To what degree? Recovering changes in the UK's graduate skill distribution

Max Schroeder
University of Glasgow

October 11, 2022
latest version

Abstract

University graduates have very differentiated skills, both compared to the general population and to other graduates. Differences arise from differences in background, course of study and individual aptitudes and interests. In this paper I study the distribution of these different skills, investigating what types of skills graduates have, and how these vary between and within broadly defined subject groups as well as across time. To this end I develop a model of occupational choice and wage determination for university graduates in the UK. Graduates differ with regards to two types of general skills: mathematical/technical and verbal/organisatorial, which are used with different intensities by different occupations in the labour market. I structurally estimate the model to find evidence of changing multivariate skill distributions over time. I find that between 1994 and 2019, the typical graduate's level of mathematical skills increased by 140% while verbal skills decreased by close to a third. While most subject categories experience increases in mathematical skills, there are particularly large gains for Business & Economics and Arts & Humanities degrees. Overall inequality in mathematical skills has also fallen significantly. This trend is suggestive of the success of governmental efforts to improve STEM education at all levels throughout the early 2000's.Counterfactual simulations suggest that in the absence of changes to the subject specific skill distributions, mean wages would be up to 8% lower, while wage inequality would be up to 5% larger. The results suggest that graduate skill supply has adjusted to changing labour market requirements.

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 Programm in Economics (2021), the Society for Computational Economics (2021), the Workshop on Applied Economics of Education (2022) and the Annual Conference of The European Association of Labour Economists (2022) for their insights. All remaining errors are my own.

1 Introduction

It is generally believed that a degree confers (or signals) certain specific skills and abilities. This belief is reflected in the economic literature on the subject of higher education - through the language of *human capital* - as well in the language that Higher Education Institutions use to describe their own function - via terms such as *graduate skills* or *attributes*. Despite this central position, there are few studies that describe what skills, or how much of a specific skill, graduates posses when they finish their university education. At a time where the value of a university degree is coming under increased scrutiny,[1] such quantitative evidence would be very valuable to prospective students, Higher Education Institutions, government departments and employers.

For about a century, a university degree has been seen as a secure route to wealth and professional success. Over the course of mastering their chosen subject, students generally acquire different skills that help them succeed in an increasingly skill biased labour market (c.f. Goldin & Katz (2009)). Recent studies emphasise the importance of subject of study ("college major") for determining the return to university for a given individual (e.g. Altonji *et al.* (2016), Andrews *et al.* (2022) and Lovenheim & Smith (2022)). Given the wide range of possible subjects students can study, it is natural to assume that differences in subject specific outcomes are due to the different types of skills that are taught in these courses and how these are rewarded in the labour market. So what types of skills do graduates have? What are the differences between an economics graduate and a medical doctor? And given large changes to the structure of the education system and the wider economy, has the distribution of different skills changed over time?

This paper provides an attempt to quantify the distribution of skills amongst recent university graduates in the UK over the last 25 years. Focussing on two relevant types of skills - mathematical/technical and verbal/organisatorial - I provide estimates of subject specific skill distributions allowing a quantitative assessment of the skills of a typical graduate as well as the degree of skill inequality between graduates. Taking a long view this paper provides separate estimates for three time periods, covering the period from 1994 until 2019, allowing an evaluation of how the graduate skill distribution has evolved during a period of significant changes in both the primary, secondary and tertiary education sector and the wider labour market. Over the same period the UK (and other developed economies) have experienced a rapid expansion of tertiary education. Since the passage of the Further and Higher Education Act 1992, university enrolment has roughly doubled to approximately 2 million in recent years. The fact, that these increases were sustained in the face of stark tuition fee increases[2] suggests that a university degree is still seen as a profitable investment by many, but rising graduate earnings inequality and underemployment (c.f. Altonji et al. (2016), Holmes & Mayhew (2016), Lindley & MacIntosh (2015)) cast some doubt on this perception. Generally, it is not clear, how the higher education sector has coped with the rapid expansion of demand and whether a degree still confers the same benefits as it did 30 years ago.

Since direct measures of graduates' skills are lacking in all but the most detailed

[1] Lately, in the UK and the US, the public discourse has revolved around the value of degrees which - according to some - do not provide relevant skills to students, resulting in high drop out rates, or poor labour market outcomes for graduates. As a response the UK government has indicated that funding for such courses might be reduced or cut completely.

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

surveys, specific 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. Hence, by specifying and estimating a corresponding structural economic model, we can make inferences about the unobserved skill endowments of university graduates only using widely available data sources.

To quantify the variation in unobserved skills, I develop a model of occupational choice for university graduates. After graduating, graduates differ with regards to their idiosyncratic endowment of two types of general skills: mathematical/technical and verbal/organisatorial. Skill endowments are modelled as draws from subject specific multivariate distributions and are thus allowed to vary between as well as within university subjects, capturing important dimensions of heterogeneity: Differences between the distributions capture differences in emphasis due to the specific subject, while each distribution encompasses a further degree of heterogeneity resulting from differences in university quality as well as inherent differences due to individual ability and aptitude conditional on subject choice. Each university subject is therefore characterised by a multivariate distribution function, and each graduate by the skill endowment which they have drawn from this distribution.

The purpose of acquiring skills is to use them in the labour market, but just like all graduates have different skills, occupations vary in what value they assign to different skills and hence the match between graduate and occupation matters for realised productivity and wages. The production sector follows the standard approach in the task-skill literature (c.f. Autor et al. (2003), Autor & Handel (2013), Sanders & Taber (2015)), and features a multitude of occupations that differ with respect to how intensely they use each type of skill in production. The combination of a worker's skills and the work task requirements of an occupation determine the worker's occupation specific productivity. Upon graduation, graduates choose their preferred occupation taking into account their idiosyncratic skill endowment as well as other preferences. It is precicely the correlation between a graduate's skills and their chosen occupation match that allows the model to be estimated with meaningful results.

I use a sample of recent university graduates from 1994-2019 together with occupation level information on work tasks, to structurally estimate the model using simulated maximum likelihood, and recover the parameters of the underlying latent skill distributions for different fields and time periods. To ensure robustness of the estimates I control for a variety of potential factors that might affect changes in skill demand. I then use the model estimates to analyse changes in the graduate skill distribution and their effects on the labour market outcomes of university graduates over two and a half decades.

The results suggest that over the time period there have been substantial changes in the distribution of skills. Over the time period, I find that the median graduate's endowment of effective mathematical/technical skills increased[3], by roughly 140%; effective verbal/organisatorial skills decreased by around a third. Across 5 major subject categories all but one - Medical and Life Sciences - saw their median mathematical/technical skills rise, most notably Business & Economics which appears to have become increasinly

[3]Beginning in the early 2000's the UK government began to strongly encourage Science, Technology, Engineering and Mathematical education at all levels of education through initiatives like the Science and Innovation Investment Framework. It is therefore not entirely surprising that these policies had a large impact on the mathematical and analytical skills of graduates from the mid-2000's onwards.

focussed on technical analysis, but also subjects in the Arts & Humanities. Correspondingly most subject categories have lost some of the verbal/organisatorial skills that used to be associated with them. Overall the increase of more technical skills has counteracted the fall in the typical graduates endowment of verbal skills, suggesting a reorientation of skill supplies in line with more demand for technical abilities.

To put these changes into perspective: in the mid 90's on average only 23% of a graduate's wage could be attributed to their mathematical/technical abilities, while verbal/organisatorial skills accounted for around 48%. In the period just before the Covid-19 pandemic, these shares had effectively reversed, with mathematical skills accounting for around 43% of hourly wages, while verbal skills only contributed around a third.

In terms of distributions, overall mathematical/technical skills inequality, as measured by the Gini coefficient decreased from 53 to 38 points, verbal skill inequality changed form 24 to 46 points. For STEM and Business & Economics, rising mathematical/technical skills were associated with shrinking within group skill inequality, while the converse held for Arts & Humanities and other degrees.

Together these changes had nontrivial effects on the labour market outcomes of university graduates. In counterfactual simulations I find that in the absence of changes to the subject specific skill distributions, mean wages would be up to 8% lower whereas wage inequality would be up to 5% larger than what is observed in the data. Additionally, I find that changes in the demographic composition of graduates had only small effects on overall labour market outcomes.

This paper adds to a large, and growing literature on the returns to higher education and specifically to a subset of this literature that investigates the return to specific fields of study (see Altonji et al. (2016), Andrews *et al.* (2022) and Lovenheim & Smith (2022) for extensive surveys). Generally, these studies estimate various latent average treatment effects, whilst trying to address the inherent difficulties caused by the existence of selection effects across dimensions of inherent ability and preference 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)). This paper retains some uniqueness by making the skills that graduates poses, the key feature of interest, thereby allowing an assessment of the mechanism underlying the returns to different fields of study.

Further, this paper contributes to the literature on how endowments of different types of skills affect labour market outcomes of graduates in an environment where occupations have different skill requirements. These papers tend to focus on the dichotomy between more general (transferable) and more specific skills leading to differences in the risk-return profiles between general and specialised degree subjects (c.f. Leighton & Speer (2020), Onozuka (2019)). Of particular importance here is the paper by Kinsler & Pavan (2015), which estimates a structural model where students acquire mathematical and verbal skills the return of which differ according to their occupation. The modeling approach taken in this paper is necessarily different from my own, but they are related in spirit.

Finally, this paper complements other attempts at eliciting the skill content of different 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 in so far as that it uses both occupation choice and wage information for estimation as well as allowing for substantial within subject skill heterogeneity. However, it shares the formers conception of a university degree as a bundle of multidimensional skills that are related to different tasks.

To the best of my knowledge this paper is the first attempt at trying to find quantitative evidence for the actual distribution of skills of university graduates in the UK. The results suggest that skill heterogeneity plays a large role in explaining the changes in the labour market outcomes of university graduates. Graduates differ in their skill endowments in accordance with the subject that they choose to study and beyond. Furthermore, the distribution of graduate's skills is changing over time meaning that graduates today might look very different from those 10-25 years ago. This finding has important implications for educational and more general economic policy going forward.

The rest of the paper is structured as follows: section 2 presents the economic model of wage setting and occupational choice; section 3 presents the econometric strategy, used to estimate the parameters of interest; section 4 presents the data sources used in the analysis; section 5 highlights the estimation procedure; section 6 covers the result; section 7 presents counterfactual experiments and section 8 concludes.

2 Model

In the following I will present an economic model of occupation choice and wage determination for recent university graduates in order to recover the skills supplied by university degrees. The economic environment in this model closely follows the literature on estimating task returns (c.f. Autor & Handel (2013), Roys & Taber (2016)). Whereas for occupation choice, I follow the methodological approach of the multinominal choice literature, where it is common to estimate unobserved parameters from the observed choices of individuals. In particular I will refer to the class of mixed logit models which seem to be particularly relevant in this context (see Train (2009), Chapter 6). For expositional simplicity, both parts are presented separately, before being combined in the next section.

2.1 Wage Determination

A worker'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 worker i is at performing task k. The exact distribution of s will be further specified below, but for now we will just take s as given.

On the firm side, the labour market consists of a large number of competitive firms of different types (henceforward referred to as "*occupations*") that 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.

A worker's human capital therefore depends on her skill-set as well as the task-productivity vector of her chosen occupation. Specifically, the human capital of worker i in occupation o is defined as:[4]

$$h_{io} = e^{\sum_K \lambda_{ok} s_{ik}} \tag{1}$$

Denote the aggregate amount of human capital in occupation o as H_o. Finally, output is produced by an aggregate production function:

$$Y = F(H_1, ..., H_O) \tag{2}$$

The marginal product of worker i in occupation o is therefore:

$$\frac{\partial Y}{\partial h_{io}} = \frac{\partial F}{\partial H_o}\frac{\partial H_o}{\partial h_{io}} = \frac{\partial F}{\partial H_o} e^{\sum_K \lambda_{ok} s_{ik}} \tag{3}$$

Denote $\frac{\partial F}{\partial H_o} = e^{\eta_o}$,[5] and assume that firms pay workers their marginal product, then the log wage of worker i in occupation o can be written as:

$$w_{io} = \eta_o + \sum_K \lambda_{ok} s_{ik} \tag{4}$$

This setup is fairly standard in the literature on tasks and skills (c.f. Autor & Handel (2013), Roys & Taber (2016)).

[4]For this exposition I am going to ignore any other factors that might influence productivity such as worker specific characteristics. Including these is a trivial extension of the model.

[5]Depending on your preferences, you might want to interpret η_o as a occupation specific demand component, or a occupation fixed effect.

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 prices" (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, simply running an OLS regression on equation (4) will not do the trick (see Autor (2013)).

In this paper however, rather than being harmful, this self selection is actually helpful as it allows us to make inferences from a worker's observed occupation to her unobserved skill-set. In order to enable this inference, we will first have to model the discrete choice behaviour of the worker directly.

2.2 Occupational Choice

A discrete choice framework presupposes, that individuals are able to assess different options that are available to them and then pick the preferred one. Here workers observe their skills, and all potentially relevant characteristics of an occupation and pick whichever occupation provides them with the highest valuation in terms of utility. In this case suppose that every graduate can observe the set O of all available occupations and attach a personal (in utility terms) valuation V_{io} to each of these options. Accordingly a worker i solves the following (static) occupational choice problem:

$$V_i = \max_{o \in O} \{V_{io}\} \tag{5}$$

Under these circumstances the individuals occupation choice o_i^* will refer to the "best" available option:

$$o_i^* = \arg\max\{V_{io}\} \tag{6}$$

In the following we will make some assumptions about the different parts affecting the worker's utility V_{io} which allow us to estimate the unobserved characteristics that we are interested in. Let us assume that the utility derived from the occupation is linear in the log wage,[6] leading to the following relationship:

$$V_{io} = w_{io} + \varepsilon_{io} \tag{7}$$

where o is one of the available occupations, w_{io} is the log wage earned by i in occupation o and ε_{io} is an individual-occupation-specific preference shock that is **i.i.d.** across all agents and all occupations.

Generally, models of this form are known as "random utility models" (RUM), since the worker's valuation of the different options V_{io} can be broken up into a "deterministic" part, w_{io} and a "random" part, ε_{io}. Importantly, the value of V_{io} is perfectly observed by the economic agent, while only o_i^* is observed by the econometrician.

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

[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.

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 preferences or frictions in the labour market.

3 Econometric Strategy

My econometric strategy combines the empirical content of the two parts 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, provided that we also have some information about the occupation task vector λ. In the later parts of this paper I will describe how this can be obtained from empirical data using standard procedures, but until then I will just take it as given.

First, I will model the worker's occupation choice using a multinominal choice model.[7] Secondly, by adding the realized wage, we not only gain additional information, but also a very nice economic interpretation for our estimates.

3.1 A mixed logit model of occupational choice

Let us recall the problem our graduate is facing. She knows her own skill set s_i, as well as the the task vectors of all occupations λ, as well as the occupation specific parameters η, and therefore perfectly knows her log wage in every occupation o:

$$w_{io} = \eta_o + \sum_K \lambda_{ok} s_{ik} \tag{8}$$

She also perfectly knows her preferences over the non-pecuniary aspects of each occupation ε_i, and is therefore able to assign to each occupation a personal valuation V_{io}. I assume, that V_{io} takes the following form:

$$V_{io} = w_{io} + \varepsilon_{io} \tag{9}$$

Finally, given this valuation the graduate chooses her preferred option:

$$o_i^* = \arg\max\{V_{io}\} \tag{10}$$

Making the standard assumption that her idiosyncratic occupation preference shocks ε_i are distributed i.i.d. Type I Extreme Value, we can express the conditional choice probability of her chosen occupation o_i^* as:

$$\Pr(o_i^*|s_i) = \frac{e^{w_{io^*}}}{\sum_{o=1}^{O} e^{w_{io}}} = \frac{e^{\eta_{o^*}+\sum_K \lambda_{o^*k} s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok} s_{ik}}} \tag{11}$$

Now, if we assume that s_i was drawn from a parametric distribution, then it is possible to identify and estimate the parameters of this distribution using a reasonable sample of individuals.[8]

[7] The following discussion presumes some familiarity with discrete choice and in particular the class of logit models. For those interested in a refresher I have added short exposition in the appendix.

[8] Assuming a parametric distribution for s turns this into a mixed logit model (see Train (2009), Chapter 6), where we are effectively treating skills as random taste parameters over the different tasks.

Let us assume that each skill vector comes from a multivariate log-normal distribution with mean μ and variance-covariance matrix Σ:

$$\log(s_i) \sim MVN(\mu, \Sigma) \tag{12}$$

the log-normal is a convenient choice here, as it ensures strictly positive support for the skill-set s, which seems like a reasonable choice for our purposes.

Using this parametric assumption, we can derive the unconditional choice probability by integrating over the distribution of s :

$$\Pr(o_i^*) = \int \frac{e^{\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o + \sum_K \lambda_{ok} s_{ik}}} f(s) d(s) \tag{13}$$

Standard results (c.f. McFadden & Train (2000)) guarantee, that we can use the unconditional choice probability in (13) to get consistent estimates for η, μ and Σ, using simulated[9] maximum likelihood. The mixed logit is an extremely flexible choice model that can indeed approximate any random utility model (c.f. McFadden & Train (2000)). Most interesting for researchers is that it naturally generates correlations in choice behaviour across similar alternatives. For example, a worker with a particular large value of some skill is going to prefer all occupations that use this skill with great intensity.

However, the model is not complete yet. As of yet there is nothing distinguishing our "skill" interpretation of s from a "taste" interpretation. Indeed, strictly speaking we would have to provide a location normalization for one of our parameters, in order to fix their relative values. In the following I will use the observed wage to address the last two points.

3.2 Adding wage information

So far the model has already made use of the wage setting equation (4), but for any draw of s_i a worker's modeled wage w_i differs from the worker's realized (observed) wage w_i^{obs},

due to the presence of other factors such as individual effort and luck. I capture these elements by adding an additional error term to the wage equation:

$$w_i^{obs} = w_i + v_i \tag{14}$$

where v_i is a random, mean zero disturbance, **independent** of the workers occupation choice. As such v_i does not impact the graduate's occupation choice, as can be easily derived from the analytic form of the occupation choice probabilities. To see this, add v_i to all potential wage outcomes w_{io}, then the logit formula implies:

$$\Pr(o_i^*) = \frac{e^{w_{io^*} + v_i}}{\sum_{o=1}^{O} e^{w_{io} + v_i}} = \frac{e^{v_i} e^{w_{io^*}}}{e^{v_i} \sum_{o=1}^{O} e^{w_{io}}} = \frac{e^{w_{io^*}}}{\sum_{o=1}^{O} e^{w_{io}}} \tag{15}$$

Hence, as long as v_i does not vary across different potential occupations, the choice probabilities remain unaffected.[10]

[9] There is no closed form solution for this integral, but integration step can be performed via simulation.

[10] The same is also true if ν_i varies across occupations, but is unanticipated by the graduate at the time she chooses her occupation.

Let us assume, that $\nu_i \sim N(0, \phi^2)$. Now it is clear that given a specific occupation choice o_i^*, and skill-set s_i, we can calculate the size of this additional error:

$$\nu_i = w_i^{obs} - \left[\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}\right] \tag{16}$$

The specific interpretation here is that ν_i gives us a measure, of how far the wage implied by the model parameters, is from a workers actual observed wage. Jumping ahead a little, we should expect the *true* model to minimize this distance.

Since ν_i is normally distributed, we have a closed form expression for the conditional probability of observing the observed wage, conditional on a certain skill set s_i and occupation choice o_i^*:

$$\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}} = \Pr(w_i^{obs}|s_i, o_i^*) \tag{17}$$

Ultimately, we are interested in finding the set of parameters, that maximizes, the unconditional joint probability that a worker chooses the occupation that she is observed choosing and that she earns the wage that she is observed earning: $\Pr(o_i^*,\, w_i^{obs})$.

To find the correct expression, we first rewrite $\Pr(o_i^*,\, w_i^{obs})$ as

$$\Pr(o_i^*, w_i^{obs}) = \int \Pr(o_i^*, w_i^{obs}|s_i) f(s) d(s) \tag{18}$$

using the law of conditional probabilities to rewrite:

$$\frac{\Pr(o_i^*, w_i^{obs}|s_i)}{\Pr(o_i^*|s_i)} = \Pr(w_i^{obs}|s_i, o_i^*) \tag{19}$$

$$\Pr(o_i^*, w_i^{obs}|s_i) = \Pr(o_i^*|s_i) * \Pr(w_i^{obs}|s_i, o_i^*) \tag{20}$$

Plugging the expression back in gives us:

$$\Pr(o_i^*, w_i^{obs}) = \int \Pr(o_i^*|s_i) \Pr(w_i^{obs}|s_i, o_i^*) f(s) d(s) \tag{21}$$

.

From (11) we know that:

$$\Pr(o_i^*|s_i) = \frac{e^{\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o + \sum_K \lambda_{ok} s_{ik}}} \tag{22}$$

and hence we can combine to write:

$$\Pr(o_i^*, w_i^{obs}|s_i) = \left(\frac{e^{\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o + \sum_K \lambda_{ok} s_{ik}}}\right)\left(\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right) \tag{23}$$

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

$$\Pr(o_i^*, w_i^{obs}) = \int \left\{ \left(\frac{e^{\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o + \sum_K \lambda_{ok} s_{ik}}} \right) \left(\frac{e^{(-\frac{\nu_i^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}} \right) \right\} f(s)d(s) \tag{24}$$

This completes the econometric specifications. By including the wage information, we have not only anchored the parameters in the occupational choice model, but also provided a neat economic interpretation for our "random coefficients" - the skill vector s. 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 (24).

3.3 Model extensions

3.3.1 Other demographic characteristics

For an empirical application, it is necessary to control for a number of observable characteristics, as well as year fixed effects. However, it is trivial to extend the model to include factors other than the skills considered above. To show this, I extend the log wage equation below:[11]

$$w_{io} = \eta_o + \sum_K \lambda_{ok} s_{ik} + \beta x_i \tag{25}$$

where x_i is a vector of observable characteristics (gender, labour market experience, etc.), β is a vector of coefficients. Clearly this equation can be inserted into the likelihood function (24), and β can be estimated as part of an extended parameter vector θ. Further, as long as neither x_i, nor β, vary across occupations (i.e. the model does not include for example either occupation specific experience (occupational tenure) among the observables, nor occupation specific coefficients in β), the additional terms do not have any impact on the occupational choice probabilities, and can therefore be ignored in the first part of the likelihood calculation.

3.3.2 Non-pecuniary aspects of occupational utility

In order to make the model more realistic, I also include non-pecuniary aspects of occupations that might affect the graduate's choice. Specifically, I augment the graduates expected payoff from choosing occupation o by a non-random occupation preference term ω_o, which is constant for all graduates and represents the (dis-) utility of working in a specific occupation. The augmented occupational valuation equation thus reads as follows:

$$V_{io} = w_{io} + \omega_o + \varepsilon_{io} \tag{26}$$

Like β, ω can be estimated as part of the extended parameter vector θ. Since ω does only affect the occupational choice probabilities, it can be ignored in the wage equation part of the likelihood function

[11] Naturally, this can be understood as an extension of the human capital equation specified above.

4 Data

4.1 Graduates

The main data source that I am using is the Quarterly Labour Force Survey (QLFS) over the period 1994-2019, which I split into three periods: 1994-2002, 2003-2011 and 2012-2019. Since 1994 the QLFS has included reasonably fine grained information on the subject of an individuals first university degree (see Lindley & MacIntosh (2015) for more details).[12] Since I am particularly interested in differences between subjects this will be a key variable of interest. Furthermore, the QLFS also contains information on an individuals current occupation, usual hourly pay and some other demographic covariates.

I restrict the sample to full time working graduates between the ages of 21 and 27, who have not graduated more than 2 years before I observe them in the sample.[13] This age restriction is put in place to make sure that we capture those graduates who are "fresh" out of university, so that their skill-set most accurately reflects their post-university endowment. A small age bracket also reduces contamination by other factors such as age and experience effects as well as on the job skill accumulation.

For each graduate in my sample I collect wages measured as usual gross hourly pay, deflated by CPI;[14] their current occupation as classified by the 1-Digit SOC Occupational classification schedule[15]; Gender; Subject of first degree; and years since graduation, which I use as a proxy for labour market experience.

I split the sample into 5 groups according to broadly defined subject degree categories: 1. Medical and Life Sciences (including Biology & Agriculture); 2. Science, Engineering, Technology & Mathematics; 3. Business Management and Economics; 4 Arts & Humanities; 5. Other Degrees. In order to avoid complications I drop all those who studied in a combined degree, as well as all those who hold any advanced degrees beyond the undergraduate level. A table summarizing the resulting sample can be found in Appendix A.

The total sample includes 10,660 graduates, with around 3,500 individuals in each time period. Over the 25 year period the most significant change is shown by the composition of subjects represented amongst graduates. STEM, Business & Economics and Arts & Humanities loose a proportion of their graduates, while Medical & Life Sciences and the category of other degrees gain in relative popularity. The differential growth of different degree subjects might also be reasonably linked to broader demographic factors, such as for example increasing participation of women in higher education as well as in the

[12] I ignore those who have more than one degree, or any further or higher degrees. Postgraduate qualifications take on a more significant role over time, as a higher percentage of graduates pursue these degrees. However, a large fraction of graduates pursues a postgraduate qualification in a subject different from their first degree, making it difficult to assign the skills they exhibit to their first or later degree. I decided to exclude postgraduates in order to keep the relationship between degree subject and skills as clean as possible. However, including higher degree holders without any further changes to the model does not change the quantitative predictions of the model with regards to skill supply or inequality (see appendix).

[13] Typically in the UK students finish highschool at 18 and enter 3 year University Courses.

[14] I also trim the top and bottom 1% of wage values to remove nonsensical values.

[15] Where necessary use crosswalks to transform other occupational classifications

	Mean hourly wage			Gini hourly wage			HHI index of occupation concentration			Share of nontypical occupations		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	10.61	11.54	10.45	0.17	0.17	0.16	0.29	0.25	0.25	12.15	18.93	20.43
STEM	10.81	11.6	11.4	0.18	0.17	0.16	0.25	0.26	0.26	12.72	16.45	15.52
Business & Economics	10.18	10.87	10.41	0.18	0.17	0.17	0.22	0.22	0.21	4.22	7.44	7.38
Arts & Humanities	9.2	9.66	9.3	0.19	0.17	0.16	0.19	0.19	0.19	10.4	14.36	16.38
Other Degrees	10.28	10.86	9.86	0.18	0.18	0.18	0.25	0.25	0.17	9.38	11.48	20.79
All	10.24	10.9	10.32	0.18	0.18	0.17	0.21	0.2	0.2	8.42	10.93	13.78

Note: Wages are CPI deflated (2014 = 100).

Table 1: Summary statistics of QLFS sample - Labour Market Outcomes

labour market. Although the share of women increased by around 4 percentage points, the relative share of women among the different subjects has remained relatively stable. Women are most underrepresented among STEM subjects - making up around a quarter to a third of all graduates, while they make up around 70% or Medical and Life Sciences.

The labour market outcomes of these recent graduates are summarised in Table 2. The average hourly real wage in the sample is relatively stable at just over £10/hour, even though it comes close to £11/hour during 2003-2011. Across subjects STEM graduates consistently earn the highest average wage, closely followed by Medical & Life Sciences and Business & Economics graduates. Arts & Humanities graduates tend to have the lowest average wages. Apart from these between subject differences, there is large variation of within subject wage inequality with within subject gini coefficients of around 18 gini points which is comparable to the overall gini coefficient. Across time there appears to be a slight decline in wage inequality: between the period 1994-2002 and 2012-2019, the overall gini coefficient of the hourly wage declines somewhat from 18 to 17 gini points.

To measure occupational outcomes, I also report the Herfindahl-Hirschman Index of occupational concentration as well as my own measure of the "share of non-typical occupations", which I define as the share of graduates that work in an occupation that used to have a share of <5% of graduates in the first period. Occupational concentration appears to somewhat decrease in Life Sciences and Other graduates, while it increases amongst STEM graduates. Overall the effect is a small reduction in occupational concentration. However, simply looking at the HHI disguises an important trend that becomes evident when we look at the share of graduates entering occupations that they would not have traditionally entered.[16] Here the share grows by approximately 60% overall, although there is some variation across different subjects. This broad trend might provide some evidence as to the theory of increasing underemployment of university graduates.

4.2 Occupations & Tasks

The one digit SOC 2000 schedule provides me with 9 occupation groups. For the task dimension I choose two broad groupings: 1. Mathematical/Technical Tasks; 2. Verbal/Organisatorial Tasks. I choose these groups since I believe that these kind of tasks are of particular relevance to university graduates.[17]

To obtain an estimate of the occupation task requirements, I use four waves of the UK Skills and Employment Survey (SES). The years of these surveys, 1997 & 2001, 2006 & 2012 and 2017 map neatly into our sample periods. Since the beginning of the Task literature with Autor et al. (2003) there have been many different approaches that try to

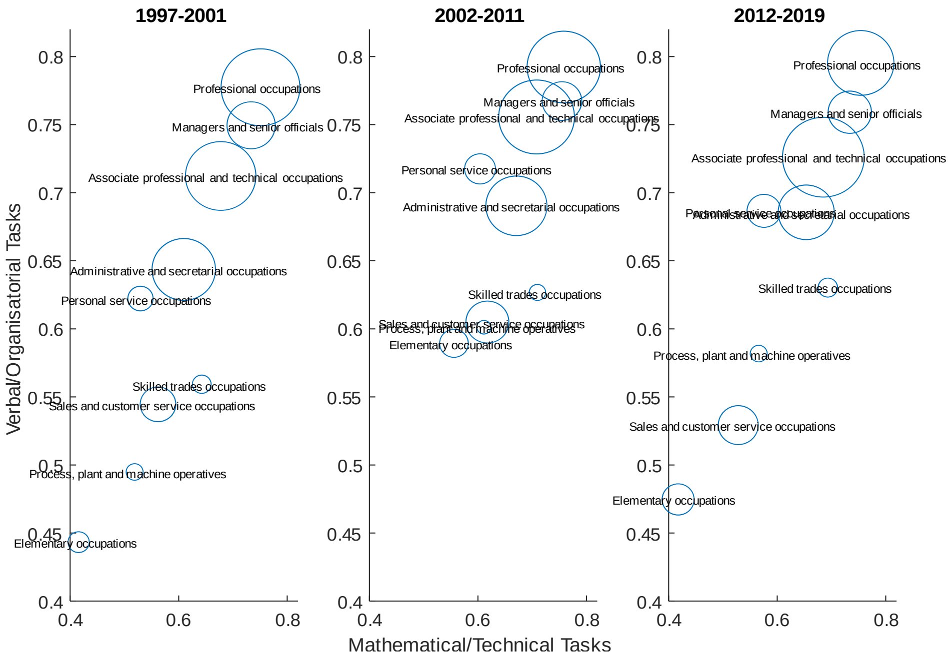
[16] These occupations typically include Skilled Trades, Service and Sales or Elementary Occupations.

[17] For example I do not include manual or routine tasks, as I do not believe that these are particularly interesting in the context of higher education.

approximate the task requirement vector λ using survey data (c.f. Autor (2013), Autor & Handel (2013), Rohrbach-Schmidt & Tieman (2013)). Here I follow the approach of Bisello (2013) who also works with the SES. In this survey, respondents answer questions related to their job and score the importance of performing certain tasks on a Likert scale. I group some of these questions into the three task dimensions and perform a dimensionality reduction using principal component analysis. I then scale the obtained values between 0 and 1 and average across occupations. The resulting task vectors summarized in Figure 1.[18]

Without going into too much detail, I would like to highlight the changes to task requirements over the time period. These have been almost exclusively positive across the two tasks between the first and the second period with a slight reversal between periods two and three. Such changes would be very much in line with any explanation emphasizing the increasing role of Cognitive and Non-Cognitive skills as a result of increased Information and Communications Technology (ICT) usage (c.f. Acemoglu & Autor (2011)). For our purposes it appears important to account for the changing task requirements over time, as they generate important variation that is useful to identify the changing parameters of the skill distributions.

Figure 1: Task Weights by 1-Digit SOC 2000 Occupation



Note: Circle size proportional to employment share in QLFS sample.

[18] More details are provided in the Appendix.

5 Estimation

Let's recall that we are interested in estimating the parameters of the degree-specific graduate skill distribution, which had been specified as:

$$\log(s_i) \sim MVN(\mu_t, \Sigma_t) \tag{27}$$

We want to recover changes of the skill distribution over time, so both μ_t and Σ_t are specified as time varying. We have specified two task dimensions and correspondingly the skill distribution also has two dimensions $k = \{1, 2\}$. Furthermore, there are 5 degree subjects, $m = \{1, .., 5\}$ and three time periods $t = \{1, 2, 3\}$, leading to $m * t = 15$, subject-period specific multivariate skill distributions. For the covariance structure, I assume that skills are uncorrelated within each subject-period distribution:

$$\Sigma_{mt} = \begin{bmatrix} \sigma^2_{1,mt} & 0 \\ 0 & \sigma^2_{2,mt} \end{bmatrix} \tag{28}$$

Note that this doesn't imply that skills are uncorrelated at the population level. If a certain subject-period combination generates high values of two different skills, it will indeed look like there exists a positive correlation between these two skills. It is only that there is no correlation within each distribution.

Occupation fixed effects (η_{ot}) are also allowed to vary between the two periods, in order to capture structural changes in the demand for their output. Across the three periods, the sample spans 26 years, and I allow for year specific aggregate conditions in the labour market, by including year fixed effects. I also include a linear term for experience and gender, both of which are allowed to vary across periods.

To summarize, we have to estimate 60 parameters (μ_{kmt} & σ^2_{kmt}) for the 30 different lognormal distributions, 24 for the occupation fixed effects η_{ot}, 24 for the occupation specific preference terms ω_{ot} , 23 year fixed effects and 3 each for gender and experience controls. Allowing for a normalization, setting the first year in each period to 0, we have a total of 137 parameters.

Setting ϕ^2, i.e. the variance of the measurement error, is a difficult task in this model, that requires some additional steps. The error term υ_i does not only capture traditional measurement error, but also any other productivity differentials that materialize over the course of the graduate's early career, such as health episodes or promotions. The standard approach to setting ϕ^2 would be to run a regression of wages on a number of observables and use the variance of the residuals as an estimate. For this model, this requires controls for mathematical, verbal abilities. Luckily, I can resort to an auxiliary data set (Understanding Society, Wave 3), providing me with an estimate for $\phi = 0.17$.[19] The details of the estimation algorithm are provided in the Appendix.

[19] For more details, see Appendix. Results are robust to increasing ϕ by 10%.

Parameter	Description	Number of Parameters	Type
μ_{kmt}	Location parameter of the subject-period-specific skill distribution.	30	Estimated
σ_{kmt}	Scale parameter of the subject-period-specific skill distribution.	30	Estimated
η_{ot}	Occupation-period specific fixed effect.	24	Estimated
ω_{ot}	Occupation-period specific occupation preferences	24	Estimated
τ	Year fixed effects	23	Estimated
γ_t	Period-specific gender coefficient	3	Estimated
ε_t	Period-specific linear experience coefficient	3	Estimated
ϕ	Standard deviation of log wage measurement errors	1	Calibrated

Table 2: Summary of Model Parameters

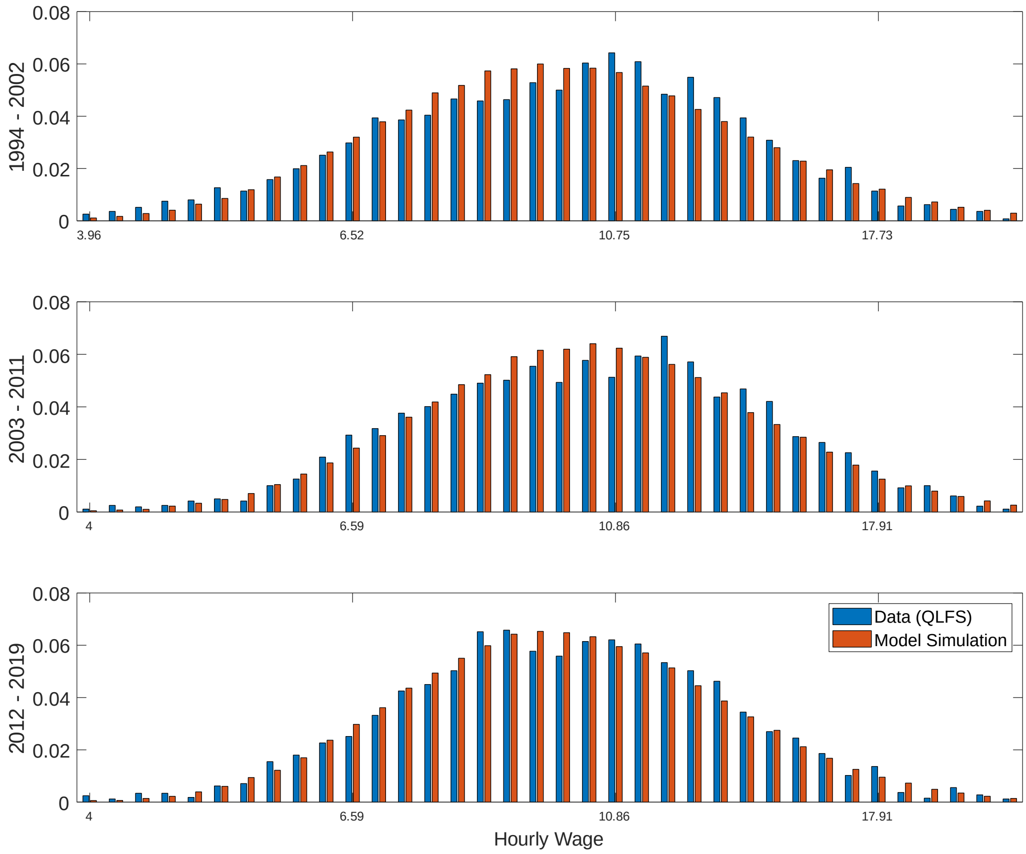
6 Results

This section presents the results of the estimated model. I first present an evaluation of the models fit to the data. Then I discuss the changes in the underlying unobserved skill distributions. Finally, I present the results of some counterfactual experiments.

6.1 Model Fit & Validation

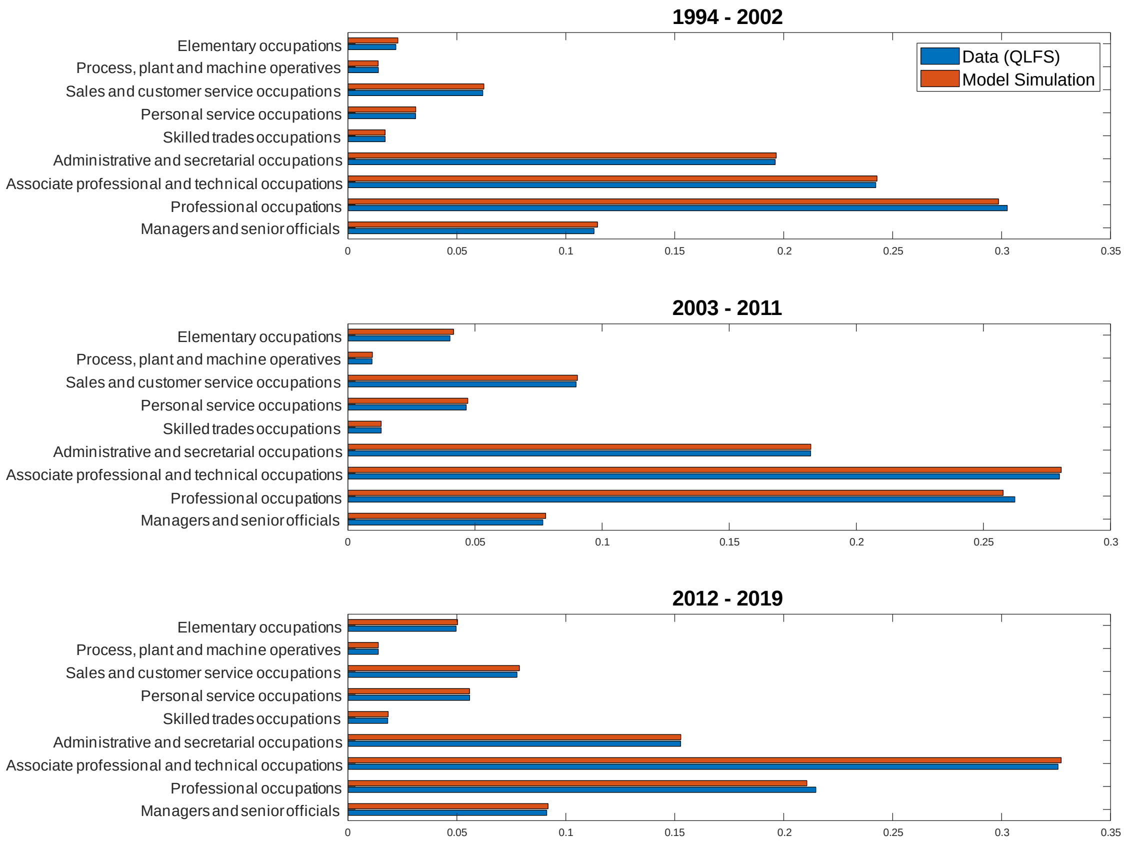
I evaluate the ability of the model to capture both the occupation choices of graduates as well as their wage outcomes. For this purpose I simulate a random, representative sample of $100,000$ graduates in each time period. Figure 2 below show the histogram of the hourly wage across all time periods, while Figure 3 highlights the model fit with respect to the occupation choices of graduates in each time period. Across both dimensions the model tracks the data very well, capturing both the shape of the wage and occupation distributions and tracking their changes across time.

Figure 2: Histogram of Hourly Wages.



Notes: Histogram of hourly wages. QLFS Data and Simulation. Wages in the data are deflated by 2014 CPI Index.

Figure 3: Occupation Distribution - QLFS Data and Model Simulation.



Notes: 1-Digit SOC 2000.

To complement the visual with some statistical evidence the table below compares the model and the data with respect to the mean wage, the Gini coefficient of the hourly wage and the occupation concentration index, as well as the share of non-typical occupations. The model fit is quite good, with the overall model predictions matching their empirical counterparts closely (usually to within 1% margin of error).

At the subject level, the predictions perform slightly worse with the wage indicators (mean and gini). The subject specific means are actually captured quite well to within a maximum deviation of around 5%. Within subject gini coefficients are generally slightly overpredicted to a maximum of around 8% above the values observed in the data. Overall, these deviations are not too concerning and I argue that the model still does a good job of capturing between and within subject differences in wages. In terms of subject specific occupational outcomes the model has slightly more difficulty matching the data indicators for both the HHI occupational concentration index, and the share of non-typical occupations. This is likely due to the fact that the model differentiates occupations only by the task weights and does not include any subject-occupation specific skills or human capital.[20] I will address this shortcoming by focussing mainly on outcomes at the level of all graduates - where the model fit is quite good - and keep discussion of subject specific outcomes to a minimum.

Mean Wage

	1994 - 2002			2003 - 2011			2012 - 2019		
	Data	Model	Difference (%)	Data	Model	Difference (%)	Data	Model	Difference (%)
Medical and Life Sciences	10.61	10.20	-3.82	11.54	11.14	-3.48	10.45	10.26	-1.87
STEM	10.81	10.51	-2.76	11.60	11.09	-4.42	11.40	10.76	-5.59
Business & Economics	10.18	10.64	4.55	10.87	11.26	3.55	10.41	10.60	1.89
Arts & Humanities	9.20	9.62	4.62	9.66	10.19	5.52	9.30	9.60	3.24
Other Degrees	10.28	10.15	-1.30	10.86	10.49	-3.41	9.86	10.09	2.34
All	10.24	10.25	0.09	10.90	10.84	-0.52	10.32	10.27	-0.45

Gini Wage

	1994 - 2002			2003 - 2011			2012 - 2019		
	Data	Model	Difference (%)	Data	Model	Difference (%)	Data	Model	Difference (%)
Medical and Life Sciences	0.171	0.185	8.328	0.171	0.177	3.351	0.160	0.161	0.452
STEM	0.180	0.190	6.032	0.169	0.177	4.991	0.155	0.167	7.704
Business & Economics	0.183	0.192	5.136	0.169	0.178	5.053	0.170	0.177	4.328
Arts & Humanities	0.188	0.190	1.144	0.167	0.172	2.844	0.157	0.168	6.759
Other Degrees	0.179	0.182	1.300	0.177	0.175	-0.958	0.176	0.178	1.490
All	0.184	0.189	3.137	0.176	0.177	0.879	0.169	0.171	1.476

HHI of Occupation Concentration

	1994 - 2002			2003 - 2011			2012 - 2019		
	Data	Model	Difference (%)	Data	Model	Difference (%)	Data	Model	Difference (%)
Medical and Life Sciences	0.29	0.21	-29.36	0.25	0.20	-21.26	0.25	0.19	-23.60
STEM	0.25	0.21	-18.88	0.26	0.19	-24.80	0.26	0.20	-25.55
Business & Economics	0.22	0.21	-6.99	0.22	0.20	-11.32	0.21	0.20	-6.68
Arts & Humanities	0.19	0.20	8.84	0.19	0.20	5.59	0.19	0.20	0.27
Other Degrees	0.25	0.20	-18.51	0.25	0.20	-20.35	0.17	0.20	15.11
All	0.21	0.21	-0.71	0.20	0.20	-0.77	0.20	0.20	-0.22

Share of nontypical Occupations

	1994 - 2002			2003 - 2011			2012 - 2019		
	Data	Model	Difference (%)	Data	Model	Difference (%)	Data	Model	Difference (%)
Medical and Life Sciences	12.15	14.45	18.87	18.93	20.16	6.49	20.43	21.93	7.34
STEM	12.72	14.72	15.66	16.45	20.18	22.71	15.52	21.60	39.18
Business & Economics	4.22	8.61	104.01	7.44	10.69	43.70	7.38	12.98	75.77
Arts & Humanities	10.40	8.92	-14.19	14.36	11.50	-19.91	16.38	13.83	-15.56
Other Degrees	9.38	11.41	21.71	11.48	15.18	32.17	20.79	16.51	-20.58
All	8.42	8.51	1.03	10.93	11.15	2.02	13.78	13.87	0.61

Note: Wages are CPI deflated (2014 = 100).

Table 3: Model Fit

6.2 Skill results

Unfortunately, the shape parameters of a log normal distribution are not particularly intuitive, and so I have presented the effective median skill levels and changes in the next

[20] For example, if the occupations of medical doctor and engineer have similar task weight vectors, then the model would predict that a graduate should be approximately indifferent in choosing either occupation, nonwithstanding that in reality there are obvious additional factors that determine whether one chooses to become one rather than the other.

table. Since the lognormal is not symmetric there is a difference between the median reported here and the mean outcome. Overall the trends and results are very similar if we take the mean instead (reported in the Appendix). For this part I will focus the discussion on the median, as it provides the convenient interpretation of the skills of a "typical" graduate.

Looking at these we can first confirm some of our initial expectations, for example STEM subjects seem to endow their students with more mathematical skills, whilst somewhat lacking in the verbal department. Further observations of this sort should convince us that the estimation is actually picking up some *real* differences between subjects.

Looking at the average *effective* skill levels, we see different trends across both dimensions: Mathematical skills increased by around 120% between 1994-2002 and 2003-2011 and by just over 140% between 1994-2002 and 2012-2019. The increase appears to be - at least in part - driven by increases in the level of technical skills of graduates from subjects that had very low levels of mathematical/technical skills at the beginning of the sample period. Particularly notable is the large increase of the median level of these skills by Business & Economics graduates, but Arts & Humanities and Other graduates also show increasing levels of these skills over the time period. The only exception to this pattern are Medical & Life Science graduates where technical skills become less prevalent over time. It is tempting to suggest that increased demand for technical skills in the labour market has provided the incentives for these observed patterns.

The increase in mathematical/technical skills is counteracted by a large decrease in verbal/organistatorial skills, which fall by roughly 33% between period one and two. This decrease remains roughly constant between the second and third period. This trend is reflected fairly uniformly across subject areas with the exception of Medical & Life Science which experiences the opposite trajectory. These results suggesting a changing skill composition amongst graduates, with a rising emphasis on hard cognitive skills that took hold particularly around the turn of the millennium. The observed increase in mathematical/technical skills, together with the fall in verbal/organisatorial skills, suggests a change in the skill composition of graduates, not necessarily an overall fall in skill levels. To assess the net effect, I sum over all three skills, to get an estimate of the overall skill level of graduates. The results suggest a modest increase in median skill levels of around 6% across the sample period. Overall, I believe the lesson to be learned here is that graduate quality has not deteriorated in the wake of the higher education expansion. This finding is consistent with Blundell et al. (2016) who suggest that a significant decline in unobserved ability of graduates was inconsistent with observed wage and employment movements.

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			All Skills		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	1.44	0.00	0.31	1.43	2.93	2.66	2.88	2.93	2.98
STEM	2.03	3.01	3.15	0.87	0.02	0.00	2.92	3.04	3.15
Business & Economics	0.60	1.59	3.13	2.28	1.40	0.00	2.89	2.99	3.14
Arts & Humanities	0.00	1.55	1.05	2.71	1.31	1.84	2.71	2.87	2.90
Other Degrees	0.00	0.39	1.65	2.82	2.46	1.34	2.82	2.86	2.99
All	0.66	1.46	1.58	2.14	1.42	1.40	2.85	2.94	3.03

Note: Based on a sample of 100,000 simulated observations.

Table 4: Median Skill Levels

So far we have looked at subject specific skill distributions. This allowed us not only to compare graduates of different subjects, but also graduates of the same subject over time. And while between and within subject heterogeneity is of much interest, I also want to consider the group of all graduates and how the underlying changes in subject specific skill supplies interact to affect the overall graduate skill distribution.

Focussing on the Gini, we can see that mathematical skills tend to be more unequally distributed than verbal skills in the mid 1990's to early 2000's, but this trend has seen a reversal over the time period over consideration. Mathematical-technical skill inequality fell from 53 gini points to 45 and then dropped further to 38, by the end of the sampled period. Conversely, verbal-organsiatorial skill inequality, which starts off at a comparatively low level comparable to mathematical skill rises from 24 to 37 and later to 46 gini points, suggesting a large increase of skill inequality across this dimension.

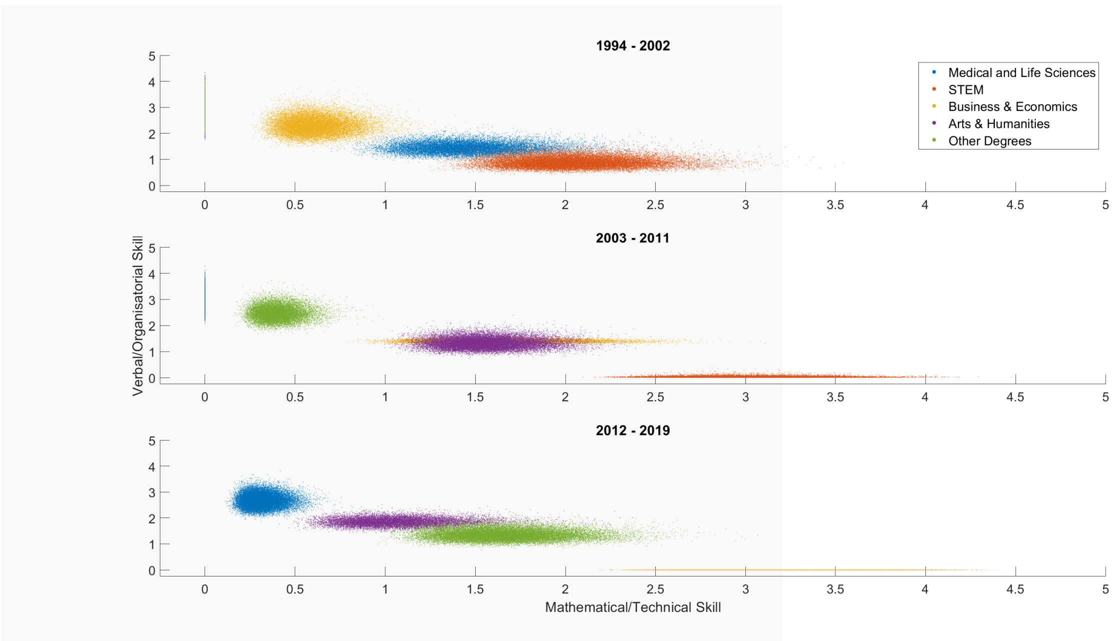
The general trend over the time period seems to go in the direction of greater equality, with overall skill inequality falling from 6 gini points to 5 gini points by the mid 2010's. This result suggests that increases in verbal skill inequality have been more than compensated for by a more equitable distribution of mathematical skills.

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			All Skills		
Gini coefficient	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	0.098	1.000	0.134	0.072	0.050	0.048	0.061	0.050	0.045
STEM	0.075	0.054	0.046	0.111	0.402	0.001	0.063	0.054	0.046
Business & Economics	0.126	0.096	0.055	0.073	0.021	0.001	0.064	0.053	0.055
Arts & Humanities	0.027	0.070	0.127	0.068	0.079	0.043	0.068	0.052	0.054
Other Degrees	0.003	0.128	0.085	0.059	0.057	0.075	0.059	0.053	0.058
All	0.531	0.451	0.379	0.243	0.370	0.456	0.065	0.054	0.054
Between	95.42%	96.44%	95.71%	89.22%	96.82%	98.17%	4.97%	6.08%	10.32%
Within	4.58%	3.56%	4.29%	10.78%	3.18%	1.83%	95.03%	93.92%	89.68%

Note: Based on a sample of 100,000 simulated observations. Between and within variance decomposition based on the square of the coefficient of variation.

Table 5: Skill Inequality

Figure 4: Visualization of the graduate skill distribution



Note: Based on a simulated sample of 100,000 observations.

6.3 Wage Decomposition

Changes in the distribution of skills and the wider structure of the economy necessarily lead to changes in how productivity accrues to different parts of a workers human capital. To assess changes over time, I decompose graduate's wages into their constituent parts. Table 6 presents the share of a workers wage that is on average due to mathematical/technical skills, verbal/organisatorial skills and their chosen occupation. The contribution of occupation fixed effects is fairly stable, both across time and across degree subjects, ranging from around 11% of hourly wages to around 14%. Changes are noticable with respect to the share due to the different types of skills. Here there is an evident trend with mathematical/technical skills gaining at the expense of verbal/organisatorial skills reflecting the changes in the underlying skill endowments. These effects are sizeable on aggregate, with verbal/organisatorial skills accounting for around half of a graduate's wage in the first time period falling to around a third by the mid 2010's. Conversely mathematical/technical skills gained in importance, increasing their share from around a quarter to around 43%.

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			Occupation Fixed Effects		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	28%	10%	13%	29%	82%	68%	12%	11%	13%
STEM	41%	80%	82%	19%	10%	10%	12%	11%	12%
Business & Economics	16%	29%	83%	51%	27%	10%	12%	11%	13%
Arts & Humanities	12%	31%	23%	76%	28%	41%	13%	12%	14%
Other Degrees	11%	14%	32%	77%	62%	28%	13%	12%	13%
All	23%	34%	43%	48%	41%	34%	12%	12%	13%

Note: Based on a sample of 100,000 simulated observations.

Table 6: Wage Decomposition

7 Counterfactuals

In this section I consider four counterfactual experiments, in order to asses the importance of different parts of the model for the changing labour market outcomes of graduates. First, I consider the role of within subject skill inequality. Secondly, I consider a counterfactual world, in which the subject specific skill distributions do not change across time. Thirdly, I turn the second counterfactual on its head by only allowing the skill distributions to vary while keeping the rest of the economic structure fixed. And finally, I assess the role that changing demographic composition of graduates has played in affecting labour market outcomes.

7.1 The role of within subject skill inequality

This model has been built to incorporate skill differences conditional on degree subject in order to better capture the considerable degree of within subject heterogeneity presented at the beginning of this paper. In this subsection I consider how large a role is played by allowing for heterogenous skills within subject categories. Ex ante, one might expect two effects: additional skill heterogeneity will **i)** increase inequality of wage outcomes, and (perhaps less obviously) **ii)** increase mean wages in an environment where skills are multidimensional and graduates are able to choose their utility maximizing occupation.

To assess the role of within subject skill inequality, I set $\sigma^2_{kmt} = 0, \forall k, m, t$ and then resimulate the model. This is equivalent to assuming that each graduate has the median amount of each skill within his degree subject cohort. Table 7 summarizes the impact of this adjustment on mean hourly wages and wage inequality.[21]

	Mean Wage (%)			Gini Wage (%)			HHI (%)			Nontypical Share (%)		
Subject	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019
Medical and Life Sciences	-4.69	-2.66	-2.95	-27.71	-22.95	-21.54	-0.15	0.04	-0.17	0.29	-0.29	0.33
STEM	-5.29	-3.83	-2.56	-28.72	-26.44	-22.25	-0.41	-0.22	-0.06	1.03	0.44	0.12
Business & Economics	-4.97	-3.71	-3.26	-29.40	-25.05	-27.63	-0.15	-0.19	-0.33	0.40	0.52	0.66
Arts & Humanities	-4.64	-3.88	-3.88	-31.04	-24.30	-26.20	-0.29	0.00	-0.13	0.81	-0.23	0.52
Other Degrees	-3.50	-3.71	-4.57	-26.34	-23.48	-29.55	-0.05	-0.22	-0.21	0.25	0.66	0.28
All	-4.70	-3.53	-3.45	-28.44	-23.89	-24.64	-0.23	-0.11	-0.17	0.61	0.24	0.33

Note: Percentage deviations are calculated relative to baseline estimates.

Table 7: Model Counterfactual I

The results suggest that within subject skill inequality plays a large role in explaining graduate wage inequality within subject and overall. The gini coefficient for hourly wages across 1994 to 2002 is around 28% lower compared to the data. In the period 2003 to 2019 this drops to around 24% which is lower but still constitutes a sizeable effect. For mean wages the effects are less pronounced as expected, but still imply a reduction of 4.7% in mean hourly wages in the first period. Later on this impact drops to around 3.5% which is still sizeable. As we saw in the preceding section, overall skill inequality has fallen across these time periods thereby reducing the impact of within subject skill heterogeneity on labour market outcomes. However, within subject skill heterogeneity does still play a sizeable role for wage dynamics even in the most recent sample.

7.2 The role of a changing skill distribution

In this next counterfactual I assess the impact of the changing skill distribution on the outcomes of graduates. For this purpose I fix both μ_{kmt} and σ^2_{kmt} at their period 1 values while allowing all other parameters to change across periods. Table 8 summarizes the effect of this adjustment on the labour market outcomes of graduates:

	Mean Wage (%)			Gini Wage (%)			HHI (%)			Nontypical Share (%)		
Subject	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019
Medical and Life Sciences	0.00	-6.57	-8.64	0.00	5.41	10.62	0.00	-1.38	-0.66	0.00	1.45	0.59
STEM	0.00	-4.52	-10.39	0.00	5.48	8.56	0.00	0.44	-0.44	0.00	-0.19	2.48
Business & Economics	0.00	-3.12	-7.06	0.00	10.41	2.74	0.00	0.20	-0.48	0.00	3.53	8.86
Arts & Humanities	0.00	-4.38	-8.76	0.00	13.10	7.71	0.00	0.59	-0.50	0.00	3.71	4.74
Other Degrees	0.00	-0.74	-5.98	0.00	4.24	-2.66	0.00	0.20	-0.29	0.00	-0.80	-0.96
All	0.00	-4.12	-8.11	0.00	7.75	4.80	0.00	0.00	-0.47	0.00	1.44	3.51

Note: Percentage deviations are calculated relative to baseline estimates.

Table 8: Model Counterfactual II

Fixing the subject specific skill distributions implies that overall changes in the labour market outcomes of graduates are due to **i)** changes in the composition of graduates, and **ii)** changes in the other structural parts of the labour market. Considering periods 2003 - 2011 and 2012 - 2019, the counterfactual analysis suggests that mean wages would have

[21] Occupational outcomes are barely affected by the removal of of within subject heterogeneity, which is likely due to the specific model setup. Results are presented nontheless for completeness.

fallen considerably across all subjects, with an average decrease of 4.12% in the second period and an drop of 8.11% in the third.

The effect is of similar size when considering wage inequality. Here the counterfactual predicts a gini coefficient that is about 7.75% higher than that observed in the data in 2003 - 2011 and 4.8% higher in 2012 -2019, suggesting that the overall decline in skill inequality that I discussed in the pervious section had a considerable impact in reducing graduate level wage inequality.

The impact of the changing skill distribution does not seem to have a large effect on occupation choices. The HH index of occupational concentration is almost unaffected compared to the baseline model.[22] Although the share of non-typical occupations is increased by 1.44% in 2003 - 2011 and 3.51% in 2012 - 2019 than predicted by the baseline model, suggesting that this mechanism has a role to play here also.

7.3 The role of a changing task weights

Complementing the preceding analysis I perform another counterfactual simulation, this time fixing the occupation specific skill weights λ_o at their $t = 1$ values while letting all other parameters change across periods. This provides an assessment of the importance of changes in the evolving skill requirements of firms. The results are presented in Table 9.

	Mean Wage (%)			Gini Wage (%)			HHI (%)			Nontypical Share (%)		
Subject	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019
Medical and Life Sciences	0.00	-9.61	-5.33	0.00	9.36	3.05	0.00	4.45	2.30	0.00	-11.87	-2.12
STEM	0.00	-8.33	-2.59	0.00	11.29	3.23	0.00	5.41	1.51	0.00	-10.55	1.12
Business & Economics	0.00	-8.78	-2.50	0.00	10.67	2.71	0.00	5.14	1.09	0.00	-17.69	-6.93
Arts & Humanities	0.00	-8.66	-4.44	0.00	11.05	2.87	0.00	4.86	1.75	0.00	-16.69	-7.17
Other Degrees	0.00	-9.15	-3.97	0.00	9.75	2.20	0.00	4.86	1.51	0.00	-11.09	3.81
All	0.00	-8.90	-3.89	0.00	10.29	3.09	0.00	4.96	1.69	0.00	-16.98	-7.37

Note: Percentage deviations are calculated relative to baseline estimates.

Table 9: Model Counterfactual III

The results suggest that had firms continued to operate with their initial technologies, average wages would be around 8.9% lower in 2003 - 2011 and 3.89% lower in 2012 - 2019. When graduate's skills are not well matched to the demands of employers wages decline, supporting the hypothesis that part of the increase in mathematical skills that I observe was driven in response to changing demands of firms.[23] Wage inequality would also have been larger in the absence of changes to the occupation task weights, with the wage gini being increased by up to 10% relative to the baseline model.

Interestingly, the changing skill demands appear to have played a larger role for occupational choices than the changing skill distributions I considered before. Particularly the share of graduates entering non-typical occupations decreases by 16.98% in 2003 - 2011 and by 7.37% in 2012 - 2019, suggesting that this channel is important for explaining this particular trend.

[22] I suggest that this is because the changes in the within subject skill distributions are not large enough to effect much of a change in preffered occupations. Instead the occupation distribution appears to be driven by chanages in the composition of graduates and other external factors.

[23] The alternative hypothesis that employers changed their production technology in response to changing skill supplies, cannot be ruled out but appears to be less plausible.

7.4 The role of changing demographics

Finally, I assess the role that changing demographics played for graduates' labour market outcomes. For this purpose I fix the distribution of demographic characteristics (experience, sex, subject choice) to 1994-2002, while allowing all parameters to vary over time. The results of this counterfactual analysis are presented in Table 10. In line with the observation, that demographic characteristics did not change significantly over the period[24], there are only minor changes to the graduates' outcomes. The one notable finding is that mean wages would be around 0.5% higher in 2003-2011 and almost 1.5% higher in 2012-2019. This is probably reflective of a higher share of high paid STEM and Business & Economics graduates in the initial period.

	Mean Wage (%)			Gini Wage (%)			HHI (%)			Nontypical Share (%)		
Subject	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019	1994 - 2001	2002 - 2011	2012 - 2019
Medical and Life Sciences	0.00	0.36	0.72	0.00	-0.73	-1.21	0.00	0.63	2.06	0.00	-3.64	-3.45
STEM	0.00	0.12	0.84	0.00	0.08	-0.29	0.00	0.40	0.13	0.00	-0.35	-1.33
Business & Economics	0.00	-0.17	0.88	0.00	0.41	1.08	0.00	-1.01	-2.35	0.00	4.74	4.72
Arts & Humanities	0.00	1.38	2.52	0.00	0.08	0.02	0.00	0.23	-0.13	0.00	-1.07	5.86
Other Degrees	0.00	0.36	0.48	0.00	0.91	-1.11	0.00	0.03	-0.64	0.00	0.28	-2.15
All	0.00	0.47	1.45	0.00	0.01	-0.03	0.00	-0.07	-0.04	0.00	-0.72	-1.09

Note: Percentage deviations are calculated relative to baseline estimates.

Table 10: Model Counterfactual IV

8 Conclusion

The formation of human capital and the acquisition of specific skills lies at the heart of a university education. Despite their centrality to both the academic and public discourse about tertiary education, there are few quantitative studies that actually investigate what skills graduates possess at the end of their university courses. This paper has been an attempt at estimating the UK's graduate skill distribution and changes to it over the last 3 decades.

My findings suggest, that while there has been a considerable change in the composition of skills that university graduates possess, this has not resulted in a large decline of the overall skill level of a typical graduate. While verbal-organisatorial skills have decreased, the typical graduate in the most recent time period has a lot more technical skills than previous cohorts. I speculate that this pattern is a result of the increase in the demand for mathematical skills in the wake of increasing use of ICT in recent decades.

The observed changes are due to both changes in the composition of graduates and changes of the skill distribution at the subject level. Overall skill inequality has declined in the wake of the expansion of higher education, making graduates more homogenous in terms of skillsets over time.

On the level of labour market outcomes, these combined changes have contributed to lower wage inequality for graduates in their early carreer. Using counterfactual simulations, I find that in the absence of changes in the subject specific skill distributions, real hourly wages would be up to 8% lower in 2012 - 2019. Additionally, wage inequality would also be larger by around 5%.

[24] See Table A3 for details.

These results are potentially of interest to government departments trying to understand the changing supply of differentiated skills at the graduate level, which is likely to be highly relevant to the development of a dynamic economy; as well as to actors in the sphere of higher education interested in making sure that their students graduate with skills that prepare them for the future.

References

ACEMOGLU, D., AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," Handbook of Labor Economics, 4, 1043–1171.

ALTONJI, J., P. ARCIDIACONO, AND A. MAUREL (2016): "The Analysis of Field Choice in College and Graduate School," Handbook of the Economics of Education, 5, 305–396.

ALTONJI, J. G., L. B. KAHN, AND J. D. SPEER (2014): "Trends in earnings differentials across college majors and the changing task composition of jobs," American Economic Review, 104, 387–393.

ANDREWS, R. J., S. A. IMBERMAN, M. F. LOVENHEIM, AND K. M. STANGE (2022): "The Returns to College Major Choice: Average and Distributional Effects, Career Trajectories, and Earnings Variability," Working Paper 30331, National Bureau of Economic Research.

AUTOR, D. H., AND M. J. HANDEL (2013): "Putting tasks to the test: Human capital, job tasks, and wages," Journal of Labor Economics, 31.

AUTOR, D. H., F. LEVY, AND R. J. MURNANE (2003): "The skill content of recent technological change: An empirical exploration," Quarterly Journal of Economics, 118, 1279–1333.

BISELLO, M. (2013): "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.

BLUNDELL, R., D. GREEN, AND W. JIN (2018): "The UK Education Expansion and Technological Change," IFS Working Paper, pp. 1–71.

GILL, J., AND G. KING (2004): "What to do when your Hessian is not invertible: Alternatives to model respecification in nonlinear estimation," Sociological Methods and Research, 33, 54–87.

GOLDIN, C. D., AND L. F. KATZ (2009): The race between education and technology. harvard university press.

GOURIEROUX, C., AND A. MONFORT (1993): "Simulation-based inference. A survey with special reference to panel data models," Journal of Econometrics, 59, 5–33.

HAJIVASSILIOU, V. A., AND P. A. RUUD (1994): "Chapter 40 Classical estimation methods for LDV models using simulation," Handbook of Econometrics, 4, 2383–2441.

HAMERMESH, D. S., AND S. G. DONALD (2008): "The effect of college curriculum on earnings: An affinity identifier for non-ignorable non-response bias," Journal of Econometrics, 144, 479–491.

HASTINGS, J. S., C. A. NEILSON, AND S. D. ZIMMERMAN (2013): "Are some degrees worth more than others? Evidence from college admission cutoffs in Chile," National Bureau of Economic Research.

Hemelt, S. W., B. Hershbein, S. M. Martin, and K. M. Stange (2021): "College Majors and Skills: Evidence from the Universe of Online Job Ads," Working Paper 29605, National Bureau of Economic Research.

Holmes, C., and K. Mayhew (2016): "The economics of higher education," Oxford Review of Economic Policy, 32, 475–496.

Kinsler, J., and R. Pavan (2015): "The specificity of general human capital: Evidence from college major choice," Journal of Labor Economics, 33, 933–972.

Kirkeboen, L. J., E. Leuven, and M. Mogstad (2016): "Field of study, earnings, and self-selection," The Quarterly Journal of Economics, 131, 1057–1111.

Lee, L. F. (1995): "Asymptotic bias in simulated maximum likelihood estimation of discrete choice models," Econometric Theory, 11, 437–483.

Leighton, M., and J. D. Speer (2020): "Labor market returns to college major specificity," European Economic Review, 128, 103489.

Lindley, J., and S. McIntosh (2015): "Growth in within graduate wage inequality: The role of subjects, cognitive skill dispersion and occupational concentration," Labour Economics, 37, 101–111.

Lovenheim, M. F., and J. Smith (2022): "Returns to Different Postsecondary Investments: Institution Type, Academic Programs, and Credentials," Working Paper 29933, National Bureau of Economic Research.

McFadden, D. (1974): "The measurement of urban travel demand," Journal of public economics, 3, 303–328.

McFadden, D., and K. Train (2000): "Mixed MNL models for discrete response," Journal of applied Econometrics, 15, 447–470.

Onozuka, Y. (2019): "Heterogeneous Skill Growth across College Majors," unpublished, pp. 1–58.

Rohrbach-Schmidt, D., and M. Tiemann (2013): "Changes in workplace tasks in Germany - evaluating skill and task measures," Journal for Labour Market Research, 46, 215–237.

Roy, A. D. (1951): "Some thoughts on the distribution of earnings," Oxford economic papers, 3, 135–146.

Roys, N., and C. Taber (2016): "Skills prices, occupations and changes in the wage structure," unpublished.

Sanders, C., and C. Taber (2012): "Life-Cycle Wage Growth and Heterogeneous Human Capital," Annual Review of Economics, 4, 399–425.

Train, K. (1999): "Halton Sequences for Mixed Logit," Working Papers, 38, 1–18.

Train, K. E. (2008): "EM algorithms for nonparametric estimation of mixing distributions," Journal of Choice Modelling, 1, 40–69.

——— (2009): Discrete choice methods with simulation, vol. 9780521816. Cambridge university press.

A Additional Tables & Results

A.1 Model Estimates

The following tables presents the SMLE results.

	Medical & Life Sciences	STEM	Business & Economics	Arts & Humanities	Other Degrees
			1994 - 2002		
μ_{Math}	0.363	0.711	-0.51	-11.409	-11.433
	(1.434)	(1.365)	(3.325)	(480.107)	(745.178)
μ_{Verbal}	0.359	-0.14	0.825	0.998	1.038
	(1.422)	(2.067)	(1.683)	(0.953)	(1.441)
σ_{Math}	0.172	0.133	0.223	0.048	0.005
	(0.47)	(0.348)	(1.159)	(191.509)	(148.002)
σ_{Verbal}	0.128	0.197	0.131	0.12	0.103
	(0.312)	(0.639)	(0.493)	(0.325)	(0.457)
			2003 - 2011		
μ_{Math}	-85.494	1.102	0.468	0.439	-0.934
	(0)	(1.426)	(1.013)	(1.116)	(2.22)
μ_{Verbal}	1.074	-3.877	0.334	0.269	0.9
	(0.191)	(16.006)	(1.078)	(1.214)	(0.888)
σ_{Math}	8.277	0.095	0.17	0.123	0.226
	(0)	(0.314)	(0.374)	(0.296)	(0.72)
σ_{Verbal}	0.089	0.746	0.036	0.139	0.102
	(0.071)	(7.171)	(0.365)	(0.357)	(0.22)
			2012 - 2019		
μ_{Math}	-1.181	1.148	1.144	0.047	0.498
	(1.751)	(0.472)	(0.616)	(0.757)	(0.525)
μ_{Verbal}	0.977	-11.606	-8.449	0.613	0.29
	(0.581)	(271.308)	(72.858)	(0.553)	(0.572)
σ_{Math}	0.238	0.082	0.097	0.224	0.149
	(0.592)	(0.133)	(0.187)	(0.317)	(0.155)
σ_{Verbal}	0.084	0.001	0.001	0.076	0.133
	(0.123)	(90.377)	(97.687)	(0.075)	(0.135)

Numerical Standard Errors in Parentheses

Table A1: Simulated Maximum Likelihood Estimates - Skill Distribution Parameters

	1994 - 2002	2003 - 2011	2012 - 2019
ε	0.076	0.057	0.045
	(0.041)	(0.033)	(0.028)
γ	-0.019	0.000	-0.043
	(0.045)	(0.031)	(0.028)
η_1	0.000	0.000	0.000
	-	-	-
η_2	0.045	0.146	0.072
	(0.187)	(0.264)	(0.085)
η_3	0.150	0.154	0.149
	(0.259)	(0.301)	(0.094)
η_4	0.116	0.148	0.137
	(0.277)	(0.176)	(0.071)
η_5	0.307	0.095	0.142
	(0.613)	(0.626)	(0.194)
η_6	0.110	0.055	0.068
	(0.517)	(0.431)	(0.238)
η_7	0.333	0.207	0.439
	(0.332)	(0.195)	(0.096)
η_8	0.445	0.236	0.280
	(0.353)	(0.218)	(0.066)
η_9	0.455	0.217	0.595
	(0.239)	(0.251)	(0.149)
ω_1	0.000	0.000	0.000
	-	-	-
ω_2	0.869	1.040	0.702
	(0.122)	(0.149)	(0.144)
ω_3	0.737	1.223	1.252
	(0.162)	(0.121)	(0.137)
ω_4	0.756	0.954	0.614
	(0.104)	(0.134)	(0.144)
ω_5	-1.743	-1.583	-1.515
	(0.287)	(0.388)	(0.195)
ω_6	-0.968	-0.272	-0.194
	(0.25)	(0.199)	(0.137)
ω_7	-0.377	0.395	0.055
	(0.142)	(0.156)	(0.161)
ω_8	-1.841	-1.867	-1.634
	(0.322)	(0.192)	(0.176)
ω_9	-1.203	-0.310	-0.286
	(0.292)	(0.139)	(0.129)
τ_1	0.000	0.000	0.000
	-	-	-
τ_2	-0.027	-0.001	-0.021
	(0.09)	(0.136)	(0.034)
τ_3	0.003	-0.006	-0.027
	(0.113)	(0.101)	(0.033)
τ_4	-0.004	0.014	0.025
	(0.112)	(0.086)	(0.039)
τ_5	0.068	0.065	0.033
	(0.112)	(0.05)	(0.044)
τ_6	0.088	0.032	0.038
	(0.121)	(0.068)	(0.044)
τ_7	0.126	0.030	0.040
	(0.125)	(0.063)	(0.032)
τ_8	0.159	-0.018	0.052
	(0.163)	(0.046)	(0.033)
τ_9	0.169	-0.044	-
	(0.114)	(0.074)	-

Numerical Standard Errors in Parentheses

Table A2: Simulated Maximum Likelihood Estimates - Other Parameters

A.2 Additional Tables

	Share of graduates (%)			Share of female of graduates (%)			Average experience		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	16.84	21.94	26.29	67.38	71.66	70.72	1.19	1.18	1.23
STEM	26.06	21.19	19.8	27.34	27.76	29.62	1.25	1.21	1.28
Business & Economics	20.26	19.49	13.87	50.38	56.22	55.26	1.25	1.28	1.2
Arts & Humanities	19.43	22.33	14.21	61.33	64.17	57.42	1.33	1.21	1.19
Other Degrees	17.41	15.05	25.82	67.71	66.85	62.62	1.25	1.22	1.28
All	3860	3587	3222	52.38	56.96	56.46	1.25	1.22	1.24

Table A3: Summary statistics of QLFS sample - Demographics

Mathematical/Technical Tasks	
Variable name	Description
cspecial	importance of: specialist knowledge or understanding
cfaults	importance of: spotting problems or faults
csolutn	importance of: thinking of solutions to problems
canalyse	importance of: analysing complex problems in depth
ccalca	importance of: arithmetic (adding, subtracting, multiplying, dividing)
cstats	importance of: advanced mathematics/ statistics
cpercent	importance of: arithmetic involving fractions (decimals, percentages, fractions)
Verbal/Organisatorial Tasks	
Variable name	Description
cteach	importance of: teaching people (individuals or groups)
cspeech	importance of: making speeches/ presentations
cteamwk	importance of: working with a team
corgwork	importance of: knowledge of how organisation works
cplanme	importance of: planning own activities
cplanoth	importance of: planning the activities of others
cmytime	importance of: organising own time
cahead	importance of: thinking ahead
cread	importance of: reading written information (eg. forms, notices, signs)
cshort	importance of: reading short documents
clong	importance of: reading long documents
cwrite	importance of: writing materials such as forms, notices or signs
cwritesh	importance of: writing short documents
cwritelg	importance of: writing long documents

Table A4: Variables used in the construction of task intensity measures

	Mathematical-Technical Tasks			Verbal-Organisatorial Tasks		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Managers and senior officials	0.73	0.75	0.73	0.75	0.77	0.76
Professional occupations	0.75	0.76	0.75	0.78	0.79	0.80
Associate professional and technical occupations	0.68	0.71	0.68	0.71	0.76	0.73
Administrative and secretarial occupations	0.61	0.67	0.65	0.64	0.69	0.69
Skilled trades occupations	0.64	0.71	0.69	0.56	0.63	0.63
Personal service occupations	0.53	0.60	0.58	0.62	0.72	0.69
Sales and customer service occupations	0.56	0.62	0.53	0.54	0.60	0.53
Process, plant and machine operatives	0.52	0.61	0.57	0.49	0.60	0.58
Elementary occupations	0.42	0.56	0.42	0.44	0.59	0.47
Mean	0.67	0.70	0.67	0.70	0.73	0.70
Standard Deviation	0.08	0.06	0.09	0.08	0.06	0.09

Table A5: Task Intensities

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			All Skills		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	1.46	0.00	0.32	1.44	2.94	2.67	2.90	2.94	2.98
STEM	2.05	3.02	3.16	0.89	0.03	0.00	2.94	3.05	3.16
Business & Economics	0.61	1.62	3.15	2.30	1.40	0.00	2.91	3.01	3.15
Arts & Humanities	0.00	1.56	1.07	2.73	1.32	1.85	2.73	2.89	2.92
Other Degrees	0.00	0.40	1.66	2.84	2.47	1.35	2.84	2.88	3.01
All	0.91	1.37	1.73	1.96	1.59	1.31	2.87	2.96	3.04

Note: Based on a sample of 100,000 simulated observations.

Table A6: Mean Skill Levels

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			All Skills		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	1.39	0.00	0.70	1.48	2.96	2.32	2.89	2.97	3.02
STEM	1.88	2.63	3.20	1.04	0.43	0.00	2.94	3.07	3.20
Business & Economics	0.00	1.79	3.15	2.88	1.23	0.00	2.88	3.04	3.15
Arts & Humanities	0.00	1.40	0.95	2.72	1.48	1.96	2.73	2.89	2.91
Other Degrees	0.01	0.00	1.50	2.83	2.88	1.50	2.84	2.89	3.02
All	0.01	1.41	1.47	2.43	1.47	1.54	2.86	2.97	3.07

Note: Based on a sample of 100,000 simulated observations.

Table A7: Median Skill Levels - Estimation including Postgraduate Qualifications

	Mathematical/Technical Skill			Verbal/Organisatorial Skill			All Skills		
	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019	1994 - 2002	2003 - 2011	2012 - 2019
Medical and Life Sciences	0.097	0.000	0.092	0.074	0.050	0.052	0.061	0.050	0.046
STEM	0.077	0.056	0.045	0.100	0.124	0.001	0.062	0.052	0.045
Business & Economics	0.001	0.080	0.057	0.066	0.055	0.001	0.066	0.053	0.057
Arts & Humanities	0.001	0.074	0.137	0.065	0.073	0.045	0.065	0.052	0.055
Other Degrees	0.000	0.000	0.093	0.059	0.051	0.080	0.059	0.051	0.061
All	0.607	0.489	0.328	0.228	0.329	0.440	0.064	0.053	0.055

Note: Based on a sample of 100,000 simulated observations.

Table A8: Gini Skill Levels - Estimation including Postgraduate Qualifications

B Technical Details

B.1 Calibration of ϕ^2

Wave 3 of the Understanding Society Survey contains a module assessing the cognitive and psychological traits of adult (16+) respondents. Questionaire items include test measuring 1. Numeric Ability, 2. A Subtraction Exercise, 3. Completion of a Number Sequence, 4. A word recall exercise & 5. Verbal Fluency. I use these items to generate 2 measures of skill, mapping into the dimensions of mathematical and verbal skill used in the model, using principal component analysis on the standardized survey responses. Then I use these measures to run a cross sectional (log) wage regression, controlling for 1 Digit SOC (2000) occupation as well as the full set of interactions with the two skill measures. The resulting regression equation is exactly the proxy analogue to the log wage equation:

$$w_i = \sum_{O} \mathbf{1}_{(o_i^*=o)}\tilde{\eta}_o + \sum_{O} \mathbf{1}_{(o_i^*=o)}\tilde{\eta}_o\tilde{s}_{i,math} + \sum_{O} \mathbf{1}_{(o_i^*=o)}\tilde{\eta}_o\tilde{s}_{i,verbal} + \tilde{\nu}_i$$

The residual variance of $\tilde{\nu}_i$ provides an estimate for ϕ^2.

I estimate the auxiliary model on a sample of full time working individuals aged 21-27, adding additional controls for sex and age as a proxy of labour market experience. The estimates are presented in the next table:

	(1) log hourly earnings
1-Digit SOC 2000=1	0
	(.)
1-Digit SOC 2000=2	-1.787**
	(0.667)
1-Digit SOC 2000=3	-0.238
	(0.488)
1-Digit SOC 2000=4	-1.928
	(0.985)
1-Digit SOC 2000=5	-0.274
	(0.996)
1-Digit SOC 2000=6	-0.980*
	(0.468)
1-Digit SOC 2000=7	-1.191**
	(0.435)
1-Digit SOC 2000=8	-1.589***
	(0.456)
1-Digit SOC 2000=9	-1.328**
	(0.436)
1-Digit SOC 2000=1XMathematical Skill	-1.113*
	(0.543)
1-Digit SOC 2000=2XMathematical Skill	0.801
	(0.653)
1-Digit SOC 2000=3XMathematical Skill	-0.716
	(0.369)
1-Digit SOC 2000=4XMathematical Skill	-0.138
	(0.342)
1-Digit SOC 2000=5XMathematical Skill	0.127
	(0.274)
1-Digit SOC 2000=6XMathematical Skill	-0.219
	(0.208)
1-Digit SOC 2000=7XMathematical Skill	-0.117
	(0.222)
1-Digit SOC 2000=8XMathematical Skill	0.546
	(0.278)
1-Digit SOC 2000=9XMathematical Skill	0.0841
	(0.206)
1-Digit SOC 2000=1XVerbal Skill	-0.218
	(0.267)
1-Digit SOC 2000=2XVerbal Skill	0.464**
	(0.165)
1-Digit SOC 2000=3XVerbal Skill	-0.158
	(0.198)
1-Digit SOC 2000=4XVerbal Skill	1.059
	(0.999)
1-Digit SOC 2000=5XVerbal Skill	-1.122
	(1.083)
1-Digit SOC 2000=6XVerbal Skill	-0.0785
	(0.246)
1-Digit SOC 2000=7XVerbal Skill	0.0634
	(0.249)
1-Digit SOC 2000=8XVerbal Skill	0.0653
	(0.172)
1-Digit SOC 2000=9XVerbal Skill	-0.0145
	(0.138)
male	0
	(.)
female	-0.0116
	(0.0273)
Age	0.0162*
	(0.00631)
Constant	2.739***
	(0.433)
Year Fixed Effects	Yes
Observations	265

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B.2 Estimation algorithm

The estimation procedure is a simple application of simulated maximum likelihood. In maximum likelihood we find a vector of parameters so that the model maximizes the probability of observing the actual outcome. The only complication, that arises here comes from the fact, that we do not have a closed form solution for the joint probability (24) 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. 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 (24) by $\Pr^{sim}(o_i^*, w_i^{obs})$ for simplicity, and let $\theta = (\eta, \mu, \Sigma)$ be the set of our parameters, we can write down the simulated log likelihood function of the as:

$$ll^{sim}(\theta;\phi^2) = \frac{1}{N}\sum_i \sum_{o=1}^{O} \mathbf{1}_{(o=o^*)} \ln \overset{sim}{\Pr}(o_i^*, w_i^{obs}) \tag{29}$$

and we can estimate θ as:

$$\hat{\theta} = \arg\max_{\theta} ll^{sim}(\theta;\phi^2) \tag{30}$$

So to specify the complete algorithm:

1. Set $q = 1$ 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 $\hat{\theta}_q$ draw a vector of s_i, R times, denoting each as s_i^r.
3. For $r = 1$ to R :
   (a) Calculate $\nu_i^r = w_i^{obs} - [\eta_{o^*} + \sum_K \lambda_{o^*k} s_{ik}^r]$.
   (b) For a given pair s_i^r, ν_i^r calculate $\Pr^r(o_i^*, w_i^{obs})$.
4. Average over all R values of $\Pr^r(o_i^*, w_i^{obs})$ to obtain:
   $\Pr^{sim}(o_i^*, w_i^{obs}) = \frac{1}{R}\sum_{r=1}^{R} \Pr^r(o_i^*, w_i^{obs})$.
5. Repeat steps $2 - 4$ for all N individuals. Calculate the log likelihood via (29) denoting it as ll_q^{sim}.
6. If $|ll_q^{sim} - ll_{q-1}^{sim}| < \epsilon$, terminate here. Otherwise increment q and find a new value $\hat{\theta}_q$ and repeat from step 2.

Finally, for the numerical evaluation of the integral I use a grid of $10,000$ 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. 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 use a two step procedure: 1. I estimate the model under the assumption of no skill heterogeneity within each subject ($\sigma^2_{kmt} = 0, \forall k, m, t$). This avoids the evaluation of the integral saving considerable computational time. Taking advantage of this speed gain, I start the optimization using $1,000,000$ random starting values. 2. I take the best of these runs as a starting value to fit the full model. Optimization is performed using Matlabs `fminunc` routine. All critical values for convergence are set to $1e-6$.

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 Asymptotic Equivalence of SML and ML

The asymptotic properties of the simulated maximum likelihood estimator have been well understood (c.f. Gourieroux and Monfort (1993), Lee (1995), and Hajivassiliou & Ruud (1994)). This short exposition here is based on the discussion in Train (2009, Chapter 10) for simplicity. Generally maximum likelihood estimation proceeds by maximizing the log likelihood function:

$$ll(\theta) = \sum_n \ln \Pr_n(\theta) \tag{31}$$

where θ is a vector of parameters and $\Pr_n(\theta)$ is the exact probability of the observed choice of observation n given θ.

Similarly, simulated maximum likelihood maximizes the simulated maximum likelihood function:

$$sll(\theta) = \sum_n \ln \Pr_n^{sim}(\theta) \tag{32}$$

where $\Pr_n^{sim}(\theta)$ is the simulated probability of the observed choice of observation n. It is known,that if $\Pr_n^{sim}(\theta)$ is an unbiased simulator for the exact probability - i.e. $E_r(\Pr_n^{sim}(\theta)) = \Pr_n(\theta)$, where the expectation is taken over r simulation draws, then there are three sources of bias in the SML estimator:

1. Sampling bias, which is the same as in the ML estimator and which goes to 0 as $N \rightarrow \infty$.

2. Simulation noise, which goes to 0 as the number of simulation draws $R \rightarrow \infty$.

3. Bias due to the fact, that $\ln \Pr_n^{sim}(\theta)$ is not an unbiased estimator of $\ln \Pr_n(\theta)$. This bias disappears if R grows faster than $\sqrt{N}$.

Hence the results, which are derived under fairly general conditions, indicate, that if $\Pr_n^{sim}(\theta)$ is an unbiased simulator, and the number of simulation draws is sufficiently larger than $\sqrt{N}$, then the MSL estimator is consistent, asymptotically normal, efficient and equivalent to traditional ML.

Therefore, the only thing that we have to show, is that our simulated joint probability $\Pr^{sim}(o_i^*, w_i^{obs}|\theta)$ is an unbiased estimator of the exact probability $\Pr(o_i^*, w_i^{obs}|\theta)$. To show this, let's remind ourselves, of how the simulated probability is obtained:

$$\overset{sim}{\Pr}(o_i^*, w_i^{obs}|\theta) = \frac{1}{R}\sum_{r=1}^{R}\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}^r}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\} \tag{33}$$

where s_i^r is the rth simulation draw of s_i. Compare this to the exact probability:

$$\Pr(o_i^*, w_i^{obs}|\theta)) = \int\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\} f(s)d(s) \tag{34}$$

Now

$$\begin{aligned}
E\left(\overset{sim}{\Pr}(o_i^*, w_i^{obs}|\theta)\right) &= \\
& E\left[\frac{1}{R}\sum_{r=1}^{R}\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}^r}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\}\right] \\
&= \frac{1}{R}\sum_{r=1}^{R} E\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}^r}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}^r]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\} \\
&= \frac{1}{R}\sum_{r=1}^{R}\int\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\} f(s)d(s) \\
&= \int\left\{\left(\frac{e^{\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}}}{\sum_{o=1}^{O} e^{\eta_o+\sum_K \lambda_{ok}s_{ik}}}\right)\left(\frac{e^{(-\frac{\left(w_i^{obs}-[\eta_{o*}+\sum_K \lambda_{o*k}s_{ik}]\right)^2}{2\phi^2})}}{\sqrt{2\pi\phi^2}}\right)\right\} f(s)d(s) \\
&= \Pr(o_i^*, w_i^{obs}|\theta))
\end{aligned} \tag{35}$$

the third line follows from the definition of the expected value operator and the fact that all s_i^r are i.i.d.

Hence, we have shown, that $\Pr^{sim}(o_i^*, w_i^{obs}|\theta)$ is an unbiased estimator for $\Pr(o_i^*, w_i^{obs}|\theta)$. Furthermore, $\Pr^{sim}(o_i^*, w_i^{obs}|\theta)$ is a continuous and twice differentiable function.

B.5 Sidenote on "effective skills"

I wanted to highlight some potential limitations of the approach taken in this paper. Namely, my estimation is only ever going to recover *effective* skills. To explain what this means for our estimation strategy, consider the following simple one dimensional example:

So far we have assumed that output was produced by the combination of worker's skills and the occupation's task requirements:

$$\ln(Y) = \lambda_o s_i \tag{36}$$

Now this expression is observationally equivalent to another expression:

$$\ln(Y) = \chi\lambda_o s_i^* \tag{37}$$

$$s_i^* = \frac{s_i}{\chi} \tag{38}$$

where χ is a general productivity parameter that is common to all occupations. In a one period case this is not particularly troublesome, but in a multi period setting one might want to take the possibility of the general productivity of a certain skill changing seriously. Unfortunately it is not possible here to address this issue and disentangle the two, which would require some information on the evolution of χ. Therefore, I just wanted to clarify that what the estimation procedure recovers is $s_i = s_i^*\chi$ which I dub *effective* skills.

Despite this caveat, I belief that there is not too much to be concerned about here - i.e. I think that I can justify that $\chi_{kt} \approx 1\ \forall k, t$ for the following reasons: i) We have observed that task requirements have changed over the period, which means that at least part of χ is actually observed and thus controlled for. ii) The inclusion of occupation-time specific fixed effects will soak up some of this aggregate change. With this small caveat out of the way, let us turn to the actual estimation results.

C Discrete Choice Modeling

Often individuals have to pick their preferred option out of a number of available alternatives (for example their choice of car model or favourite brand of breakfast cereal). Discrete choice analysis is concerned with developing tools that help researchers analyze behaviour in these types of choice scenarios. In the following I will quickly sketch the basic multivariate logit model (c.f. McFadden (1974), Train (2009)), but much of the intuition and technical language is transferable to other discrete choice models.

Suppose a researcher observes a large number of individuals choosing between a number of J available alternatives. Now the researcher supposes that each individual i would derive utility V_{ij} from choosing option $j = \{1, .., J\}$, where V_{ij} has the following form:

$$V_{ij} = \underbrace{\beta X_j}_{\textit{deterministic}} + \underbrace{\varepsilon_{nj}}_{\textit{random}} \tag{39}$$

Here β is a vector of population preferences for different attributes, and X_j is a vector of observable attributes of option j. Commonly, βX_j is referred to as the *deterministic* part of utility, since it is common for all members of the population, while ε_{ij} is called the *random* part of utility, which represents individual $i's$ idiosyncratic preference for option j. Each agent picks the alternative that provides them with the highest utility:

$$V_{ij^*} = \max\{V_{i1}, ...V_{iJ}\} \tag{40}$$
$$j^* = \arg\max\{V_{i1}, ...V_{iJ}\} \tag{41}$$

Given the presence of the random component of utility, every alternative has (at least in principle) the chance of getting chosen by some individual in the population. Denote the probability of option k being chosen by individual i as P_{ik} :

$$P_{ik} = \Pr(\beta X_k + \varepsilon_{nk} > \beta X_j + \varepsilon_{nj} \forall j \neq k) \tag{42}$$
$$P_{ik} = \Pr(\varepsilon_{nj} < \beta X_j - \beta X_k + \varepsilon_{nk} \forall j \neq k) \tag{43}$$

P_{ik} can be interpreted as the ex ante (before ε_{nj} is realized) probability of i choosing option k. Alternatively, in a large sample of individuals P_{ik} can be interpreted as the proportion of individuals who choose option k.

In order to estimate the values of the population preferences β, we need to specify a distribution for the random component of utility. In the case of the logit model, we assume that each ε_{ij} is distributed i.i.d. type I extreme value. The probability density function of this distribution is given by:

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{ij}e^{-e^{-\varepsilon_{ij}}}} \tag{44}$$

and the cumulative density function by:

$$F(\varepsilon_{ij}) = e^{-e^{-\varepsilon_{ij}}} \tag{45}$$

Using this functional form, one can derive the choice probabilities in closed form (McFadden (1974)):

$$P_{ik} = \frac{e^{\beta X_k}}{\sum_j e^{\beta X_j}} \tag{46}$$

This is the famous logit formula, which provides the starting point for the discussion in section 3.