Skills, Tasks and Degrees

Max Schroeder
University of Durham

August 20, 2024

Preliminary and incomplete
Comments welcome

Abstract

Increasingly, young university graduates do not join typical graduate occupations after completing their studies. But is this by itself evidence of a considerable decline in the labour market skill of recent cohorts of university graduates? To investigate this question I propose a model of the labour market for young graduates, featuring heterogeneous skill supply on the side of graduates and differentiated skill demand on the side of occupations. The model allows the seperation of three different mechanisms: i) changing quality of graduates or skill supply; ii) changes in the returns to skill in different occupations; iii) different non-pecuniary preferences.

I structurally estimate the model using data on young university graduates in the UK. The estimation suggests that between the 2001-2010 and 2011-2019 cohorts, the average level of skill among graduates has fallen by around 25% of a standard deviation, explaining around 50% of the fall in employment in professional occupations. Sorting according to ability, implies that these reductions in the share of professional occupations are concentrated in the lower end of the skill distribution.

Furthermore, increases in the amenity value of routine and service occupations explain why graduates increasingly these non-traditional career options.

Keywords: skills, tasks, degree, tertiary education, occupation choice
JEL Classification: I24, J24

I would like to thank S. Lazarakis, M. Leighton and R. Foltyn for comments on earlier versions of this paper. I further thank the seminar attendants at the Annual meetings of the Scottish Economic Society (2021), the Scottish Graduate Programme in Economics (2021), the Society for Computational Economics (2021), the Workshop on Applied Economics of Education (2022), the Annual Conference of The European Association of Labour Economists (2022), 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) and the Society of Labor Economists (2024) for their insights. All remaining errors are my own.

1 Introduction

Over the last 30 years, the UK (and other developed economies) has experienced a rapid expansion of tertiary education participation. Since the passage of the Further and Higher Education Act 1992, university enrolment has roughly doubled to approximately 2.5 million in 2019/20, a trend that was sustained in the face of stark tuition fee increases.[1] 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. Altonji et al. (2016), Lindley & MacIntosh (2015)). 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 to this expensive investment made by so many young individuals.

But the changing supply of skills is only one part of the ever-developing labour market. Over the past decades there have been important changes in the demand for the skills that young graduates possess, driven by changes in technology and the wider macroeconomic environment (see Acemoglu & Autor (2011) for a survey). Furthermore, current generations also have different preferences along non-pecuniary job characteristics, which evolve over time and influence occupation decisions. The observed patterns of labour market outcomes for young graduates are the result of the interactions between these forces, making monocausal explanations doubtful.

The question therefore is to what degree are the observed patterns in the labour market outcomes of young graduates due to changing distribution of graduate skills, the evolving demand for these skills or changing preferences and demographic factors?

To answer this question, I develop a model of occupational choice for university graduates to quantify the importance of the demand and supply of graduates' skills for the labour market success of young graduates. Skills are generally unobservable for an econometrician working with individual level survey data. In order to address this issue, I take another approach - framing the question as a latent variable problem: skills are unobserved but related to observable choices and labour market outcomes. By specifying and estimating a corresponding structural economic model, we 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 university graduates from 2001-2019 and recover the parameters of the underlying latent skill distributions for different cohorts of university graduates. I then use the model estimates to analyse changes in the graduate skill distribution, the changing demand for these skills and their combined effects on the labour market outcomes of university graduates over the first two decades of the 21st century by running counterfactual simulation decompositions.

I find that between 2001-2010 and 2011-2019 the mean of the distribution of graduates' skills has decreased by about 25% of a standard deviation. This has a direct effect on the share of young graduates entering professional occupations, explaining around 50% 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 to not 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 of service occupations as an alternative.

This paper adds to the literature on graduate underemployment with a specific focus on the UK (see

[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.

Chevalier & Lindner (2006), Green & Heneske (2016), 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 the three main drivers of these trends: i) changes in the distribution of graduates' skills as a result of the higher education expansion, ii) changes in the return to graduates' skills in different occupations due to technological change, iii) changing preferences. My main findings suggest that differences in the returns to skills in non-graduate occupations and changes in graduates' preferences over occupations play a bigger role relative to between cohort differences in the supply of skills.

This paper also contributes to a large literature on the returns to higher education with an emphasis on heterogeneity of returns (see Altonji et al. (2016), Andrews et al. (2022) and Lovenheim & Smith (2023) for extensive surveys). Generally, these studies aim to estimate the returns to post-secondary education, whilst trying to address the inherent difficulties caused by the existence of selection effects across dimensions of inherent ability and preferences using administrative cutoff rules (see for example Kirkeboen et al. (2016), Hastings et al. (2013)); or try to control for observable factors (Hamermesh & Donald (2008)). A general finding of this literature is that returns to a college degree vary according to a number of factors, such as field of study, degree classification earned, or institution attended. In this paper I focus on the interaction between the supply of and demand for graduate's skills in a partial equilibrium framework, thereby providing a potential mechanism for the observed differences in labour market outcomes for different cohorts of graduates.

Further, this paper also speaks to a large literature that investigates the way in which technological change affects the sorting of different workers across occupations and correspondingly the wage distribution. Key here is the task framework based on the seminal work of Autor et al. (2013) and formalised in Acemoglu & Autor (2011). The task framework shifts emphasis onto specific job tasks and therefore 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. The approach I take in this paper is trying to estimate both changing supply and demand factors, which is closely related to Diegert (2024), who develops an econometric framework to estimate the supply of skills as well as occupation specific skill prices for the US over 30 years. My paper is closely related to this work, although with a different geographic and demographic focus.

Finally, this paper complements previous work that aims to elicit the skill content of university degrees. Altonji et al. (2014), create measures of the task content of different subject by mapping task measures from the Dictionary of Occupational Titles to graduate's occupation choices. Similarly, Hemelt et al. (2021) collect information from online job postings, to associate desired skills with different degree subjects. My paper differs as that it occupation choice and wage information to estimate a continuous distribution of graduate skills.

To the best of my knowledge this paper is the first to estimate the demand for 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 the counterfactual decompositions; finally, section 7 concludes.

2 Motivating empirics

The following section presents some evidence as to the changing labour market outcomes of young graduates over the last 20 years. 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 5 consecutive quarters and 20% of the sample is replaced in every wave. In the following section I will use the cross sectional QLFS to motivate the analysis in the rest of the paper, while for the estimation I will use the 5-quarterly version of the Longitudinal QLFS.

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

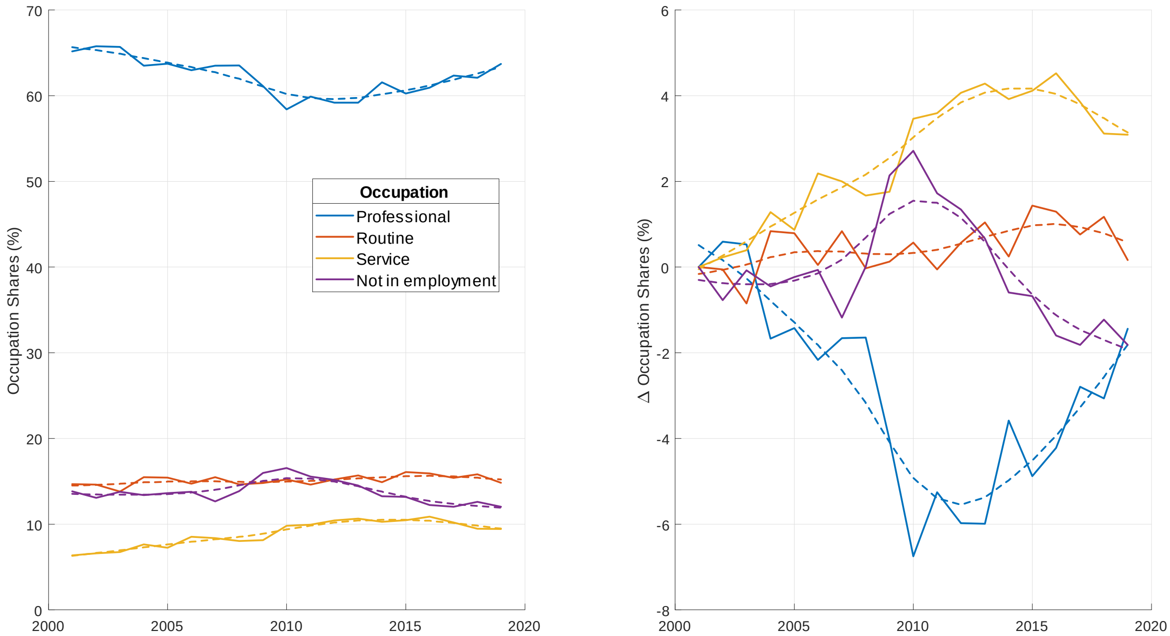
Figure 1 presents the occupation shares of young graduates between 2001 and 2019 disaggregated by 3 major occupation groups - Professional, Routine and Services, as well as non-employment.[4] By most accounts, a "graduate job" is a key marker of success for a young graduate, which is commonly identified with employment in professional occupations such as Law, Medicine or Financial Services. The data reveals that a majority of young graduates are employed in professional occupations with a high of around 65% in the early 2000s, although exhibiting a somewhat declining trend into the post financial crisis years after which the numbers rebound somewhat. In contrast to this the share of young graduates working in service occupations increases steadily over the first 15 years of the sample, increasing by almost 50% relative to it's low baseline. Employment in routine occupations is relatively steady over the sample period, although there is a slight increase in the years post 2013. 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.

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

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

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

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



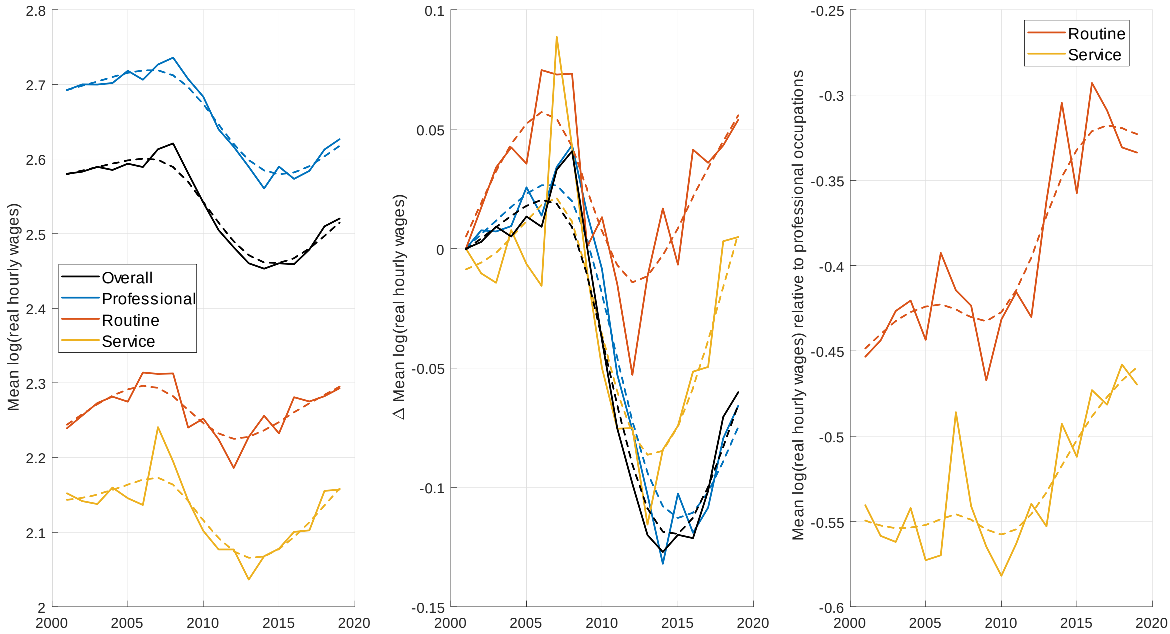
Note: Graduates aged between 21 and 30 years. Change relative to 2001 level.

Broken lines represent HP-filtered trends with a smoothing parameter of 6.25.

Source: Quarterly Labour Force Survey

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 of 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 were growing in the run up to 2008 and then only experienced a small decline before making a strong recovery. Wages for service occupations lie somewhere in the middle between these two cases. 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 hourly wages, deflated by 2014 CPI. Working graduates aged between 21 and 30 years.

Broken lines represent HP-filtered trends with a smoothing parameter of 6.25. Change relative to 2001 level.

Source: Quarterly Labour Force Survey

The data presented in this section highlight some interesting trends in the labour market outcomes of young graduates. Over the two decades preceding the Covid19 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 and particularly relative to wages earned in routine and service occupations. While economic logic dictates that changes in relative wages should result in a reallocation of labour from one occupation to another, the underlying drivers of these changes are not well understood. In the following section I will outline a quantitative economic model that might shed some light on the deeper, structural forces behind these trends.

3 Model

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

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

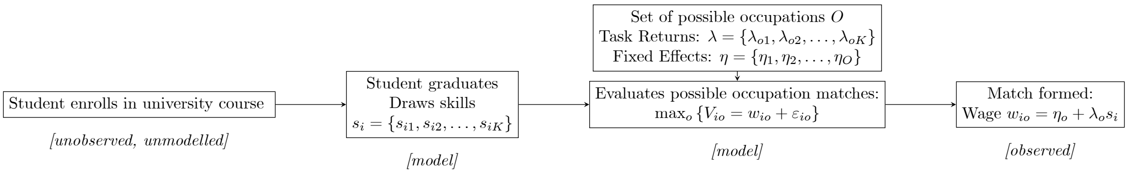


Figure 3: Model Summary

For ease of exposition, I first present a cross sectional version of the model, with suppressed time subscripts. The extension to a more dynamic framework, where parameters are allowed to vary over time is straightforward. A graduate's multidimensional skill-set is summarized by a K dimensional vector $s_i = \{s_{i1}, s_{i2}, .., s_{iK}\}$ where each element $s_{ik} \geq 0$ describes how effective graduate i is at performing task k. I assume that s_i is a unobserved, random vector that is drawn from a parametric distribution $s_i \sim S$ that is the same for all individuals $i \in I$. On the demand side, the labour market consists of a large number of *occupations* that each use the different skills supplied to them in different proportions. Specifically, every occupation $o \in O$ has an associated vector $\lambda_o = \{\lambda_{o1}, \lambda_{o2}, ...\lambda_{oK}\}$ where each element $\lambda_{ok} \geq 0$ summarizes the productivity of task k in occupation o.[5]

A worker'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 non-skill characteristics (gender, labour market experience, etc.) and β is a vector of coefficients. This setup is common in the literature on tasks and skills (c.f. Autor & Handel (2013), Roys & Taber (2016)).

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

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

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

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

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

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

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

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

Generally, economies of the type described above are characterized by the sorting of workers according to comparative advantage (see Roy (1951)). This self-selection of workers into different occupations according to their different abilities, poses the main obstacle that is faced by the literature that is concerned with estimating *task returns* (i.e. the set λ), since there will be a positive correlation between an occupations' task prices λ_o and the skills supplied by workers selecting into this occupation (see Autor (2013)). In this setup, however, rather than being harmful, self-selection is actually helpful as it allows us to make inferences from a worker's observed occupation o_i^* to her unobserved skill-set s_i.

4 Econometric strategy & estimation

The econometric strategy combines the empirical content of the economic model described above. The key ingredient is that both, a worker's occupation choice and her realized wage are informative about her skill set, given the economic structure described in the last section. The key parameters of interest are those relating to the determination of the log wage η, λ, β, the occupation specific preferences ω, as well as the distribution of graduates' skills S.

Normalising the option of not working to 0 for convenience, and making the standard assumption that idiosyncratic occupation preference shocks ε_i are distributed i.i.d. Type I Extreme Value with variance $\frac{\pi^2}{\rho 6}$, where $\rho > 0$ is an additional scale parameter, we can express the conditional choice probability of the graduate's chosen occupation o_i^* as:

$$\Pr(o_i^*|s_i) = \frac{e^{\rho(w_{io^*}+\omega_{o^*})}}{1+\sum_{o=1}^{O} e^{\rho(w_{io}+\omega_o)}} = \frac{e^{\rho\left(\eta_{o^*}+\lambda'_{o^*}s_i+\beta' x_i+\omega_{o^*}\right)}}{1+\sum_{o=1}^{O} e^{\rho\left(\eta_o\lambda'_o s_i+\beta' x_i+\omega_o\right)}}. \tag{4}$$

Further, assume that observed wages w_i^* are subject to measurement error

$$w_i^* = w_i + v_i, \tag{5}$$

where v_i is a random disturbance term, **independent** of the workers occupation choice and distributed $\nu_i \sim N(0, \phi^2)$.[8] 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_i^*:

$$\Pr(w_i^*|s_i, o_i^*) = \frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}. \tag{6}$$

Hence, we can write the joint conditional probability as:

$$\Pr(o_i^*, w_i^*|s_i) = \left(\frac{e^{\rho\left(\eta_{o^*}+\lambda'_{o^*}s_i+\beta' x_i+\omega_{o^*}\right)}}{1+\sum_{o=1}^{O} e^{\rho\left(\eta_o\lambda'_o s_i+\beta' x_i+\omega_o\right)}}\right)\left(\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right). \tag{7}$$

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

$$\Pr(o_i^*, w_i^*) = \int \left(\frac{e^{\rho\left(\eta_{o^*}+\lambda'_{o^*}s_i+\beta' x_i+\omega_{o^*}\right)}}{1+\sum_{o=1}^{O} e^{\rho\left(\eta_o\lambda'_o s_i+\beta' x_i+\omega_o\right)}}\right)\left(\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right) f(s)d(s) \tag{8}$$

[8] I interpret this additional error term as an idiosyncratic wage shock, that is not anticipated by the graduate at the time when she makes her occupational choice.

Standard results (c.f. McFadden & Train (2000)) guarantee, that we can use the unconditional choice probability in (8) to get consistent estimates for η, λ, β, ω and S using simulated maximum likelihood.[9] 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 (8). Diegert (2024) provides non-parametric identification results for a more general model, showing that the parameters are identified if the econometrician has panel data with length greater than $2 * K$. The longitudinal panel of the QLFS contains 5 quarterly observations for each individual, allowing me to implement a repeated panel estimation strategy, as showcased in Diegert (2024). For this purpose I separate the sample of graduates into two cohorts denoted by $c = \{1, 2\}$ depending on the year they are first observed in the sample.[10]

The final dataset contains $12,084$ graduates, totalling $60,420$ observations. 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 LQLFS dataset contains longitudinal weights to account for attrition bias.

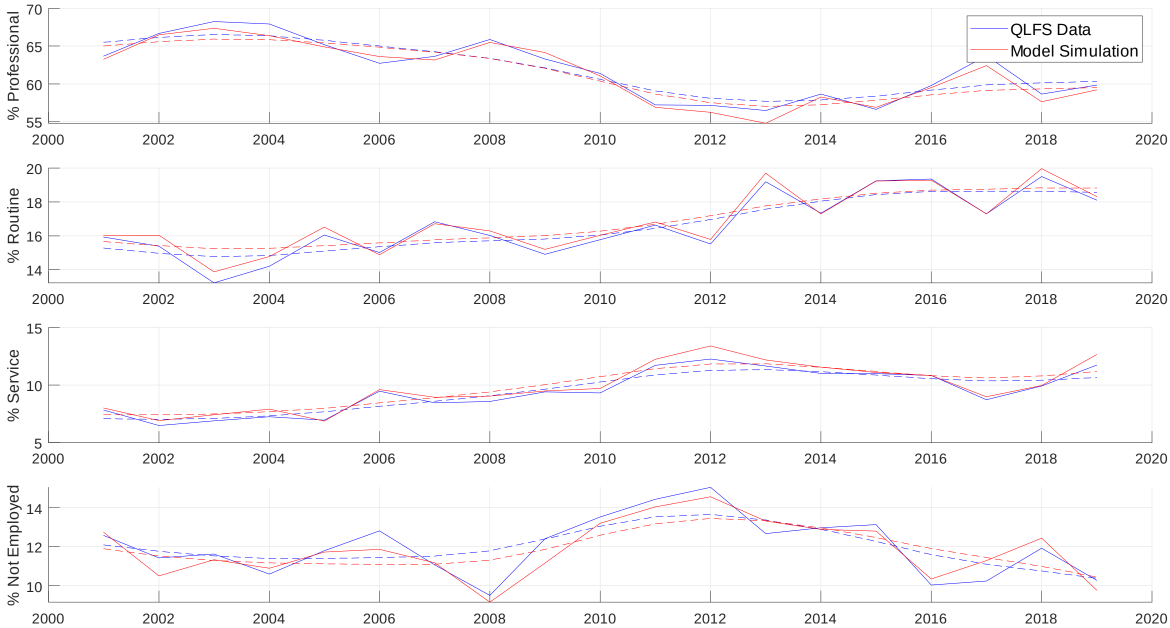
For the estimation I consider a model with a single type of skill, and I specify the skill distribution as a log normal distribution with cohort specific mean and variance parameters: $\log(s_i) \sim MVN(\mu_c, \sigma_c^2)$, which allows the skill distribution to vary for the different cohorts, accounting for the changing distribution of skill. To account for changes in the demand for skills, I further allow the occupation fixed effects η_{ot}, the occupation specific returns to skill λ_{ot} and occupation specific preference terms ω_{ot} to vary at an annual frequency. I estimate the parameters of the model using simulated maximum likelihood.

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

The model fit for the occupational choices of young university graduates is generally robust with respect to the actual data, effectively capturing long-run trends.[11] 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. However, there are some discrepancies: on average, the model underpredicts the share of graduates in professional occupations by approximately 0.5%. Conversely, it slightly overpredicts the proportions in routine jobs by an average of 0.3%, service occupations by 0.4%, and underpredicts the share of non-employed graduates by 0.2%. Despite these minor variances, the model demonstrates a strong overall fit, successfully reflecting the broader trends in occupational choices among university graduates.

[9] There is no closed form solution for this integral, but integration step can be performed via simulation.
[10] The cut-off year in this case is 2010.
[11] For details see table A1 in the appendix.

Figure 4: Model fit - occupation shares

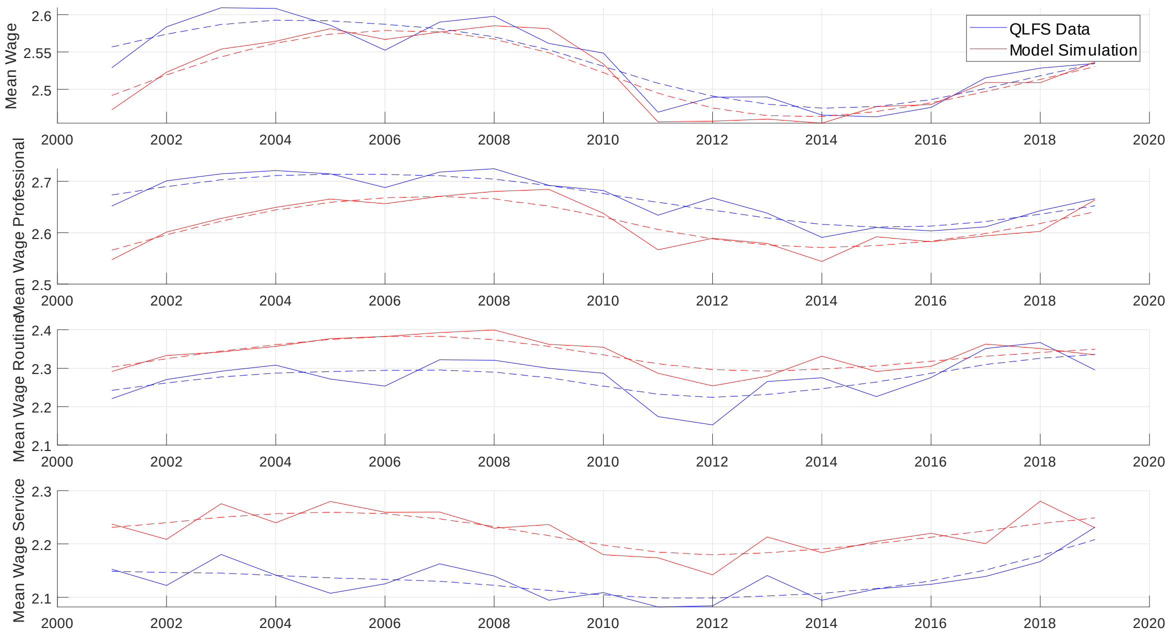


Notes: Broken lines represent HP-filtered trends with a smoothing parameter of 6.25

Based on a simulated sample of 1,000,000 graduates.

Figure 5 provides a comparison of mean log wages across the occupational categories. The model's simulations align closely with the observed data, reproducing the relative wage differences between occupations and the overarching trends. The model shows a slight underprediction in the professional category, with an average difference of 0.05 log points per year. In contrast, the model slightly overpredicts wages in routine and service occupations by 0.06 and 0.09 points, respectively.

Figure 5: Model fit - mean log wages.



Notes: Broken lines represent HP-filtered trends with a smoothing parameter of 6.25

Based on a simulated sample of 1,000,000 graduates.

5 Results

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

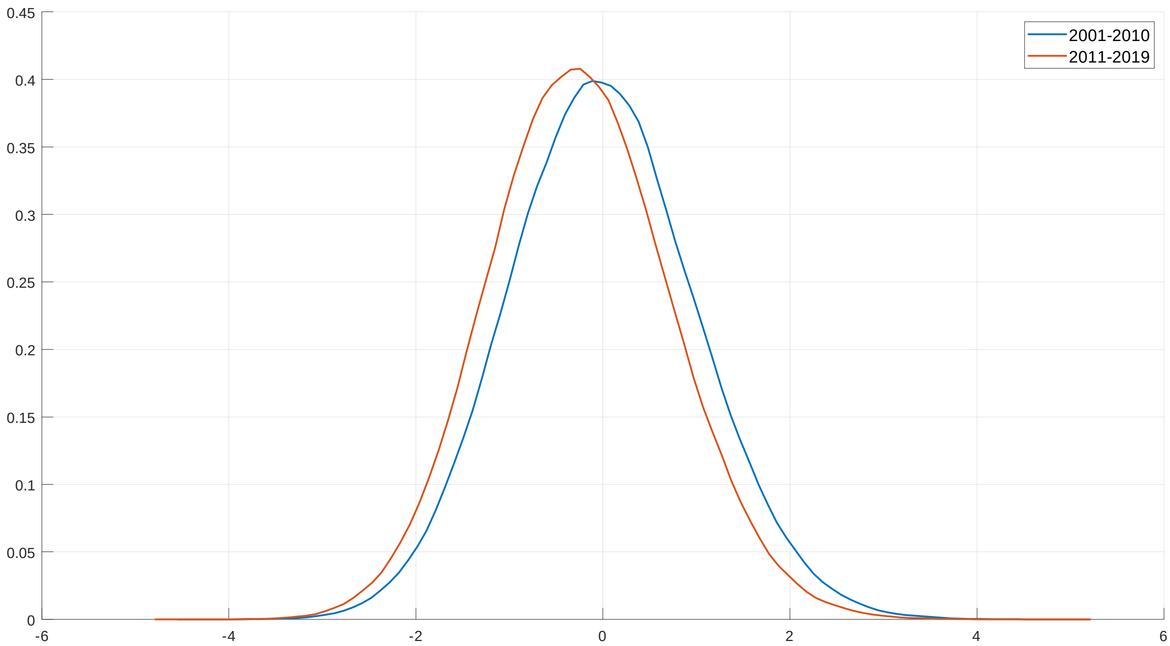
5.1 Graduate skills & the returns to skill

Figure 6 presents kernel density estimates of the distribution of graduates' skills, comparing the 2001-2010 and 2011-2019 cohorts. For ease of comparison the distributions are standardised to have mean 0 and a standard deviation of 1 in the first cohort. The main observation with respect to the changing graduate skill distribution across the two cohorts is that between 2001-2010 and 2011-2019 the mean (median) of the skill distribution fell by 0.25 (0.25) points, equivalent to 25% (25%) of a standard deviation among the first cohort, a moderate, but not insignificant reduction. [12] Anecdotally, this shift is consistent with the main reforms to the production of university graduates that occurred over the late 1990's and 2000's: The rapid expansion of higher education participation at the intensive (increased number of students) and extensive (increased number of HE providers) margins, would likely have negatively affected the average ability of high school leavers entering university, as well as the quality of the instructions received while at university (see Carneiro & Lee (2011)). Curiously, however the variance of the skill distribution has not been markedly affected by these developments. The standard deviation of the second cohort is just

[12] See also figure A2 in the appendix.

2.5% below the first cohort, suggesting that the decrease of graduate's skills is somewhat uniform across the distribution.

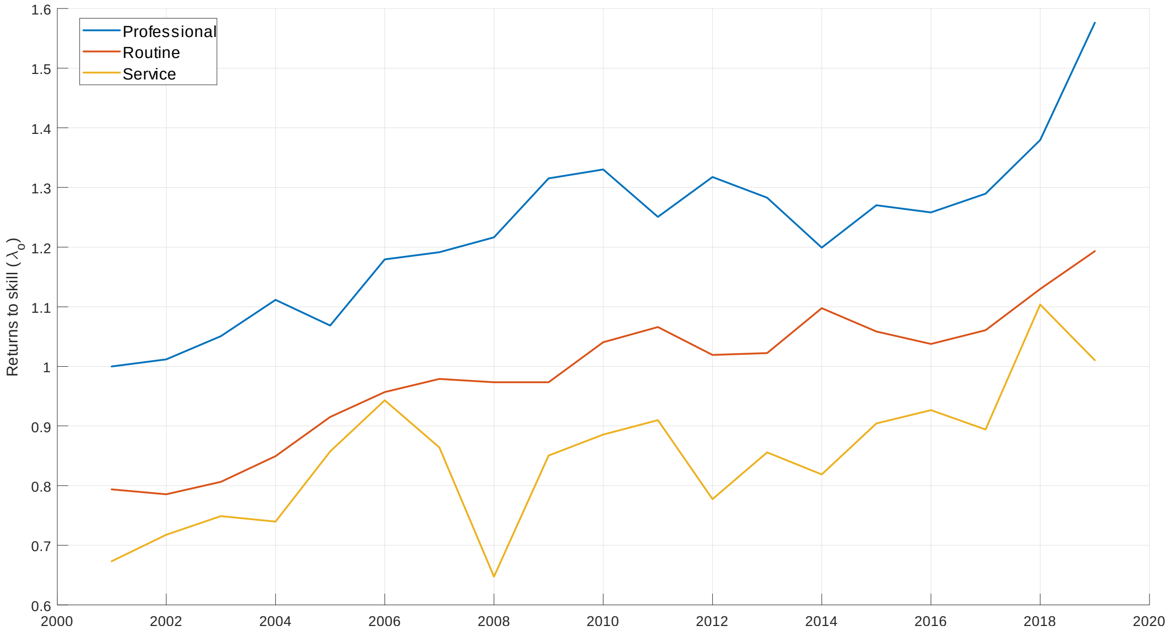
Figure 6: Distribution of graduate skill



Note: Based on a simulated sample of 1,000,000 graduates. Distributions normalised to have mean 0 and variance 1 in 2001-2011.

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. Figure 7 provides the models 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. Over the 19-year period the is a pronounced increase in the return to skills across all occupational groups, consistent with the theory of within occupation task upgrading. The trend appears to stall somewhat around the time of the financial crisis and its aftermath. While it is not easy to make out on the plot, the relative premium that professional occupations offer falls slightly as a result of a faster catch-up of the other occupations. Throughout the 2000's the return to skills in routine (service) occupations is 79% (69%) of that of professional occupations, while in the 2010's the gap has decreased to 82% (69%) respectively. The loss of comparative advantage by professional occupations is a potential contributing factor in the changing occupational destinations of young university graduates.

Figure 7: Returns to skill by occupation

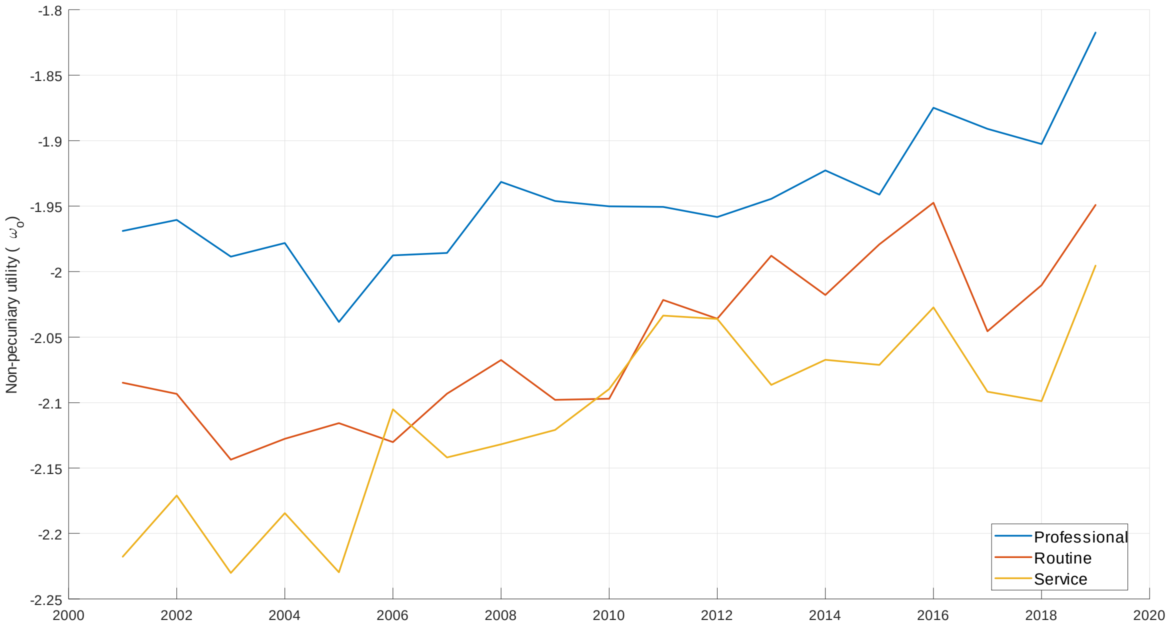


Note: Returns relative to professional occupations in 2001. Source: Model Estimates.

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

Figure 8 below 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 normalised to 0 for all periods. The estimated values are all negative, since any work carries a degree of disutility relative to the option of not joining the labour force. Since the numeraire of the model is the log wage, we can deduce that on average, graduates require a log wage of between 1.9 and 2.2 for them to consider working more attractive than remaining outside the labour market. While it is generally true that labour as such comes with a utility penalty, there are important variations across occupations and across time. First, there appears to be an ordering amongst the different occupations, with professional occupations providing the lowest level of disamenity followed by routine occupations which weakly dominate the service occupations. This pattern is consistent with expectations - professional jobs provide more interesting work, while also providing a greater degree of autonomy to graduates and are also less likely to involve unpleasant or dangerous activities. Secondly, while the relative rankings remain stable over time, there is a general trend of increasing amenity values across all occupations from about 2006 until the end of the sample period. During this general rise the gap between professional and routine and service occupations appears to narrow somewhat, suggesting that the latter are becoming more attractive relative to professional occupations. A general increase in amenity values could suggest an overall improvement in working conditions, or since the values are normalised relative to not working, a depreciation of the outside option. A reduction in the relative gap between professional and routine/service occupations hints at a relative appreciation of the utility associated with these occupations by for example improved working conditions. It could also be indicative of a more general shift in attitudes towards such "non-graduate" jobs. In either case this reduction in the disamenities gap is likely a contributing factor to the increasing share of graduates in non-typical occupations.

Figure 8: Occupation preferences over time

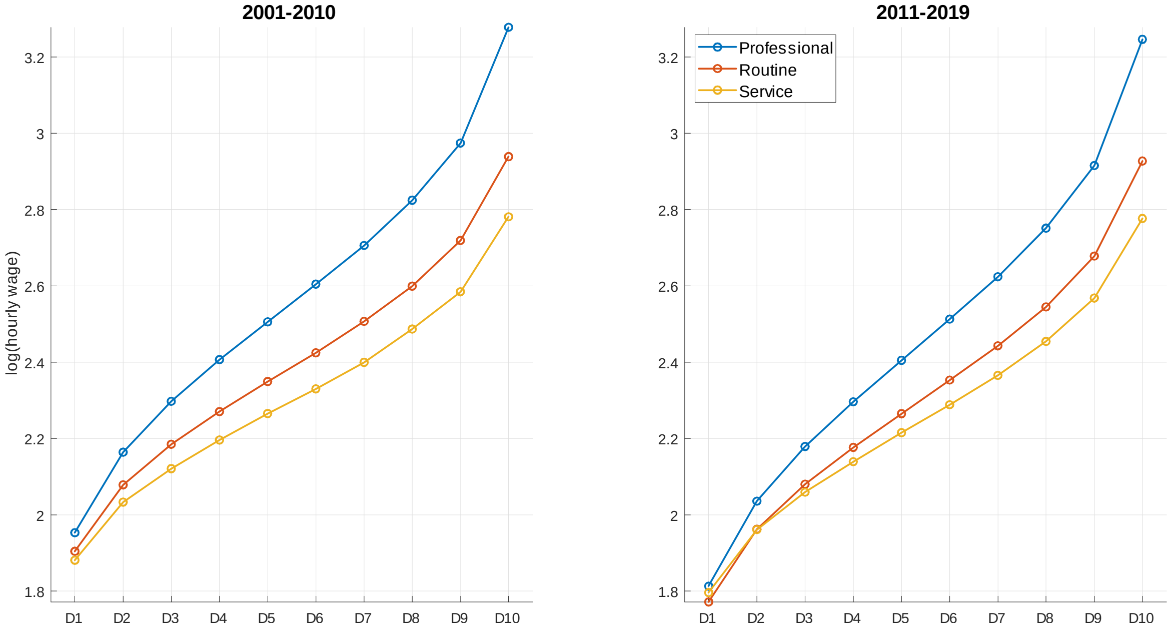


Note: Value of non-employment normalised to 0. Source: Model Estimates.

5.2 Skills and sorting

The economy described by the model in section 3 is characterised by the endogenous sorting of graduates into different occupations. Graduates with high levels of skill will - ceteris paribus - enter occupations where the return to these skills is high, while those with low levels of skill will be drawn to occupations where the skill-based compensation is lower. Figure 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. This can be clearly seen across both cohorts, as the gap in log wages, is small at the first two or three deciles, but then opens up, becoming larger as the comparative advantage of professional occupations relative to routine and service occupations begins to assert itself.

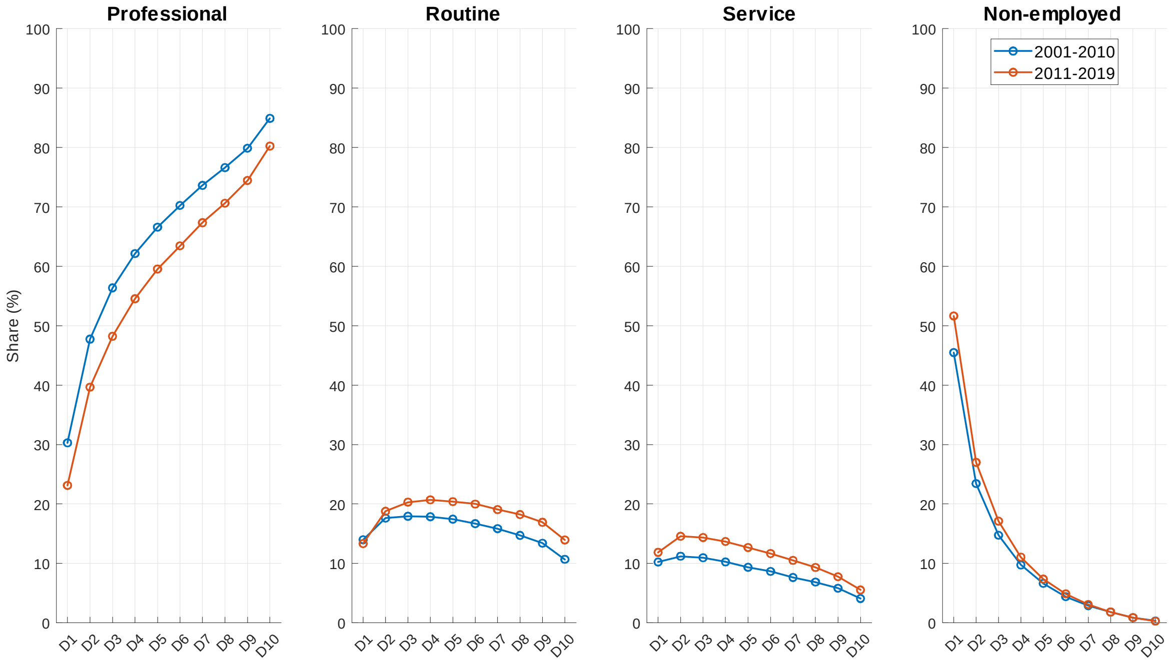
Figure 9: Hourly log wage by skill deciles



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles based on respective period.

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 occupation at all. However, as skill levels increase, the probability of being employed in a professional occupation increases rapidly. For example a graduate in the 5th decile of the skill distribution is more than twice as likely to be employed in a professional occupation than a graduate in the first decile, and for those in the top decile, the probability is around 80%. For routine and service occupation the general pattern is approximately hump shaped. Initially, moving into higher skill deciles graduates are more likely to be employed in these occupations - primarily because they are more likely to be employed at all - but this trend generally peaks around the 3rd or 4th decile and then slowly decreases, as higher skilled graduates sort into professional jobs. Comparing the two cohorts we can see how the changing skill distributions and returns to skills have affected the sorting across the skill distribution. Generally, the patterns remain the same across both cohorts, although graduates are slightly less likely to enter professional occupations and slightly more likely to join routine or service occupations at every skill decile.

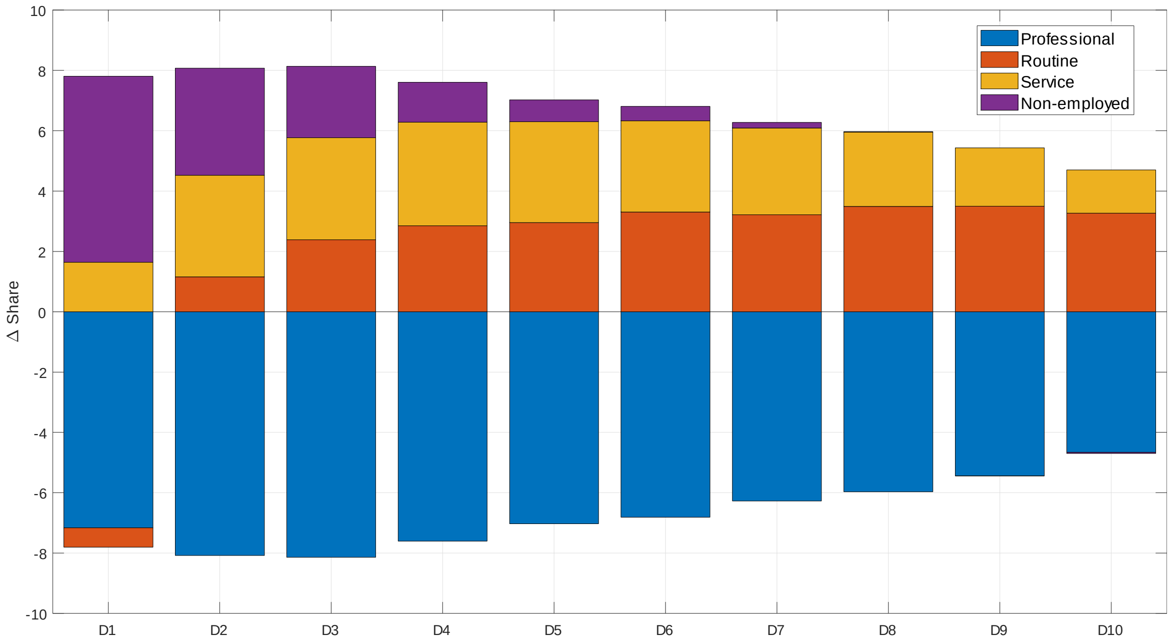
Figure 10: Occupational sorting by skill deciles



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles based on respective period.

Figure 11 provides a more detailed assessment of the differences in sorting patterns across the skill distribution. As indicated in figure 10, 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 -8% in the second skill decile to just over -4% in the 10th decile. Accordingly, the alternative career paths chosen by graduates also vary across the skill distribution: At the lower end of the skill distribution, graduates are more likely to be out of the labour market entirely, while between the 3rd and 5th decile they are increasingly likely to join service occupations. From the 6th decile onwards, the displaced graduates are increasingly moving into routine occupations instead. These patterns highlight that the observed changes to the labour market destinations of graduates are heterogeneous across the skill distribution and that therefore interventions to address perceived imbalances need to be carefully targeted.

Figure 11: Changes in occupational sorting by skill deciles



Note: Changes in occupation shares, based on a simulated sample of 1,000,000 graduates. Skill deciles based on respective period.

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

6 Counterfactual decompositions

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

There are three main structural forces that I have modelled and that I consider in this exercise: i) the skill distribution of young graduates (represented by μ and σ); ii) the occupation specific skill prices and returns to skill (represented by λ_o and η_o); and iii) the non-pecuniary component of utility (ω_o). Each of these three represents a potential mechanism for the changing labour market outcomes of young graduates. I will assess the importance of each channel by fixing all parameters except for those associated with the considered mechanism at their 2001-2010 values, only allowing the parameters of interest to take on their estimated 2011-2019 values.[13] I then simulate the model and compare the outcomes with those from the baseline model. In the following I will refer to these three counterfactuals as the *skills* counterfactual, the *prices* counterfactual and the *preferences* counterfactual respectively.

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

Table 1 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 2001-2010 and 2011-2019 based on the baseline model.[14] 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.

Baseline Model		Skills Counterfactual		Prices Counterfactual		Preferences Counterfactual	
Occupation Shares	Total Δ	Δ	Explained (%)	Δ	Explained (%)	Δ	Explained (%)
Professional	-6.58	-3.47	52.77	-0.71	10.80	-1.99	30.22
Routine	2.51	0.24	9.44	-0.54	-21.54	3.02	120.60
Service	3.03	0.35	11.71	0.58	19.10	1.61	53.01
Non-employed	1.05	2.88	274.46	0.67	63.97	-2.64	-251.02
Log Wages							
Mean	-0.07	-0.08	105.77	0.04	-58.01	-0.02	22.84
Mean Professional	-0.05	-0.08	151.98	0.05	-92.59	0.00	2.95
Mean Routine	-0.05	-0.05	114.59	0.03	-60.22	0.00	7.49
Mean Service	-0.03	-0.05	132.06	0.03	-79.58	-0.01	21.30

Note: Simulations based on a representative sample of 1,000,000 graduates.

Table 1: Counterfactual Decomposition

The baseline model suggests an average reduction of around 6.5% in the share of graduates in professional occupations between the two periods. A hypothesised driver of this change was the changing skill distribution of graduates. In the previous section we saw that the average skill level of graduates declined between 2001-2010 and 2011-2019, providing a plausible mechanism for the reduction in the share of professional occupations. The counterfactual simulation suggests that the shifting skill distribution explains around 50% of the decline in the share of professional occupations: if only the skill distribution had changed between the two periods, the observed reduction in professional occupations would be around 3.5%. As anticipated, the reduction in average skill levels reduces the number of graduates for whom it is profitable to join professional occupations, reducing the share by around half of the observed fall in the data. The remaining contribution comes from changes in prices (11%) and preferences (30%), suggesting that apart from changing skills, new attitudes towards different occupations are an important diver of the observed decline of professional occupations.

Both routine and service occupations saw an increase of around 3% between the periods. The fall in the skill level, accounts for around 9% and 11% of this increase respectively, as the re-sorting of graduates into these middle skilled occupations is only moderate under this counterfactual. Instead, the counterfactual predicts most graduates that fail to enter professional jobs, would choose not to participate in the labour market at all, increasing the share of non-employed by around 3% - almost 3 times the increase observed in the data. This suggests, that while a decline in average skill levels can be considered a major factor in the decline of professional occupations, it is unlikely to explain the increasing shares of routine and service occupations.

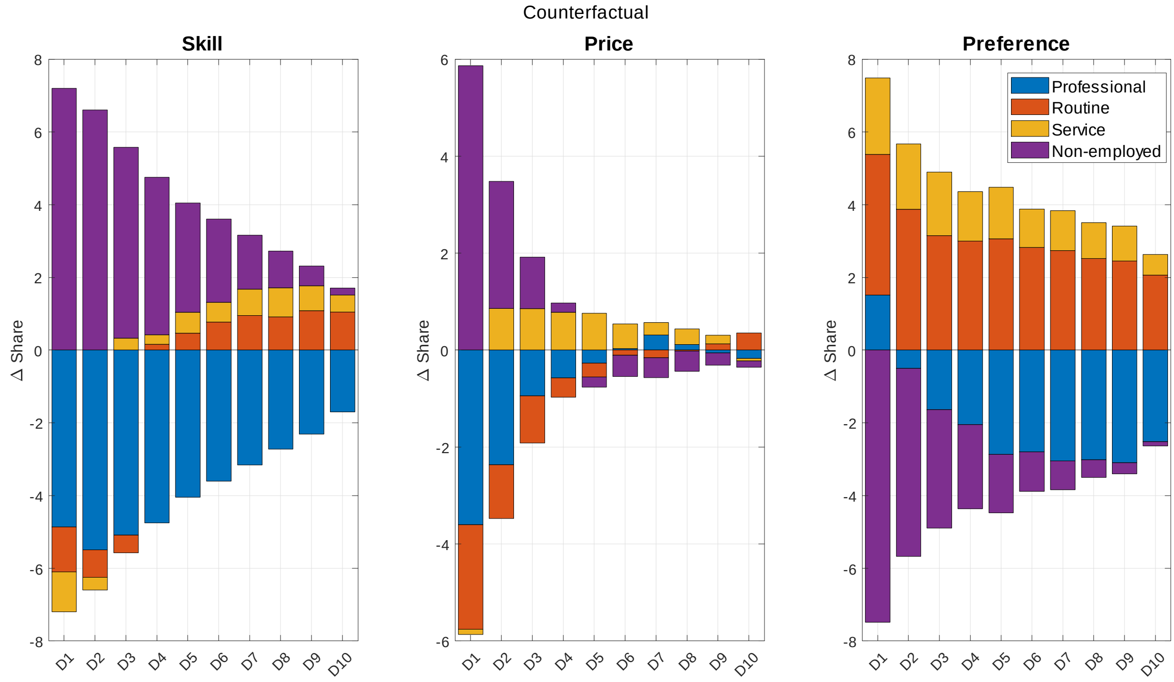
Turning to the prices counterfactual, we note that the contribution of prices to the decline in the share of professionals is moderate at 11% - an increase in the return to skill makes these occupations less attractive for graduates at the lower end of the skill distribution, while leaving the top graduates mostly unaffected. Quantitatively, however the effect does not appear very important. The counterfactual, however, explains around 20% of the increase in service occupations and over 60% of the increase in non-employed graduates. Graduates with relatively lower skill levels are discouraged from entering occupations, with a high degree of skill related compensation. A rise of variable compensation across all

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

occupations has the effect of pushing out graduates at the bottom of the occupation specific skill distribution, into those occupations with less skill-based compensation and ultimately out of the labour market all together. The total effect depends on the initial distribution of skills, and the level of skill prices in each occupation. For the current counterfactual the prediction suggests that the increase in skill returns pushed out graduates from professional and to some degree routine occupations, into service occupations, but particularly out of the labour market into non-employment. As with the skill counterfactual, price moves by themselves can explain - to some degree at least - some features of the changing occupation landscape among graduates, in this case the rise in non-employment, but it fails to account for other key phenomena, such as the increasing share of routine occupations.

Preferences play the main role in explaining the shift towards routine and service occupations, but also make a considerable contribution to the decline of professional occupations. The counterfactual simulation suggests that in the absence of any other changes, differences in the non-pecuniary preferences associated with the different occupations would account for 30% in the fall in the share of professional occupations and predict a 3% increase in the share of routine occupations, representing around 120% of the observed total change. The most dramatic impact, however, is on the share of non-employed which is predicted to decline by 2.6%, a large fall relative to the baseline prediction on a slight increase. As we have seen in the previous section, the models estimates for all occupations was increasing over the sample period, suggesting a fall in the value of the outside option. When remaining outside the labour market becomes less attractive more graduates are driven to seek employment, and since non-employed graduates are more likely to come from the lower end of the skill distribution, they have a higher chance of joining routine and service occupations. This by itself does however not explain the large increase in the share of routine and service occupations. Figure 12 outlines the counterfactual sorting patterns across the skill distribution. Subplot 3 suggest that under the preference counterfactual, graduates at all deciles of the skill distribution have a higher likelihood of entering routine and service occupations. This is caused by these occupations catching up to professionals in relative terms, indicated by the narrowing of the gap in non-pecuniary amenity values. This narrowing then can explain the remainder of the observed increase in the share of these occupations.

Figure 12: Occupational sorting under different counterfactuals 2011-2019



Notes: Change in occupational choice probability relative to baseline model. Based on a simulated samples of 1,000,000 graduates.

Turning to the decomposition with respect to wages, we can note that the skill counterfactual does a good job in explaining the overall fall of wages by around 7 log points. It however overstates the reduction among professionals by about 50%, while overshooting for routine and service occupations by around 10%. and 30%. Ceteris paribus, a reduction in the average skill level will reduce wages across the board, which is consistent with the patterns observed in the data. Also consistent is the observation that wages fall less in routine and service occupations relative to professional, since the sorting of graduates means that now relatively higher ability graduates will consider these occupations, as is indicated in the first panel of Figure 12. The price counterfactual suggests that wages should have increased across all occupations, which is contrary to the observed patterns in the data. Holding the skill level fixed, while increasing the returns to skill should have boosted average wages by 4 log points. Again, price changes by themselves do not appear to be a major explanation for the observed patterns in the data, but rather a factor that interacts with the distribution of skills and changes therein. Changes in preferences by themselves have only a minor effect on observed wages. This is because the only way in which they affect wages is by changing the sorting patterns across occupations, particularly by moving some graduates into the labour market who would otherwise be not employed. Since these are more likely to come from the lower tail of the skill distribution that negatively affects the average level of skill in the occupations they sort into.

7 Conclusion

The formation of human capital and the acquisition of specific skills lies at the heart of a university education. With more and more graduates in the UK failing to obtain "graduate jobs" the public is increasingly concerned that universities are failing to equip graduates with the skills they need to succeed in the labour market. But 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 evolving preferences that jointly determine the distribution of graduates across occupations.

My estimation has found that while the average level of graduates' skills has reduced by about 25% of a standard deviation, this explains around 50% of the decline in the share of graduates entering professional occupations. Sorting between graduates and occupations based on skill, means that those failing to get 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 increasingly sort into routine and service occupations, where an increase in variable compensation has made these occupations more attractive to graduates with higher levels of skill. This is complemented by a change in non-pecuniary preferences, particularly a reduction in the value of the outside option, which makes employment within these occupations a more attractive option.

The overall message of this paper is that no single trend - skills, prices or preferences - can easily account for the evolution of graduates' labour market outcomes. Instead, they are the outcome of intricate interactions between many different factors. This should provide a note of caution to all those interested in "solving" the graduate underemployment problem. While trying to align graduates' skills better with the demands of the labour market is likely a good idea it is important to consider where in the distribution these concerns are brought to bear. Giving already highly achieving students a leg up is likely not going to affect their chances at a graduate job all that much. Rather proper emphasis should be put on 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 demand and supply of skills is not only a matter of HEIs and education policy, but should also involve broader industrial policy stakeholders, to be able to anticipate future demands and meet them by providing solid foundations for future cohorts of graduates. Finally, while having a grad job can be very rewarding, not all underemployment is necessarily involuntary, jobs evolve over time becoming more or less attractive to graduates across a variety of dimensions not all of which are purely monetary.

This paper has tried to address a small part of a big research question, and as such necessarily leaves many questions unanswered, some of which suggest themselves as extensions or variants of the model explored in this paper: Firstly, 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, by including selection into university education is likely going to be a fruitful avenue for research. Secondly, while the increasing presence of graduates in non-traditional roles is well documented, I am not aware of any research showing how this influences these occupations. A larger potential pool of high skilled applicants might facilitate technology adoption within these occupations, leading to the increase in skill prices estimated for these occupations in this paper. Lastly, this paper has focussed on a narrow population - young university graduates. Extending the scope of the model to include other populations and sub-populations might provide interesting insights into the interplay of skill supply for different sections of society. I leave these and further questions for future research.

References

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

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

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

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

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

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

[7] Autor, David, and F. Levy. "R. Murnane, 2003, The Skill Content of Recent Technological Change: An Empirical Exploration." Quarterly Journal of Economics 118, no. 4.

[8] Bisello, Martina. "Job polarization in Britain from a task-based perspective. Evidence from the UK Skills Surveys." Department of Economics and Management, University of Pisa Discussion Paper 160 (2013).

[9] Blundell, Richard, David Green, and Wenchao Jin. "The UK education expansion and technological change." Institute for Fiscal Studies Working Paper (2018).

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

[11] Diegert, Paul. "The Role of Skills and Sorting in Explaining Wage Inequality." (2022).

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

[13] Goldin, Claudia, and Lawrence F. Katz. The race between education and technology. harvard university press, 2010.

[14] Gourieroux, Christian, and Alain Monfort. "Simulation-based inference: A survey with special reference to panel data models." Journal of Econometrics 59, no. 1-2 (1993): 5-33.

[15] Hajivassiliou, Vassilis A., and Paul A. Ruud. "Classical estimation methods for LDV models using simulation." Handbook of econometrics 4 (1994): 2383-2441.

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

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

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

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

[20] Hou, Shihang. "The Role of Uncertainty in Education Mismatch." (2022).

[21] Kinsler, Josh, and Ronni Pavan. "The specificity of general human capital: Evidence from college major choice." Journal of Labor Economics 33, no. 4 (2015): 933-972.

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

[23] Lee, Lung-Fei. "Asymptotic bias in simulated maximum likelihood estimation of discrete choice models." Econometric Theory 11, no. 3 (1995): 437-483.

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

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

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

[27] McFadden, Daniel. "The measurement of urban travel demand." Journal of public economics 3, no. 4 (1974): 303-328.

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

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

[30] Onozuka, Yuki. "Heterogeneous Skill Growth across College Majors." (2019): 1-58.

[31] Rohrbach-Schmidt, Daniela, and Michael Tiemann. "Changes in workplace tasks in Germany—evaluating skill and task measures." Journal for Labour Market Research 46, no. 3 (2013): 215-237.

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

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

[34] Sanders, Carl, and Christopher Taber. "Life-cycle wage growth and heterogeneous human capital." Annu. Rev. Econ. 4, no. 1 (2012): 399-425.

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

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

[37] Train, Kenneth E. "EM algorithms for nonparametric estimation of mixing distributions." Journal of Choice Modelling 1, no. 1 (2008): 40-69.

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

A Additional Tables & Results

A.1 Additional Tables

	Data Average	Simulation Average	Difference	% Difference
Log Wage				
Mean	2.54	2.52	-0.02	-0.66
Mean Professional	2.67	2.62	-0.05	-1.84
Mean Routine	2.28	2.34	0.06	2.66
Mean Service	2.13	2.22	0.09	4.25
Variance	0.15	0.17	0.02	11.27
Variance Professional	0.11	0.17	0.05	48.08
Variance Routine	0.11	0.11	0.00	1.43
Variance Service	0.09	0.08	-0.01	-9.17
P10	2.03	2.01	-0.01	-0.69
P50	2.56	2.49	-0.06	-2.44
P90	3.02	3.06	0.04	1.42
Occupation Shares				
Share Professional	0.62	0.61	0.00	-0.76
Share Routine	0.17	0.17	0.00	1.54
Share Service	0.09	0.10	0.00	3.74
Share Non-employed	0.12	0.12	0.00	-1.17

Note: Model simulations based on a sample of 1,000,000 graduates.

Table A1: Details of Model Fit

	2001-2010	2011-2019
Mean	0.00	-0.25
Median	-0.02	-0.27
Standard Dev.	1.00	0.97
P10	-1.27	-1.48
P25	-0.69	-0.92
P50	-0.02	-0.27
P75	0.66	0.40
P90	1.29	1.01
Mean Professional	0.24	0.03
Mean Routine	-0.07	-0.25
Mean Service	-0.23	-0.43
Mean Non-employed	-1.04	-1.27
StdDev. Professional	0.93	0.90
StdDev. Routine	0.90	0.87
StdDev. Service	0.89	0.85
StdDev. Non-employed	0.80	0.75
StdDev. Non-employed	0.89	0.68

Note: Skill distributions normalised to have mean 0 and standard deviation 1 in 2001-2010.

Table A2: Details of graduate skill distribution

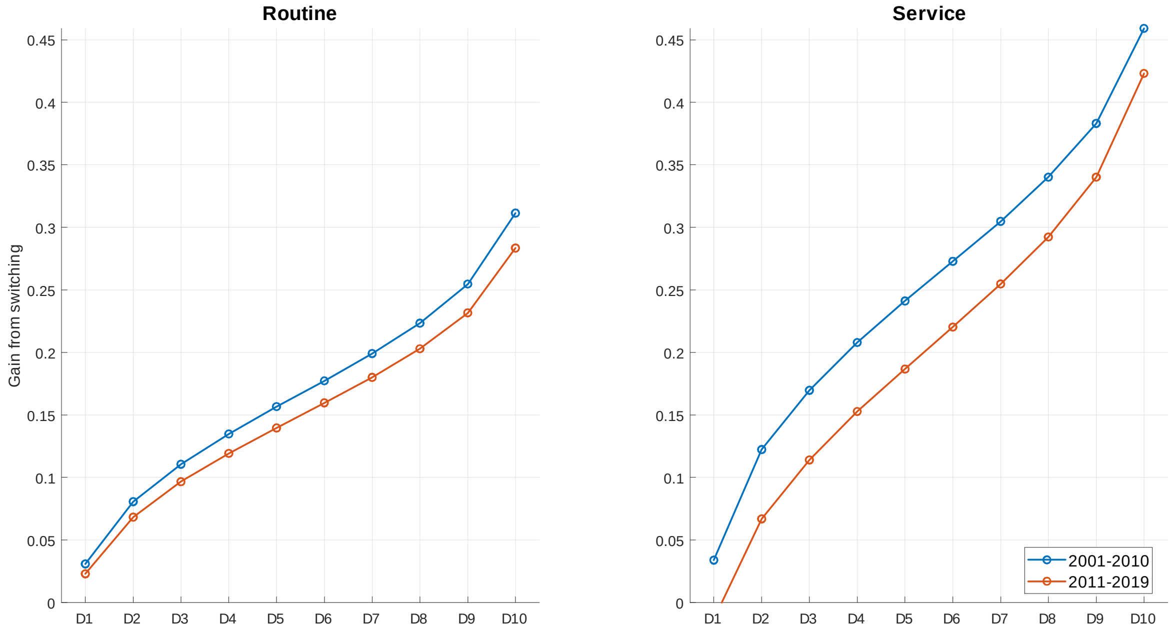
	Data				Baseline Model			
	2001-2010	2011-2019	Difference	% Difference	2001-2010	2011-2019	Difference	% Difference
Occupation Shares								
Professional	64.870	58.680	-6.190	-9.543	64.589	58.004	-6.584	-10.194
Routine	15.331	18.018	2.687	17.529	15.672	18.177	2.505	15.985
Service	8.062	10.995	2.933	36.377	8.370	11.399	3.028	36.180
Non-employed	11.737	12.307	0.570	4.859	11.369	12.420	1.051	9.241
Log Wages								
Mean	2.577	2.492	-0.085	-3.280	2.554	2.482	-0.072	-2.809
Mean Professional	2.701	2.630	-0.071	-2.641	2.642	2.591	-0.052	-1.958
Mean Routine	2.284	2.265	-0.020	-0.868	2.358	2.311	-0.047	-1.998
Mean Service	2.134	2.131	-0.003	-0.119	2.240	2.205	-0.035	-1.545

Note: Comparision of average occupation shares between two periods. Model simulations based on a sample of 1,000,000 graduates.

Table A3: Period Comparision

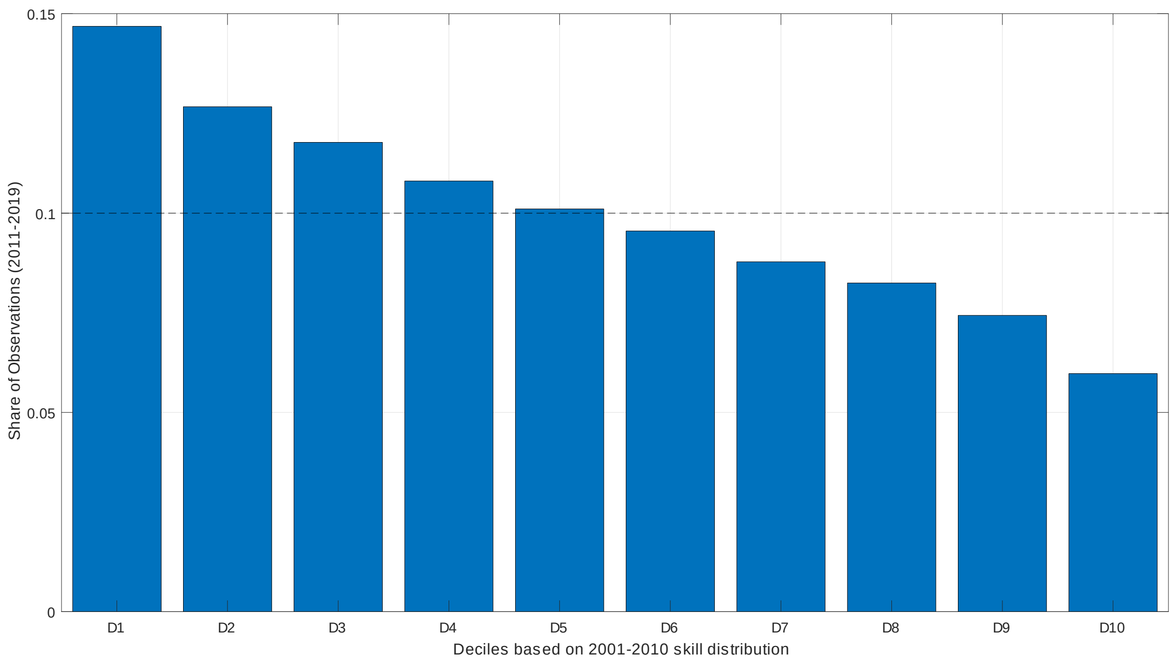
A.2 Additional Figures

Figure A1: Wage gains from switching to professional occupations



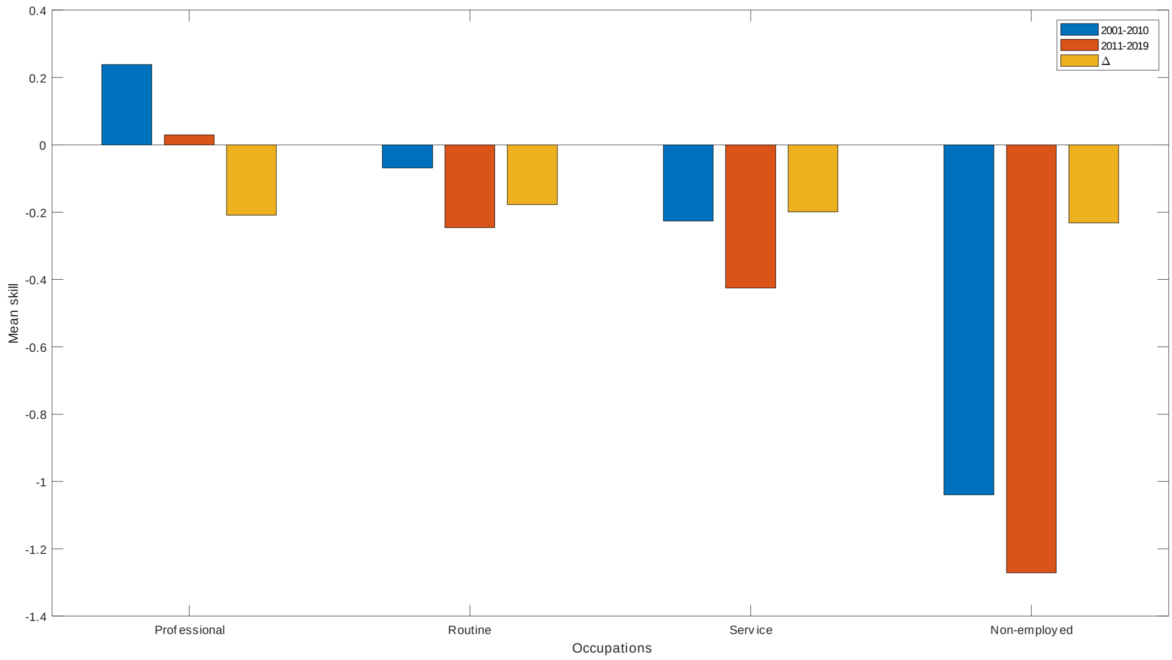
Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles based on respective period.

Figure A2: Histogram of 2011-2019 skill distribution



Note: Based on a simulated sample of 1,000,000 graduates. Skill deciles based on 2001-2010 distribution.

Figure A3: Average skill levels by occupation



Note: Based on a simulated sample of 1,000,000 graduates. Distributions normalised to have mean 0 and variance 1 in 2001-2011.

B Technical Details

B.1 Imputation of missing wage values

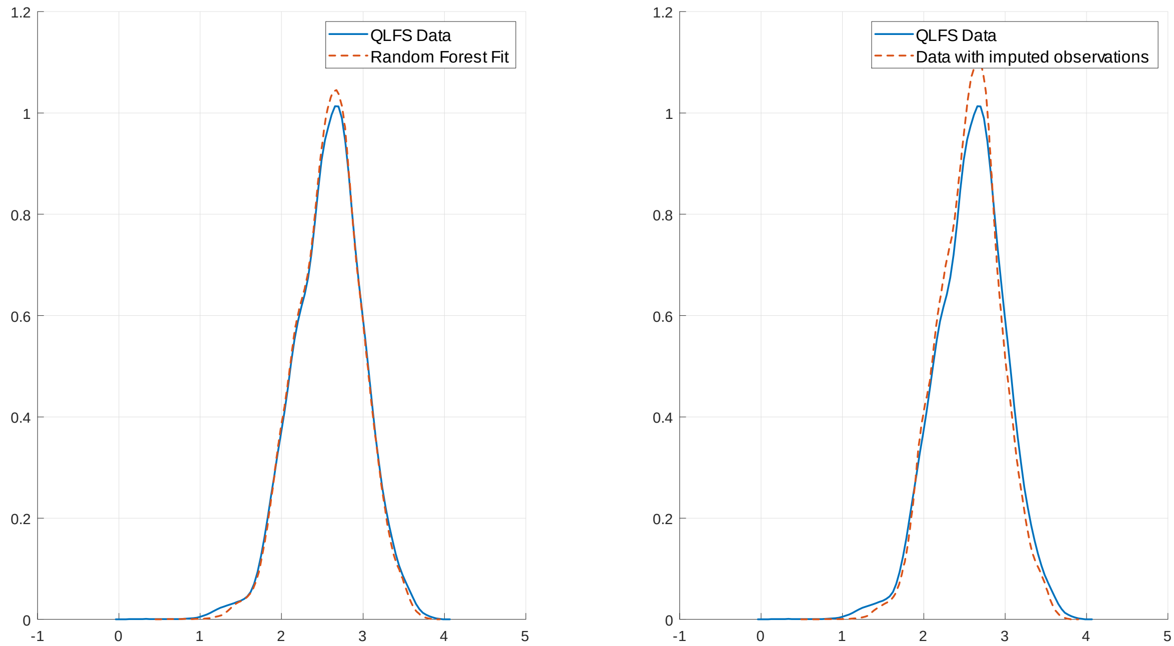
In the QLFS, wage data is only collected during the first and last interviews, resulting in three missing wage observations for each individual. To impute these missing wage values, I employ a two-step methodology that leverages both fixed effects to account for individual-specific skills and a Random 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 imputatiuon 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 teh imputation procedure. The first panel depicts the observed wages, as well as the predictions based on the random forrest 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.

B.2 Estimation algorithm

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

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

Table B1: Summary of Model Parameters

The only complication, that arises here comes from the fact, that we do not have a closed form solution for the joint probability (8) and thus have to evaluate the integral via simulation. This can be done by taking draws from the distribution of s, evaluating $\Pr(o_i^*, w_i^{obs}|s_i)$ at each of these draws and then averaging over the results. Since we are dealing with a panel of graduates, $\Pr(o_i^*, w_i^*|s_i)$ here denotes the conditional probability of a sequence of 5 occupation choices o_i^* and observed wages w_i^*. Standard results suggest, that as long as one uses a large enough number of draws to approximate the integral, the Maximum Simulated Likelihood Estimation (MSLE) is asymptotically equivalent to classical Maximum Likelihood Estimation (MLE) (c.f. McFadden & Train (2000)). For a proof that the MSL estimator is unbiased and efficient see the appendix.

Denote the simulated counterpart of (8) by $\Pr^{sim}(o_i^*, w_i^*)$ for simplicity, and let θ be the set of our parameters, we can write down the simulated log likelihood function of the as:

$$ll^{sim}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\sum_{o=1}^{O}\mathbf{1}_{(o=o^*)} \ln \overset{sim}{\Pr}(o_i^*, w_i^*) \tag{9}$$

and we can estimate θ as:

$$\hat{\theta} = \arg\max_{\theta} ll^{sim}(\theta) \tag{10}$$

So to specify the complete algorithm:

1. Set $p = 1$ and make a guess for $\hat{\theta}_1$. Specify a tolerance criterion ϵ. Set R, the number of draws used to approximate the integral to a reasonably high number.
2. For each individual i, given $\hat{\theta}_p$ draw a vector of s_i, R times, denoting each as s_i^r.
3. For $r = 1$ to R :
   (a) For $q = 1$ to 5 : Calculate $\nu_{iq}^r = w_{iq}^* - \left[\eta_{o^*} + \lambda_{o^*}^{'} s_i^r\right]$. For a given pair s_i^r, ν_{iq}^r calculate $\Pr_q^r(o_{iq}^*, w_{iq}^*)$.
   (b) Calculate $\Pr^r(o_i^*, w_i^*)$ as $\prod_{q=1}^{5} \Pr_q^r(o_{iq}^*, w_{iq}^*)$
4. Average over all R values of $\Pr^r(o_i^*, w_i^*)$ to obtain:
   $\Pr^{sim}(o_i^*, w_i^*) = \frac{1}{R}\sum_{r=1}^{R} \Pr^r(o_i^*, w_i^*)$.
5. Repeat steps $2 - 4$ for all N individuals. Calculate the log likelihood via (9) denoting it as ll_p^{sim}.
6. If $|ll_p^{sim} - ll_{p-1}^{sim}| < \epsilon$, terminate here. Otherwise increment p and find a new value $\hat{\theta}_p$ and repeat from step 2.

For the numerical evaluation of the integral I use a grid of $1,000$ quasi random Halton draws, which have been shown to provide about an order of magnitude more accuracy than simple random draws (Train (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 $1e^{-6}$.

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^θ that span the parameter space θ. For each point in G^θ 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 point K^θ. 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 K^θ to maximize the log-likelihood function.

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 neighborhood 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.

B.3 Standard errors

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

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

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.

B.4 Alternative interpretation of ϕ

In the main section 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 be accepted by some readers.

To illustrate let us return to the joint probability (8):

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

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

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

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