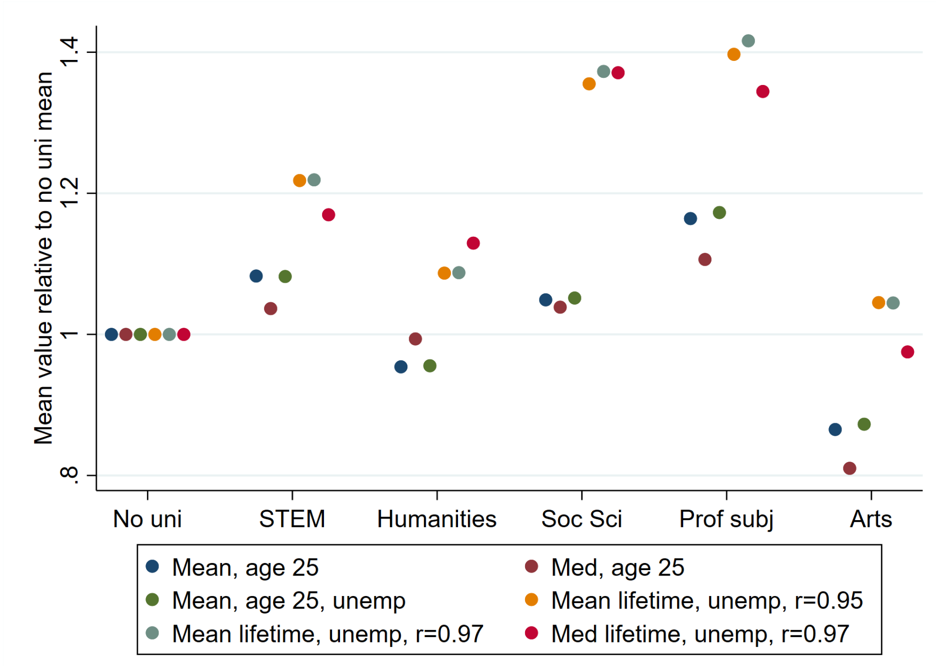
Subject	Mean emp-weighted lifetime earnings (r = 0.97)	Median emp-weighted lifetime earnings (r = 0.97)
No university	5.80e+05	6.10e+05
STEM	7.07e+05	7.13e+05
Humanities and languages	6.30e+05	6.89e+05
Social sciences	7.96e+05	8.36e+05
Professional subjects	8.21e+05	8.20e+05
Arts	6.05e+05	5.95e+05

1 This table summarises the predicted lifetime earnings for each subject, with the 2nd column showing the average earnings and the 3rd column showing the median earnings.
Source: Analysis on Next Steps data.

3.4.2 Unemployment and life-cycle Wage profiles

As mentioned in the previous subsection, earnings at age 25 are not representative of earnings over the entire life-cycle, which is arguably what rational decision-makers should consider when making education decisions. In this subsection, I consider the effect of unemployment and earnings over the life-cycle on the relative attractiveness of various subjects at university. The computation of these elements was discussed in section 3.2.

Figure 3.1: Illustration of the effect of modelling assumptions on relative pecuniary incentives



Notes: The graph plots the mean predicted value for each subject relative to the mean predicted value for not going to university, for five wage specifications.

Figure 3.1 examines the relative predicted pecuniary return of various subjects relative to not going to university under six different specifications; the corresponding table is included in the appendix as table 3.17. Each dot in the graph represents the ratio of predicted earnings under that specification relative to not attending university.

I make the following observations. First, it seems that the choice of the mean or the median as the relevant moment substantially affects the relative attractiveness of subjects; it increases the attractiveness of the humanities, and decreases the attractiveness of the other four subjects relative to the no university baseline. Second, accounting for unemployment seems to be relatively unimportant quantitatively. Third, I find that accounting for the life-cycle is very important and increases the relative attractiveness of going to university substantially for almost all subjects and particularly for social science and professional subjects. Finally, I find that the choice of the discount rate is of second-order importance, and that mechanically it affects the relative attractiveness of subjects where earnings are accrued later in life.

Given these considerations, I use unemployment-weighted, mean life-cycle earnings with a discount rate of 0.97 as my main earnings variable in the subsequent choice analysis, and report estimates for models using the median lifetime earnings and mean age 25 earnings as robustness checks.

3.5 Preference for Expected Labour Market Income and Implied Elasticities

In this section, I discuss the results on the relation between expected wage and subject choice (denoted by λ_i in the choice model). I first present results from models 1, 2, 3, and 4, which cover models with the mean age 25, both mean and median lifetime expected annual earnings, without interactions, for both stated and observed choice. These results are presented in table 3.18, included in the appendix. In the main text, I will focus on the elasticities implied by the estimates.

The point estimates of the parameters in a nonlinear multinomial logit model can be difficult to interpret quantitatively. To assess the quantitative impact of expected earnings on subject choice, I calculate the elasticity of choice probability of each subject to a change in its own expected wage. As discussed in Train (2009), this can be interpreted as the % change of the subject probability in each subject given a 1% increase in expected earnings in that subject (for the respective earnings variable). These elasticities can be calculated as follows:

$$\text{choice elasticity}_{ij} = \frac{\partial U_{ij}}{\partial \hat{W}^l_{ij}} \times \hat{W}^l_{ij} \times (1 - Pr_{ij}) \tag{3.14}$$

Table 3.10 summarises the implied elasticities of an increase in the associated wage variable for each specification on the choice probability for 6 different subjects, with the standard errors in parentheses.

Table 3.10: Mean elasticity of choice probability to changes in own-alternative expected lifetime earnings under 4 specifications

Subject	(1)	(2)	(3)	(4)
No university	0.890	1.415	0.651	1.350
	(0.296)	(0.351)	(0.251)	(0.646)
STEM	1.848	2.871	1.735	2.399
	(0.620)	(0.720)	(0.675)	(1.154)
Humanities and languages	1.758	2.910	1.673	2.386
	(0.593)	(0.739)	(0.654)	(1.151)
Social sciences	2.227	3.517	2.230	2.647
	(0.756)	(0.902)	(0.878)	(1.283)
Professional subjects	1.917	2.918	2.037	2.413
	(0.642)	(0.735)	(0.793)	(1.161)
Arts	1.732	2.591	1.654	2.306
	(0.597)	(0.683)	(0.671)	(1.127)
Log-likelihood	-5877.5	-5873.9	-5215.7	-5879.7

I find average elasticities ranging from 0.89 to 2.2, with students most responsive to earnings changes in the social sciences and least responsive to earnings changes in the no-university option. These estimates are larger than those estimated in Beffy et al. (2012), although comparing them is not straightforward given the differences in the earnings definitions. Although these elasticities seem to be large, note that a 1% increase in lifetime earnings is typically a large change in absolute terms. For example, 1% of average discounted lifetime earnings conditional on studying the social sciences is about 8000 pounds. Nevertheless, these estimates suggest that the government has some leeway to change the relative demand for various subjects using pecuniary policies, like subject-specific tuition caps.

I find that using observed choice rather than stated choice attenuates the implied elasticities for most subjects besides social sciences and professional subjects, suggesting that the unmodelled supply side is relevant in determining the actual subject mix that occurs in practice. Using age 25 earnings instead increases the implied elasticities, although the model fit worsens marginally. Using median lifetime earnings instead of mean lifetime earnings also increases the implied elasticities and improves the model fit.

Table 3.19 specifies the full pattern of substitution (averaged over the population) implied by specification 1, and is included in the appendix. The multinomial logit functional form assumption implies that the cross-elasticity of choice is the same for all other alternatives for each person, and as such substitution patterns in the case where the preference parameter is identical for all students is of relatively less interest. This issue is ameliorated in the case of models where the preference parameter varies by individual characteristics.

3.5.1 Heterogeneity of Elasticities

Table 3.11 shows the estimated heterogeneity on the relative importance of wage variables (specifications 5 and 6) for subject choice, with t-statistics in parentheses. I do not find many statistically significant cases of heterogeneity, although it seems that this seems to be an issue of power[10] and the estimated heterogeneous components seem to be relatively large in size. I also find that including these interactions substantially affects my estimates of the elasticity

[10] Since the expected earnings variable varies by characteristics specified in table 3.1, it will not vary conditional on those variables. This means that the variation in expected earnings which identifies the coefficients in λ_i comes from the variables not included in the λ_i term. These variables include the IDACI score and the specific composition of the GCSE score conditional on the total normalised average KS4 score (i.e. whether students achieved their score by being stronger in maths, english or other subjects).

of subject choice to earnings.

In both cases, the most important dimension of heterogeneity is that higher ability students, especially those in the top quartile of GCSE scorers, are more responsive to expected earnings. I find relatively few dimensions of heterogeneity in terms of personal characteristics, a notable exception being that being from a higher socio-economic status is associated with a lower elasticity of subject choice to earnings.

Table 3.11: Interacted coefficients on wage variable for 2 specifications of choice model

	Choice inferred from stated preferred subject	Choice inferred from actually observed univ subject
Wagevar	9.8×10^{-7}	-1.14×10^{-5}
	(0.19)	(-1.41)
Female $\times$ Wagevar	-6.99×10^{-7}	-4.22×10^{-7}
	(-0.70)	(-0.37)
Upper $\times$Wagevar	-5.94×10^{-7}	$-1.91 \times 10^{-6**}$
	(-0.99)	(-2.77)
Asian $\times$Wagevar	-9.83×10^{-7}	-1.23×10^{-7}
	(-0.92)	(-0.11)
Black $\times$Wagevar	2.41×10^{-7}	-1.23×10^{-6}
	(0.15)	(-0.68)
Mixed $\times$Wagevar	-1.16×10^{-6}	2.69×10^{-7}
	(-0.76)	(0.14)
Other $\times$Wagevar	-2.08×10^{-6}	-1.69×10^{-6}
	(-1.27)	(-0.96)
Ability_qrt=2 $\times$ Wagevar	-9.75×10^{-7}	9.86×10^{-6}
	(-0.21)	(1.27)
Ability_qrt=3 $\times$Wagevar	3.51×10^{-7}	1.31×10^{-5}
	(0.08)	(1.69)
Ability_qrt=4 $\times$ Wagevar	2.14×10^{-6}	$1.53 \times 10^{-5*}$
	(0.46)	(1.97)
Observations	29034	29034
ll	-5862.0	-5184.3
choicevar	Stated choice	Observed choice

1 This table presents coefficients on the individual and subject-specific expected earnings variable in models 5-6 as specified in table 3.2.
Source: Analysis on Next Steps data.

Table 3.12 shows the implied own- and cross- of elasticities of subject choice to changes in earnings for model 5, using stated choices. The first-order finding is that the elasticities are much smaller relative to the models without interactions on the wage variable coefficient. The standard errors are also larger given the high-dimensional specification estimated. The cross-elasticities are also relatively more asymmetric relative to the models without interactions. In particular, STEM, humanities and languages, and social sciences are subjects where enrolment is most responsive to changes in their pecuniary attractiveness.

Table 3.12: Average substitution patterns between subjects (specification 5)

Subject	No university	STEM	Humanities and languages	Social sciences	Professional subjects	Arts
No university	0.275	-0.211	-0.168	-0.135	-0.147	-0.108
	(0.603)	(0.332)	(0.293)	(0.343)	(0.387)	(0.503)
STEM	-0.118	0.925	-0.321	-0.227	-0.235	-0.114
	(0.183)	(1.172)	(0.325)	(0.269)	(0.279)	(0.187)
Humanities and languages	-0.045	-0.153	0.909	-0.128	-0.093	-0.068
	(0.078)	(0.157)	(1.202)	(0.168)	(0.122)	(0.122)
Social sciences	-0.049	-0.160	-0.187	0.912	-0.108	-0.072
	(0.123)	(0.193)	(0.246)	(1.533)	(0.192)	(0.165)
Professional subjects	-0.133	-0.401	-0.326	-0.265	0.737	-0.149
	(0.345)	(0.485)	(0.422)	(0.477)	(1.277)	(0.376)
Arts	-0.010	-0.020	-0.024	-0.019	-0.016	0.423
	(0.053)	(0.035)	(0.046)	(0.047)	(0.045)	(1.281)

3.6 Unexplained Preferences for Subjects and Correlations with Observables

The models estimated produce alternative-specific constants and subject-specific coefficients on individual-specific observable variables which do not differ by alternative. These coefficients can be interpreted as capturing the preference for each alternative that is not captured by the alternative-varying variable, in this case, the expected lifetime earnings conditional on studying each subject, and how this preference is correlated with individual attributes. These coefficients should be interpreted descriptively as correlations, and the intention is that these results can motivate further research into the deeper causal factors that underlie these relations.

I examine correlations in four areas: sex, ethnicity, family background, and prior schooling performance. I find that sex, ethnicity and prior grades are extremely important factors in subject choice. Women select into higher education at a higher rate even after controlling for ability and select into STEM (excluding biological sciences) at a far lower rate than men. Asian, Black and other ethnicity students select into higher education at a far higher rate than white students do, and typically have very different subject choice patterns. I find that there is only a small effect of family background, in particular, family socio-economic status, whether the student is eligible for free school meals, and whether the student lived in an area with high income-deprivation, on subject choice, after controlling for other factors. I also find that schooling performance matters substantially for subject choice. Students with greater language ability (measured by performance on the GCSE English exam) are more likely to select humanities subjects and less likely to select into STEM, and students with greater mathematical ability are more likely to select into STEM and medicine and less likely to choose the humanities. Students with greater KS4 performance overall are more likely to attend university, more likely to choose traditional academic subjects like STEM and the humanities, and less likely to choose business and administration.

As with the coefficient on the earnings variable, the quantitative importance of each coefficient in a multinomial discrete choice model is difficult to interpret. For each factor, I present two summary indications of the factors' quantitative importance. To give a sense of how important each factor is, I provide calculations of the average choice probabilities for each subject over the whole population if the value for the factor were changed but the student was otherwise the same. For example, to consider the impact of sex on subject choice probabilities, I calculate the average choice probability in the sample if every student were female and if they were male. I then present the mean difference between the two probabilities as percentage points. This is intended not as a realistic counterfactual but as an illustration of the magnitude of the choice substitutions correlated with these personal characteristics variables.

3.6.1 Personal Characteristics

Table 3.13 below presents calculations of the relative change in choice probabilities in the population by personal characteristics. For example, the first row calculates the population-wide differences in choice probabilities for a world where all students and men and a world where all students were women. The section on ethnicity calculates differences relative to a world where all students are white.

Table 3.13: Difference in Choice Probabilities (in pp) by Characteristic

Subject	No university	STEM	Humanities and languages	Social sciences	Professional subjects	Arts
Female	-5.402	-8.176	-0.408	3.924	11.177	-1.115
	(1.366)	(1.127)	(1.056)	(0.908)	(2.102)	(1.280)
Higher socio-economic status	-4.881	5.585	1.923	-2.679	0.033	0.018
	(1.626)	(1.554)	(0.903)	(1.511)	(1.489)	(0.665)
Free school meals	-2.655	1.677	1.734	-1.812	-0.441	1.496
	(2.844)	(2.624)	(2.306)	(1.763)	(3.089)	(1.714)
IDACI Score	0.500	1.470	-0.053	-0.910	0.249	-1.256
	(1.278)	(1.142)	(0.838)	(0.893)	(1.366)	(0.609)
Ethnicity relative to white						
Asian	-25.592	10.197	-3.307	15.473	6.047	-2.819
	(2.228)	(4.199)	(1.900)	(3.706)	(5.993)	(0.732)
Black	-28.456	6.944	-2.102	9.406	13.836	0.373
	(2.747)	(4.449)	(2.686)	(3.643)	(5.116)	(2.677)
Mixed	-7.764	4.665	0.952	1.237	-2.941	3.851
	(4.502)	(4.704)	(3.106)	(2.894)	(4.394)	(3.599)
Other	-24.275	10.551	-7.749	-0.837	22.742	-0.432
	(4.268)	(4.980)	(1.226)	(2.644)	(5.913)	(2.567)

I find that women seem to prefer STEM subjects less than men (even controlling for differences in expected earnings) and also prefer to attend university at higher rates. They relatively prefer the social sciences and professional subjects. This coheres with other research on women's major choices, particularly the relatively lower rate at which they choose STEM majors.

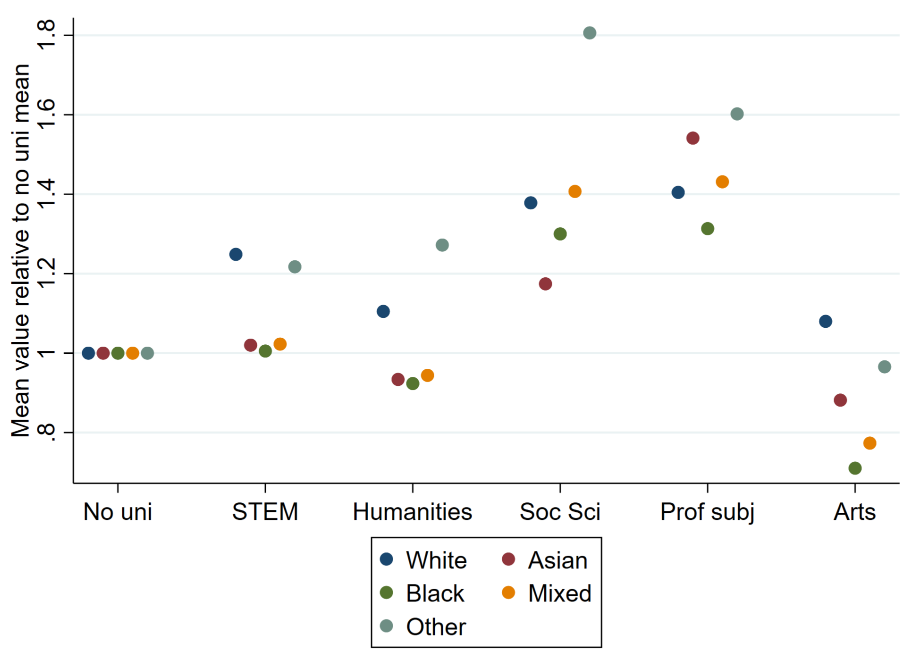
I find that being from a higher socio-economic class is largely associated with attending university at a higher rate and being more likely to choose STEM fields. Relative to socio-economic class, being eligible for free school meals and being from more deprived areas seems less important quantitatively for subject choice.

I find that in the UK context, ethnicity seems to correlate very strongly with subject choice. First, all ethnicities are far more likely to attend university than white students. Asian students are likely to choose social science and STEM subjects, while black students are more likely to choose professional subjects, and to a lesser extent, STEM. The size of these relative preferences are far larger than those correlating to sex or measures of class and deprivation. These ethnicity-based differences do not exist to the same magnitude in the US context, and if anything, black Americans are less likely to attend university than white

Americans (Patnaik et al. (2020b)).

What drives these ethnic differences? To decompose this result further, I first consider the relative earnings between subjects, split by ethnicity. Figure 3.2 plots the predicted relative unemployment-weighted lifetime earnings relative to no university by ethnicity. One striking finding is that ethnic minorities are not predicted to experience a graduate premium when studying STEM, the humanities and the arts. This would suggest that if they select into college at a similar rate to white students, mechanically, the model would load this preference on the alternative-specific coefficient. In other words, these students are not predicted to earn more in the STEM and the humanities, yet they are observed to select into it; thus, this is interpreted in my model as an unexplained preference for such subjects.

Figure 3.2: Relative predicted lifetime earnings by subject and ethnicity



Notes: The graph plots the mean predicted value for each subject relative to the mean predicted value for not going to university, for five wage specifications.

Table 3.14 plots the raw stated choice probabilities by ethnicity. I find that the relative preference of ethnic minorities for attending university is also reflected in raw subject choice differences. Asian and black students are far more likely to study for degrees in professional subjects than their white counterparts. Thus, the estimated subject choice preferences is a combination of raw differences in subject choice, and ethnic minorities experiencing a relatively smaller graduate premium in STEM, humanities and the arts.

Table 3.14: Raw choice probabilities by ethnicity

Population	(1)	(2)	(3)	(4)	(5)	(6)
All	37.26	16.71	8.73	9.84	23.94	3.51
White	40.15	16.44	9.36	8.83	21.44	3.79
Asian	16.39	18.54	(x)	17.95	42.76	(x)
Black	14.97	17.42	(x)	19.78	40.30	(x)
Mixed	34.45	14.83	9.29	13.08	23.84	4.51
Other	11.71	23.26	(x)	12.10	50.13	(x)

1 This table presents the raw subject choices broken down by ethnicity. Instances where the cell-count is smaller than 10 is withheld for statistical disclosure concerns. In cases where only one cell per row is withheld, the statistic for the next smallest cell is also withheld.
Source: Analysis on Next Steps data.

It is interesting and suggestive that much of the differences in raw stated subject choice conditional on attending university lies in minorities' stated preference for professional subjects. To my knowledge, it has not been explored why this might be the case. It might be interesting for future work to explore, for example, if this could be due to minorities being unable to access informal apprenticeship networks that are more dominated by native white students.

3.6.2 Relation between subject choice and KS4 performance

KS4 variables are the main grade variables that students have access to when making their subject choice decision. To represent the impact of the KS4 grade variables in affecting subject choice, I calculate the average choice probability if all students were average in that score, if all students were 1 standard deviation above average, and if all students were 1 standard deviation below average. I then calculate the differences of the +1 SD and -1 SD counterfactuals against the mean counterfactual[11].

Table 3.15 presents the results of the calculation described above. As found in the American context by Arcidiacono (2004), ability sorting seems to play a large role in subject choice, even beyond the effect built into the pecuniary motive. If every student had a normalised KS4 English grade of 1, selection into humanities would increase by 9.9 percentage points, relative to 2.3 pp for STEM. If every student had a normalised KS4 Mathematics grade of -1, selection into STEM would fall by 5.2 pp, relative to a fall of 1.1 for humanities. Thus, students seem to have preferences for subjects that they are relatively good at.

Selection on ability in individual subjects is only part of the story and the results from the estimation also suggest that subject choice varies by the overall academic ability of the student as measured by KS4 grades. Again, to represent the quantitative importance of the coefficients in the appendix, I calculate the changes in choice probability if all students had a normalised KS4, excluding English and maths, of 1 and -1, against the counterfactual scenario where everyone has a score of 0. I find that selection into STEM would increase when the average student quality increases and selection into professional subjects would fall. This coheres with the notion that more able students are more likely to select into more 'academic' subjects at university while less well-performing students select into courses which prepare them for jobs.

[11] This exercise is limited in a sense because students' KS4 grades in English and Maths are highly correlated. These numbers are intended to give a sense of the magnitude of the effects on choice probability of the highlighted KS4 variables.

Table 3.15: Difference in Choice Probabilities (in pp) by ability variables

Subject	-1 SD	+1 SD
KS4 Maths Score		
No university	8.217	-12.883
	(2.117)	(1.854)
STEM	-5.154	14.231
	(0.278)	(0.950)
Humanities and languages	-1.063	-0.543
	(0.881)	(1.021)
Social sciences	1.154	-2.861
	(1.505)	(1.249)
Professional subjects	-6.762	4.686
	(0.921)	(1.517)
Arts	3.607	-2.630
	(1.537)	(0.622)
KS4 English Score		
No university	15.346	-18.779
	(1.640)	(1.611)
STEM	-4.478	2.285
	(0.763)	(1.187)
Humanities and languages	-1.794	9.900
	(0.204)	(0.684)
Social sciences	-3.803	5.851
	(0.199)	(0.944)
Professional subjects	-3.717	-1.342
	(1.327)	(1.472)
Arts	-1.554	2.085
	(0.189)	(0.885)
Av KS4 Score excl English and Maths		
No university	0.674	-0.937
	(1.481)	(1.404)
STEM	-2.934	3.428
	(0.703)	(0.997)
Humanities and languages	0.495	-0.564
	(0.977)	(0.903)
Social sciences	-0.430	0.348
	(0.753)	(0.810)
Professional subjects	2.268	-2.321
	(1.563)	(1.409)
Arts	-0.074	0.047
	(0.474)	(0.491)

3.7 Conclusion

In this paper, I contribute to the literature by estimating a multinomial logit model to analyse the determinants of major choice in the UK context. I find that earnings, defined as earnings at age 25 or as lifetime earnings, are positively associated with choice probability, although the quantitative impact of expected earnings on subject choice is small, once heterogeneous weights on the pecuniary motive are accounted for. I find that in that case, the elasticity of subject choice to a change in lifetime earnings ranges from 0.28 to 0.93. This is driven largely by high ability students who seem to be more responsive to earnings in their higher education choices.

I find that there are substantial differences in choice probability between students of different sexes, ethnicities and to a lesser extent, socio-economic classes. While there has been a lot of work on explaining the gender gap in subject choice, from differences in risk attitudes, differences in degrees of competitiveness, and differences in preferences for various job attributes, there is not a lot of work explaining differences in subject choice between students of different ethnicities in part because this difference does not seem to be present in the American context. This may be an interesting avenue for future research, especially in the British context. A particular question that is of practical importance is whether minorities attend university at higher rates because they are substituting away from non-university further education options, like apprenticeships. This points to the need for research addressing substitution between different post-compulsory schooling options and how access to particular option may differ by student characteristics.

Finally, I find evidence that there are differences in estimated choice probabilities between models that use stated preference and observed choice as the relevant choice variable. I find that the elasticity of choice probabilities to changes in expected earnings is higher in models using stated preference than models using observed choice, and argue that this should be interpreted as evidence hinting at supply constraints in the subject choice decision. This points to the need to model the supply side when considering structural policy counterfactuals. For example, policies like differential tuition rates that are designed to increase the supply of graduates in STEM degrees may be ineffective because increased demand for STEM courses may not be met by sufficient increase in supply.

3.8 Appendix

A Additional Tables

Table 3.16: Quantile regression coefficients for log annual gross pay at age 25

	(1) No uni	(2) STEM	(3) Humanities	(4) Social science	(5) Prof subjects	(6) Arts
Normalised KS4 maths point score	0.073	0.165	0.077	0.145	0.185	0.162
	(0.021)	(0.045)	(0.086)	(0.074)	(0.054)	(0.088)
Normalised KS4 English point score	0.102	0.084	0.145	0.115	0.100	0.035
	(0.022)	(0.050)	(0.087)	(0.073)	(0.052)	(0.106)
Average GCSE point score (excluding English and Maths)	-0.003	0.015	0.042	0.028	0.022	0.036
	(0.017)	(0.043)	(0.112)	(0.082)	(0.049)	(0.111)
Female	-0.248	-0.164	-0.007	-0.172	-0.095	0.208
	(0.023)	(0.048)	(0.080)	(0.069)	(0.049)	(0.102)
Upper SEC	0.027	-0.037	-0.076	0.115	0.006	0.119
	(0.025)	(0.048)	(0.083)	(0.072)	(0.050)	(0.098)
Whether eligible for Free School Meal	-0.058	-0.172	0.054	0.123	-0.149	-0.274
	(0.042)	(0.096)	(0.189)	(0.147)	(0.113)	(0.243)
Asian	0.097	0.064	-0.083	-0.086	0.181	0.176
	(0.099)	(0.125)	(0.355)	(0.151)	(0.117)	(1.159)
Black	0.144	0.234	0.569	0.080	0.171	0.041
	(0.098)	(0.183)	(0.846)	(0.220)	(0.161)	(0.740)
Mixed	0.082	0.083	0.087	0.099	0.139	0.015
	(0.101)	(0.180)	(0.455)	(0.179)	(0.186)	(0.562)
Other	-0.001	0.310	0.079	0.050	0.150	0.174
	(0.099)	(0.101)	(1.220)	(0.560)	(0.107)	(0.432)
IDACI score	-0.155	-0.514	-0.187	-0.160	-0.348	-0.171
	(0.075)	(0.162)	(0.337)	(0.262)	(0.156)	(0.308)
Constant	10.007	0.021	-0.144	-0.119	-0.028	-0.532
	(0.024)	(0.064)	(0.121)	(0.106)	(0.067)	(0.113)
Observations	1294	587	214	305	578	126

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$

1 This table presents regression coefficients of a median regression of log annual earnings at age 25 on the variables listed. The regression was performed on the entire dataset by interacting the subject of study with the specified variables as follows. Robust standard errors are calculated.
Source: Analysis on Next Steps data.

Table 3.17: Relative Earnings Ratio to No Degree Under Different Specifications

Specification	No Uni	STEM	Hum and Lang	Soc Sci	Prof Subj	Arts
Predicted mean annual earnings at age 25	1	1.083	0.954	1.049	1.164	0.865
Predicted median annual earnings at age 25	1	1.037	0.994	1.039	1.106	0.810
Predicted mean employment weighted earnings at age 25	1	1.082	0.956	1.052	1.173	0.873
Predicted median employment weighted earnings at age 25	1	1.036	0.995	1.041	1.115	0.818
Predicted mean lifetime employment-weighted earnings, r = 0.95	1	1.220	1.096	1.356	1.399	1.016
Predicted mean lifetime employment-weighted earnings, r = 0.97	1	1.221	1.096	1.373	1.418	1.014
Predicted median lifetime employment-weighted earnings, r = 0.95	1	1.170	1.139	1.353	1.328	0.949
Predicted median lifetime employment-weighted earnings, r = 0.97	1	1.171	1.139	1.372	1.346	0.947

1 This table summarises the raw data used to plot figure 3.1. It calculates the ratio of earnings relative to not attending university for the other five subjects under different specifications of the earnings variable
Source: Analysis on Next Steps and Labour Force Survey data.

Table 3.18: Coefficients on wage variable for 4 specifications of choice model

	(1)	(2)	(3)	(4)
Mean, emp-weighted, lifetime earnings	0.00000318**		0.00000301**	
	(3.00)		(2.59)	
Median, emp-weighted, lifetime earnings		0.00000482***		
		(4.02)		
Mean, emp-weighted, age 25 earnings				0.000143*
				(2.10)
Observations	29034	29034	29034	29034
ll	-5877.5	-5873.9	-5215.7	-5879.7
choicevar	Stated choice	Stated choice	Observed choice	Stated choice

1 This table presents coefficients on the individual and subject-specific expected earnings variable in models 1-4 as specified in table 3.2.
Source: Analysis on Next Steps data.

Table 3.19: Average Substitution Patterns Between Subjects (Specification 1)

Subject	No university	STEM	Humanities and languages	Social sciences	Professional subjects	Arts
No university	0.890	-0.523	-0.421	-0.488	-0.566	-0.686
	(0.296)	(0.175)	(0.143)	(0.164)	(0.188)	(0.237)
STEM	-0.288	1.848	-0.508	-0.417	-0.444	-0.282
	(0.097)	(0.620)	(0.172)	(0.144)	(0.150)	(0.098)
Humanities and languages	-0.111	-0.243	1.758	-0.237	-0.177	-0.167
	(0.038)	(0.083)	(0.593)	(0.081)	(0.060)	(0.059)
Social sciences	-0.174	-0.299	-0.345	2.227	-0.278	-0.222
	(0.059)	(0.104)	(0.119)	(0.756)	(0.094)	(0.077)
Professional subjects	-0.505	-0.777	-0.620	-0.693	1.917	-0.522
	(0.169)	(0.264)	(0.210)	(0.233)	(0.642)	(0.180)
Arts	-0.071	-0.050	-0.062	-0.061	-0.061	-0.061
	(0.025)	(0.018)	(0.022)	(0.021)	(0.021)	(0.597)

1 This table summarises the average implied substitution patterns across subjects given the estimation results for model 1. The calculation of the elasticities is described in full in the full text.
Source: Analysis on Next Steps data.

Conclusion

This thesis comprises three essays, each of which contributes to the intersection between the economics of higher education and labour economics.

My first essay analyses the alleged phenomenon of 'overeducation', characterised by the prevalence of graduates who work in so-called low-skill occupations. I interpret the phenomenon through the lens of a matching model with an endogenous education decision with uncertainty and heterogeneity to suggest two senses in which overeducation may occur - (i) in the sense that workers may make ex-post unprofitable decisions despite the choice being ex-ante rational and (ii) in the sense aggregate welfare for workers would increase to a point if future workers invested in education. I estimate my model on British data and conduct policy counterfactuals. I emphasise two main findings relating to policy. First, a planner aiming to adjust the education profile by implementing a graduate tax or subsidy has to trade off the interests of workers and firms. Second, reducing worker uncertainty about their return to education is unambiguously good for them but may make firms worse off by increasing wages.

My second essay argues that it is highly likely that the pay premium for workers in specific occupations should vary between firms and this variation should be accounted for in the literature on the sources of income heterogeneity. We propose a flexible framework where workers' income is log-additive in an individual-specific fixed effect, a job-specific fixed effect and other controls. We use this model to (i) characterise the share of income inequality due to heterogeneity between occupations, (ii) characterise how the extent of heterogeneity in firm pay premia may vary between occupations, and (iii) show that the size of the geographical unit is relevant for the extent of firm-related income heterogeneity but not occupation-related income heterogeneity. Technically, our model can be seen as an analysis of Abowd et al. (1999) type models where the individual fixed effect can vary across the workers' working life.

My final essay attempts to estimate the responsiveness of students' higher education subject choices to changes in the wage structure. I conduct a revealed preference analysis using variation in expected earnings with respect to plausibly exogenous personal characteristics. I largely confirm the primary finding in the literature that the elasticities are quantitatively small. I also document that there is substantial heterogeneity in subject choice by sex, ethnicity and to a smaller extent socio-economic status.

The questions left unanswered in this thesis point to intriguing possible themes and avenues for future work. First, there is a relative paucity of work that considers the optimality of higher education policy in general equilibrium. There is also relatively little work incorporating selection and uncertainty into popular macroeconomic models of the relation between education and the wage structure, although this is an area of active research. Second, although there is a burgeoning literature, based on survey and hypothetical choice methods, on the analysis of expectations in education choice, there is still relatively little work on how expectations can be affected by policy. This constitutes a missing link between the microeconomic focus on expectations in choice analysis and the macroeconomic focus on the role of shifting wage premia in affecting the education profile. Third, given recent theoretical developments such as the task framework and the increasing availability of matched employer-employee data, there is renewed interest in the role of occupations in wage determination and how matching to occupations occurs. Although this has been a long-standing research question, we still lack tractable ways of characterising the relations between worker training and the occupations they work in. The development of such a tractable characterisation remains a priority in the intersection between the economics of education and labour economics.

Bibliography

Abowd, J. M., Creecy, R. H., & Kramarz, F. (2002). Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data. Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau.

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–333.

Abramitzky, R., Lavy, V., & Segev, M. (2019). The effect of changes in the skill premium on college degree attainment and the choice of major. Tech. rep., National Bureau of Economic Research.

Acemoglu, D. (1999). Changes in unemployment and wage inequality: An alternative theory and some evidence. *American economic review*, *89*(5), 1259–1278.

Acemoglu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. In *Handbook of Labor Economics*, vol. 4, (pp. 1043–1171). Elsevier.
URL https://linkinghub.elsevier.com/retrieve/pii/S0169721811024105

Adda, J., & Cooper, R. W. (2003). *Dynamic economics: quantitative methods and applications*. MIT press.

Aghion, P., Bergeaud, A., Blundell, R. W., & Griffith, R. (2019). The innovation premium to soft skills in low-skilled occupations. *Available at SSRN 3489777*.

Agrawal, S., & Phillips, D. (2020). Catching up or falling behind? geographical inequalities in the uk and how they have changed in recent years. *The Institute for Fiscal Studies*, (p. p2).

Albrecht, J., & Vroman, S. (2002). A matching model with endogenous skill requirements. *International Economic Review*, *43*(1), 283–305. Publisher: Wiley Online Library.

Altonji, J. (1993). The Demand for and Return to Education When Education Outcomes Are Uncertain. *Journal of Labor Economics*, *11*(1), 48–83.
URL https://EconPapers.repec.org/RePEc:ucp:jlabec:v:11:y:1993:i:1:p:48-83

Altonji, J., Arcidiacono, P., & Maurel, A. (2015). The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects. Tech. rep., National Bureau of Economic Research, Inc. Issue: 21655.
URL https://EconPapers.repec.org/RePEc:nbr:nberwo:21655

Altonji, J., Blom, E., & Meghir, C. (2012). Heterogeneity in Human Capital Investments: High School Curriculum, College Major, and Careers. Tech. rep., National Bureau of Economic Research, Inc. Issue: 17985.
URL https://EconPapers.repec.org/RePEc:nbr:nberwo:17985

Alvarez, J., Benguria, F., Engbom, N., & Moser, C. (2018). Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics*, *10*(1), 149–89.

Andrews, M. J., Gill, L., Schank, T., & Upward, R. (2012). High wage workers match with high wage firms: Clear evidence of the effects of limited mobility bias. *Economics Letters*, *117*(3), 824–827.

Andrews, R. J., Imberman, S. A., Lovenheim, M. F., & Stange, K. M. (2022). The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability. Working Paper 30331, National Bureau of Economic Research. Series: Working Paper Series.
URL http://www.nber.org/papers/w30331

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, *121*(1-2), 343–375.
URL https://EconPapers.repec.org/RePEc:eee:econom:v:121:y:2004:i:1-2:p:343-375

Arcidiacono, P., Aucejo, E., Maurel, A., & Ransom, T. (2016). College Attrition and the Dynamics of Information Revelation. Tech. rep., National Bureau of Economic Research, Inc. Issue: 22325.
URL https://EconPapers.repec.org/RePEc:nbr:nberwo:22325

Arcidiacono, P., Hotz, V. J., & Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, *166*(1), 3 – 16.
URL http://www.sciencedirect.com/science/article/pii/S0304407611001151

Athey, S. (2002). Monotone Comparative Statics under Uncertainty. *The Quarterly Journal of Economics*, *117*(1), 187–223.
URL https://academic.oup.com/qje/article-lookup/doi/10.1162/003355302753399481

Athey, S., & Wager, S. (2021). Policy learning with observational data. *Econometrica*, *89*(1), 133–161. Publisher: Wiley Online Library.

Avery, C., Gurantz, O., Hurwitz, M., & Smith, J. (2018). Shifting college majors in response to advanced placement exam scores. *Journal of Human Resources*, *53*(4), 918–956. Publisher: University of Wisconsin Press.

Bagnoli, M., & Bergstrom, T. (2005). Log-concave probability and its applications. *Economic Theory*, *26*(2), 445–469.
URL http://link.springer.com/10.1007/s00199-004-0514-4

Barnichon, R., & Zylberberg, Y. (2019). Underemployment and the trickle-down of unemployment. *American Economic Journal: Macroeconomics*, *11*(2), 40–78. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2425.

Baum-Snow, N., & Pavan, R. (2013). Inequality and city size. *Review of Economics and Statistics*, *95*(5), 1535–1548.

Becker, G. S. (1973). A theory of marriage: Part I. *Journal of Political economy*, *81*(4), 813–846. Publisher: The University of Chicago Press.

Beffy, M., Fougere, D., & Maurel, A. (2012). Choosing the Field of Study in Postsecondary Education: Do Expected Earnings Matter? *The Review of Economics and Statistics*, *94*(1), 334–347.

URL https://EconPapers.repec.org/RePEc:tpr:restat:v:94:y:2012:i:1:p:334-347

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, *122*(3), 507–553.

Belfield, C., Britton, J., Buscha, F., Dearden, L., Dickson, M., van der Erve, L., Sibieta, L., Vignoles, A., Walker, I., & Zhu, Y. (2018a). The impact of undergraduate degrees on early-career earnings. Tech. rep., Department for Education.
URL https://www.ifs.org.uk/publications/13731

Belfield, C., Crawford, C., & Sibieta, L. (2018b). Long-run comparisons of spending per pupil across different stages of education. Publisher: Institute for Fiscal Studies.

Berger, M. C. (1988). Predicted Future Earnings and Choice of College Major. *ILR Review*, *41*(3), 418–429.
URL https://doi.org/10.1177/001979398804100306

Bilal, A. (2021). The geography of unemployment. Tech. rep., National Bureau of Economic Research.

Bloesch, J., Larsen, B., & Taska, B. (2021). Which workers earn more at productive firms? position specific skills and individual worker hold-up power.

Blundell, R., Costa Dias, M., Goll, D., & Meghir, C. (2019). Wages, Experience and Training of Women over the Lifecycle. Tech. rep., Cowles Foundation for Research in Economics, Yale University. Issue: 2174R.
URL https://EconPapers.repec.org/RePEc:cwl:cwldpp:2174r

Blundell, R., Costa Dias, M., Meghir, C., & Shaw, J. (2016). Female Labor Supply, Human Capital, and Welfare Reform. *Econometrica*, *84*(5), 1705–1753.
URL https://doi.org/10.3982/ECTA11576

Blundell, R., Green, D. A., & Jin, W. (2022). The UK as a Technological Follower: Higher Education Expansion and the College Wage Premium. *The Review of Economic Studies*, *89*(1), 142–180. Publisher: Oxford University Press.

Blundell, R., et al (2000). The Returns to Higher Education in Britain: Evidence from a British Cohort. *Economic Journal*, *110*(461), F82–99.
URL https://EconPapers.repec.org/RePEc:ecj:econjl:v:110:y:2000:i:461:p:f82-99

Bolton, P. (2019). Student loan statistics. Publisher: UK Parliament.

Boneva, T., & Rauh, C. (2020). Socio-Economic Gaps in University Enrollment: The Role of Perceived Pecuniary and Non-Pecuniary Returns. Tech. rep., CESifo. Issue: 6756.
URL https://drive.google.com/file/d/1rFGjj43JEDsDM57vyc11QLwsKIGi8y8P/view

Bonhomme, S., Holzheu, K., Lamadon, T., Manresa, E., Mogstad, M., & Setzler, B. (2020). How much should we trust estimates of firm effects and worker sorting? Tech. rep., National Bureau of Economic Research.

Bonhomme, S., Lamadon, T., & Manresa, E. (2019). A distributional framework for matched employer employee data. *Econometrica*, *87*(3), 699–739.

Britton, J., van der Erve, L., & Higgins, T. (2019). Income contingent student loan design: Lessons from around the world. *Economics of Education Review*, *71*, 65–82.
URL https://linkinghub.elsevier.com/retrieve/pii/S0272775717301668

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices*. *The Quarterly Journal of Economics*, *129*(3), 1409–1447.
URL https://academic.oup.com/qje/article/129/3/1409/1817134

Card, D. (1999). The causal effect of education on earnings. *Handbook of labor economics*, *3*, 1801–1863. Publisher: Elsevier.

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018a). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, *36*(S1), S13–S70. Publisher: University of Chicago Press Chicago, IL.

Card, D., Cardoso, A. R., Heining, J., & Kline, P. (2018b). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics*, *36*(S1), S13–S70.

Card, D., Cardoso, A. R., & Kline, P. (2016). Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women. *The Quarterly Journal of Economics*, *131*(2), 633–686.

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of west german wage inequality. *The Quarterly journal of economics*, *128*(3), 967–1015.

Carneiro, P., & Lee, S. (2011). Trends in Quality-Adjusted Skill Premia in the United States, 1960—2000. *The American Economic Review*, *101*(6), 2309–2349.
URL www.jstor.org/stable/23045644

Chade, H., Eeckhout, J., & Smith, L. (2017). Sorting through Search and Matching Models in Economics. *Journal of Economic Literature*, *55*(2), 493–544.
URL https://pubs.aeaweb.org/doi/10.1257/jel.20150777

Chade, H., & Lindenlaub, I. (2021). Risky Matching.

Cunha, F., & Heckman, J. (2016). Decomposing Trends in Inequality in Earnings into Forecastable and Uncertain Components. *Journal of Labor Economics*, *34*(S2), S31–S65.
URL https://www.journals.uchicago.edu/doi/10.1086/684121

Cunha, F., Heckman, J., & Navarro, S. (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers*, *57*(2), 191–261.
URL https://EconPapers.repec.org/RePEc:oup:oxecpp:v:57:y:2005:i:2:p:191-261

Dauth, W., Findeisen, S., Moretti, E., & Suedekum, J. (2018). Matching in cities. Tech. rep., National Bureau of Economic Research.

Davis, D. R., & Dingel, J. I. (2019). A spatial knowledge economy. *American Economic Review*, *109*(1), 153–70.

D'Costa, S., & Overman, H. G. (2014). The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics*, *48*, 168–179.

Delavande, A., & Zafar, B. (2019). University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints. *Journal of Political Economy*, *127*(5), 2343–2393. Publisher: The University of Chicago Press Chicago, IL.

Deming, D. J. (2023). Multidimensional Human Capital and the Wage Structure. Publisher: National Bureau of Economic Research.

Devereux, P. J., & Fan, W. (2011). Earnings returns to the British education expansion. *Economics of Education Review*, *30*(6), 1153–1166.
URL https://linkinghub.elsevier.com/retrieve/pii/S0272775711000471

Dolado, J. J., Jansen, M., & Jimeno, J. F. (2009). On-the-job search in a matching model with heterogeneous jobs and workers. *The Economic Journal*, *119*(534), 200–228. Publisher: Oxford University Press Oxford, UK.

Dominitz, J., & Manski, C. (1994). Eliciting Student Expectations of the Returns to Schooling. Tech. rep., National Bureau of Economic Research, Inc. Issue: 4936.
URL https://EconPapers.repec.org/RePEc:nbr:nberwo:4936

Education, D. F., & University Of Essex, I. F. S. (2022). United Kingdom Household Longitudinal Study; UKHLS; NPDUnderstanding Society: Waves 1- , 2008-: Secure AccessUnderstanding Society: Linked Education Administrative Datasets (National Pupil Database), England, 1995-2018: Secure Access.
URL https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=7642#3

Ekeland, I., Heckman, J. J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of political economy*, *112*(S1), S60–S109. Publisher: The University of Chicago Press.

Elias, P., & Purcell, K. (2004). Is Mass Higher Education Working? Evidence from the Labour Market Experiences of Recent Graduates. *National Institute Economic Review*, *190*, 60–74. Publisher: Cambridge University Press.

Faggio, G., Salvanes, K. G., & Van Reenen, J. (2010). The evolution of inequality in productivity and wages: panel data evidence. *Industrial and Corporate Change*, *19*(6), 1919–1951.

Galichon, A. (2016). *Optimal transport methods in economics*. Princeton University Press.

Gaubert, C. (2018). Firm sorting and agglomeration. *American Economic Review*, *108*(11), 3117–53.

Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of labor economics*, *19*(2), 316–342.

Gottschalk, P., & Hansen, M. (2003). Is the proportion of college workers in noncollege jobs increasing? *Journal of Labor Economics*, *21*(2), 449–471. Publisher: The University of Chicago Press.

Gourieroux, C., Gourieroux, M., Monfort, A., & Monfort, D. A. (1996). *Simulation-based econometric methods*. Oxford university press.

Green, F., & Zhu, Y. (2010). Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education. *Oxford Economic Papers*, *62*(4), 740–763. URL https://academic.oup.com/oep/article-lookup/doi/10.1093/oep/gpq002

Greene, W. H. (2003). *Econometric analysis*. Pearson Education India.

Guvenen, F., Kuruscu, B., Tanaka, S., & Wiczer, D. (2020). Multidimensional skill mismatch. *American Economic Journal: Macroeconomics*, *12*(1), 210–44.

Haanwinckel, D. (2020). Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution. *Unpublished manuscript*.

Hanushek, E. A., & Woessmann, L. (2010). Education and economic growth. *Economics of education*, *60*, 67. Publisher: Elsevier San Diego, CA.

Heckman, J., Lochner, L., & Todd, P. (2005). Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. Tech. rep., Institute of Labor Economics (IZA). Issue: 1700.
URL https://EconPapers.repec.org/RePEc:iza:izadps:dp1700

Heckman, J. J., Humphries, J. E., & Veramendi, G. (2018). Returns to Education: The Causal Effects of Education on Earnings, Health, and Smoking. *Journal of Political Economy*, *126*(S1), S197–S246.
URL https://www.journals.uchicago.edu/doi/10.1086/698760

Heckman, J. J., & Sattinger, M. (2015). Introduction to the distribution of earnings and of individual output, by AD Roy. *The Economic Journal*, *125*(583), 378–402. Publisher: Oxford University Press Oxford, UK.

Heckman, J. J., & Vytlacil, E. (2001). Policy-Relevant Treatment Effects. *The American Economic Review*, *91*(2), 107–111. Publisher: American Economic Association.
URL http://www.jstor.org/stable/2677742

Hou, S., & Milsom, L. (2021). The butcher, the brewer, or the baker: The role of occupations in explaining wage inequality. Publisher: University of Oxford.

Jackson, P. (2021). Equilibrium Underemployment.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, *125*(2), 515–548. Publisher: MIT Press.

Jochmans, K., & Weidner, M. (2019). Fixed-effect regressions on network data. *Econometrica*, *87*(5), 1543–1560.

Katz, L. F., & Goldin, C. (2008). *The Race Between Education and Technology*. Belknap Press for Harvard University Press.

Katz, L. F., & Murphy, K. M. (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics*, *107*(1), 35–78.
URL https://academic.oup.com/qje/article-lookup/doi/10.2307/2118323

Kirkeboen, L. J., Leuven, E., & Mogstad, M. (2016). Field of study, earnings, and self-selection. *The Quarterly Journal of Economics*, *131*(3), 1057–1111. Publisher: Oxford University Press.

Kline, P., Saggio, R., & Sølvsten, M. (2020). Leave-out estimation of variance components. *Econometrica*, *88*(5), 1859–1898.

Lange, F. (2007). The speed of employer learning. *Journal of Labor Economics*, *25*(1), 1–35. Publisher: The University of Chicago Press.

Lee, N., Sissons, P., & Jones, K. (2016). The geography of wage inequality in british cities. *Regional Studies*, *50*(10), 1714–1727.

Lemieux, T. (2006). Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *American Economic Review*, *96*(3), 461–498.
URL https://pubs.aeaweb.org/doi/10.1257/aer.96.3.461

Lemieux, T. (2014). Occupations, fields of study and returns to education. *Canadian Journal of Economics/Revue canadienne d'économique*, *47*(4), 1047–1077. Publisher: Wiley Online Library.

Li, H., & Xia, X. (2022). Grades as Signals of Comparative Advantage: How Letter Grades Affect Major Choices. *SSRN Electronic Journal*.
URL https://www.ssrn.com/abstract=4028958

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies*, *84*(2), 718–789.

Lise, J., & Postel-Vinay, F. (2020a). Multidimensional Skills, Sorting, and Human Capital Accumulation. *American Economic Review*, *110*(8), 2328–2376.
URL https://pubs.aeaweb.org/doi/10.1257/aer.20162002

Lise, J., & Postel-Vinay, F. (2020b). Multidimensional skills, sorting, and human capital accumulation. *American Economic Review*, *110*(8), 2328–76.

Long, M. C., Goldhaber, D., & Huntington-Klein, N. (2015). Do completed college majors respond to changes in wages? *Economics of Education Review*, *49*, 1 – 14.
URL http://www.sciencedirect.com/science/article/pii/S027277571500103X

Machin, S. (2011). Changes in uk wage inequality over the last forty years. *The Labour Market in Winter-The State of Working Britain*.

Manski, C. (1993). Adolescent Econometricians: How Do Youth Infer the Returns to Schooling? In *Studies of Supply and Demand in Higher Education*, (pp. 43–60). National Bureau of Economic Research, Inc.
URL https://EconPapers.repec.org/RePEc:nbr:nberch:6097

McFadden, D. (1974). The measurement of urban travel demand. *Journal of Public Economics*, *3*(4), 303–328.
URL https://linkinghub.elsevier.com/retrieve/pii/0047272774900036

Montmarquette, C., Cannings, K., & Mahseredjian, S. (2002). How do young people choose college majors? *Economics of Education Review*, *21*(6), 543 – 556.
URL http://www.sciencedirect.com/science/article/pii/S0272775701000541

Mortensen, D. (2003). *Wage dispersion: why are similar workers paid differently?*. MIT press.

Mueller, H. M., Ouimet, P. P., & Simintzi, E. (2017). Wage inequality and firm growth. *American Economic Review*, *107*(5), 379–83.

Nord, S. (1980). Income inequality and city size: An examination of alternative hypotheses for large and small cities. *The Review of Economics and Statistics*, (pp. 502–508).

O'Leary, N., & Sloane, P. (2016). Too many graduates? An application of the Gottschalk–Hansen model to young British graduates between 2001–2010. *Oxford Economic Papers*, *68*(4), 945–967.
URL https://academic.oup.com/oep/article-lookup/doi/10.1093/oep/gpw027

Patnaik, A., Venator, J., Wiswall, M., & Zafar, B. (2020a). The Role of Heterogeneous Risk Preferences, Discount Rates, and Earnings Expectations in College Major Choice. Tech. rep., National Bureau of Economic Research, Inc. Issue: 26785.
URL https://EconPapers.repec.org/RePEc:nbr:nberwo:26785

Patnaik, A., Wiswall, M., & Zafar, B. (2020b). College Majors. Tech. Rep. w27645, National Bureau of Economic Research, Cambridge, MA.
URL http://www.nber.org/papers/w27645.pdf

Petrongolo, B., & Pissarides, C. (2006). Scale effects in markets with search. *The Economic Journal*, *116*(508), 21–44.

Postel-Vinay, F., & Robin, J.-M. (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, *70*(6), 2295–2350.

Rice, P., Venables, A. J., & Patacchini, E. (2006). Spatial determinants of productivity: analysis for the regions of great britain. *Regional science and urban economics*, *36*(6), 727–752.

Roy, A. D. (1951). SOME THOUGHTS ON THE DISTRIBUTION OF EARNINGS 1. *Oxford Economic Papers*, *3*(2), 135–146. _eprint: https://academic.oup.com/oep/article-pdf/3/2/135/6966554/3-2-135.pdf.
URL https://doi.org/10.1093/oxfordjournals.oep.a041827

Sattinger, M. (1993). Assignment models of the distribution of earnings. *Journal of economic literature*, *31*(2), 831–880. Publisher: JSTOR.

Schaefer, D., & Singleton, C. (2020). Recent changes in british wage inequality: Evidence from large firms and occupations. *Scottish Journal of Political Economy*, *67*(1), 100–125.

Shapley, L. S., & Shubik, M. (1971). The assignment game I: The core. *International Journal of game theory*, *1*(1), 111–130. Publisher: Springer.

Shephard, A., & Sidibe, M. (2019). Schooling Investment, Mismatch, and Wage Inequality. *SSRN Electronic Journal*.
URL https://www.ssrn.com/abstract=3422854

Shimer, R., & Smith, L. (2000). Assortative matching and search. *Econometrica*, *68*(2), 343–369. Publisher: Wiley Online Library.

Snedecor, G. W., & Cochran, W. G. (1989). Statistical methods, eight edition. *Iowa state University press, Ames, Iowa*, *1191*.

Song, J., Price, D. J., Guvenen, F., Bloom, N., & Von Wachter, T. (2019). Firming up inequality. *The Quarterly journal of economics*, *134*(1), 1–50.

Statistics, O. F. N. (2020). ASHE; New Earnings Survey; NESAnnual Survey of Hours and Earnings, 1997-2020: Secure Access.
URL https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6689#15

Statistics, O. F. N. (2022). BSDBusiness Structure Database, 1997-2021: Secure Access.
URL https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6697#14

Tan, B. J. (2022). The Consequences of Letter Grades on Labor Market Outcomes and Student Behavior. *Journal of Labor Economics*, (p. 719994).
URL https://www.journals.uchicago.edu/doi/10.1086/719994

Topkis, D. M. (1998). *Supermodularity and Complementarity*. Princeton University Press.

Torres, S., Portugal, P., Addison, J. T., & Guimaraes, P. (2018). The sources of wage variation and the direction of assortative matching: Evidence from a three-way high-dimensional fixed effects regression model. *Labour Economics*, *54*, 47–60.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

University Of Essex, I. F. S. (2022). United Kingdom Household Longitudinal Study; UKHLSUnderstanding Society: Waves 1- , 2008-: Secure AccessUnderstanding Society:

Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009: Secure Access.
URL https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=6676#12

Venables, A. J. (2011). Productivity in cities: self-selection and sorting. *Journal of Economic Geography*, *11*(2), 241–251.

Walker, I., & Zhu, Y. (2008). The College Wage Premium and the Expansion of Higher Education in the UK*. *Scandinavian Journal of Economics*, *110*(4), 695–709.
URL http://doi.wiley.com/10.1111/j.1467-9442.2008.00557.x

Waltmann, B., van der Erve, L., Dearden, L., & Britton, J. (2020). The impact of undergraduate degrees on lifetime earnings. Tech. rep., The IFS.
URL https://www.ifs.org.uk/publications/14729

Willis, R., & Rosen, S. (1979). Education and Self-Selection. *Journal of Political Economy*, *87*(5), S7–36.
URL https://EconPapers.repec.org/RePEc:ucp:jpolec:v:87:y:1979:i:5:p:s7-36

Wiswall, M., & Zafar, B. (2014). Determinants of College Major Choice: Identification using an Information Experiment. *The Review of Economic Studies*, *82*(2), 791–824.
URL https://doi.org/10.1093/restud/rdu044

Wiswall, M., & Zafar, B. (2017). Preference for the Workplace, Investment in Human Capital, and Gender*. *The Quarterly Journal of Economics*, *133*(1), 457–507.
URL https://doi.org/10.1093/qje/qjx035

Zafar, B. (2013). College Major Choice and the Gender Gap. *Journal of Human Resources*, *48*(3).
URL https://EconPapers.repec.org/RePEc:uwp:jhriss:v:48:y:2013:iii:1:p:545-595