Skills, Tasks and Degrees

Max Schroeder*

April 14, 2025

latest version

Abstract

An increasing number of young university graduates are not entering professional occupations, raising concerns about a decline in their skills compared to earlier cohorts. This perspective may overlook evolving occupational returns to skill and job amenities. I develop an economic model featuring heterogeneous skill supply, differentiated returns to skill across occupations, and changing preferences over non-wage job characteristics. Estimating the model using UK data from 2001–2019, I find a significant decline in average graduate skill levels - about 18% of a standard deviation - which explains 58% of reduced professional employment. Additionally, changes in amenity values push graduates into routine and service occupations.

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 Durham, 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, like 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 & McIntosh, 2015, Altonji, Arcidiacono, 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 costly investment made by 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. Over the past decades, there have been significant changes in the demand for the skills of young graduates, 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 choices. 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 approach - framing the question as a latent variable problem: skills are unobserved but

[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.

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.

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 estimated model 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 18% of a standard deviation, with a particularly strong increase in the share of low-skilled graduates. This had a direct effect on the share of young graduates entering professional occupations, explaining around 58% 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 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, Arcidiacono, 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, Levy, 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, Kahn, 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.[2] 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 and 30 years.[3] 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.[4]

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 as non-employment.[5] 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

[2]In this respect the QLFS is similar to the CPS, albeit with a quarterly frequency.

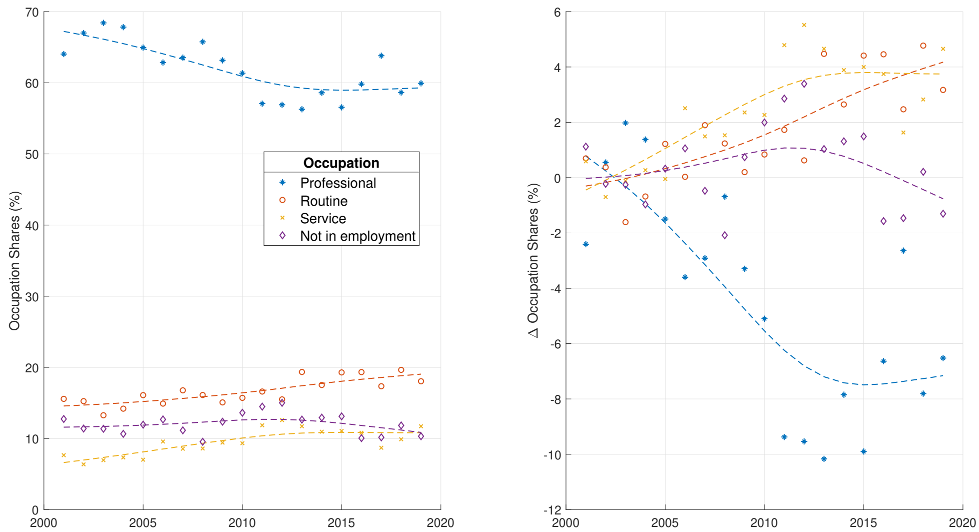
[3]Typically in the UK students finish high school at 18 and enter 3-year University Courses.

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

[5]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.

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 4pp 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 4pp 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.

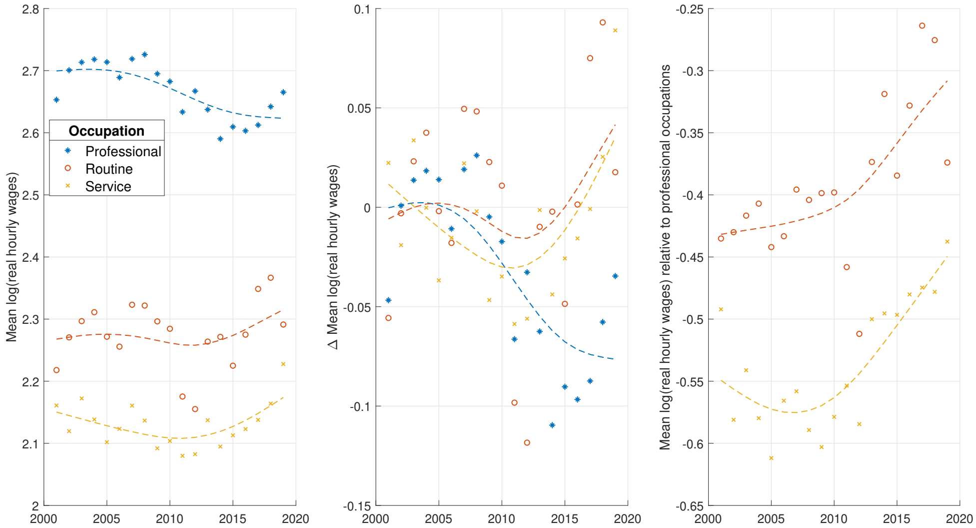
Figure 1: Trends in occupation shares of young graduates in the UK (2001–2019)



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 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 2000s 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.

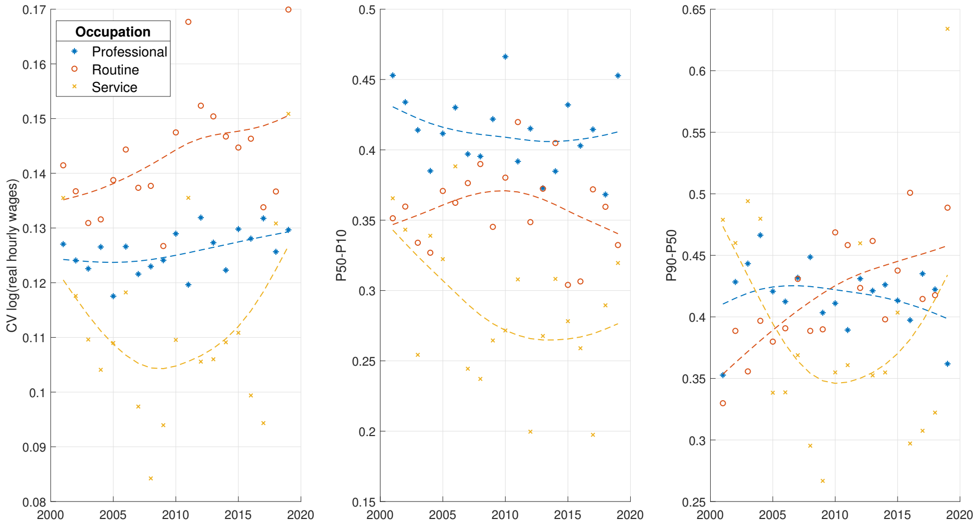
Figure 2: Trends in real wages among young graduates in the UK (2001–2019)



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).

Apart from changes in relative wages between different occupations, there also have been substantial movements in the distribution of wages within occupations. Figure 3 plots three measures of wage inequality for the three occupation groups, namely the coefficient of variation of log hourly wages, and the P50-P10 and P90-P50 measures of lower and upper tail inequality. Over the time period in question, overall wage inequality has risen slightly for professional occupations, and quite strongly for routine occupations. Service occupations, on the other hand, appear to exhibit a strong U-shaped trend, where wage inequality initially falls until around 2010 after which it begins to rise again. These trends in overall wage inequality are related to changes in the lower and upper tail measures of wage inequality. Lower tail inequality, as measured by the P50-P10 measure, trends downwards for all occupations over the sample period, although there is a distinct hump-shaped pattern for routine occupations. This suggests that inequality at the lower end of the within-occupation wage distribution is broadly falling over the sample period, even though this trend is not very pronounced among professional and routine occupations. Looking at changes in upper tail inequality, there is a slight downward trend among professionals, but a pronounced upward trend among routine workers. Service workers see a strong U-shaped pattern that somewhat mirrors the picture already documented for the coefficient of variation.

Figure 3: Trends in wage inequality 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 multinomial choice literature (see K. E. 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 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 4 below.

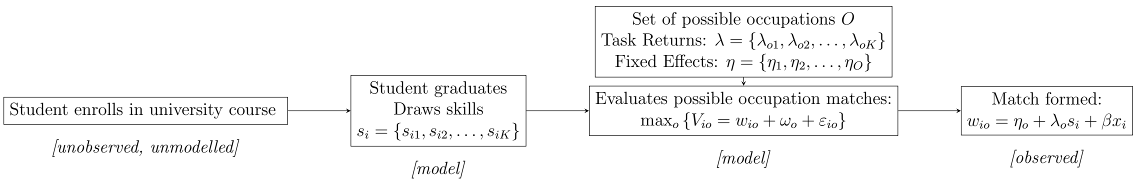


Figure 4: Model Timeline

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.[6]

A graduate's productivity therefore depends on her skill-set s_i as well as the task-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 of the graduate and β is a vector of coefficients.[7] This setup is common in the literature on tasks and skills (c.f. Autor & Handel, 2013, Roys & Taber, 2016).

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

[7] In principle the unobserved skill vector s_i should contain all relevant information for the determination of the graduates' wage, but in order to ensure robustness in the empirical application I further control for the sex of the graduate, a second order polynomial of their labour market experience and whether they have an advanced degree.

After graduation, graduates observe their own 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 and other occupation characteristics. This is likely to be the case for an economic agent with a suitably defined utility function (e.g. logarithmic), who is borrowing constrained.[8] This leads to the following potential valuation:

$$V_{io} = w_{io} + \omega_o + \frac{1}{\rho} * \varepsilon_{io}, \tag{2}$$

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.[9] $\rho > 0$ is a scalar that governs the relative importance of the idiosyncratic preference shock in the agent's valuation. 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. For the remainder of this paper, I will denote the observed occupation choice and wage as o_i^* and w_i^* respectively, while their model implied counterparts are denoted without the asterisk.

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). 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

[8] I believe it reasonable to assume that this situation applies to the sample population studied in this paper.

[9] 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.

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 in the previous section. 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 task ($K = 1$), 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.[10] Each cohort of graduates then corresponds to a cohort-specific skill distribution S_c. Without loss of generality, assume that graduates' skills can take values in the set $s_i \in [0, 1]$. I make no further assumptions about the distribution of S_c, instead, I rely on a novel identification result by Diegert (2024) to identify S_c non-parametrically.

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 final dataset contains 12,084 graduates, totalling 60,420 observations.[11] The QLFS records

[10]The cut-off year is 2010.

[11]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.

wages only at the first and last interview, so I impute missing wage observations. The details of the imputation are described in Appendix B.

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 $\frac{\pi^2}{6\rho^2}$, 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 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:

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

Since I 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:

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

Finally, integrating over the distribution of s leads to the unconditional joint probability:[12]

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

Standard results (c.f. McFadden & K. 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 (see also Heckman et al., 2006 or

[12]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 100 pseudo-random Halton draws.

Aucejo & James, 2021).

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.[13] This allows me to identify the wage equation up to a scalar. I introduce a normalization by fixing the range of s_i to be between 0 and 1. This normalization enables the identification of the remaining parameters relative to this benchmark. Because I further observe the same individual in different time periods τ, I 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 inDiegert, 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) and preferences, not

[13] 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.

on the time-varying productivity shocks. Finally, since $K = 1$ and $T = 5$, I satisfy the condition $K < 2T$, which is necessary for identification.

4.3 Estimation and Model Fit

For estimation, I approximate the distribution of skills S_c using a histogram.[14] I separate the unit interval into $L = 10$ equally spaced bins, each with an associated probability p_{lc}. I assume that graduates draw a bin with the associated probability and then draw a skill s_i from the range of values within that bin with uniform probability.[15] I maximize the likelihood function using a custom global multi-start algorithm, described in Appendix B.

After estimating the model, I evaluate its ability to capture graduates' occupation choices and wage outcomes. For this purpose, I simulate a representative sample of graduates. Figure 5 below highlights the model fit with respect to the occupation choices of graduates while Figure 6 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.[16] 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 6 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 0.04 in log wages on average. In contrast, the model slightly overpredicts wages in routine and service occupations by 0.06 and 0.1, respectively.

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.

[14] By using a histogram one can approximate any function arbitrarily closely, by increasing the number of bins L.

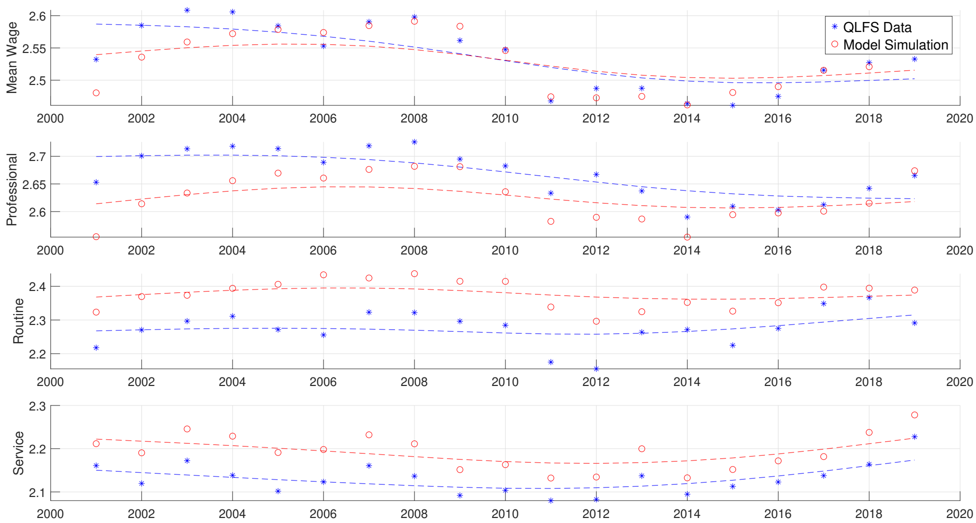
[15] This approach to approximating the skill distribution is based on Train's (2016) treatment of the distributions of latent taste parameters.

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

Figure 5: Model fit - occupation shares

Note: Broken lines represent HP-filtered trends with a smoothing parameter of 100, based on a simulated sample of graduates. *Source:* Baseline Model.

Figure 6: 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 100, based on a simulated sample of graduates. *Source:* Baseline Model.

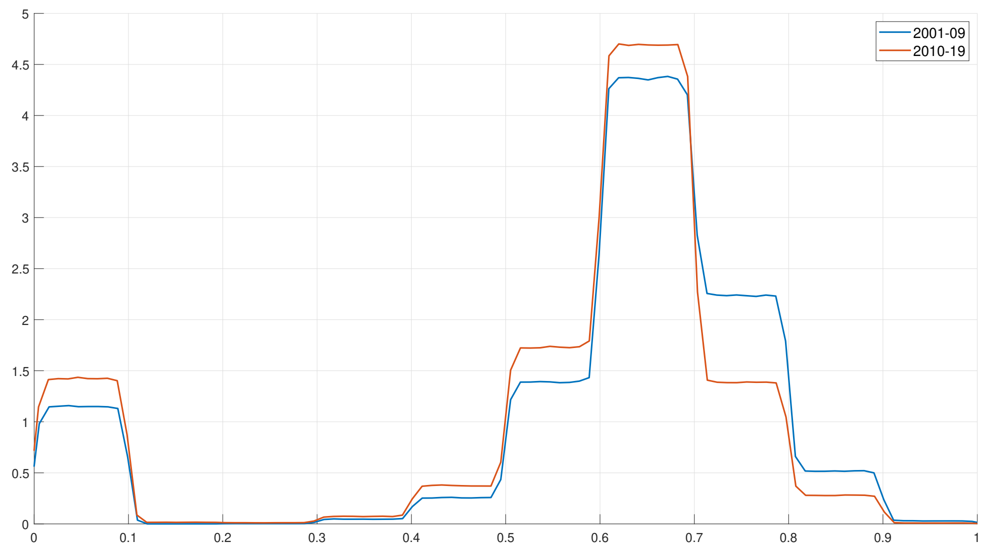
5 Results

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

Figure 7 shows the estimated skill distributions for graduates of the 2001-2009 and 2010-2019 cohorts.[17] Across cohorts, the distribution of skills shows two notable features: i) The skill distributions appear to be approximately left skewed, with the majority of the probability mass between 0.5 and 0.8. ii) There is an excess mass at the very bottom of the skill distribution. The concentration of graduates among the middle and upper-middle deciles of the distribution of skills, suggests a reasonably homogenous concentration of skill amongst university graduates. Further, the elevated share of graduates in the first decile is indicative of a small number of individuals who somehow fail to acquire skills on par with the majority of their contemporaries. A comparison of the two cohorts suggests that between 2001-09 and 2011-19, the distribution of skills experienced a shift downwards, with a strong fall in the share of graduates in the top 3 skill deciles. Conversely the share of graduates in the middle third of the skill distribution has increased as well as a considerable increase in the share of graduates in the first decile of the distribution.

Figure 7: Graduate Skill Distributions



Note: Kernel density estimates of cohort-specific skill distributions. *Source:* Baseline Model.

Panel A of table 1 allows a more detailed comparison of the 2001-2009 and 2010-2019 cohorts. 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 6.5% (3.1%) points, equivalent to 18% (9%) of a standard deviation among the first cohort—a moderate but not insignificant reduction. Anecdotally, this shift

[17]See also Panel B of table 1.

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, Lindley & McIntosh, 2015).

Concurrent with the reduction of average skill level the degree of skill inequality between graduates increased between the two cohorts. Considering measures of overall variation, the coefficient of variation and the gini index increased by 11% and 13% respectively. However, these changes are most notable at the bottom of the skill distribution. While the mean-to-median ratio indicates a slight reduction of inequality, the p50-to-p10 ratio suggests a large increase in inequality at the lower tail of the distribution, which is not mirrored by changes in the p90-to-p50 ratio which remains quite stable. Other percentile measures support this, as the changes are heavily skewed towards the lower-end of the skill distribution. While the 90th percentile of the 2011-19 cohort is only 4% below the previous cohort, the 10th percentile has fallen by a fifth. These distributional observations lend further credence to the theory that the main changes to the skill distribution are due to the addition of additional graduates at the left tail of the distribution.

Table 1: Details of Graduate Skill Distribution

Panel A: Moments of the graduate skill distribution

Period	*Mean*	*Median*	$\frac{Mean}{Median}$	*Std.Dev.*	*Coef.Var.*	*Gini*	$P10$	$P25$	$P50$	$P75$	$P90$	$\frac{P50}{P10}$	$\frac{P90}{P50}$
2001-2009	0.594	0.649	0.915	0.216	0.363	0.175	0.087	0.575	0.649	0.713	0.780	7.432	1.201
2010-2019	0.556	0.629	0.883	0.224	0.404	0.198	0.070	0.535	0.629	0.682	0.749	8.979	1.190
Δ	-0.038	-0.020	-0.032	0.009	0.041	0.023	-0.017	-0.040	-0.020	-0.030	-0.031	1.547	-0.011
Δ%	-6.454	-3.093	-3.468	4.149	11.334	13.298	-19.793	-7.040	-3.093	-4.252	-3.992	20.821	-0.927

Panel B: Histogram of the graduate skill distribution

Period	[0–0.1]	[0.1–0.2]	[0.2–0.3]	[0.3–0.4]	[0.4–0.5]	[0.5–0.6]	[0.6–0.7]	[0.7–0.8]	[0.8–0.9]	[0.9–1]
2001-2009	0.115	0.000	0.000	0.005	0.026	0.139	0.437	0.224	0.052	0.003
2010-2019	0.143	0.002	0.001	0.007	0.038	0.173	0.469	0.139	0.028	0.001
Δ	0.028	0.002	0.001	0.003	0.012	0.034	0.032	-0.085	-0.024	-0.002
Δ%	24.598	9510.460	214.761	54.724	46.162	24.151	7.367	-37.994	-45.944	-71.074

Note: Simulations based on a representative sample of graduates. *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.

5.2 Returns to skill and occupational amenities

Figure 8, Subplot 1, provides the model's estimates for the returns to skill for the different occupations. In line with expectations, professional occupations provide the highest return to skills, followed by routine and then service occupations. Overall the return to skill in routine occupations is much closer to professional returns while for service occupations the model estimates close to zero returns over the majority of the sample. Negligible returns among service occupations suggest that skilled graduates do not have a particular advantage in these occupations which might be in line with our expectations given the type of work in these occupational fields.[18] 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.

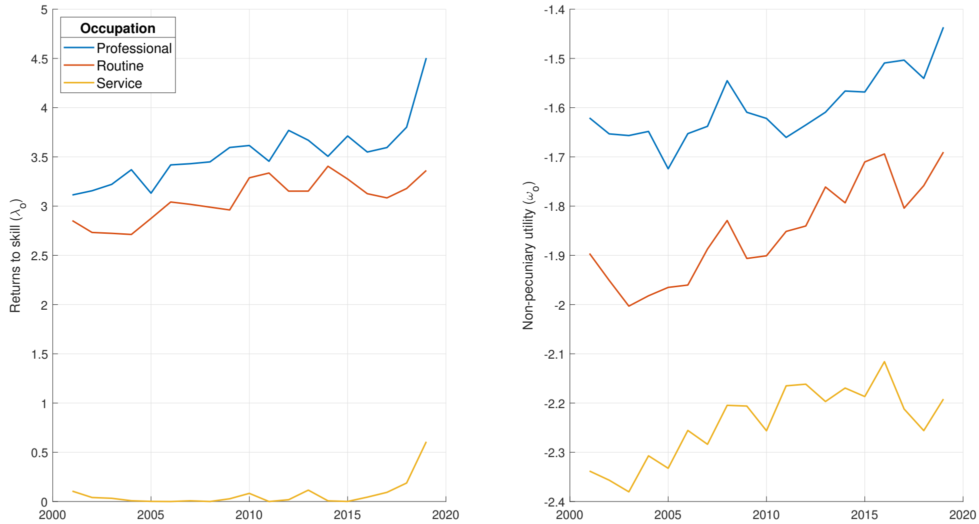
Over the 19-year period, there is a pronounced increase in the return to skills across professional and routine occupations, consistent with the theory of within-occupation task upgrading. The trend appears to be stronger in the lead-up to the 2008 financial crisis and then flatten out afterwards. The growth of the return to skill between 2001 and 2010 is 15% for professional and 14% for routine occupations. In the years between 2010 and 2019, the return to skills in professional occupations increased by 22%, but this is mainly driven by a very high estimate for 2019. Excluding this likely outlier the increase is more modest at 5% - so about a third of the growth experienced in the previous decade. For routine occupations the trend is even negative, with a fall of 3.3% between 2010-19.

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 8, 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

[18] A flat skill-wage profile might also be a result of the wage-setting practices in the majority of service sector jobs, where wages are generally quite rigid, while performance is rewarded through other channels that are unmodelled in this exercise.

Figure 8: Returns to skill and non-pecuniary utility components by occupation



Note: Subplot 1: Returns to skill. Subplot 2: Non-pecuniary utility; value of non-employment normalized to 0.
Source: Baseline Model Estimates.

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 1.45 and 2.35 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 disamenity, followed by routine occupations, which 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. This is particularly true for routine occupations, where the average gap to professional occupations reduces from -0.29 in 2001-10 to -0.22 in 2011-19.

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. These explanations are broadly in line with the changes to the welfare system in the UK under the coalition government, such as the introduction of universal credit in 2010. Further, the experience of the financial crisis and the Great Recession would have made job security and stability more relevant to the decision-making of young graduates.

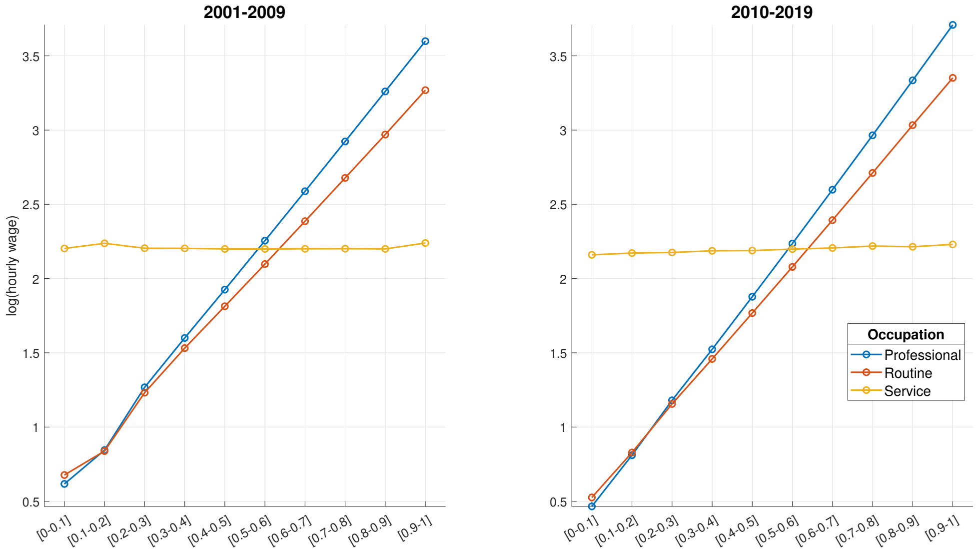
A reduction in the relative gap between professional and routine/service occupations hints at a relative appreciation of the utility associated with these occupations, possibly through improved working conditions such as better health and safety standards or more flexible working hours. It could also indicate a more general shift in graduates' 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.

5.3 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 9 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, and almost 0 for services. Below the mean of the skill distribution, service occupations pay higher wages than both professional and routine occupations, yet above this threshold and by extension for a majority of graduates, professional occupations pay comfortably more than services and slightly more than routine jobs.

Figure 10 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 form of employment. If they consider employment at all, they would likely favour service occupations where the flat wage profile offers at least some monetary compensation. However, as skill levels increase, the probability of being employed in a professional occupation increases rapidly. For example, a graduate with a skill level between 0.4 and 0.5 is almost 50 percentage points more likely to be employed in a professional occupation than a graduate with a skill level between 0 and 0.1, and for those in the highest skill bin, the probability is above 80%. For routine occupations, the general pattern is approximately hump-shaped. Initially,

Figure 9: Log hourly wage by skill deciles



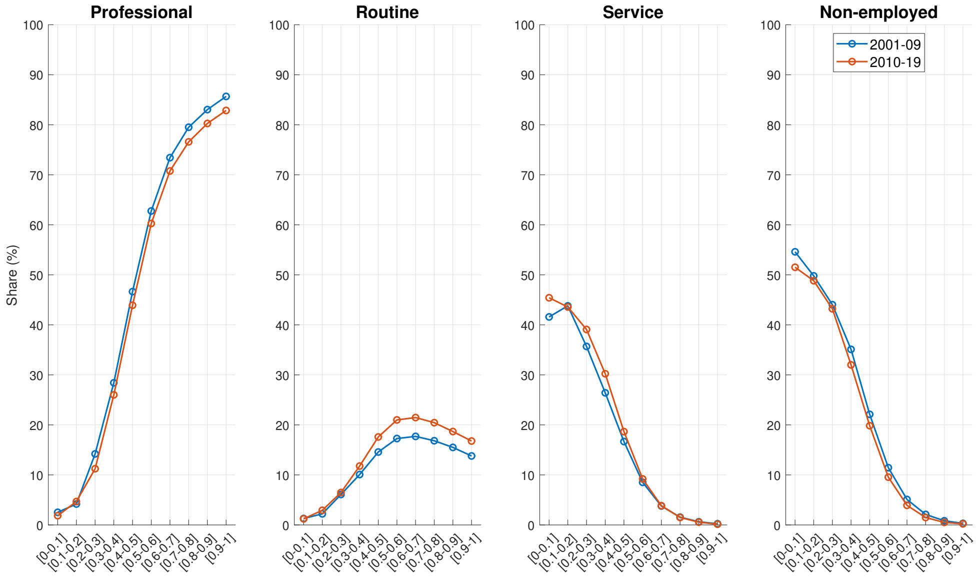
Note: Based on a simulated sample of graduates. *Source:* Baseline Model.

moving into higher-skill deciles graduates are more likely to be employed in routine jobs - primarily because they are more likely to be employed at all - but this trend peaks at skill levels around 0.5 to 0.7 and then slowly decreases, as higher-skilled graduates sort into professional jobs. Service jobs are most prominent among the set of graduates with skills ranging from 0 to 0.3, with their popularity veining as skill levels increase. In this respect, service occupations resemble an alternative to being out of the labour force for low-skilled graduates. 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 11 provides a more detailed assessment of the differences in sorting patterns across the two periods. As indicated in Figure 11, 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 -3% for the most skilled graduates to just over -0.5% among the lowest skilled.

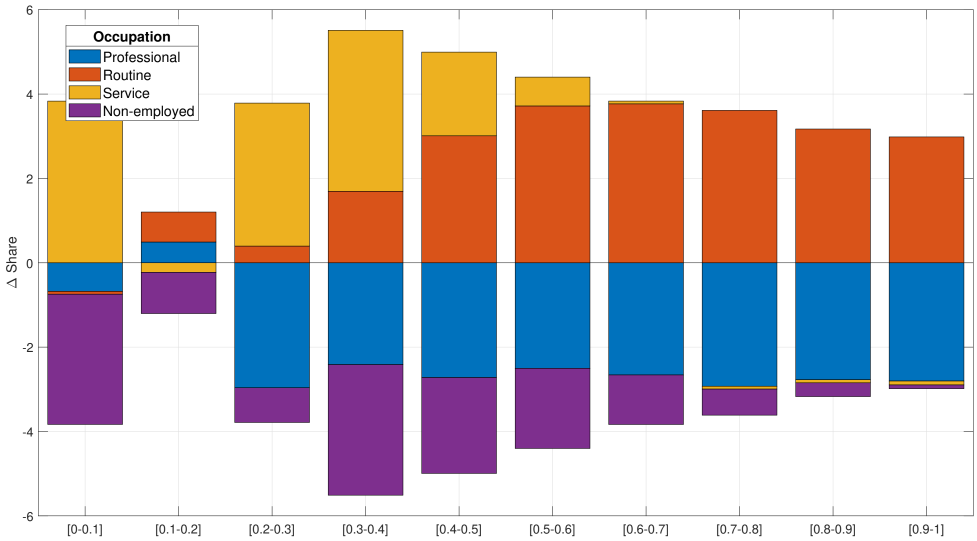
In this section, I have outlined the changes that have taken place in the labour market for young graduates. I have focussed on the changing skill distribution, the changing return structure for these skills and non-pecuniary aspects of different occupations and how these might contribute to the observed 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.

Figure 10: Occupational sorting by skill deciles



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

Figure 11: Changes in occupational sorting by skill deciles



Note: Changes in occupation shares, based on a simulated sample of graduates. *Source:* Baseline Model.

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 S_c; ii) the occupation-specific skill prices and returns to skill (represented by η_o and λ_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.[19] 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 my model, the effects of skills, preferences, and prices on labour market outcomes are not simply 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, returns* and *preferences* counterfactual 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.[20] 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 6pp 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 18% of a standard deviation, which we hypothesized to

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

[20]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**		**Returns Counterfactual**		**Preferences Counterfactual**	
Occupation Shares	**Total Δ**	**Δ**	**Explained (%)**	**Δ**	**Explained (%)**	**Δ**	**Explained (%)**
Professional	-5.984	-3.487	58.272	-0.115	1.926	-2.382	39.802
Routine	2.767	-0.437	-15.799	-0.207	-7.494	3.412	123.293
Service	2.682	1.788	66.665	-0.250	-9.337	1.144	42.672
Non-employed	0.535	2.137	399.275	0.573	107.077	-2.174	-406.352
Log Wages							
Mean	-0.069	-0.077	111.604	0.025	-35.635	-0.017	24.031
Mean Professional	-0.050	-0.078	155.232	0.031	-61.464	-0.003	6.232
Mean Routine	-0.051	-0.067	132.191	0.021	-41.223	-0.005	9.032
Mean Service	-0.026	0.000	1.620	-0.022	83.629	-0.004	14.751

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

have negatively impacted graduates' ability to secure professional jobs.

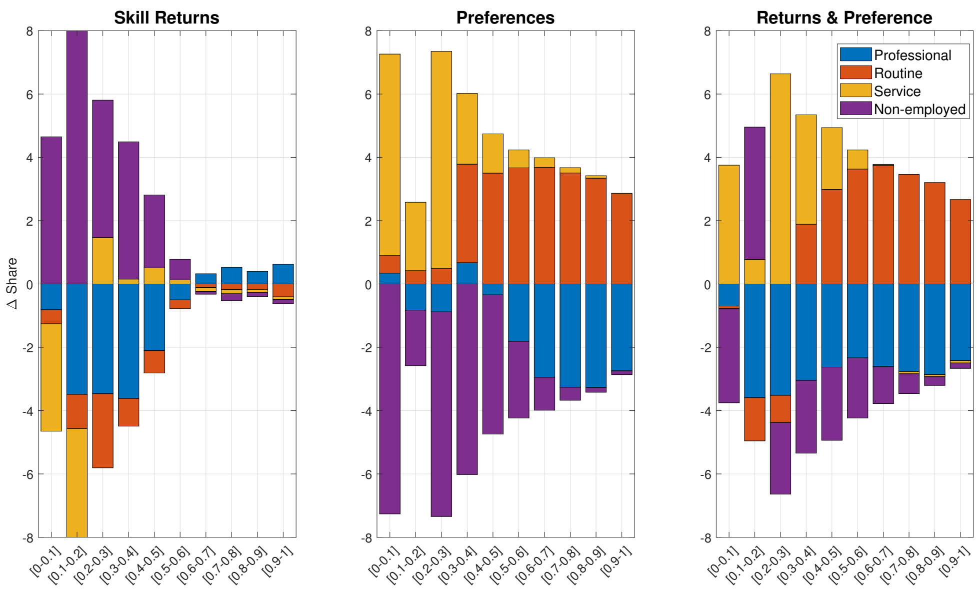
The simulation suggests that changes in the skill distribution explain around 58% of the decline in the share of professional occupations. If only skills had changed between these two periods, the share of graduates in professional roles would have decreased by roughly 3.5pp. As we have discussed in a previous section, this is due to the reduced likelihood of lower-skilled graduates entering professional occupations. As graduates become less skilled, they sort downward in the ranking of occupations into less skill-intensive ones. Since the drop in the probability of joining professional occupations is particularly steep at low skill levels, the large increase in the mass of graduates at the lower tail of the skill distribution exacerbates the reduction in the share of graduates in these occupations.

With respect to routine occupations one might expect a fall in the average skill level to increase the share of graduates in routine roles, but since the decline is driven to a large extend by the left tail, the counterfactual predicts a decline of roughly 0.4pp. Since this is contrary to what we observed in the data, this already indicates that another factor plays a role in explaining both, the fall in the share of professional occupations and the increase in routine jobs. Instead the falling share of professional graduates is absorbed by service occupations, explaining around 67% of the increase, and a large increase in the share of non-employed graduates. Again, while the increase in the share of service occupations is in line with expectations - lower skilled graduates have a higher propensity to enter service occupations - the large increase in non-employed graduates is out of line with the moderate increase observed in the data, suggesting additional forces are at play.

Considering the returns counterfactual, we notice that changes in the returns to skill between 2001 and 2019 are unlikely to reconcile the skills counterfactual with the data. The counterfactual suggests a small fall in employment across all occupations with a disproportionate effect on service sector employment. However, combined these produce a small increase in the share of non-employed graduates that closely matches the data. Figure 12 subplot 1 sheds some additional light on this phenomenon. We note that the effects of this counterfactual are particularly strung below skill levels of 0.5, where graduates are considerably less likely to be employed in professional occupations and considerably more likely to be not in the labour force at all. An overall higher level of returns to skill, induces a small degree of substitution between professional and routine occupations for above-median-skilled graduates. While the same induces a large increase in the probability of being out of the labour force for lower-skilled graduates. However, as these graduates already have a low rate of participation, the overall effect is muted.

Turning to the preferences counterfactual, we had noted two general trends for the non-pecuniary occupation-specific utilities: i) the closing of the relative gap between professional and routine occupations; and ii) a decrease in the relative value of not working. Subplot 2 of Figure 12 shows how these trends affect graduates at different skill levels. Firstly, we note that at the top of the skill distribution, there is a strong reallocation from professional occupations into routine occupations, while in the bottom half of the skill distribution, we see a strong decrease in the probability of being out of the labour force, commensurate with the fall in the value of non-employment. Combined the preference counterfactual induces a decline of around 2.4pp in the share of professional occupations and a 3.4pp increase in the share of routine occupations and a 1.1pp increase in the share of service occupations. This is counterpointed by an approximately 2.2pp fall in the share of graduates out of the labour force.

Figure 12: Occupational sorting under different counterfactuals (2010-2019)



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

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 returns, 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 jobs. By themselves, they also imply an unrealistically large increase in the share of non-employed graduates.

2. Changes in returns to skill tend to push low-skilled graduates out of the labour market, 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 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 18% of a standard deviation, this explains around 58% 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 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 in service 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 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

Acemoglu, Daron & David H. Autor (2011). "Skills, Tasks and Technologies: Implications for Employment and Earnings". In: *Handbook of Labor Economics.* Ed. by Orley Ashenfelter & David Card. Vol. 4. Elsevier, pp. 1043–1171.

Altonji, Joseph G., Peter Arcidiacono & Arnaud Maurel (2016). "The Analysis of Field Choice in College and Graduate School: Determinants and Wage Effects". In: *Handbook of the Economics of Education.* Ed. by Eric A. Hanushek, Stephen J. Machin & Ludger Woessmann. Vol. 5. Elsevier, pp. 305–396.

Altonji, Joseph G., Lisa B. Kahn & Jamin D. Speer (2014). "Trends in Earnings Differentials across College Majors and the Changing Task Composition of Jobs". In: *American Economic Review* 104.5, pp. 387–393.

Andrews, Rodney J. et al. (2022). *The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability.* Tech. rep. w30331. National Bureau of Economic Research.

Aucejo, Esteban & Jonathan James (2021). "The Path to College Education: The Role of Math and Verbal Skills". In: *Journal of Political Economy* 129.10, pp. 2905–2946.

Autor, David H. (2013). "The "Task Approach" to Labor Markets: An Overview". In: *Journal for Labour Market Research* 46.3, pp. 185–199.

Autor, David H. & Michael J. Handel (2013). "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages". In: *Journal of Labor Economics* 31.S1, S59–S96.

Autor, David H., Frank Levy & Richard J. Murnane (2003). "The Skill Content of Recent Technological Change: An Empirical Exploration". In: *The Quarterly Journal of Economics* 118.4, pp. 1279–1333.

Carneiro, Pedro & Sokbae Lee (2011). "Trends in Quality-Adjusted Skill Premia in the United States, 1960–2000". In: *American Economic Review* 101.6, pp. 2309–2349.

Deming, David J. (2017). "The Growing Importance of Social Skills in the Labor Market". In: *The Quarterly Journal of Economics* 132.4, pp. 1593–1640.

Diegert, Paul (2024). "The Role of Skills and Sorting in Explaining Wage Inequality". Unpublished manuscript.

Firpo, Sergio, Nicole M Fortin & Thomas Lemieux (2011). *Occupational tasks and changes in the wage structure.* Tech. rep. IZA Discussion Papers.

Gill, Jeff & Gary King (2004). "What to Do When Your Hessian Is Not Invertible: Alternatives to Model Respecification in Nonlinear Estimation". In: *Sociological Methods & Research* 33.1, pp. 54–87.

Goos, Maarten, Alan Manning & Anna Salomons (2014). "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring". In: *American Economic Review* 104.8, pp. 2509–2526.

Green, Francis & Yu Zhu (2010). “Overqualification, Job Dissatisfaction, and Increasing Dispersion in the Returns to Graduate Education”. In: *Oxford Economic Papers* 62.4, pp. 740–763.

Hamermesh, Daniel S. & Stephen G. Donald (2008). “The Effect of College Curriculum on Earnings: An Affinity Identifier for Non-Ignorable Non-Response Bias”. In: *Journal of Econometrics* 144.2, pp. 479–491.

Hastings, Justine S., Christopher A. Neilson & Seth D. Zimmerman (2013). *Are Some Degrees Worth More Than Others? Evidence from College Admission Cutoffs in Chile.* Tech. rep. w19241. National Bureau of Economic Research.

Heckman, James J., Jora Stixrud & Sergio Urzua (2006). “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior”. In: *Journal of Labor Economics* 24.3, pp. 411–482.

Hemelt, Steven W. et al. (2021). *College Majors and Skills: Evidence from the Universe of Online Job Ads.* Tech. rep. w29605. National Bureau of Economic Research.

Holmes, Craig & Ken Mayhew (2016). “The Economics of Higher Education”. In: *Oxford Review of Economic Policy* 32.4, pp. 475–496.

Hou, Shihang (2023). “The Role of Uncertainty in Education Mismatch”. Unpublished manuscript.

Kirkeboen, Lars J., Edwin Leuven & Magne Mogstad (2016). “Field of Study, Earnings, and Self-Selection”. In: *The Quarterly Journal of Economics* 131.3, pp. 1057–1111.

Leighton, Margaret & Jamin D. Speer (2020). “Labor Market Returns to College Major Specificity”. In: *European Economic Review* 128, p. 103489.

Lindley, Joanne & Steven McIntosh (2015). “Growth in Within Graduate Wage Inequality: The Role of Subjects, Cognitive Skill Dispersion and Occupational Concentration”. In: *Labour Economics* 37, pp. 101–111.

Lovenheim, Michael F. & Jonathan Smith (2022). *Returns to Different Postsecondary Investments: Institution Type, Academic Programs, and Credentials.* Tech. rep. w29933. National Bureau of Economic Research.

McFadden, Daniel & Kenneth Train (2000). “Mixed MNL Models for Discrete Response”. In: *Journal of Applied Econometrics* 15.5, pp. 447–470.

O’Leary, Nigel & Peter Sloane (2016). “Too Many Graduates? An Application of the Gottschalk–Hansen Model to Young British Graduates between 2001–2010”. In: *Oxford Economic Papers* 68.4, pp. 945–967.

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

Roy, Andrew Donald (1951). “Some Thoughts on the Distribution of Earnings”. In: *Oxford Economic Papers* 3.2, pp. 135–146.

Roys, Nicolas & Christopher Taber (2016). “Skills Prices, Occupations and Changes in the Wage Structure”. Unpublished manuscript.

— (2022). “Skill Prices, Occupations, and Changes in the Wage Structure for Low Skilled Men”. Unpublished manuscript.

Shorrocks, Anthony F. (1982). “Inequality Decomposition by Factor Components”. In: *Econometrica*, pp. 193–211.

Sorkin, Isaac (2018). “Ranking Firms Using Revealed Preference”. In: *The Quarterly Journal of Economics* 133.3, pp. 1331–1393.

Train, Kenneth (2000). “Halton Sequences for Mixed Logit”. Unpublished manuscript.

— (2016). “Mixed Logit with a Flexible Mixing Distribution”. In: *Journal of Choice Modelling* 19, pp. 40–53.

Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

A Additional Tables & Results

A.1 Additional Tables

Table A1: Details of Model Fit

Occupation Shares	**Data Average**	**Simulation Average**	**Difference**	**% Difference**
Share Professional	0.619	0.617	-0.002	-0.347
Share Routine	0.166	0.166	0.000	-0.270
Share Service	0.095	0.098	0.004	3.868
Share Non-employed	0.120	0.119	-0.001	-0.893
Log Wage	**Data Average**	**Simulation Average**	**Difference**	**% Difference**
Mean Overall	2.536	2.529	-0.007	-0.257
Mean Professional	2.667	2.624	-0.043	-1.603
Mean Routine	2.275	2.377	0.102	4.486
Mean Service	2.130	2.192	0.062	2.905
Variance Overall	0.152	0.151	0.000	-0.266
Variance Professional	0.112	0.136	0.024	20.962
Variance Routine	0.106	0.122	0.016	14.892
Variance Service	0.090	0.088	-0.002	-2.462
P10 Overall	2.025	2.066	0.041	2.032
P10 Professional	2.257	2.201	-0.056	-2.463
P10 Routine	1.901	1.998	0.097	5.123
P10 Service	1.823	1.810	-0.013	-0.708
P50 Overall	2.558	2.539	-0.018	-0.723
P50 Professional	2.670	2.627	-0.043	-1.593
P50 Routine	2.261	2.394	0.133	5.875
P50 Service	2.112	2.193	0.081	3.859
P90 Overall	3.021	2.999	-0.021	-0.706
P90 Professional	3.088	3.069	-0.018	-0.591
P90 Routine	2.680	2.772	0.093	3.452
P90 Service	2.500	2.571	0.071	2.856
P50-P10 Overall	0.532	0.473	-0.060	-11.209
P50-P10 Professional	0.413	0.426	0.013	3.159
P50-P10 Routine	0.360	0.395	0.035	9.842
P50-P10 Service	0.289	0.383	0.094	32.695
P90-P50 Overall	0.463	0.460	-0.003	-0.609
P90-P50 Professional	0.418	0.442	0.024	5.812
P90-P50 Routine	0.419	0.379	-0.040	-9.622
P90-P50 Service	0.388	0.378	-0.010	-2.604

Note: Model simulations based on a sample of 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	0.673	0.659	0.293	0.296	0.099	0.108	0.278	0.279
2010-2019	0.650	0.638	0.264	0.258	0.095	0.103	0.263	0.261
Δ	-0.023	-0.021	-0.029	-0.038	-0.005	-0.005	-0.015	-0.017
Δ%	-3.348	-3.136	-10.050	-12.974	-4.598	-4.183	-5.524	-6.250

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

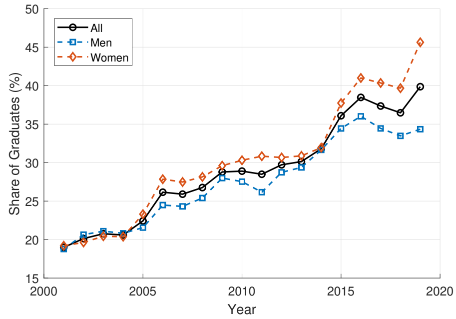
Table A3: Period Comparison

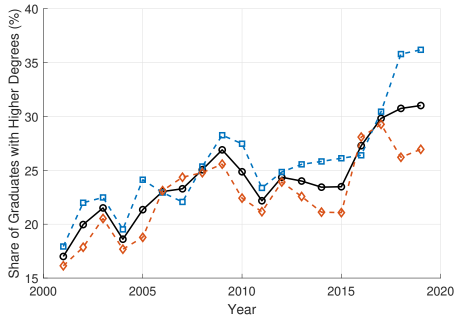
	Data				Baseline Model			
	2001-2009	**2010-2019**	**Difference**	**% Difference**	**2001-2009**	**2010-2019**	**Difference**	**% Difference**
Occupation Shares								
Professional	65.281	58.894	-6.387	-9.784	64.498	58.601	-5.898	-9.144
Routine	15.246	17.832	2.586	16.961	15.247	18.024	2.777	18.214
Service	7.939	10.860	2.921	36.796	8.561	11.267	2.706	31.605
Non-employed	11.534	12.414	0.880	7.627	11.694	12.109	0.415	3.549
Log Wages								
Mean	2.580	2.497	-0.083	-3.231	2.561	2.495	-0.066	-2.566
Variance	0.151	0.153	0.002	1.038	0.146	0.157	0.011	7.410
P10	2.068	1.987	-0.082	-3.955	2.106	2.022	-0.084	-3.985
P50	2.600	2.519	-0.081	-3.100	2.568	2.507	-0.061	-2.357
P90	3.064	2.982	-0.082	-2.682	3.028	2.967	-0.061	-2.029
Mean Professional	2.703	2.634	-0.069	-2.545	2.646	2.599	-0.047	-1.775
Mean Routine	2.285	2.266	-0.019	-0.846	2.399	2.352	-0.047	-1.956
Mean Service	2.134	2.126	-0.008	-0.357	2.202	2.180	-0.022	-1.017

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

A.2 Additional Figures

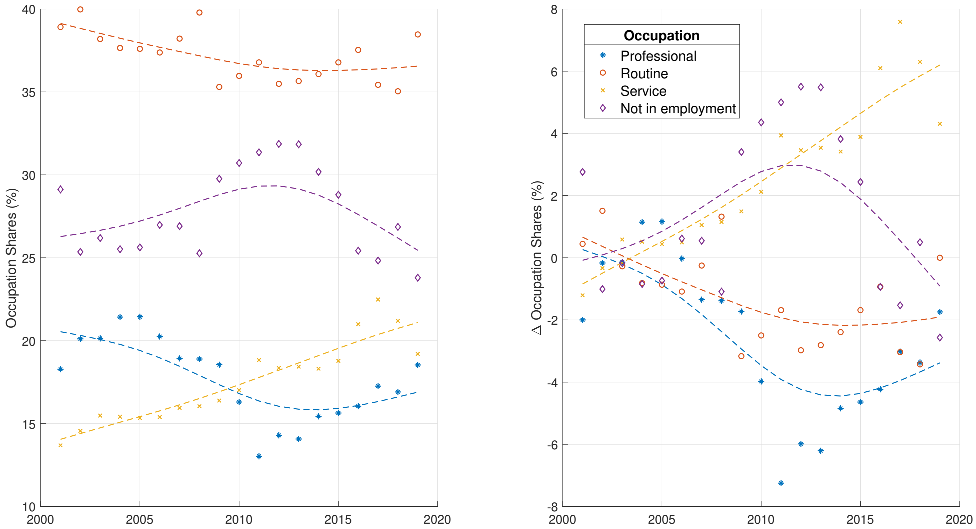
Figure A1: Trends in higher education attainment among young adults in the UK (2001–2019)





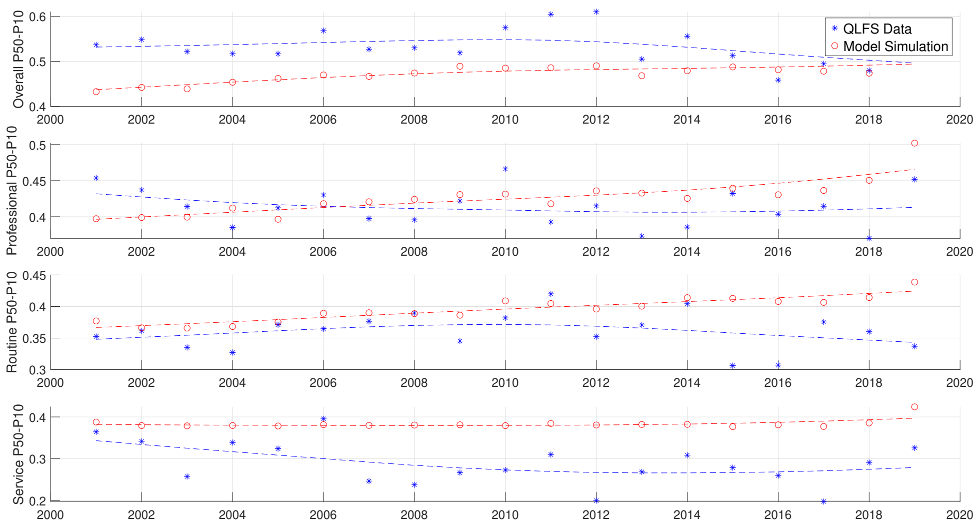
Note: Subplot 1: Share of young adults with a university degree, aged 21-30 years. Subplot 2: Share of graduates with a higher degree, aged 21-30 years. *Source:* Quarterly Labour Force Survey (2001-2019).

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



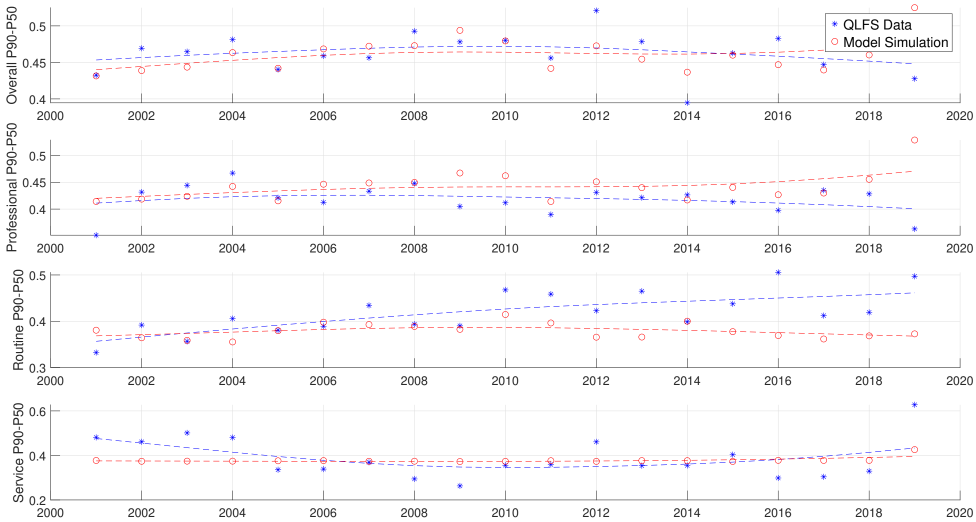
Note: Adults without a university degree, aged 21-30 years; change relative to 2001-05 average. Broken lines are HP-filtered trends (smoothing parameter = 100). SOC 2000 1-Digit Occupation Classification. Professional includes codes 1-3; Routine includes codes 4,5,8,9; Service includes codes 6,7; Not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

Figure A3: Model fit - P50-P10



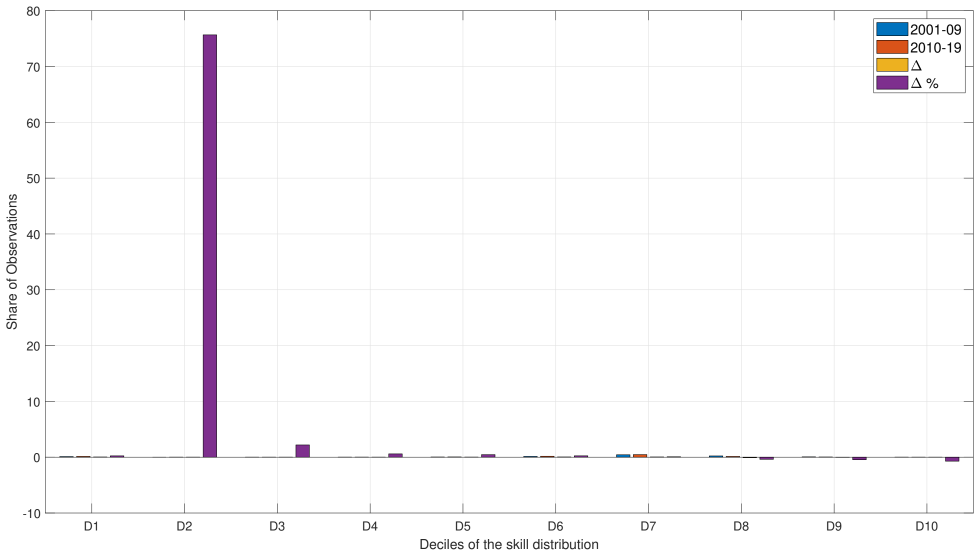
Note: Broken lines represent HP-filtered trends with a smoothing parameter of 100, based on a simulated sample of graduates. *Source:* Baseline Model.

Figure A4: Model fit - P90-P50



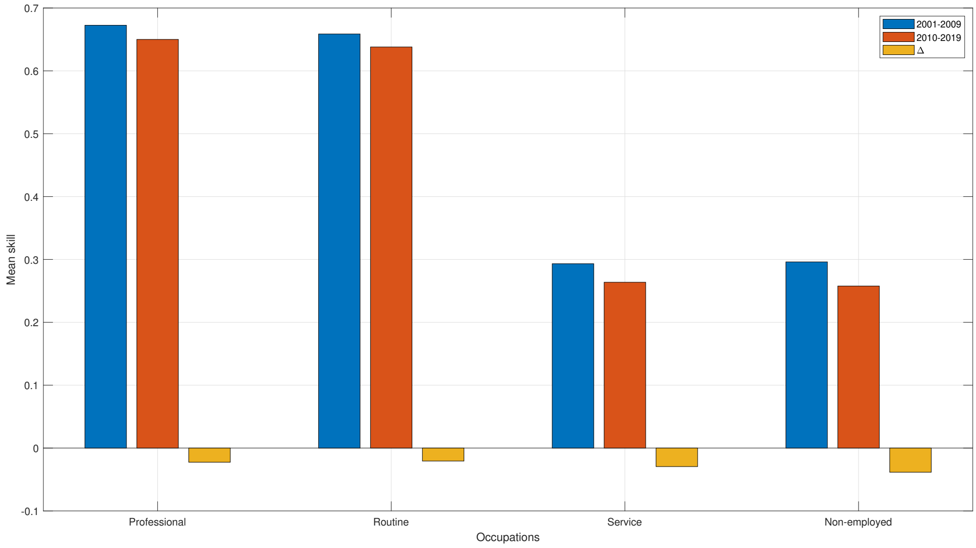
Note: Broken lines represent HP-filtered trends with a smoothing parameter of 100, based on a simulated sample of graduates. *Source:* Baseline Model.

Figure A5: Histogram of 2010-2019 skill distribution



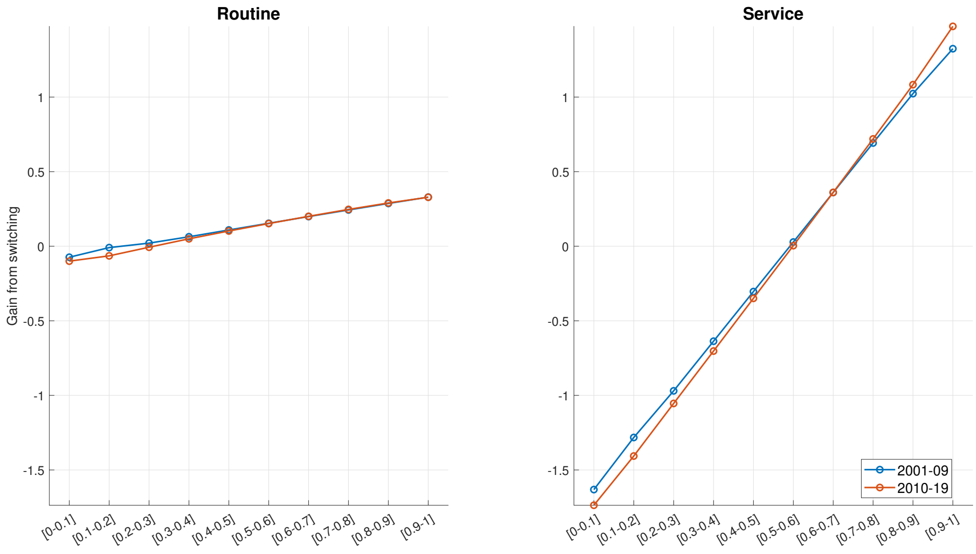
Note: Based on a simulated sample of graduates.
Source: Baseline Model.

Figure A6: Average skill levels by occupation



Note: Based on a simulated sample of graduates. *Source:* Baseline Model.

Figure A7: Wage gains from switching to professional occupations



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

B Technical Details

B.1 Imputation of missing wage values

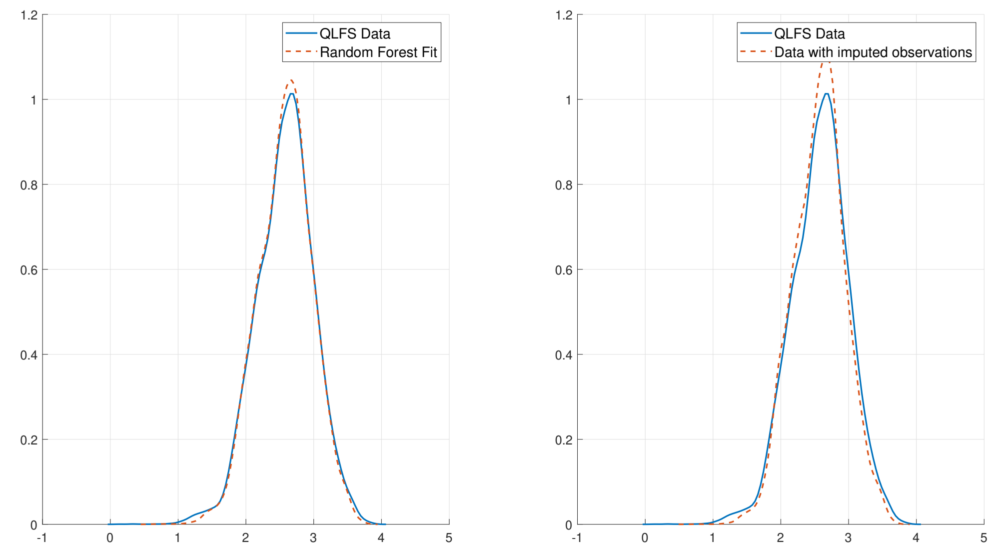
In the QLFS, wage data is only collected during the first and last interviews, resulting in three missing wage observations for each individual. To impute these missing wage values, I employ a two-step methodology that leverages both fixed effects to account for individual-specific skills and a Random 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. The objective is to find a vector of parameters θ that maximizes the probability of observing the actual outcomes. The only complication, that arises in this case is that we do not have a closed-form solution for the joint probability (9) and thus have to approximate the integral via quadrature or simulation.

Take graduate i with skills s_i, then, the conditional joint probability of observing the occupation choice o^*_{it} and the wage w^*_{it} is given by 7. Since we observe graduates for $T = 5$ periods the probability of observing the sequences o^*_i and w^*_i is given by $\Pr(o^*_i, w^*_i|s_i) = \prod_{t=1}^{T} \Pr(o^*_{it}, w^*_{it}|s_i)$. Following from K. Train, 2016, the unconditional probability can be approximated by: $\Pr(o^*_i, w^*_i) = \sum_{r=1}^{R} \Pr(o^*_i, w^*_i|s_r)W(s_r|\alpha)$, where $W(s_r|\alpha)$ is a weighting scheme that describes the probability of drawing s_r from some distribution S and α is a vector of parameters that describes that distribution. As described in the main part of the paper, I approximate S with a histogram on the unit interval with L evenly spaced bins B_l. Hence for a random s_r the probability $W(s_r|\alpha)$ is equal to the probability of selecting bin B_l, into which s_r falls. I specify $Pr(B_l) = \frac{e^{\alpha_l}}{\sum_{l=1}^{L} e^{\alpha_l}}$, with α_L normalised to 0. This procedure provides a fast, and importantly numerically stable way of calculating the weights and hence approximate $\Pr(o^*_i, w^*_i)$. 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 & K. Train, 2000).

For ease of notation let's assume that α is a subset of the extended parameter vector θ. We can write down the simulated log-likelihood function as:

$$ll^{sim}(\theta) = \frac{1}{N}\sum_{i=1}^{N} ln[\sum_{r=1}^{R}\sum_{o=1}^{O}\mathbf{1}_{(o=o^*)}\Pr(o_i^*, w_i^*|s_r)W(s_r|\alpha)] \tag{B1}$$

and we can estimate θ as:

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

The complete algorithm proceeds as follows:

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

   (b) Calculate $W(s_r|\alpha)$.

   (c) 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. Evaluate the integral over all R values of $\Pr^r(o_i^*, w_i^*|s_i^r)$ to obtain:

   $\Pr(o_i^*, w_i^*) = \sum_{r=1}^{R}\Pr^r(o_i^*, w_i^*|s_i^r)W(s_r|\alpha)$.

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 100 quasi-random Halton draws, which have been shown to provide about an order of magnitude more accuracy than simple random draws (K. Train, 2000, K. E. Train, 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. K. E. 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:

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

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**
α_c	Parameters describing the skill distribution.	18
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 a better fit on one dimension against a 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.