Skills, Tasks and Degrees

Max Schroeder*

October 30, 2024

latest version

Abstract

Increasingly, young university graduates do not join traditional graduate occupations after completing their studies. This phenomenon has led to concerns about a possible decline in recent graduate's skills relative to earlier cohorts. However, these explanations often overlook the simultaneous effects of the evolving nature of occupations, both with respect to the return to skill and other job amenities. To address this gap, I propose an economic model of the labour market for young graduates, featuring heterogeneous skill supply and differentiated return to skill across occupations, while also accounting for changing preferences over non-wage job characteristics. Using data on young graduates in the UK from 2001-2019, I estimate the model to quantify changes in the graduate skill distribution, returns to skill, and preferences over time. The estimation reveals a significant decline in the average skill level of graduates, around 21% of a standard deviation, explaining approximately 46% of the reduction in professional employment. Concurrently, changing amenity values are driving the increase in routine and service occupations, particularly due to a reduction in the value of not working, which plays a critical role in pushing graduates into these non-traditional occupations. My findings suggest that addressing graduate underemployment requires a holistic approach that considers skill supply, labour demand, and graduates' preferences.

Keywords: skills, tasks, tertiary education, occupation choice

JEL Classification: I23, I24, I26, J24

*Durham University Business School, Email: max.j.schroeder@durham.ac.uk

I would like to thank Konstantinos Angelopoulos, Richard Foltyn, Spyridon Lazarakis, Margaret Leighton, and Jim Malley for comments on earlier versions of this paper. I further thank the seminar attendants at the Annual meetings of the Scottish Economic Society (2021), the Scottish Graduate Programme in Economics (2021), the Society for Computational Economics (2021), the Workshop on Applied Economics of Education (2022), the Annual Conference of The European Association of Labour Economists (2022, 2024), the Second Annual Scotland and Northern England Conference in Applied Microeconomics (2023), the Royal Economic Society Annual Conference (2023), Lancaster University (2024), the IAAEU Workshop on Labour Economics (2024), the Society of Labor Economists (2024), and the EEA-ESEM Congress (2024) for their insights. All remaining errors are my own.

1 Introduction

Over the last 30 years, the UK (and other developed economies) has experienced a rapid expansion of tertiary education participation. Since the passage of the Further and Higher Education Act 1992, university enrolment has roughly doubled to approximately 2.5 million in 2019/20, a trend that was sustained in the face of stark tuition fee increases.[1] This has had a strong impact on the composition of the workforce in the UK, particularly among those who are just entering the labour market: Between 2001 and 2019, the share of young adults, aged 21 to 30 with a university degree approximately doubled from around 20% to almost 40% (see Figure A1).

At the same time, increasing numbers of degree holders fail to secure employment in a job where a degree is a required qualification (c.f. Hou (2023)). Such instances of underemployment, together with rising graduate earnings inequality, cast some doubt on whether Higher Education Institutions (HEIs) are delivering on their promises of "graduate skills" and "graduate jobs" for their alumni (c.f. Lindley & MacIntosh (2015), Altonji et al. (2016)). In the UK and elsewhere, these developments have meant that the value of a university degree is coming under increased public scrutiny, with some questioning the return on this expensive investment made by so many young individuals.

However, the changing supply of skills is only one part of the ever-developing labour market. Equally important are the shifts in demand for these skills, driven by technological advancements and macroeconomic changes, as well as evolving preferences among graduates regarding job characteristics. Over the past decades, there have been significant changes in the demand for the skills that young graduates possess, driven by changes in technology and the wider macroeconomic environment (see Acemoglu & Autor (2011) for a survey). Furthermore, current generations likely have different preferences regarding non-pecuniary job characteristics, which evolve over time and influence occupational decisions. The observed patterns of labour market outcomes for young graduates are the result of the interactions between these forces, making mono-causal explanations unlikely.

The question, therefore, is to what degree are the observed patterns in the labour market outcomes of young graduates due to the changing distribution of graduate skills, the evolving patterns of returns to these skills, or changing preferences over non-wage attributes of different occupations? Understanding these different factors is crucial for developing effective policy interventions, as they help identify the key drivers behind graduate underemployment and wage inequality, allowing for more targeted policy interventions.

A key challenge when it comes to answering this question is that labour market skills are generally unobservable for the econometrician. To address this issue, I take another

[1]The cap on the amount that universities can charge was increased nearly threefold in England in 2012, leading to a large increase in tuition fees, with most institutions charging the maximum amount.

approach—framing the question as a latent variable problem: skills are unobserved but related to observable choices and labour market outcomes. I develop a model of occupational choice for university graduates and use it to quantify the importance of the demand for and supply of graduates' skills for the labour market outcomes of young graduates. By specifying and estimating a corresponding structural economic model, one can make inferences about the unobserved skill endowments of university graduates, as well as the returns to these skills in different occupations. A structural economic model is particularly suitable for this type of analysis because it allows us to account for the underlying relationships between skills, occupational choices, and wages, providing a more complete picture of how these factors interact within the labour market.

I structurally estimate the model using a sample of recent UK university graduates from 2001-2019 and recover the parameters of the underlying latent skill distributions for different cohorts of university graduates, as well as the changing returns to skills across different occupations. I then use the model estimates to analyse changes in the graduate skill distribution, the changing returns to these skills, and their combined effects on labour market sorting among university graduates over the first two decades of the 21st century.

I find that between 2001-2009 and 2010-2019, the mean of the distribution of graduates' skills has decreased by about 21% of a standard deviation. This has a direct effect on the share of young graduates entering professional occupations, explaining around 46% of the decline observed in the data. Sorting along the skill distribution means that most graduates who do not enter professional occupations come from the bottom of the skill distribution and are more likely not to enter the labour market as a result. Above the median, graduates also face a lower probability of being employed in professional occupations but increasingly choose employment in routine or service occupations as an alternative.

This paper adds to the literature on graduate underemployment with a specific focus on the UK (see Green & Zhu (2010), Holmes & Mayhew (2016), O'Leary & Sloane (2016)). These studies find that graduates in the UK are increasingly likely to be employed in roles that were not traditionally considered graduate occupations. I complement their findings by providing a flexible occupational choice framework that can accommodate three potential drivers of these trends: i) changes in the distribution of graduates' skills as a result of higher education expansion; ii) changes in the return to graduates' skills in different occupations due to technological change; iii) changing preferences over different occupations. My main findings suggest that while changes in the distribution of skills play a major role in explaining a reduction in the share of graduates entering professional occupations, they are not the main driver of the reallocation of these graduates towards routine and service occupations. Differences in preferences for non-pecuniary aspects of occupations, on the other hand, contribute just under half of the decline in professional occupations and can explain a substantial amount of the reallocation towards less skill-intensive occupations.

This paper also contributes to a large literature on the returns to higher education with an emphasis on the heterogeneity of returns (see Altonji et al. (2016), Leighton & Speer (2020), Andrews et al. (2022) and Lovenheim & Smith (2022) for extensive surveys). Generally, these studies aim to estimate the returns to post-secondary education while trying to address the inherent difficulties caused by selection effects across dimensions of inherent ability and preferences, using administrative cut-off rules (see for example Hastings et al. (2013), Kirkeboen et al. (2016)) or attempting to control for observable factors (Hamermesh & Donald (2008)). A general finding of this literature is that returns to a college degree vary according to several factors, such as the field of study, degree classification earned, or institution attended. In this paper, I focus on the interaction between the supply of and return to graduates' skills in a partial equilibrium framework, thereby providing a potential mechanism for the observed differences in labour market outcomes for different cohorts of graduates.

Furthermore, this paper also speaks to a large body of literature that investigates how technological change affects the sorting of different workers across occupations and, correspondingly, the wage distribution. Exemplified by the task framework based on the seminal work of Autor et al. (2003) and further developed in Acemoglu & Autor (2011), the task framework shifts the emphasis onto specific job tasks and provides an explanatory framework in which a worker's skill set and the tasks to which they are assigned are jointly important for individual productivity and earnings. This provides an incentive to consider changes in the supply of and the demand for skills as important drivers of labour market sorting. While earlier work has focused primarily on the demand for skills (c.f. Firpo et al. (2011), Autor & Handel (2013), Goos et al. (2014), Deming (2017)), recent work has begun to look at the supply and demand of skills jointly (see Roys & Taber (2022), Diegert (2024)). Considering both skill supply and demand allows for a more comprehensive understanding of labour market dynamics, as it captures not only the availability of skills but also how they are valued across different occupations. Accounting for these forces separately is particularly relevant when considering long time horizons where the skill distribution might have changed over time, and is important for assessing and shaping policy that influences the distribution of skills. My paper represents another attempt in this growing literature, based closely on the approach outlined in Diegert (2024), who looks at the population of workers in the US, while this paper is focussed on the specific question related to recent university graduates in the UK.

Finally, this paper complements previous work that aims to elicit the skill content of university degrees. Altonji et al. (2014) create measures of the task content of different subjects by mapping task measures from the Dictionary of Occupational Titles to graduates' occupation choices. Similarly, Hemelt et al. (2021) collect information from online job postings to associate desired skills with different degree subjects. My paper differs in that, it uses occupation choice and wage information to estimate a continuous distribution of graduate skills. By focusing on wage outcomes and occupational sorting, my approach provides a more dynamic picture of how skills translate into labour market performance, taking into account both monetary rewards and the heterogeneity in how skills are valued across occupations.

To the best of my knowledge, this paper is the first to estimate the returns to and supply of university graduates' skills in the UK and apply the model to the question of graduate underemployment.

The rest of the paper is structured as follows: Section 2 presents some motivating facts about the labour market outcomes of young graduates in the UK over the last 20 years; Section 3 outlines the economic model of wage-setting and occupational choice; Section 4 presents the econometric strategy used to estimate the parameters of interest; Section 5 covers the discussion of results and Section 6 presents the counterfactual decompositions; finally, Section 7 concludes.

2 Motivating empirics

The following section provides evidence of how labour market outcomes for young graduates have evolved over the last 20 years, highlighting significant shifts that have implications for current debates on the effectiveness of higher education and the value of a university degree in today's economy. My main data source is the Quarterly Labour Force Survey (QLFS), which is a representative household survey in the UK, surveying approximately 40,000 responding UK households per quarter. The survey features a staggered longitudinal design, where households are interviewed for five consecutive quarters and 20% of the sample is replaced in every wave. In this section, I will use the five-quarter longitudinal version of the QLFS to motivate the analysis in the rest of the paper.

In my analysis, I focus on the outcomes of "young" graduates aged between 21[2] and 30 years. This group is likely more homogenous, and also has less labour market experience, meaning that their skill endowments will be more closely related to their post-graduation endowments, which are of key interest to this paper. Furthermore, the outcomes of this age group are shaping the public debate around university effectiveness.[3]

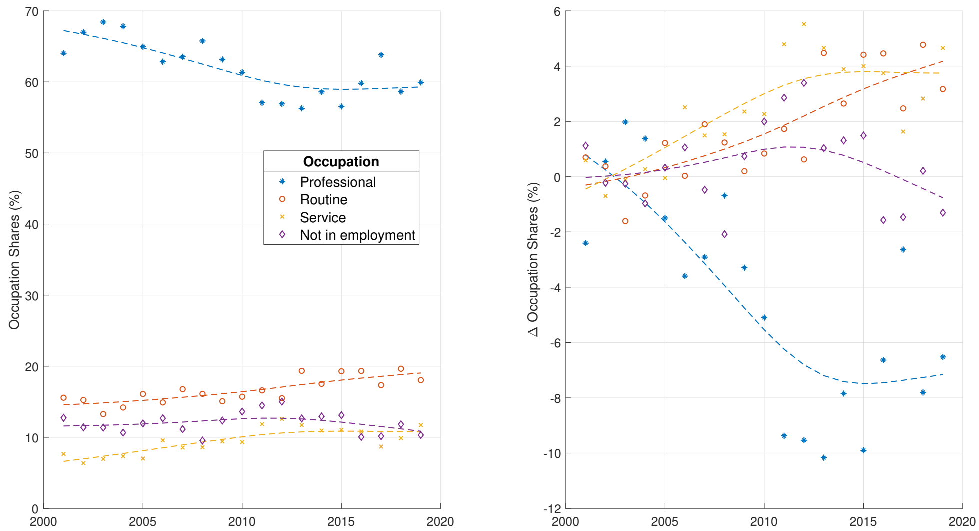
Figure 1 presents the occupation shares of young graduates between 2001 and 2019 disaggregated by 3 major occupation groups - Professional, Routine and Services, as well

[2]Typically in the UK students finish high school at 18 and enter 3-year University Courses.

[3]Often outcomes 5 years post-graduation are seen as an important milestone to judge the "success" of young graduates.

as non-employment.[4] By most accounts, a "graduate job" is a key marker of success for a young graduate, commonly identified with employment in professional occupations such as Law, Medicine, or Financial Services. The data reveals that a majority of young graduates are employed in professional occupations, with a high of around 65% in the early 2000s. The share exhibits a declining trend into the post-financial crisis years, bottoming out around 56% in the early 2010s, after which the numbers rebound somewhat. In contrast, the share of young graduates working in routine and service occupations increases steadily over the sampled period. Employment in routine occupations begins at around 15% in the early 2000s, but rises steadily over the sample period, increasing by around 4% by the end of the 2010s. The rise in the share of routine workers among graduates is particularly interesting, because over the same time period the share of routine workers amongst young adults without a degree is falling (see Figure A2). The share of graduates in service occupations likewise rises by around 4% over the 19-year period, although the lower initial baseline means that this represents an increase of close to 50%. Finally, the share of young graduates not in employment peaks around the years of the Great Recession (2010-2012) but then falls quite quickly to below its initial level.

Figure 1: Trends in occupation shares of young graduates in the UK (2001 - 2019)



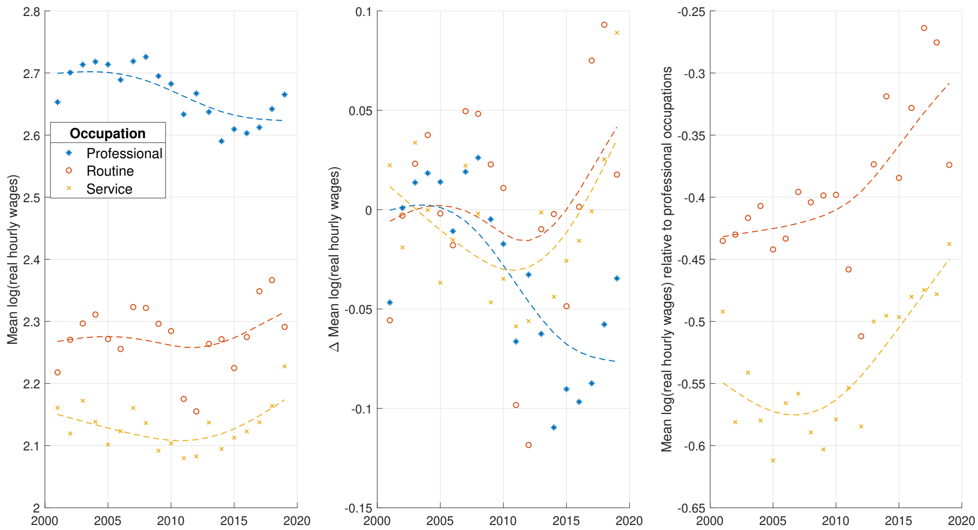
Note: Graduates aged 21-30 years; change relative to 2001-05 average. Broken lines are HP-filtered trends (smoothing parameter = 100). SOC 2000 1-Digit Occupation Classification. Professional includes codes 1-3; Routine includes codes 4,5,8,9; Service includes codes 6,7; Not in employment includes unemployed and those out of the labour force for other reasons. *Source:* Quarterly Labour Force Survey (2001-2019).

Figure 2 shows trends in hourly real wages of young graduates over the same period. Those employed in professional occupations earn by far the highest wages, as expected, but there is a general decline in hourly wages across all occupation groups in the wake of

[4]Occupational groups are based on 1-digit SOC 2000 classifications. Professionals include codes 1-3, Routine includes codes 4, 5, 8 & 9 and Service includes codes 6 & 7. Non-employed combines the unemployed and those not in the labour force.

the financial crisis. The wages of professionals contract most sharply between 2009 and 2015, which is in stark contrast to the wages of those employed in routine occupations which only experienced a small decline before making a strong recovery. Wages for service occupations fall quite notably in the first decade of the sample but then rebound fairly strongly in the second half. Indeed, while wages in professional occupations have not recovered their early 2000's levels, graduates in routine and service occupations have seen their wages recover much more quickly from the financial crisis and its aftermath. As a result of these dynamics, the relative pay premium enjoyed by those in professional occupations has fallen, by about a third relative to routine occupations and about a fifth relative to service occupations.

Figure 2: Trends in real wages among young graduates in the UK (2001 - 2019)



Note: Wages are log hourly wages, deflated by 2014 CPI. Working graduates aged 21-30 years. Broken lines represent HP-filtered trends (smoothing parameter = 100). Change relative to 2001-05 average. SOC 2000 1-Digit Occupation Classification. Professional includes codes 1-3; Routine includes codes 4,5,8,9; Service includes codes 6,7. *Source:* Quarterly Labour Force Survey (2001-2019).

The data presented in this section highlight some interesting trends in the labour market outcomes of young graduates. Over the two decades preceding the COVID-19 pandemic, the share of graduates entering professional occupations has declined while those of routine and service occupations have grown. At the same time, the wages of graduates in professional occupations have fallen post-2008 particularly relative to wages earned in routine and service occupations. While economic logic dictates that changes in relative wages should result in a reallocation of labour from one occupation to another, the underlying drivers of these changes are not well understood. In the following section, I will outline a quantitative economic model that might shed some light on the deeper, structural forces behind these trends.

3 Model

In this section, I present an economic model of occupation choice and wage determination for university graduates. The economic environment in this model closely follows the literature on task-based occupational choice and wage determination (c.f. Autor & Handel (2013), Roys & Taber (2016)), while the decision framework of the graduate is based on the approach of the multinominal choice literature (see Train (2009), Chapter 6).

My model begins at labour market entry, and abstracts from the decision to enter university, and other decisions taken during higher education. Instead, I assume that all unobserved heterogeneity among graduates can be summarized by a latent vector of unobserved skills s that captures all relevant differences in labour market skills among graduates. Upon graduation, graduates draw a stochastic realisation of their skill set from a parametric distribution. Given a set of skills a graduate then forms expectations about the wage they can earn across all possible occupations, which depends on the occupation-specific return to skill. They then choose an occupation match taking into account their expected wage and preferences over non-pecuniary job-attributes. This match is then observed by the econometrician. The detailed timeline assumed to hold in the model is specified in Figure 3 below.

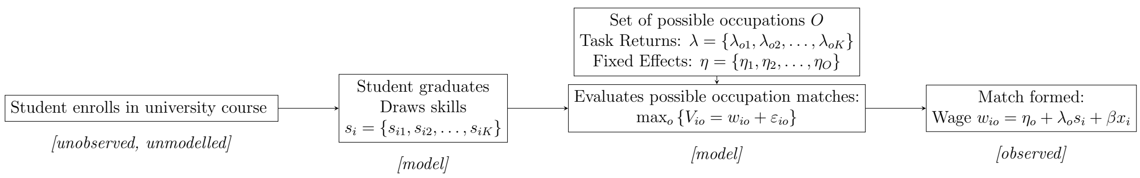


Figure 3: Model Summary

For ease of exposition, I first present a cross-sectional version of the model, with suppressed time subscripts. The extension to a more dynamic framework, where parameters are allowed to vary over time is straightforward.

A graduate's multidimensional skill set is summarized by a K dimensional vector $s_i = \{s_{i1}, s_{i2}, .., s_{iK}\}$ where each element $s_{ik} \geq 0$ describes how effective graduate i is at performing task k. I assume that s_i is an unobserved, random vector that is drawn from a parametric distribution $s_i \sim S$ that is the same for all individuals $i \in I$. On the demand side, the labour market consists of a large number of *occupations* that each use the different skills supplied to them in different proportions. Specifically, every occupation $o \in O$ has an associated vector $\lambda_o = \{\lambda_{o1}, \lambda_{o2}, ...\lambda_{oK}\}$ where each element $\lambda_{ok} \geq 0$ summarizes the productivity of task k in occupation o.[5]

A graduate's productivity therefore depends on her skill-set s_i as well as the task-

[5]In the rest of the paper I will refer to these as the "task" or "skill-specific" returns or prices of occupation o.

specific returns λ_o of the occupation she is matched with. Specifically, I assume that the potential log wage of graduate i in occupation o can be written as:

$$w_{io} = \eta_o + \lambda_o s_i + \beta x_i, \tag{1}$$

where η_o is an occupation-specific fixed effect, x_i is a vector of other characteristics (gender, labour market experience, etc.) and β is a vector of coefficients. This setup is common in the literature on tasks and skills (c.f. Autor & Handel (2013), Roys & Taber (2016)).

After graduation, graduates observe their skills, and all potentially relevant characteristics of all occupations and pick whichever occupation provides them with the highest utility. That is to say that every graduate can observe the set O of all available occupations and attach a personal valuation V_{io} to each of these options. I assume that the utility derived from the occupation is linear in the log wage[6] and other occupation characteristics leading to the following potential valuation:

$$V_{io} = w_{io} + \omega_o + \varepsilon_{io}, \tag{2}$$

where o is one of the available occupations, w_{io} is the expected log wage earned by i in occupation o, ω_o is an occupation-specific preference term that is common among all graduates and ε_{io} is an individual-specific preference shock that is **i.i.d.** across all agents and all occupations.[7] Accordingly, a worker i solves the following (static) occupational choice problem:

$$V_i = \max_{o \in O} \{V_{io}\} \tag{3}$$

Under these circumstances the individual's occupation choice o_i^* will refer to the occupation with the highest valuation. Importantly, the value of V_{io} is observed by the economic agent, while only o_i^* is observed by the econometrician.

Generally, economies of the type described above are characterized by the sorting of workers according to comparative advantage (see Roy (1951)). This self-selection of workers into different occupations according to their different abilities poses the main obstacle that is faced by the literature that is concerned with estimating *task returns* (i.e. the set λ), since there will be a positive correlation between an occupations' task prices λ_o and the skills supplied by workers selecting into this occupation (see Autor (2013)).

[6] This is likely to be the case for an economic agent with a suitably defined utility function (e.g. logarithmic), who is borrowing constrained. I believe it reasonable to assume that this situation applies to the sample population studied in this paper.

[7] The random component ε_{io} is random in an idiosyncratic sense. Two workers with the same deterministic wage may have different preferences over the set of occupations. This differentiation in choice behaviour is important since otherwise, the utility-maximizing choice would be the same for every worker, leading to unrealistic predictions. Furthermore, the introduction of this random term allows us to capture other factors that influence occupation choice besides the desire to maximize wages, such as other preferences or frictions in the labour market.

In this paper, however, rather than being harmful, self-selection is actually helpful as it allows us to make inferences from a worker's observed occupation o_i^* to her unobserved skill-set s_i.

4 Econometric strategy & estimation

4.1 Econometric Strategy

The primary goal of my econometric strategy is to estimate the structural parameters of the model by leveraging both the observed occupation choices and the wages of graduates. The key insight is that a graduate's occupation choice o_i^* and realized wage w_i^* are both informative about their unobserved skill level s_i, given the economic structure described earlier. The parameters of interest include those related to the determination of the log wage (η, λ, β), the occupation-specific preferences (ω), and the distribution of graduates' skills (S).

For estimation, I focus on a model with a single type of skill ($K = 1$), which simplifies the analysis while capturing the essential variation in the data. The labour market is divided into the three major occupation groups, along with the outside option of not working, resulting in four possible occupation choices ($O = 4$). To account for potential changes in the skill distribution over time, I separate the sample of graduates into two cohorts, denoted by $c = 1, 2$, based on the year they are first observed in the sample.[8] I specify the skill distribution S as a log-normal distribution with cohort-specific mean and variance parameters: $\log(s_i) \sim \mathcal{N}(\mu_c, \sigma_c^2)$. This allows the skill distribution to vary between cohorts, capturing shifts in the underlying skill levels of graduates. The log-normal is further a convenient choice, as it has positive support and naturally generates a right-skewed distribution of wages.

To account for changes in the returns to skill and job amenities over time, I allow the occupation fixed effects (η_o^τ), the occupation-specific returns to skill (λ_o^τ), and the occupation-specific preference terms (ω_o^τ) to vary annually, where $\tau = 2001, \ldots, 2019$. This temporal variation accommodates potential changes in the labour market across different years, that might have influenced graduates' labour market choices and outcomes.

I exploit the panel structure of the Quarterly Labour Force Survey (QLFS) dataset to identify the model parameters. Each individual's choices are observed over $T = 5$ consecutive quarters, providing multiple observations per individual. This panel data structure is crucial, as it allows me to observe within-individual variation over time, which aids in disentangling the effects of unobserved skills from the returns to skill. The

[8]The cut-off year is 2010.

final dataset contains 12,084 graduates, totalling 60,420 observations.[9]

Given the short time span of the panel, I assume that skills remain constant over the five quarters, i.e., $s_{it} = s_i$ for all t. This assumption is reasonable, as significant changes in skill levels are unlikely over such a brief period.

I now reintroduce time subscripts into the notation to reflect the panel data structure. The observed occupation and wage outcomes are vectors of length T: $o_i^* = \{o_{i1}^*, o_{i2}^*, ..., o_{iT}^*\}$, $w_i^* = \{w_{i1}^*, w_{i2}^*, ..., w_{iT}^*\}$. This notation emphasizes that I observe a sequence of choices and wages for each individual over time, which is essential for the estimation strategy.

To estimate the structural parameters of the model, I proceed in several steps. First, I derive the conditional probability of a graduate choosing a particular occupation, given their unobserved skill level. Second, I incorporate measurement errors in observed wages. Third, I derive the joint probability of observing both the occupation choice and the wage, conditional on the unobserved skill. Finally, I integrate over the distribution of unobserved skills to obtain the unconditional likelihood function, which depends only on observable variables.

I begin by normalizing the utility of the outside option (not working, $o = O$) to zero for convenience. I also make the standard assumption that the idiosyncratic occupation preference shocks ε_{it} are independently and identically distributed (i.i.d.) Type I Extreme Value with variance $\frac{\pi^2}{6\rho}$, where $\rho > 0$ is a scale parameter. Under these assumptions, the conditional probability that graduate i chooses occupation o_{it}^* at time t, given their skill level s_i, is given by:

$$\Pr(o_{it}^*|s_i) = \frac{e^{\rho\left(\eta_{o^*}^{\tau} + \lambda_{o^*}^{\tau} s_i + \beta x_{it} + \omega_{o^*}^{\tau}\right)}}{1 + \sum_{o=1}^{O-1} e^{\rho\left(\eta_o^{\tau} + \lambda_o^{\tau} s_i + \beta x_{it} + \omega_o^{\tau}\right)}}, \tag{4}$$

which is the familiar logit probability, having substituted the expected wage 1 into the graduates' decision problem 2.

Next, I account for measurement errors in observed wages. Suppose the observed wage w_{it}^* is subject to measurement error:

$$w_{it}^* = w_{it} + \nu_{it}, \tag{5}$$

where ν_{it} is an i.i.d. disturbance term, independent of the worker's occupation choice and distributed normally with mean zero and variance ϕ_o^2, i.e., $\nu_{it} \sim \mathcal{N}(0, \phi_o^2)$. This measurement error can be interpreted as an idiosyncratic wage shock that is not anticipated by the graduate at the time of occupational choice.

As a result, we can express the probability of observing the observed wage, conditional

[9] Because the QLFS is not designed to be a panel survey the 5-quarter longitudinal panel suffers from considerable attrition, resulting in a dataset that is significantly smaller than the purely cross-sectional QLFS. However, the L-QLFS dataset contains longitudinal weights to account for attrition bias, which I include in my estimation.

on a the graduate's skill-set s_i and occupation choice o^*_{it}:

$$\Pr(w^*_{it}|s_i, o^*_{it}) = \frac{e^{(-\frac{\nu^2_{it}}{2\phi^2})}}{\sqrt{2\pi\phi^2_o}}. \tag{6}$$

Combining the conditional choice probability with this expression, the joint probability of observing both the occupation choice o^*_{it} and the wage w^*_{it}, conditional on s_i, is:

$$\Pr(o^*_{it}, w^*_{it}|s_i) = \left(\frac{e^{\rho\left(\eta^\tau_{o*}+\lambda^\tau_{o*}s_i+\beta x_{it}+\omega^\tau_{o*}\right)}}{1+\sum_{o=1}^{O-1} e^{\rho\left(\eta^\tau_o+\lambda^\tau_o s_i+\beta x_{it}+\omega^\tau_o\right)}}\right)\left(\frac{e^{(-\frac{\nu^2_{it}}{2\phi^2})}}{\sqrt{2\pi\phi^2_o}}\right). \tag{7}$$

Since we observe multiple periods for each individual, the joint probability of observing the full sequence of choices and wages over T periods, conditional on s_i, is given by the product over time:

$$\Pr(o^*_i, w^*_i|s_i) = \prod_{t=1}^{T}\left(\left(\frac{e^{\rho\left(\eta^\tau_{o*}+\lambda^\tau_{o*}s_i+\beta x_{it}+\omega^\tau_{o*}\right)}}{1+\sum_{o=1}^{O-1} e^{\rho\left(\eta^\tau_o+\lambda^\tau_o s_i+\beta x_{it}+\omega^\tau_o\right)}}\right)\left(\frac{e^{(-\frac{\nu^2_{it}}{2\phi^2})}}{\sqrt{2\pi\phi^2_o}}\right)\right). \tag{8}$$

Finally, integrating over the distribution of s[10] leads to the unconditional joint probability:

$$\Pr(o^*_i, w^*_i) = \int\prod_{t=1}^{T}\left(\left(\frac{e^{\rho\left(\eta^\tau_{o*}+\lambda^\tau_{o*}s_i+\beta x_{it}+\omega^\tau_{o*}\right)}}{1+\sum_{o=1}^{O-1} e^{\rho\left(\eta^\tau_o+\lambda^\tau_o s_i+\beta x_{it}+\omega^\tau_o\right)}}\right)\left(\frac{e^{(-\frac{\nu^2_{it}}{2\phi^2})}}{\sqrt{2\pi\phi^2_o}}\right)\right) f(s)d(s) \tag{9}$$

Standard results (c.f. McFadden & Train (2000)) guarantee, that we can use the unconditional choice probability in (9) to get consistent estimates for η, λ, β, ω and S using simulated maximum likelihood. In the appendix, I describe a complete algorithm that can be used to estimate the parameters of interest from this model, using the likelihood function implied by (9).

4.2 Identification

In this subsection, I provide an intuitive explanation of how the key parameters of the model are identified. My identification strategy is underpinned by the formal results presented in Diegert (2024), who provides a non-parametric identification theorem for a model that nests mine. Like him, I exploit the factor structure inherent in the model, which facilitates the identification of its key parameters.

[10] There is no closed-form solution for this integral, but the integration step can be performed via simulation. In the estimation, I use a set of 1,000 pseudo-random Halton draws.

Specifically, the unobserved skill s_i serves as a latent variable influencing both wages and occupational choices. Observed wages can therefore be viewed as measurements of these unobserved skills, with the time-varying shocks ν_{it} acting as i.i.d. measurement errors. To identify the parameters of the log wage equation, I exploit the properties of these measurement errors. Since ν_{it} is assumed to be i.i.d. and independent of both the unobserved skills s_i and the occupational choices, the variation in observed wages across different occupations and time periods provides the necessary information to disentangle the effects of the unobserved skills, the occupation-specific returns to skill λ_o^τ, and the occupation fixed effects η_o^τ.[11] This allows me to identify the wage equation up to a scalar. I introduce a normalization by fixing one of the occupation-specific returns to skill. Specifically, I set $\lambda_1^{2001} = 1$, which anchors the scale of s_i. This normalization enables the identification of the remaining parameters relative to this benchmark. Because we further observe the same individual in different time periods τ, we can further identify changes in λ_o^τ and η_o^τ using changes in individual wages over time.

Intuitively, since s_i is fixed for each individual, observing changes in wages, as individuals switch occupations over time, helps identify λ_o^τ and η_o^τ separately from s_i. This is similar to estimating occupation-specific fixed effects in a panel-data setting, where within-individual variation aids in separating the effects of permanent and idiosyncratic heterogeneity.

Once the parameters of the log wage equation are identified, differences in occupational choices among individuals help pin down the occupation-specific preference parameters ω_o^τ. Given that the expected wages are determined by the estimated λ_o^τ and η_o^τ, any systematic variation in occupational choices beyond what wages can explain is attributed to the non-pecuniary preferences captured by ω_o^τ.

For a more formal verification, one can show that my model adheres to the assumptions I.1–I.4 in Diegert (2024), which are sufficient for identification in his framework: Firstly, the specification of the log wage as linear in skill with additive errors satisfies Assumptions I.1 and I.4 in Diegert (2024). The factors affecting productivity are separated into the time-invariant skill s_i and the time-varying shocks ν_{it}, as shown in the log wage equation 1. Moreover, when expressed in exponential form, the productivity function is multiplicatively separable in ν_{it} and s_i, this satisfies the multiplicative separability required by Assumption I.4. Secondly, the assumption that $\nu_{i,t}$ is i.i.d. and independent of s_i and the occupational choices satisfies Assumption I.2, ensuring that the productivity shocks are serially independent and independent of skills and preferences. Thirdly, since the productivity shocks ν_{it} do not enter into the utility function, Assumption I.3 is satisfied. Utilities depend only on expected wages (which are functions of s_i and x_i)

[11] In particular, the distribution of wages within a given occupation, conditional on x_{it}, identifies the quantity $s_i\lambda_o^\tau$, while the difference in average wages across occupations helps identify η_o^τ.

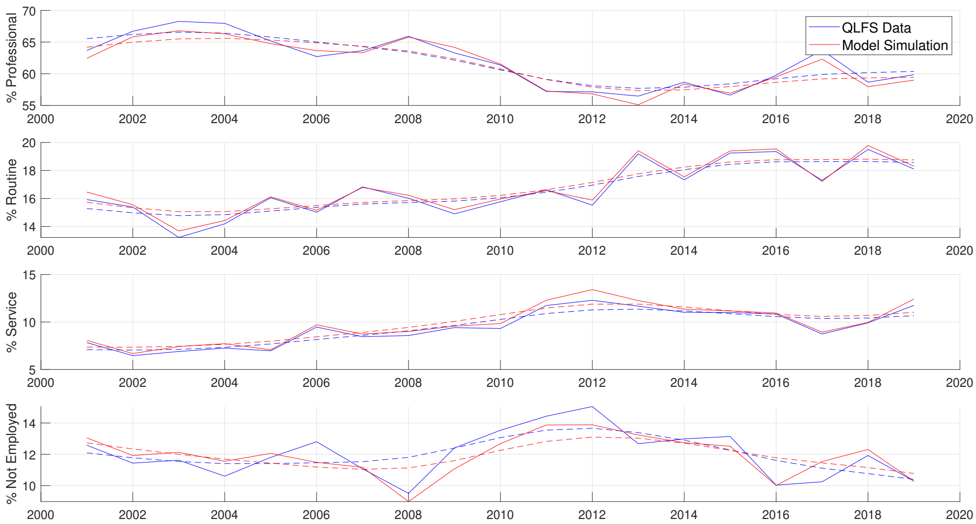
and preferences, not on the time-varying productivity shocks. Finally, since $K = 1$ and $T = 5$, I satisfy the condition $K < 2T$, which is necessary for identification in Diegert's framework.

4.3 Estimation

After estimating the model, I evaluate its ability to capture graduates' occupation choices and wage outcomes. For this purpose, I simulate a random, representative sample of 1,000,000 graduates. Figure 4 below highlights the model fit with respect to the occupation choices of graduates while Figure 5 shows the model fit with respect to average log wages.

The model fit for the occupational choices of young university graduates is generally close to the actual data, effectively capturing long-run trends.[12] Specifically, the model performs well in aligning with long-term patterns in employment across different occupations, such as the fall of the share of professional occupations and the steady rise of routine and service occupations.

Figure 4: Model fit - occupation shares

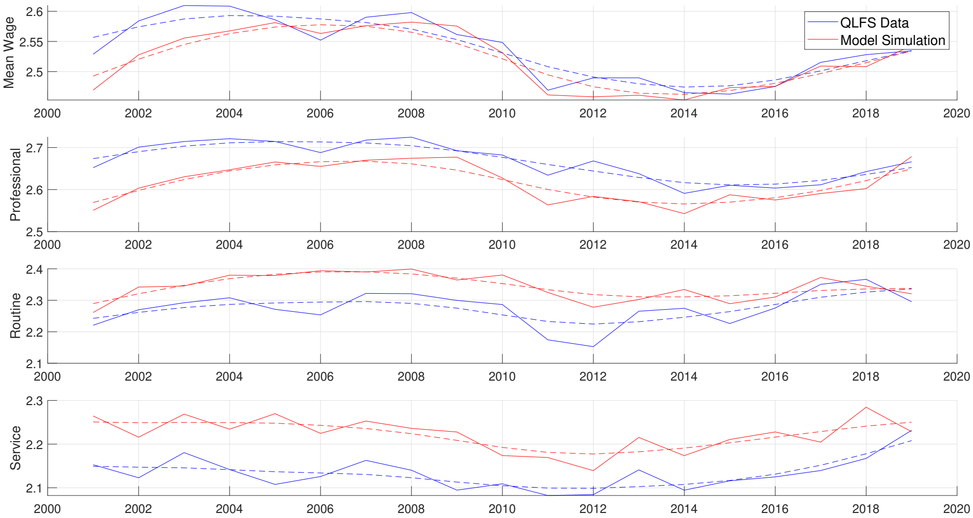


Note: Broken lines represent HP-filtered trends with a smoothing parameter of 6.25, based on a simulated sample of 1,000,000 graduates. *Source:* Baseline Model.

Figure 5 provides a comparison of mean log hourly wages across the occupational categories. The model's simulations align closely with the observed data, reproducing the relative wage differences between occupations and the overarching trends. The model shows a slight underprediction in the professional category, with an average difference of 0.05 in log wages on average. In contrast, the model slightly overpredicts wages in routine and service occupations by 0.07 and 0.09, respectively.

[12] For details see Table A1 in the appendix.

Figure 5: Model fit - mean log wages



Note: Mean of log hourly wages, deflated by 2014 CPI. Broken lines represent HP-filtered trends with a smoothing parameter of 6.25, based on a simulated sample of 1,000,000 graduates. *Source:* Baseline Model.

Overall, the model does a good job of reproducing the main elements for both occupational distribution and wage outcomes. It effectively captures the observed shifts in employment shares across professional, routine, and service categories, as well as the relative wage dynamics between these groups. In the next section, I will outline the model's results in some additional detail.

5 Results

This section presents the results of the estimated model. I first present the model estimates for the graduate skill distribution, the returns to skill and the non-pecuniary benefits, accruing to different occupations, as these are the underlying drivers of the observed dynamics in graduates' labour market outcomes. I then analyse, what the changes in these structural determinants mean for the sorting of graduates into different occupations. Finally, I decompose the importance of the different factors, using counterfactual decompositions in the following section.

5.1 Graduate skills & the returns to skill

Table 1 presents estimates of the distribution of graduates' skills, comparing the 2001-2009 and 2010-2019 cohorts. For ease of comparison, the distributions are standardized to have a mean of 0 and a standard deviation of 1 in the first cohort. The main observation regarding the changing graduate skill distribution across the two cohorts is that between 2001-2009 and 2010-2019, the mean (median) of the skill distribution fell by 0.21

(0.21) points, equivalent to 21% (21%) of a standard deviation among the first cohort—a moderate but not insignificant reduction.[13] Anecdotally, this shift is consistent with the main reforms to the production of university graduates that occurred over the late 1990s and 2000s. The rapid expansion of higher education participation at both the intensive (increased number of students) and extensive (increased number of HE providers) margins likely negatively affected the average ability of high school leavers entering university, as well as the quality of the instruction they received while at university (see Carneiro & Lee (2011)). Curiously, however, the variance of the skill distribution has not been markedly affected by these developments. The standard deviation of the second cohort is 6% below that of the first cohort, suggesting a small contraction of the skill distribution.

The percentiles of the skill distribution provide additional insight into how the contraction is distributed across different parts of the distribution. The 10th percentile (P10) and the 25th percentile (P25) show smaller declines compared to the median, with P10 falling by 0.14 points and P25 by 0.18 points, whereas the median fell by 0.21 points. In contrast, the 75th percentile (P75) and the 90th percentile (P90) experienced larger declines, with P75 falling by 0.25 points and P90 by 0.28 points. This suggests that the contraction primarily affects the upper tail of the distribution, indicating that the decline in skills has been more pronounced among higher-skilled individuals. One possible explanation for this pattern could be that universities are focusing more on supporting lower-performing students, which may have led to a shift in resources and attention away from the most talented students, ultimately contributing to a disproportionate decline in their skill levels.

Table 1: Details of Graduate Skill Distribution

	Mean	**Median**	**Std. Dev.**	**P10**	**P25**	**P50**	**P75**	**P90**
2001-2009	0.000	-0.004	1.000	-1.278	-0.678	-0.004	0.668	1.285
2010-2019	-0.212	-0.217	0.944	-1.416	-0.853	-0.217	0.421	1.002
Δ	-0.212	-0.213	-0.056	-0.138	-0.175	-0.213	-0.247	-0.283

Note: Simulations based on a representative sample of 1,000,000 graduates. Mean standardized to 0 and standard deviation to 1 in 2001-2009.
Source: Baseline Model.

Put into context, these results can already provide a partial explanation as to why the share of professional occupations has fallen among university graduates. Assuming that professional occupations provide a higher return to skills than service or routine occupations, a fall in the average level of skills would—ceteris paribus—make these occupations less attractive to graduates, leading to a decline in the share of graduates entering these roles.

Figure 6, Subplot 1, provides the model's estimates for the returns to skill for the

[13] See also Figure A4 in the appendix.

different occupations. In line with expectations, professional occupations provide the highest return to skills, followed by routine and then service occupations. This hierarchy of returns means that as the overall skill level of graduates decreases, the incentive to pursue professional occupations diminishes, especially for those whose skills no longer meet the higher threshold required. Instead, these graduates may opt for routine or service occupations, where the returns to skill, while lower, are more attainable given their abilities.

Over the 19-year period, there is a pronounced increase in the return to skills across all occupational groups, consistent with the theory of within-occupation task upgrading. Before the financial crisis, the trend is somewhat stable for professional occupations, with a slight increase for routine occupations, and a strong downward trend for service occupations. Following 2009-2010, there is a notable level shift upward for professional and routine occupations, with a possible continuing trend of slight growth. In contrast, the returns for service occupations exhibit strong growth in the post-crisis period, highlighting a significant shift in the return to skills in these roles.

While it is not easy to discern from the plot, the relative premium that professional occupations offer falls slightly as a result of a faster catch-up by the other occupations. Throughout the 2000s, the return to skills in routine (service) occupations is 82% (65%) of that of professional occupations, while in the 2010s, the gap decreased to 86% (68%), respectively. The loss of comparative advantage by professional occupations is a potential contributing factor in the changing occupational destinations of young university graduates.

Figure 6: Returns to skill and non-pecuniary utility components by occupation

Note: Subplot 1: Returns to skill relative to professional occupations in 2001 (=1). Subplot 2: Non-pecuniary utility; value of non-employment normalized to 0. *Source:* Baseline Model Estimates.

When choosing an occupation, money is rarely the only objective that matters. Occupations provide important (dis-)amenities to their workers such as attractive workplace conditions, flexible workdays or a general feeling of prestige associated with a particular job. These additional perks enter the decision framework of workers and can play an important factor in explaining occupation choices in conjunction with differences in earnings and wages (see Sorkin (2018)). My model captures these non-pecuniary preferences in two distinct ways: Firstly, the idiosyncratic preference term ε_{io} that reflects the (random) preferences of individual graduates, and secondly the set of occupation-specific general preferences ω_{ot} that reflect the general prevailing tastes of the population of graduates as a whole.

Figure 6, Subplot 2, plots the estimated values of the fixed preferences over the sample period. The plotted values are relative to the outside option of not working, which has been normalized to 0 for all periods. The estimated values are all negative since any work carries a degree of disutility relative to the option of not joining the labour force. Since the numéraire of the model is the log wage, we can deduce that, on average, graduates require a log wage of between 1.80 and 2.25 for them to consider working more attractive than remaining outside the labour market.

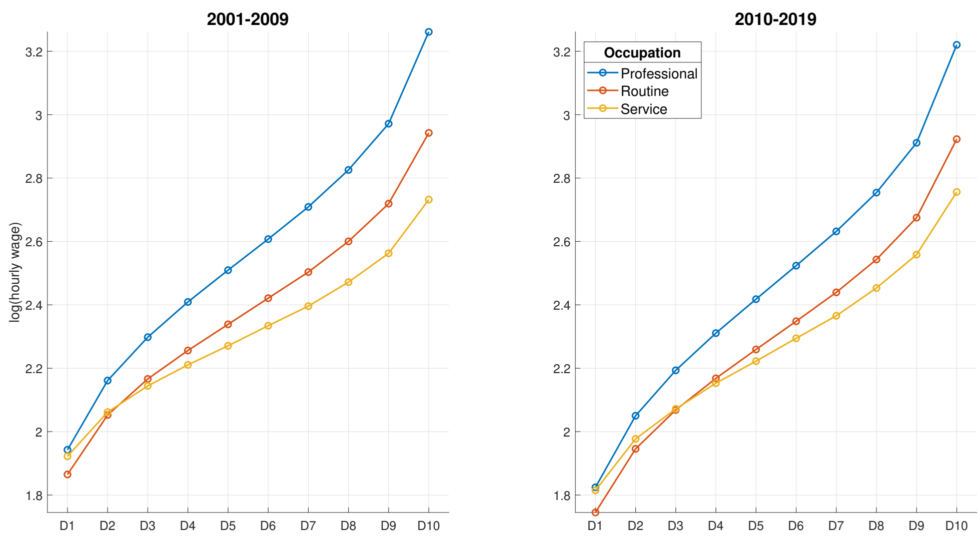
While it is generally true that labour comes with a utility penalty, there are important variations across occupations and over time. First, there appears to be an ordering among the different occupations, with professional occupations providing the lowest level of dis-amenity, followed by routine occupations, which weakly dominate the service occupations. This pattern is consistent with expectations—professional jobs provide more interesting work, greater autonomy, and are less likely to involve unpleasant or dangerous activities. Second, while the relative rankings remain stable over time, there is a general trend of increasing amenity values across all occupations from about 2008 until the end of the sample period. During this general rise, the gap between professional and routine/service occupations appears to narrow somewhat, suggesting that the latter are becoming more attractive relative to professional occupations.

A general increase in amenity values could suggest an overall improvement in working conditions, or—since the values are normalized relative to not working—a depreciation of the outside option. Possible drivers of this depreciation could include economic insecurity, such as a lack of stable income for those not in employment, reductions in welfare benefits, or an increase in the perceived risks associated with being out of the labour force for an extended period. Such factors make the option of not working less attractive, thus increasing the relative appeal of available job opportunities. This improvement might include factors such as better job security, enhanced work-life balance, or improved workplace environments, making these jobs more appealing over time. A reduction in the relative gap between professional and routine/service occupations hints at a relative appreciation of the utility associated with these occupations, possibly through improved working conditions such as better health and safety standards or more flexible working hours. It could also indicate a more general shift in societal attitudes towards such "non-graduate" jobs, where these roles are increasingly seen as viable and respectable career options. In either case, this reduction in the disamenities gap is likely a contributing factor to the increasing share of graduates in non-typical occupations, as the improved perception and conditions of these roles make them more attractive to graduates who may otherwise have pursued traditional professional careers.

5.2 Skills and sorting

The economy described by the model in section 3 is characterised by the endogenous sorting of graduates into different occupations. Graduates with high levels of skill will - ceteris paribus - enter occupations where the return to these skills is high, while those with low levels of skill will be drawn to occupations where the skill-based compensation is lower. Figure 7 shows the mean hourly wage by skill deciles for the 3 different occupation groups. Wages increase monotonically by skill deciles for all occupations; however, the rate of increase depends on the value of λ_o and hence the increase is steepest for professional occupations. This can be clearly seen across both cohorts, as the gap in log wages, is small at the first two or three deciles, but then opens up, becoming larger as the comparative advantage of professional occupations relative to routine and service occupations begins to assert itself.

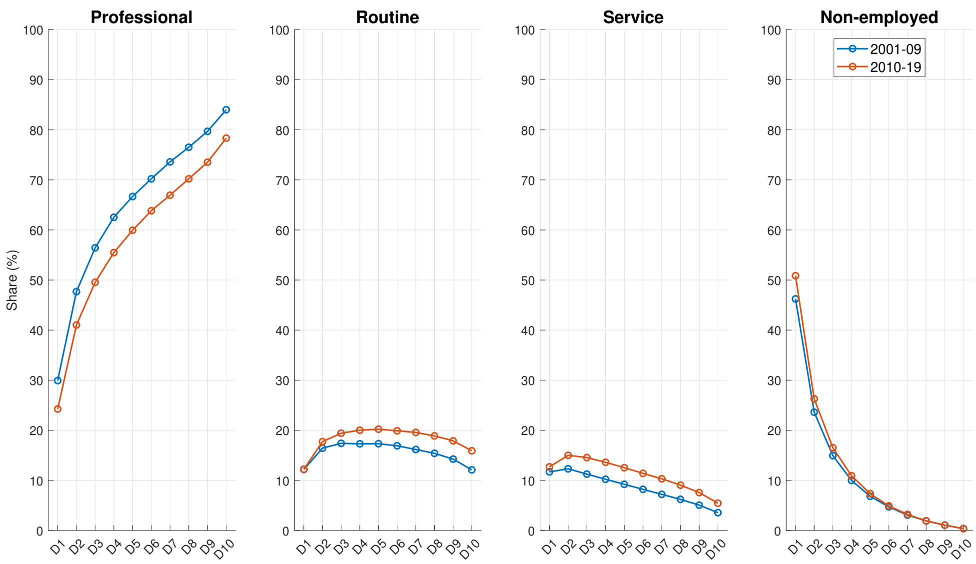
Figure 7: Log hourly wage by skill deciles



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles are based on the respective period. *Source:* Baseline Model.

Figure 8 illustrates how this dynamic results in the sorting behaviour of graduates into different occupations. At the lowest skill deciles, graduates do not have a strong incentive to enter professional occupations, or indeed any occupation at all. However, as skill levels increase, the probability of being employed in a professional occupation increases rapidly. For example, a graduate in the 5th decile of the skill distribution is more than twice as likely to be employed in a professional occupation than a graduate in the first decile, and for those in the top decile, the probability is around 80%. For routine and service occupations, the general pattern is approximately hump-shaped. Initially, moving into higher skill deciles graduates are more likely to be employed in these occupations - primarily because they are more likely to be employed at all - but this trend generally peaks around the 3rd or 4th decile and then slowly decreases, as higher-skilled graduates sort into professional jobs. Comparing the two cohorts we can see how the changing skill distributions and returns to skills have affected the sorting across the skill distribution. Generally, the patterns remain the same across both cohorts, although graduates are slightly less likely to enter professional occupations and slightly more likely to join routine or service occupations at every skill decile.

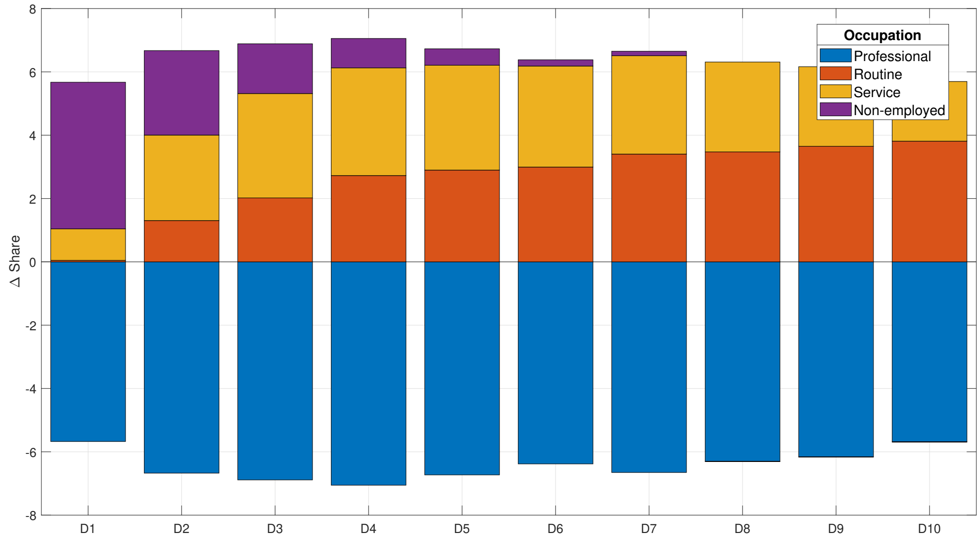
Figure 8: Occupational sorting by skill deciles



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles are based on the respective period.
Source: Baseline Model.

Figure 9 provides a more detailed assessment of the differences in sorting patterns across the skill distribution. As indicated in Figure 9, graduates are less likely to be employed in professional occupations at every point of the skill distribution, but the effect is not homogenous and ranges from -7% in the fourth skill decile to just over -5.5% in the 10th decile. Accordingly, the alternative career paths chosen by graduates also vary across the skill distribution: At the lower end of the skill distribution, graduates are more likely to be out of the labour market entirely, while between the 3rd and 5th decile they are increasingly likely to join service occupations. From the 6th decile onwards, the displaced graduates are increasingly moving into routine occupations instead. These patterns highlight that the observed changes to the labour market destinations of graduates are heterogeneous across the skill distribution and that therefore interventions to address perceived imbalances need to be carefully targeted.

Figure 9: Changes in occupational sorting by skill deciles



Note: Changes in occupation shares, based on a simulated sample of 1,000,000 graduates. Skill deciles are based on the respective period. *Source:* Baseline Model.

In this section, I have outlined the changes that have taken place in the labour market for young graduates. I have focussed on the changing skill distribution, the changing return structure for these skills and non-pecuniary aspects of different occupations and how these might contribute to the observed labour market dynamics. In the next section, I will explore the question of what changes are driving the observed patterns of occupational choices for both cohorts, by running counterfactual decompositions.

6 Counterfactual decompositions

In this section, I consider a number of counterfactual experiments, in order to assess the importance of different structural forces in driving the changing labour market outcomes of young graduates. The model allows me to decompose the observed changes, by fixing certain parameters at their earlier values and simulating the model into later periods. The differences between these counterfactual simulations with respect to the estimated full model will provide some insight into the underlying factors driving the observed patterns in the data.

There are three main structural forces that I have modelled and that I consider in this exercise: i) the skill distribution of young graduates (represented by μ and σ); ii) the occupation-specific skill prices and returns to skill (represented by λ_o and η_o); and iii) the non-pecuniary component of utility (ω_o). Each of these three represents a potential mechanism for the changing labour market outcomes of young graduates. I will assess the importance of each channel by fixing all parameters except for those associated with the considered mechanism at their 2001-2009 values, only allowing the parameters of interest to take on their estimated 2010-2019 values.[14] I then simulate the model and compare the outcomes with those from the baseline model. To assess the relative contributions of each factor, I employ the Shapley decomposition method (see Shorrocks (1982)). The Shapley decomposition calculates the marginal contribution of each factor by considering all possible combinations, ensuring a fair allocation of the total effect among the factors. Unlike a simple decomposition, the Shapley method accounts for interactions between factors, providing a more accurate and unbiased estimate of each factor's impact. This is particularly relevant in cases like this, where there are nonlinear interactions between different structural forces. In our model, the effects of skills, preferences, and prices on labour market outcomes are not merely additive; they interact in complex ways that can amplify or mitigate each other's impact. The Shapley decomposition captures these nonlinear interactions by considering all possible combinations of factors, ensuring that the contribution of each factor is properly attributed even when their effects are interdependent. In the following, I will refer to these three counterfactuals as the *skills,*the *prices*counterfactual and the *preferences* respectively.

Table 2 shows the results from the counterfactual decomposition exercises. The first column restates the average change in the share of the different occupations and occupation average wages between the periods 2001-2009 and 2010-2019 based on the baseline model.[15] The remaining columns show the average period difference based on the relevant counterfactual as well as the percentage of the change in a given quantity explained by the counterfactual.

The baseline model shows a reduction of approximately 6.3% in the share of graduates entering professional occupations between the periods 2001-2009 and 2010-2019. The skills counterfactual aims to isolate how much of this decline is explained by changes in the skill distribution of graduates. Specifically, during this period, the mean level of graduate skills declined by about 21% of a standard deviation, which we hypothesized to have negatively impacted graduates' ability to secure professional jobs.

The simulation suggests that changes in the skill distribution explain around 46% of the decline in the share of professional occupations. If only skills had changed between these two periods, the share of graduates in professional roles would have decreased by roughly

[14]For those parameters that vary year on year, I fix them at the average of the earlier period.

[15]The changes based on the baseline model are very close to those exhibited in the data. See Table A3 for details.

Table 2: Shapley Decompositions

Baseline Model		**Skills Counterfactual**		**Prices Counterfactual**		**Preferences Counterfactual**	
Occupation Shares	**Total Δ**	**Δ**	**Explained (%)**	**Δ**	**Explained (%)**	**Δ**	**Explained (%)**
Professional	-6.313	-2.885	45.693	-0.930	14.724	-2.671	42.302
Routine	2.623	0.175	6.659	-0.633	-24.128	3.048	116.193
Service	3.034	0.580	19.123	0.366	12.079	2.195	72.341
Non-employed	0.656	2.130	324.801	1.196	182.399	-2.573	-392.319
Log Wages							
Mean	-0.070	-0.079	111.954	0.027	-38.104	-0.023	32.156
Mean Professional	-0.052	-0.083	159.949	0.032	-61.339	-0.005	10.452
Mean Routine	-0.044	-0.059	133.740	0.019	-44.386	-0.007	15.636
Mean Service	-0.031	-0.039	127.173	0.019	-61.782	-0.008	25.128

Note: Simulations based on a representative sample of 1,000,000 graduates. *Source:* Baseline Model and Counterfactual Model Simulations.

2.9%. As shown in Figure 9, subplot 1, most of this reduction in professional occupations is concentrated among graduates in the lower skill deciles, and there is little evidence of re-sorting into routine and service occupations; instead, graduates predominantly exit the labour force altogether. This is due to a cascading effect caused by the almost uniform reduction in the skill level of graduates. As graduates become less skilled, they sort downward in the ranking of occupations into less skill-intensive ones. At the conclusion of this process, approximately the same mass of graduates that have left professional occupations has left the labour force altogether, leading to a large increase in the share of non-employed graduates. This is an important observation since it shows that a fall in the skill level by itself cannot account for the observed patterns of graduate's occupation choices, particularly the reallocation to routine and service roles.

The skill counterfactual also predicts changes in average wages that closely match the majority of the baseline simulation. However, it tends to overshoot, particularly for professional occupations, where it overpredicts the decline in wages by around 160%. This indicates that while declining skills can explain much of the observed wage decline, other forces are at play in moderating these effects.

The prices counterfactual aims to understand how changes in occupation-specific returns to skills have shaped labour market outcomes. The results indicate that changes in skill prices have not had a major impact on the share of graduates in professional or routine and service occupations.

The prices counterfactual explains approximately 15% of the decline in the share of professional occupations. It incorrectly predicts a decline in the share of routine occupations, while it accounts for about 12% of the increase in service occupations. Overall, the price counterfactual does not predict a large reallocation of graduates across different occupational categories.

The exception to this is the share of non-employed, where the counterfactual predicts a large increase, amounting to 180% of the baseline level. This significant increase in non-employment highlights the strong disemployment effect that rising skill prices have on graduates, especially those at the lower end of the skill distribution. The inability of these individuals to meet the higher skill thresholds leads to their exclusion from the labour market, resulting in a substantial rise in non-employment. Figure 9, subplot 2, shows that the increase in returns to skill has raised the skill threshold at which a graduate is indifferent between working and not working. As a result, there is a strong disemployment effect at the lower end of the skill distribution, while the rest of the skill distribution remains largely unaffected.

The prices counterfactual makes a prediction contrary to the baseline model, suggesting that wages should increase as a result of higher returns to skill. This discrepancy underscores the importance of the interaction between changes in skill levels and skill prices in matching the observed data. Without accounting for the decline in average skills, the rise in skill prices alone would imply a wage increase, which does not align with the observed trends.

The preferences counterfactual examines how changes in non-pecuniary preferences influence occupational choices. The results suggest that changes in preferences account for approximately 42% of the decline in the share of professional occupations and almost all of the observed increase in routine (116%) and service (72%) occupations. Preferences also play a significant role in reducing non-employment, a trend that mostly offsets the disemployment tendencies of the skills and price counterfactuals.

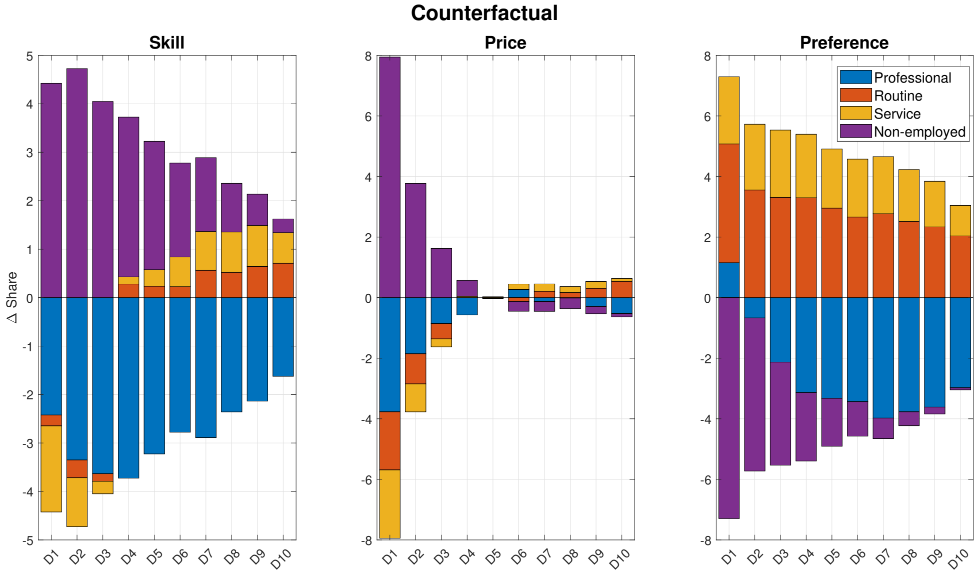
Figure 9, subplot 3, shows that a lower value of non-employment makes it more likely that low-skilled graduates are employed, particularly in routine and service occupations. This interaction between skill levels and preferences is crucial in explaining why non-employment has not risen more substantially in response to the decline in skill levels. An increase in amenities in routine and service roles relative to the outside option make them more attractive options for graduates who might otherwise remain out of the labour force.

The increase in the share of routine and service occupations can also be linked to changes in relative amenity values. As the perceived disamenities of these roles have decreased relative to professional roles, graduates are more likely to enter these occupations. The effects on wages are generally small and arise mainly through selection into lower-paying occupations, such as routine and service jobs.

The decomposition has shown that each factor plays a distinct role in shaping the observed changes in labour market outcomes for young graduates. Specifically, it highlights how changes in skill distribution, skill prices, and non-pecuniary preferences contribute differently to the shifts in occupational choices and employment patterns:

1. Changes in the skill distribution explain a significant portion of the decline in professional occupations but have a limited impact on the rise of routine and service jobs. By themselves, they also imply an unrealistically large increase in the share of non-employed graduates.

Figure 10: Occupational sorting under different counterfactuals (2010-2019)



Note: Change in occupational choice probability relative to the baseline model. Based on a simulated sample of 1,000,000 graduates. *Source:* Counterfactual Models.

2. Changes in skill prices tend to push graduates out of the labour market or into less skill-intensive roles, but their impact on routine and service occupations is limited. In isolation, they also imply increases in average wages, which are contrary to the observed patterns.

3. Changes in non-pecuniary preferences are the main driver behind the observed rise in routine and service occupations, suggesting that shifts in job characteristics and broader social attitudes have made these roles more appealing. Also, the decrease in the value of non-employment dampens the disemploying tendencies of the other two channels.

The overall message from these decompositions is that the changes in graduates' labour market outcomes cannot be attributed to a single factor alone. Instead, they result from the complex interaction of declining skill levels, shifting demand for skills, and evolving preferences. Each of these forces plays a role, and their combined effects shape the observed trends.

The findings suggest that policy interventions aimed at addressing graduate underemployment need to consider all three factors. Improving the skill levels of graduates is necessary but not sufficient—attention must also be given to enhancing the attractiveness of professional roles and addressing the exclusionary effects of rising skill prices. Moreover, policies aimed at improving job quality in routine and service occupations could further influence the sorting of graduates into these roles.

The interaction between these channels is particularly important. For example, changes in skill prices without corresponding changes in skill levels can lead to disemployment effects, while shifts in preferences can mitigate some of the negative impacts of declining skill levels. Understanding these interactions is key to designing effective policies that support graduates in finding suitable employment.

7 Conclusion

The formation of human capital and the acquisition of specific skills lie at the heart of a university education. With an increasing number of graduates in the UK failing to obtain "graduate jobs," there is growing public concern that universities are not equipping graduates with the skills they need to succeed in the labour market. However, it is not only the supply of skills that determines the labour market outcomes of graduates. In this paper, I have developed an economic model that accounts for the changing skill distribution of graduates, as well as the evolving demand for these skills and the evolving preferences that jointly determine the distribution of graduates across occupations.

My estimation has found that while the average level of graduates' skills has fallen by about 21% of a standard deviation, this explains around 46% of the decline in the share of graduates entering professional occupations. Sorting between graduates and occupations based on skill means that those failing to secure a "graduate job" are heavily concentrated in the lower deciles of the skill distribution, where they are increasingly unlikely to participate in the labour market. Those in higher-skill deciles are increasingly sorting into routine and service occupations, where an increase in variable compensation has made these occupations more attractive to graduates with higher levels of skill. This is complemented by a change in non-pecuniary preferences, particularly a reduction in the value of the outside option of not working, which makes employment within these occupations a more attractive choice.

The overall conclusion of this paper is that no single trend—skills, prices, or preferences—can easily account for the evolution of graduates' labour market outcomes. Instead, these are the result of intricate interactions between multiple factors and channels. This should serve as a note of caution to all those interested in "solving" the graduate underemployment problem. While aligning graduates' skills more closely with the demands of the labour market is likely a good idea, it is important to consider where in the distribution these concerns are most pressing. Giving already highly achieving students a leg up is unlikely to affect their chances of obtaining a graduate job substantially. Rather, proper emphasis should be placed on supporting those in the lower tail of the skill distribution to ensure more equitable opportunities for all.

Furthermore, since the demand for different skills can evolve rapidly, aligning the supply and demand of skills is not only a matter for Higher Education Institutions and education policy, but it should also involve broader industrial policy stakeholders. Industrial policy plays a crucial role in shaping the types of skills that are needed in the economy, and collaboration between education providers and industry can ensure a more dynamic response to labour market needs. By engaging industry stakeholders in the process of curriculum development, HEIs can help ensure that the skills being taught are those that are in the highest demand, thereby improving the employability of future graduates.

It is also important to recognize that skills are not static; they evolve in response to changes in technology, industry requirements, and broader socio-economic conditions. This highlights the importance of lifelong learning and continuous skills development as essential components of addressing graduate underemployment. Governments, educational institutions, and employers all have roles to play in fostering a culture of lifelong learning. Providing access to retraining and upskilling opportunities, especially for graduates who may have initially entered lower-skill roles, can help them transition to higher-value occupations as their careers progress.

Finally, while securing a graduate job can be rewarding, not all underemployment is necessarily involuntary. Jobs evolve over time, becoming more or less attractive to graduates across various dimensions—not all of which are purely monetary. For some graduates, taking up a position in a routine or service role might offer valuable experience, opportunities for growth, or other non-monetary benefits that are difficult to quantify. Therefore, policy responses to underemployment should be nuanced, recognizing that career paths are varied and that the concept of a "graduate job" is not universally static.

This paper has sought to address a small part of a larger research question, and as such necessarily leaves many questions unanswered, some of which suggest themselves as extensions or variants of the model explored here. First, while I provide estimates of the graduate skill distribution, I remain agnostic about the causes of its drift over time. Investigating the drivers of the changing graduate skill distribution, including selection into university education, is likely to be a fruitful avenue for future research. Factors such as changing access to higher education, demographic shifts, and economic incentives all potentially contribute to changes in the skill composition of graduates, and understanding these drivers is key to developing effective policy interventions.

Second, while the increasing presence of graduates in non-traditional roles is well documented, there is little research on how this presence influences these occupations. A larger potential pool of highly skilled applicants might facilitate technology adoption within these roles, contributing to the increase in skill prices estimated for these occupations in this paper. Moreover, the increased presence of graduates in these roles could lead to a transformation of the nature of the work itself, as employers adjust job responsibilities to better match the capabilities of their more highly skilled workforce. Understanding these dynamics could help policymakers and industry leaders maximize the benefits of a highly educated workforce, even in sectors not traditionally associated with graduate employment.

Lastly, this paper has focused on a narrow population—young university graduates. Extending the scope of the model to include other populations and sub-populations might provide interesting insights into the interplay of skill supply for different sections of society. I leave these and further questions for future research.

References

[1] Acemoglu, Daron, and David Autor. ”Skills, tasks and technologies: Implications for employment and earnings.” In Handbook of labor economics, vol. 4, pp. 1043-1171. Elsevier, 2011.

[2] Altonji, Joseph G., Peter Arcidiacono, and Arnaud Maurel. ”The analysis of field choice in college and graduate school: Determinants and wage effects.” In Handbook of the Economics of Education, vol. 5, pp. 305-396. Elsevier, 2016.

[3] Altonji, Joseph G., Lisa B. Kahn, and Jamin D. Speer. ”Trends in earnings differentials across college majors and the changing task composition of jobs.” American Economic Review 104, no. 5 (2014): 387-393.

[4] Andrews, Rodney J., Scott A. Imberman, Michael F. Lovenheim, and Kevin M. Stange. The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability. No. w30331. National Bureau of Economic Research, 2022.

[5] Autor, David, H. ”The “task approach” to labor markets: an overview.” Journal for Labour Market Research 46, no. 3 (2013): 185-199.

[6] Autor, David H., Frank Levy, and Richard J. Murnane. ”The skill content of recent technological change: An empirical exploration.” The Quarterly journal of economics 118, no. 4 (2003): 1279-1333.

[7] Autor, David H., and Michael J. Handel. ”Putting tasks to the test: Human capital, job tasks, and wages.” Journal of labor Economics 31, no. S1 (2013): S59-S96.

[8] Carneiro, Pedro, and Sokbae Lee. ”Trends in quality-adjusted skill premia in the United States, 1960–2000.” American Economic Review 101, no. 6 (2011): 2309-2349.

[9] Deming, David J. ”The growing importance of social skills in the labor market.” The quarterly journal of economics 132, no. 4 (2017): 1593-1640.

[10] Diegert, Paul. ”The Role of Skills and Sorting in Explaining Wage Inequality.” (2024).

[11] Firpo, S. ”Occupational tasks and changes in the wage structure.” (2011).

[12] Gill, Jeff, and Gary King. ”What to do when your Hessian is not invertible: Alternatives to model respecification in nonlinear estimation.” Sociological methods & research 33, no. 1 (2004): 54-87.

[13] Green, Francis, and Yu Zhu. "Overqualification, job dissatisfaction, and increasing dispersion in the returns to graduate education." Oxford economic papers 62, no. 4 (2010): 740-763.

[14] Goos, Maarten, Alan Manning, and Anna Salomons. "Explaining job polarization: Routine-biased technological change and offshoring." American economic review 104, no. 8 (2014): 2509-2526.

[15] Hamermesh, Daniel S., and Stephen G. Donald. "The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias." Journal of Econometrics 144, no. 2 (2008): 479-491.

[16] Hastings, Justine S., Christopher A. Neilson, and Seth D. Zimmerman. Are some degrees worth more than others? Evidence from college admission cutoffs in Chile. No. w19241. National Bureau of Economic Research, 2013.

[17] Hemelt, Steven W., Brad Hershbein, Shawn M. Martin, and Kevin M. Stange. College majors and skills: evidence from the universe of online job ads. No. w29605. National Bureau of Economic Research, 2021.

[18] Holmes, Craig, and Ken Mayhew. "The economics of higher education." Oxford review of economic policy 32, no. 4 (2016): 475-496.

[19] Hou, Shihang. "The Role of Uncertainty in Education Mismatch." (2023).

[20] Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. "Field of study, earnings, and self-selection." The Quarterly Journal of Economics 131, no. 3 (2016): 1057-1111.

[21] Leighton, Margaret, and Jamin D. Speer. "Labor market returns to college major specificity." European Economic Review 128 (2020): 103489.

[22] Lindley, Joanne, and Steven McIntosh. "Growth in within graduate wage inequality: The role of subjects, cognitive skill dispersion and occupational concentration." Labour Economics 37 (2015): 101-111.

[23] Lovenheim, Michael F., and Jonathan Smith. Returns to different postsecondary investments: Institution type, academic programs, and credentials. No. w29933. National Bureau of Economic Research, 2022.

[24] McFadden, Daniel, and Kenneth Train. "Mixed MNL models for discrete response." Journal of applied Econometrics 15, no. 5 (2000): 447-470.

[25] Office for National Statistics. (2024). Labour Force Survey. [data series]. 11th Release. UK Data Service. SN: 2000026, DOI: `http://doi.org/10.5255/UKDA-Series-2000026`

[26] O'Leary, Nigel, and Peter Sloane. "Too many graduates? An application of the Gottschalk–Hansen model to young British graduates between 2001–2010." Oxford Economic Papers 68, no. 4 (2016): 945-967.

[27] Roy, Andrew Donald. "Some thoughts on the distribution of earnings." Oxford economic papers 3, no. 2 (1951): 135-146.

[28] Roys, Nicolas, and Christopher Taber. "Skills prices, occupations and changes in the wage structure." (2016).

[29] Roys, Nicolas, and Christopher Taber. "Skill Prices, Occupations, and Changes in the Wage Structure for Low Skilled Men (Preliminary and Incomplete)." (2022).

[30] Shorrocks, Anthony F. "Inequality decomposition by factor components." Econometrica: Journal of the Econometric Society (1982): 193-211.

[31] Sorkin, Isaac. "Ranking firms using revealed preference." The quarterly journal of economics 133, no. 3 (2018): 1331-1393.

[32] Train, Kenneth. "Halton sequences for mixed logit." (2000).

[33] Train, Kenneth E. Discrete choice methods with simulation. Cambridge university press, 2009.

A Additional Tables & Results

A.1 Additional Tables

Table A1: Details of Model Fit

Log Wage	**Data Average**	**Simulation Average**	**Difference**	**% Difference**
Mean	2.537	2.520	-0.017	-0.672
Mean Professional	2.667	2.616	-0.051	-1.927
Mean Routine	2.275	2.343	0.068	2.974
Mean Service	2.133	2.222	0.090	4.198
Variance	0.152	0.165	0.013	8.844
Variance Professional	0.113	0.161	0.048	42.730
Variance Routine	0.106	0.112	0.006	5.373
Variance Service	0.091	0.085	-0.006	-6.726
P10	2.026	2.012	-0.014	-0.696
P50	2.557	2.500	-0.058	-2.253
P90	3.021	3.056	0.035	1.173
Occupation Shares				
Share Professional	0.619	0.615	-0.005	-0.766
Share Routine	0.166	0.168	0.002	1.183
Share Service	0.095	0.098	0.004	3.863
Share Non-employed	0.120	0.119	-0.001	-0.724

Note: Model simulations based on a sample of 1,000,000 graduates. *Source:* QLFS (2001-2019) and Baseline Model.

Table A2: Skill levels by occupation

	Mean				**Standard Deviation**			
	Professional	**Routine**	**Service**	**Non-employed**	**Professional**	**Routine**	**Service**	**Non-employed**
2001-2009	0.243	-0.026	-0.290	-1.054	0.922	0.903	0.895	0.826
2010-2019	0.038	-0.153	-0.407	-1.211	0.864	0.851	0.833	0.769
Δ	-0.205	-0.127	-0.117	-0.157	-0.058	-0.052	-0.062	-0.057

Note: Simulations based on a representative sample of 1,000,000 graduates. Means standardized to 0 and standard deviations to 1 in 2001-2009. *Source:* Baseline Model.

Table A3: Period Comparison

	Data				**Baseline Model**			
	2001-2009	**2010-2019**	**Difference**	**% Difference**	**2001-2009**	**2010-2019**	**Difference**	**% Difference**
Occupation Shares								
Professional	65.257	58.950	-6.307	-9.665	64.454	58.140	-6.313	-9.795
Routine	15.282	17.793	2.510	16.425	15.557	18.181	2.623	16.863
Service	7.923	10.827	2.904	36.653	8.379	11.413	3.034	36.210
Non-employed	11.537	12.430	0.893	7.739	11.610	12.265	0.656	5.648
Log Wages								
Mean	2.580	2.498	-0.082	-3.180	2.553	2.483	-0.070	-2.758
Mean Professional	2.703	2.635	-0.068	-2.520	2.640	2.588	-0.052	-1.960
Mean Routine	2.284	2.267	-0.017	-0.763	2.363	2.320	-0.044	-1.854
Mean Service	2.137	2.129	-0.008	-0.352	2.237	2.206	-0.031	-1.380

Note: Comparison of average occupation shares between two periods. Model simulations based on a sample of 1,000,000 graduates. *Source:* QLFS (2001-2019) and Baseline Model.

Table A4: Parameter Estimates I

Parameter	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\sigma}_1^2$	$\hat{\sigma}_2^2$	$\hat{\beta}_{\text{Female}}$	$\hat{\beta}_{\text{Exp}}$	$\hat{\beta}_{\text{Exp}^2}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\phi}_3$	$\hat{\rho}$
Estimate	3.5458	3.5433	0.0118	0.0111	-0.0165	0.0740	-0.0028	0.0896	0.0798	0.1252	4.5517
Std. Error	(0.0002)	(0.0002)	(0.0000)	(0.0000)	(0.0002)	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0002)	(0.0009)
Sample Size	60420										
Log Likelihood	34957.8681										

Note: Baseline model estimates. Numerical standard errors are in parentheses. *Source:* Baseline Model Estimates.

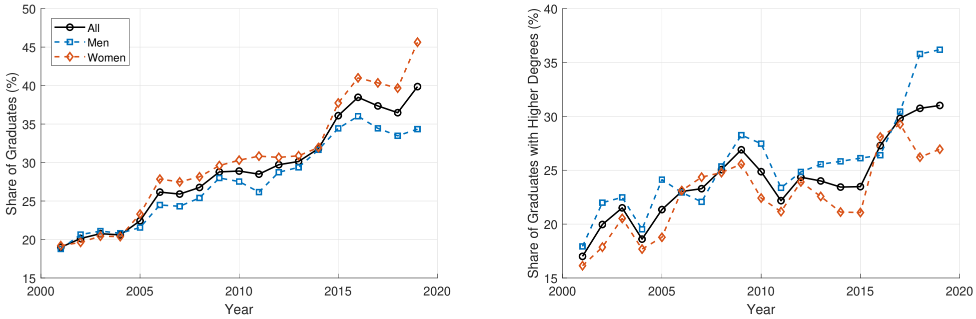
Table A5: Parameter Estimates II

Year	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\lambda}_3$	$\hat{\eta}_1$	$\hat{\eta}_2$	$\hat{\eta}_3$	$\hat{\omega}_1$	$\hat{\omega}_2$	$\hat{\omega}_3$
2001	1.0000	0.7749	0.8115	-32.5014	-24.8619	-26.1393	-1.8999	-2.0196	-2.1691
	(-)	(0.0009)	(0.0044)	(0.0010)	(0.0053)	(0.0259)	(0.0004)	(0.0005)	(0.0032)
2002	0.9156	0.7507	0.6876	-29.5196	-23.9770	-21.8903	-1.9595	-2.0911	-2.1846
	(0.0006)	(0.0006)	(0.0034)	(0.0033)	(0.0036)	(0.0198)	(0.0003)	(0.0003)	(0.0022)
2003	0.9207	0.7460	0.6578	-29.6786	-23.8126	-20.7906	-1.9813	-2.1331	-2.2335
	(0.0005)	(0.0004)	(0.0018)	(0.0031)	(0.0026)	(0.0108)	(0.0003)	(0.0003)	(0.0010)
2004	0.9497	0.7935	0.5687	-30.6654	-25.4293	-17.7067	-1.9768	-2.1268	-2.2044
	(0.0004)	(0.0005)	(0.0014)	(0.0024)	(0.0028)	(0.0080)	(0.0004)	(0.0003)	(0.0009)
2005	0.8976	0.7844	0.6583	-28.8321	-25.1283	-20.8184	-2.0351	-2.1151	-2.2369
	(0.0003)	(0.0003)	(0.0012)	(0.0019)	(0.0016)	(0.0072)	(0.0002)	(0.0002)	(0.0008)
2006	0.9573	0.8024	0.6512	-30.9208	-25.7190	-20.5852	-1.9829	-2.1221	-2.1144
	(0.0003)	(0.0004)	(0.0012)	(0.0017)	(0.0024)	(0.0072)	(0.0003)	(0.0003)	(0.0007)
2007	0.9638	0.8045	0.6032	-31.1332	-25.7951	-18.8818	-1.9934	-2.0921	-2.1833
	(0.0003)	(0.0004)	(0.0010)	(0.0017)	(0.0023)	(0.0058)	(0.0002)	(0.0003)	(0.0005)
2008	0.9786	0.7991	0.4485	-31.6476	-25.5964	-13.5159	-1.9326	-2.0622	-2.1573
	(0.0003)	(0.0004)	(0.0008)	(0.0016)	(0.0022)	(0.0048)	(0.0002)	(0.0002)	(0.0004)
2009	1.0699	0.8287	0.6009	-34.8325	-26.6349	-18.8114	-1.9453	-2.0879	-2.1375
	(0.0003)	(0.0004)	(0.0007)	(0.0015)	(0.0021)	(0.0041)	(0.0002)	(0.0002)	(0.0004)
2010	1.0857	0.9400	0.6557	-35.3849	-30.4925	-20.7494	-1.9436	-2.0785	-2.0960
	(0.0003)	(0.0004)	(0.0007)	(0.0014)	(0.0026)	(0.0042)	(0.0002)	(0.0003)	(0.0004)
2011	1.0273	0.9745	0.7081	-33.3643	-31.7409	-22.5693	-1.9589	-2.0219	-2.0488
	(0.0003)	(0.0006)	(0.0009)	(0.0014)	(0.0036)	(0.0051)	(0.0002)	(0.0004)	(0.0005)
2012	1.0809	0.9073	0.5884	-35.2336	-29.4363	-18.4396	-1.9602	-2.0231	-2.0429
	(0.0003)	(0.0004)	(0.0009)	(0.0015)	(0.0022)	(0.0054)	(0.0002)	(0.0002)	(0.0004)
2013	1.0506	0.9045	0.6786	-34.1797	-29.3116	-21.4944	-1.9505	-1.9865	-2.0994
	(0.0003)	(0.0003)	(0.0008)	(0.0014)	(0.0016)	(0.0045)	(0.0001)	(0.0002)	(0.0004)
2014	1.0084	0.9514	0.6531	-32.7196	-30.9262	-20.6456	-1.9211	-2.0053	-2.0708
	(0.0002)	(0.0003)	(0.0009)	(0.0013)	(0.0016)	(0.0050)	(0.0002)	(0.0002)	(0.0005)
2015	1.0810	0.9136	0.7719	-35.2100	-29.6193	-24.7397	-1.9312	-1.9604	-2.0653
	(0.0002)	(0.0004)	(0.0008)	(0.0011)	(0.0023)	(0.0044)	(0.0002)	(0.0002)	(0.0004)
2016	1.0604	0.9228	0.8139	-34.4827	-29.9018	-26.1755	-1.8644	-1.9271	-2.0156
	(0.0002)	(0.0003)	(0.0006)	(0.0013)	(0.0019)	(0.0037)	(0.0002)	(0.0002)	(0.0003)
2017	1.0987	0.9483	0.7661	-35.8103	-30.7329	-24.5304	-1.8892	-2.0388	-2.0951
	(0.0003)	(0.0003)	(0.0007)	(0.0020)	(0.0018)	(0.0044)	(0.0002)	(0.0002)	(0.0004)
2018	1.1766	0.9580	0.9522	-38.5140	-31.0786	-30.9313	-1.8987	-2.0003	-2.0947
	(0.0004)	(0.0003)	(0.0007)	(0.0024)	(0.0020)	(0.0042)	(0.0002)	(0.0002)	(0.0004)
2019	1.3945	1.0237	0.8398	-46.0360	-33.3248	-26.9993	-1.8184	-1.9484	-2.0162
	(0.0004)	(0.0007)	(0.0008)	(0.0024)	(0.0038)	(0.0049)	(0.0003)	(0.0004)	(0.0004)

Note: Baseline model estimates. Numerical standard errors are in parentheses. $\hat{\lambda}_1^{2001}$ normalised to 1. *Source:* Baseline Model Estimates.

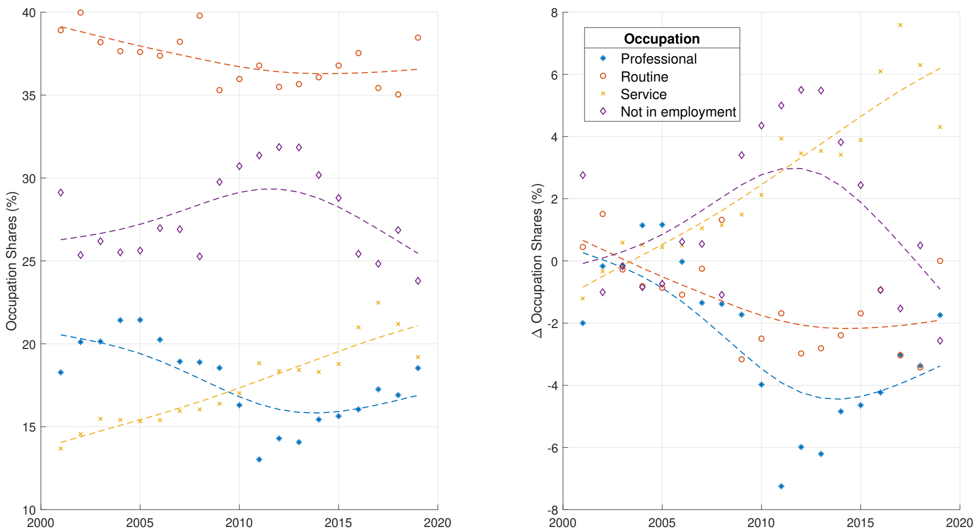
A.2 Additional Figures

Figure A1: Trends in higher education attainment among young adults in the UK (2001–2019)



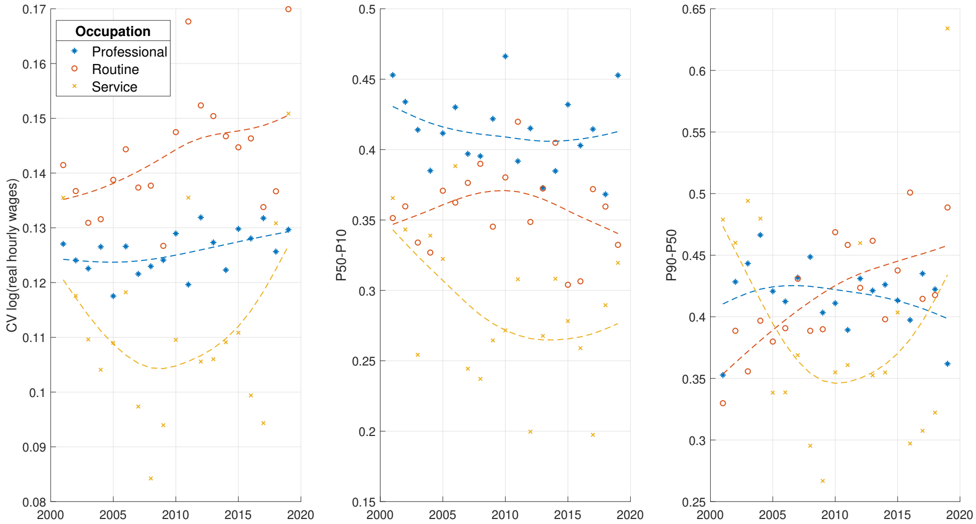
Note: Subplot 1: Share of young adults with a university degree, aged 21-30 years. Subplot 2: Share of graduates with a higher degree, aged 21-30 years. *Source:* Quarterly Labour Force Survey (2001-2019).

Figure A2: Trends in occupation shares of young adults without a degree in the UK (2001–2019)



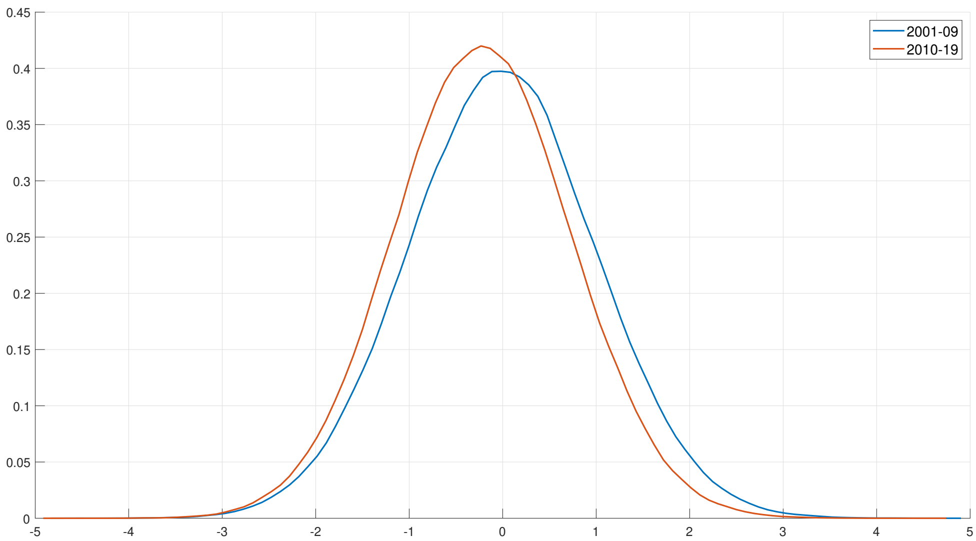
Note: Adults without a university degree, aged 21-30 years; change relative to 2001-05 average. Broken lines are HP-filtered trends (smoothing parameter = 100). SOC 2000 1-Digit Occupation Classification. Professional includes codes 1-3; Routine includes codes 4,5,8,9; Service includes codes 6,7; Not in employment includes unemployed and those out of the labour force for other reasons. *Source:* Quarterly Labour Force Survey (2001-2019).

Figure A3: Trends in real wages among young graduates in the UK (2001–2019)



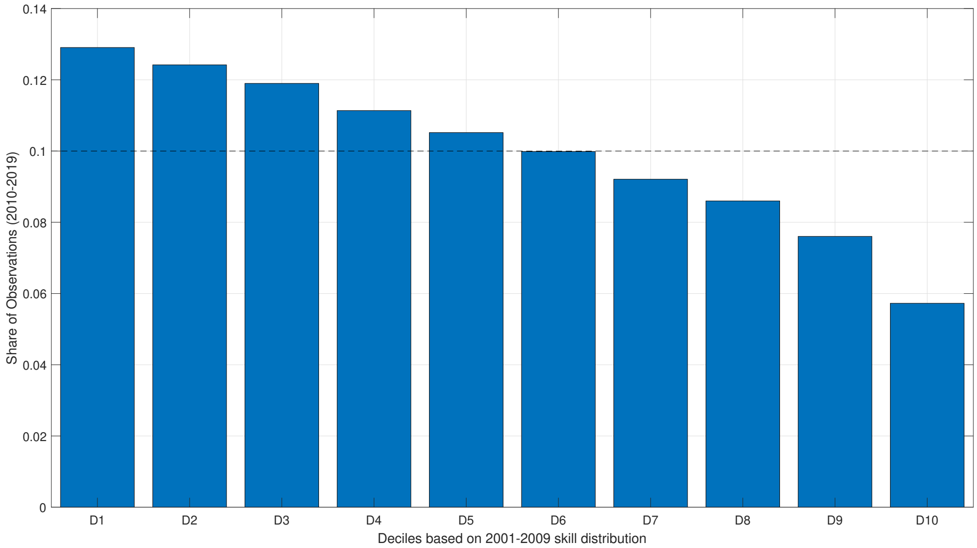
Note: Wages are log hourly wages, deflated by 2014 CPI. Working graduates aged 21-30 years. Broken lines represent HP-filtered trends (smoothing parameter = 100). Change relative to 2001-05 average.
Source: Quarterly Labour Force Survey (2001-2019).

Figure A4: Distribution of graduate skill



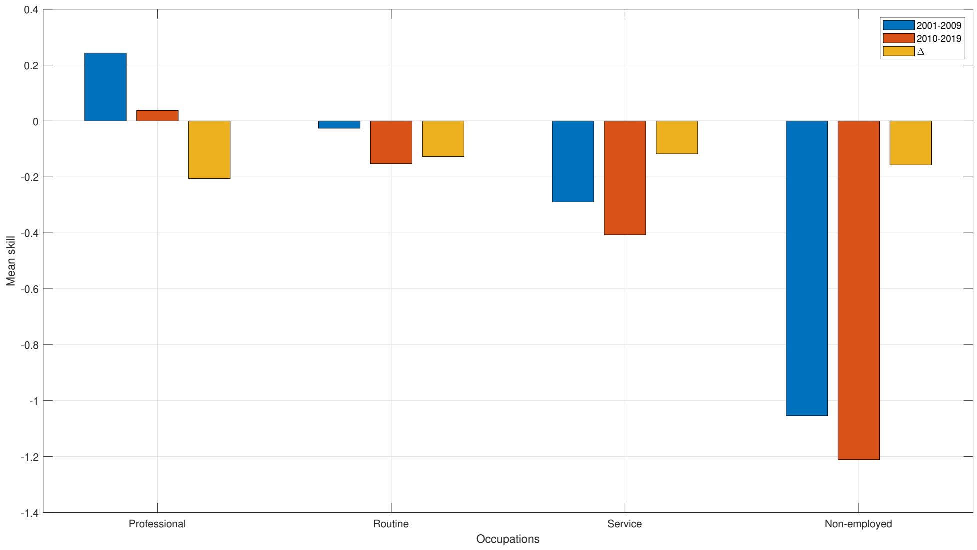
Note: Based on a simulated sample of 1,000,000 graduates. Distributions normalized to have mean 0 and variance 1 in 2001-2009. *Source:* Baseline Model.

Figure A5: Histogram of 2010-2019 skill distribution



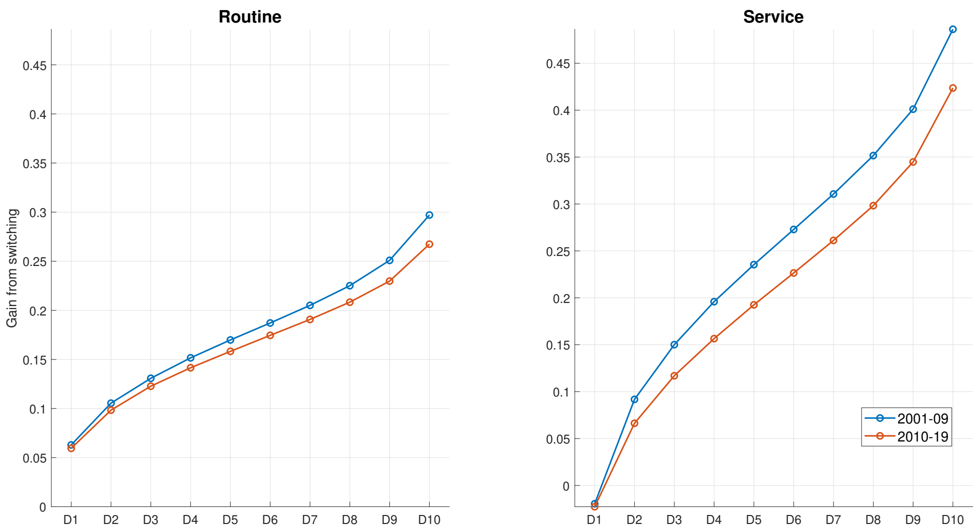
Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles based on 2001-2009 distribution. *Source:* Baseline Model.

Figure A6: Average skill levels by occupation



Note: Based on a simulated sample of 1,000,000 graduates. Distributions normalized to have mean 0 and variance 1 in 2001-2009. *Source:* Baseline Model.

Figure A7: Wage gains from switching to professional occupations



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles are based on the respective period.
Source: Baseline Model.

B Technical Details

B.1 Imputation of missing wage values

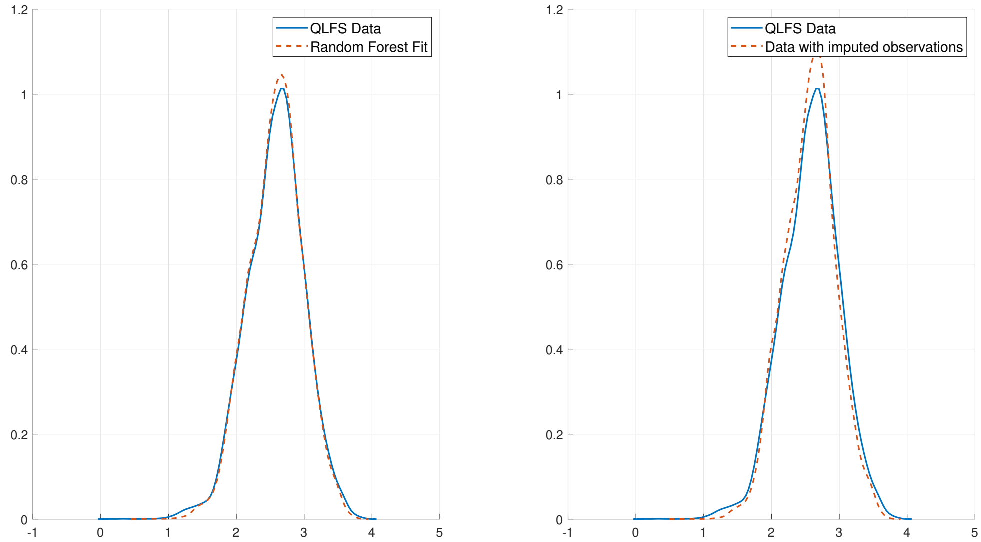
In the QLFS, wage data is only collected during the first and last interviews, resulting in three missing wage observations for each individual. To impute these missing wage values, I employ a two-step methodology that leverages both fixed effects to account for individual-specific skills and a Random Forest model to predict the missing wages.

A key problem for the imputation of wage data based on observable characteristics is that these methods do not account for unobservable heterogeneity among individuals, which is key in a panel setting. To account for individual-specific abilities, I incorporate individual fixed effects into the imputation procedure. This is done by specifying a linear fixed-effects regression where the hourly pay is regressed on variables such as experience, its square (to capture non-linear effects), year, quarter, sex, occupation, and government office region, including a fixed effect for each individual. This model helps us extract individual-specific effects which are then used in the second stage of the imputation procedure. For the imputation phase, I use a Random Forest regression model. The predictors include the fixed effects extracted from the fixed-effects model along with experience, year, quarter, sex, occupation, and government office region.

The Random Forest model is configured with 200 trees, allowing for robust predictions by averaging over multiple decision trees to reduce overfitting. I ensure that the model is fine-tuned by setting the maximum number of splits to the number of observations in the training data minus one and a minimum leaf size of one. This configuration helps in capturing the complex relationships within the data. The predicted wages are then integrated back into the dataset, replacing the missing values.

Figure B1 below showcases the results of the imputation procedure. The first panel depicts the observed wages, as well as the predictions based on the random forest model. The close alignment between the actual data and the predicted values indicates a good fit of the model to the observed values. This suggests that the model captures the underlying patterns in the wage data effectively, validating the robustness of the imputation methodology. The second panel shows the distribution of hourly wages with and without imputed values. The overall shape of the distribution remains consistent, demonstrating that the imputed values align well with the observed data distribution. This suggests, that the imputed values are not arbitrary but rather grounded in the underlying data patterns, preserving the integrity of the dataset.

Figure B1: Distribution of observed wages and imputed values



Note: Kernel density estimates of CPI (2014) deflated log hourly wages. *Source:* QLFS (2001-2019) and authors' calculations.

B.2 Estimation algorithm

The estimation procedure is a simple application of simulated maximum likelihood. In maximum likelihood, we find a vector of parameters so that the model maximizes the probability of observing the actual outcome.

The only complication, that arises here comes from the fact, that we do not have a closed-form solution for the joint probability (9) and thus have to evaluate the integral via simulation. This can be done by taking draws from the distribution of s, evaluating $\Pr(o_i^*, w_i^{obs}|s_i)$ at each of these draws and then averaging over the results. Since we are dealing with a panel of graduates, $\Pr(o_i^*, w_i^*|s_i)$ here denotes the conditional probability of a sequence of 5 occupation choices o_i^* and observed wages w_i^*. Standard results suggest, that as long as one uses a large enough number of draws to approximate the integral, the Maximum Simulated Likelihood Estimation (MSLE) is asymptotically equivalent to classical Maximum Likelihood Estimation (MLE) (c.f. McFadden & Train (2000)).

Denote the simulated counterpart of (9) by $\Pr^{sim}(o_i^*, w_i^*)$ for simplicity, and let θ be the set of our parameters, we can write down the simulated log-likelihood function of the as:

$$ll^{sim}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\sum_{o=1}^{O}\mathbf{1}_{(o=o^*)}\ln \overset{sim}{\Pr}(o_i^*, w_i^*) \tag{B1}$$

and we can estimate θ as:

$$\hat{\theta} = \arg\max_{\theta} ll^{sim}(\theta) \tag{B2}$$

So to specify the complete algorithm:

1. Set $p = 1$ and make a guess for $\hat{\theta}_1$. Specify a tolerance criterion ϵ. Set R, the number of draws used to approximate the integral to a reasonably high number.

2. For each individual i, given $\hat{\theta}_p$ draw a vector of s_i, R times, denoting each as s_i^r.

3. For $r = 1$ to R :

   (a) For $q = 1$ to 5 : Calculate $\nu_{iq}^r = w_{iq}^* - (\eta_{o^*} + \lambda_{o^*}^\tau s_i^r + \beta x_i)$. For a given pair s_i^r, ν_{iq}^r calculate $\Pr_q^r(o_{iq}^*, w_{iq}^* | s_i^r)$.

   (b) Calculate $\Pr^r(o_i^*, w_i^* | s_i^r)$ as $\prod_{q=1}^{5} \Pr_q^r(o_{iq}^*, w_{iq}^* | s_i^r)$

4. Average over all R values of $\Pr^r(o_i^*, w_i^* | s_i^r)$ to obtain:
   $\Pr^{sim}(o_i^*, w_i^*) = \frac{1}{R}\sum_{r=1}^{R} \Pr^r(o_i^*, w_i^* | s_i^r)$.

5. Repeat steps $2 - 4$ for all N individuals. Calculate the log-likelihood via (B1) denoting it as ll_p^{sim}.

6. If $|ll_p^{sim} - ll_{p-1}^{sim}| < \epsilon$, terminate here. Otherwise, increment p and find a new value $\hat{\theta}_p$ and repeat from step 2.

For the numerical evaluation of the integral, I use a grid of $1,000$ quasi-random Halton draws, which have been shown to provide about an order of magnitude more accuracy than simple random draws (Train (2000, 2009)). To ensure stochastic equicontinuity I use the same set of points for each agent at each iteration. For updating $\hat{\theta}_p$ in step 6, I use Matlab's `fminunc` routine, using central numerical derivatives and critical values of $1e^{-6}$.

I calculate numerical standard errors following the well-known (c.f. Train (2009)) relationship between the hessian of the likelihood function and the information identity: For the correctly specified model, the error of the MLE estimate $\hat{\theta}$ is distributed according to:

$$\sqrt{N}(\hat{\theta} - \theta^*) \rightarrow N(0, -\mathbf{H}^{-1}) \tag{B3}$$

where θ^* is the true parameter vector, and $-\mathbf{H}$ is the information matrix. To avoid complications due to the numerical procedure and the high dimensionality of the problem, I calculate a numerical hessian of the likelihood function at the SMLE estimate and then use a pseudo-inverse (c.f. Gill & King (2004)) to obtain the standard errors for the estimated parameters.

Table B1: Summary of Model Parameters

Parameter	Description	Number of Parameters
μ_c	Location parameter of the cohort-specific skill distribution.	2
σ_c	Scale parameter of the cohort-specific skill distribution.	2
η_{ot}	Occupation-year specific fixed effect.	57
λ_{ot}	Occupation-year specific return to skill.	56
ω_{ot}	Occupation-year specific occupation preferences.	57
β	Gender coefficient, experience and $experience^2$ coefficients.	3
ρ_o	Scaler of idiosyncratic preference shock.	1
ϕ	Standard deviation of log wage measurement errors.	3

B.2.1 The cluster refinement global optimization algorithm

The likelihood function generated by this problem is smooth, but not globally concave which makes it difficult for gradient-based optimization routines that are prone to converge to local minima. This is a general problem for the class of discrete choice models, but especially here given the high dimensionality of the parameter space. To maximize the log-likelihood function, I therefore develop a novel global optimization algorithm that utilizes machine learning to effectively search through the high dimensional parameter space. The algorithm proceeds as follows:

1. Define a grid of initial starting points G that span the parameter space θ. For each point in G evaluate the log-likelihood function.
2. Discard points where the log-likelihood is below a certain threshold criterion.
3. Use a clustering algorithm to cluster the remaining points into K clusters.
4. From each cluster select a set of points P^k. The selection can either be the point with the best log-likelihood value in the cluster, or a weighted average of all points in the cluster or both.
5. Use a local solver starting at each point in P^k to maximize the log-likelihood function using a nonlinear optimization routine. Repeat steps 2-5 as required.

The main idea behind the algorithm is that the clustering algorithm will group points that are *similar* together. Points that are close together in the parameter space are likely in the neighbourhood of the same local maximum, so it is unnecessary to run local solvers from each of these points. The computational savings can be used to explore further regions of the parameter space.

For further refinement, steps 2-5 can be repeated using the local maxima found by the nonlinear solvers in step 5 and so on. Using this method it is practical to start with a large number of clusters in the beginning and reduce this number in each successive iteration. In doing so, it is advised to initially set the convergence criteria to relatively high values or limit the number of iterations for the local solvers in the beginning and tighten the criteria over successive iterations.

For the estimation, I begin with a grid of 100,000 points where the likelihood function is evaluated once on each point. I then run 2 iterations of the cluster refinement algorithm on the set of these points, using a cluster number of 50 in the first round and 10 in the second round. In each round, I select two points from each cluster, namely the point with the best likelihood value and the cluster's weighted midpoint. For the final run, I select the point with the best likelihood value.

B.3 Alternative interpretation of ϕ

In the main part of this paper, I introduced ϕ as the standard deviation of an idiosyncratic shock to the graduates' wage that was assumed to be independent of the graduates' occupation choice. In this subsection, I want to quickly outline an alternative interpretation of ϕ that doesn't rely on the structural interpretation and therefore might be easier to accept for some readers.

To illustrate let us return to the joint probability (9):

$$\Pr(o_i^*, w_i^*) = \int \Pr(o_i^*|s_i) \Pr(w_i^*|s_i, o_i^*) f(s) d(s).$$

Note that this formulation shows that the estimation is essentially trying to match two conditional probabilities: i) the conditional probability of choosing occupation o_i^*; ii) the conditional probability of the observed distribution of wages w_i^*. The hope is that if the model is flexible enough (i.e. has enough free parameters) there will be no conflict between these two objectives: the same parameter vector θ^* that maximizes the joint probability also maximizes the individual conditional probabilities. However in reality we might not be close to θ^* and particularly during the estimation the estimator will encounter points where trade-offs have to be made between the two counteracting objectives. In other words, the estimator needs to have an exchange rate to trade off better fit on one dimension against worse fit on another.

By looking at the way that ϕ enters the likelihood function to see that it provides an implicit weight for making this trade-off: $\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}$.

If ϕ is small then values of ν_i away from 0 will lead to large losses in terms of likelihood. In other words, there is a high priority on matching the wage distribution, even at the expense of the occupation distribution. If ϕ is large, then the estimator is more forgiving with respect to large deviations from the observed wage and puts relatively more weight on matching the conditional occupation choice probabilities. In this interpretation, ϕ

is simply a tuning parameter that helps us find the right balance among our different objectives.