To what degree? Recovering changes in the UK's graduate Skill Distribution

Max Schroeder

University of Glasgow

24th of November 2020

Introduction

- For about a century, a university degree has been seen as a clear route to professional success (e.g. Goldin & Katz (2009)).
- However, recently the value of a degree has become a little less obvious:
    - Rising graduate earnings inequality (Altonji *et al.* (2016)).
    - Increasing *underemployment* of graduates (Holmes & Mayhew (2016)).

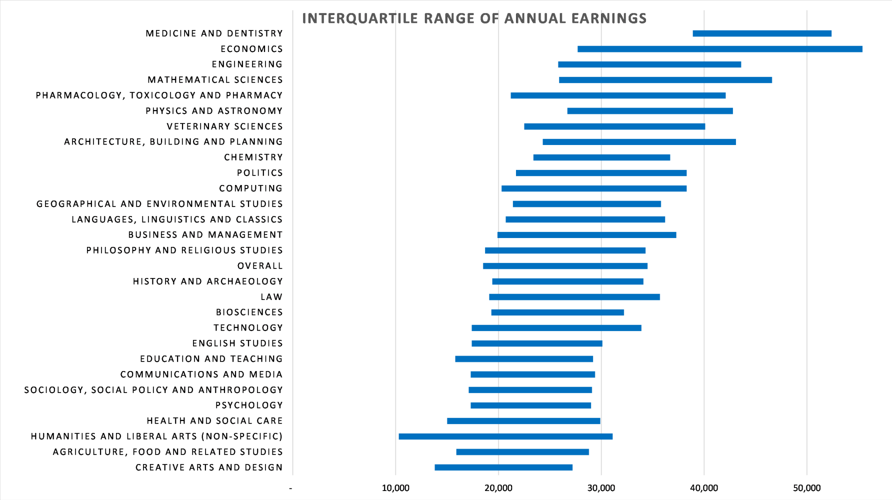

Figure: 5 years after graduation; Source: DfE LEO database, authors own calculations

Motivation

- The dispersion of graduate's earnings appears to be large. But what accounts for these differences?
- Possible Factors considered by the literature (non exhaustive):
  - University Quality (Drydakis (2016))
  - Student's performance (Feng & Graetz (2017))
  - Degree Subject (Lindley & MacIntosh (2015))

How this paper fits in

- Common thread through this literature: Earnings differentials **between** and **within** identifiable groups of graduates.
- **Differentials of Skills (Human Capital) between** and **within** identifiable groups of graduates.
- Importance of studying the graduate skill distribution: Not only do graduates account for an increasing part of the population, but their skills are also crucial for economic development and growth.

Research Question

In the absence of any direct measures -
What can we say about the (changing) distribution of graduates' skills?

Aims of this paper

- Estimate a structural occupational choice model based on the task-skill framework to recover the unobserved parameters of the **multidimensional** graduate skill distribution.
- Establish a set of stylized facts about the changing skill distribution of university graduates, and highlight the extend of skill inequality **between** and **within** university subjects.
- Be *agnostic* about the origins of these skills (training, selection, innate ability, etc.).

The Economic Environment

- There is a finite number of *Occupations*, that differ in their demand for a finite number of general skills according to the tasks that are required to produce their output in each case.
- Similarly, workers are heterogenous with respect to the multidimensional set of skills that they have.
- Workers choose which occupation they want to enter workers are paid their marginal product

Occupations and Wage Setting

- Occupations have different production technologies, summarised by a K dimensional task vector λ, that determines how productive different type of workers' skills are.
- They act perfectly competetive and pay workers their marginal product: ▸ Derivation

Wage Setting

$$\ln(w_{i,o}) = \underbrace{\eta_o}_{FE} + \sum_K \underbrace{\lambda_{o,k}}_{tasks} * \underbrace{s_{i,k}}_{skills}$$

The Graduate's Occupational Choice Problem

- The graduate perfectly knows her own skills and preferences and the task demands of all occupations an therefore expected earnings in each occupation.
- The value of an occupation choice is determined by **a)** the consumption value of expected earnings and **b)** a random, idiosyncratic preference shock.

The Occupation Choice Problem

$$o_i^* = \arg\max_{o \in O} \left\{ V_{i,o} = \underbrace{\ln(w_{i,o})}_{deterministic} + \underbrace{\varepsilon_{i,o}}_{idiosyncratic} \right\}$$

Econometric Specification

- I make 3 assumptions in order to estimate the model:

1. $\varepsilon_{i,o}$ are distributed **i.i.d.** Type I EV (**Logit Assumption**).
2. Skills follow a multivariate log-normal distribution:

$$\log(s_i) \sim MVN(\mu, \Sigma)$$

3. Wages are imprecicely measured:

$$\begin{aligned} ln(w_i^{obs}) &= \ln(w_i) + v_i \\ v_i &\sim N(0, \phi^2) \end{aligned}$$

Likelihood Function

- By combining the elements of the model, we can form the likelihood function: ▸ Derivation

Unconditional Likelihood Function

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

Estimation Strategy

- The estimation procedure works via *simulated maximum likelihood* - we find the vector of economic parameters θ, that maximizes the probability of observing the actual outcomes.
- Necessary because of the complexity of the processes involved, as well as the presence of unobservables.
- Focus on computation and simulation.

▸ Estimation Algorithm

Data Sources

- Quarterly Labour Force Survey (1994 - 2011)
  - Degree Subject
  - Gross hourly wages (CPI deflated)
  - Occupation (ISCO 88)
- Skills and Employment Survey (1997, 2001, 2006, 2012)
  - Task intensities
- Focus on *fresh* graduates ($<$ 24 years old). Resulting in a sample of $\sim 5,000$ individuals.

Degree Subjects

Subject	1994 – 2002	2003 – 2011	Change (%)
Medical & Life Sciences	339	523	54.3
STEM	578	550	-4.8
Economics & Business	449	475	5.8
Arts & Humanities	419	630	50.4
Education	127	175	37.8
Other	248	282	13.7
Total	2160	2635	22.0

Tasks (with Examples)

1. Mathematical-Technical Tasks
    - Specialist knowledge and understanding
    - Analyzing complex problems in depth
    - Advanced Mathematics/Statistics
2. Verbal-Organisatorial Tasks
    - Making Speeches/Presentations
    - Write long documents
    - Planning activities of others
3. Interpersonal Tasks
    - Dealing with people
    - Selling a product or service
    - Counselling, advising or caring

Task Intensities (Table)

	Mathematical/Technical			Verbal/Organisatorial			Interpersonal		
2-Digit ISCO88 OCCUPATION	1997 & 2001	2006 & 2012	Change	1997 & 2001	2006 & 2012	Change	1997 & 2001	2006 & 2012	Change
ARMED FORCES	0.67	0.65	-0.02	0.76	0.74	-0.02	0.70	0.70	0.00
LEGISLATORS AND SENIOR OFFICIALS	0.69	0.68	-0.02	0.80	0.80	0.00	0.63	0.79	0.16
CORPORATE MANAGERS	0.74	0.76	0.02	0.76	0.78	0.02	0.81	0.83	0.02
GENERAL MANAGERS	0.72	0.75	0.03	0.68	0.73	0.05	0.81	0.84	0.03
PHYSICAL, MATHEMATICAL AND ENGINEERING	0.80	0.83	0.03	0.71	0.73	0.02	0.69	0.71	0.02
LIFE SCIENCE AND HEALTH PROFESSIONALS	0.73	0.77	0.04	0.75	0.75	0.00	0.74	0.79	0.05
TEACHING PROFESSIONALS	0.72	0.72	0.00	0.83	0.85	0.02	0.74	0.76	0.02
OTHER PROFESSIONALS	0.71	0.69	-0.02	0.78	0.78	0.01	0.77	0.77	0.01
PHYSICAL AND ENGINEERING SCIENCE ASSOCIATE PROFESSIONALS	0.73	0.73	0.00	0.67	0.69	0.02	0.66	0.70	0.04
LIFE SCIENCE AND HEALTH ASSOCIATE PROFESSIONALS	0.66	0.73	0.08	0.76	0.80	0.04	0.80	0.82	0.02
TEACHING ASSOCIATE PROFESSIONALS	0.64	0.67	0.03	0.79	0.79	0.00	0.77	0.82	0.05
OTHER ASSOCIATE PROFESSIONALS	0.68	0.71	0.03	0.70	0.75	0.05	0.78	0.78	0.00
OFFICE CLERKS	0.61	0.67	0.06	0.65	0.69	0.05	0.64	0.68	0.03
CUSTOMER SERVICE CLERKS	0.62	0.66	0.05	0.64	0.66	0.02	0.75	0.79	0.04
PERSONAL AND PROTECTIVE SERVICES WORKERS	0.51	0.59	0.08	0.59	0.70	0.11	0.69	0.75	0.06
MODELS, SALESPERSONS AND DEMONSTRATORS	0.50	0.59	0.08	0.47	0.58	0.11	0.75	0.80	0.05
MARKET-ORIENTED SKILLED AGRICULTURAL AN	0.55	0.66	0.11	0.53	0.65	0.13	0.59	0.61	0.02
EXTRACTION AND BUILDING TRADE WORKERS	0.65	0.72	0.08	0.57	0.61	0.04	0.62	0.65	0.03
METAL, MACHINERY AND RELATED TRADES WORKERS	0.68	0.74	0.05	0.58	0.65	0.06	0.59	0.65	0.06
PRECISION, HANDICRAFT, PRINTING AND RELATED TRADES WORKERS	0.60	0.72	0.12	0.52	0.64	0.12	0.51	0.62	0.12
OTHER CRAFT AND RELATED TRADES WORKERS	0.54	0.63	0.08	0.45	0.57	0.12	0.54	0.67	0.13
STATIONARY PLANT AND RELATED OPERATORS	0.59	0.71	0.12	0.51	0.62	0.10	0.50	0.57	0.07
MACHINE OPERATORS AND ASSEMBLERS	0.55	0.65	0.10	0.47	0.61	0.14	0.49	0.60	0.11
DRIVERS AND MOBILE PLANT OPERATORS	0.45	0.52	0.08	0.49	0.57	0.08	0.56	0.60	0.04
SALES AND SERVICES ELEMENTARY OCCUPATIONS	0.42	0.55	0.13	0.46	0.62	0.16	0.57	0.70	0.13
AGRICULTURAL, FISHERY AND RELATED LABOURERS	0.43	0.65	0.22	0.42	0.59	0.17	0.45	0.56	0.11
LABOURERS IN MINING, CONSTRUCTION, MANUFACTURING & TRANSPORT	0.50	0.60	0.10	0.50	0.60	0.10	0.50	0.62	0.12

Estimates Period 1994 - 2002

	Medical & Life Sciences	STEM	Economics & Business	Arts & Humanities	Education	Other
μ_{Math}	-0.20	-0.17	-0.76	-0.77	-0.28	-0.71
	(0.0244)	(0.1130)	(0.0225)	(0.0187)	(0.0526)	(0.0469)
μ_{Verbal}	-0.22	-0.44	-0.08	-0.75	-0.48	-0.32
	(0.0231)	(0.0930)	(0.0104)	(0.0267)	(0.0280)	(0.0303)
μ_{Inter}	-0.32	-0.47	-0.29	0.09	-0.16	-0.19
	(0.0348)	(0.0114)	(0.0048)	(0.0019)	(0.0106)	(0.0042)
σ_{Math}	0.40	0.39	0.53	0.54	0.48	0.64
	(0.0049)	(0.0053)	(0.0197)	(0.0183)	(0.0221)	(0.0258)
σ_{Verbal}	0.29	0.33	0.33	0.53	0.50	0.45
	(0.0036)	(0.0054)	(0.0091)	(0.0122)	(0.0242)	(0.0095)
σ_{Inter}	0.41	0.48	0.39	0.26	0.54	0.39
	(0.0042)	(0.0046)	(0.0088)	(0.0059)	(0.0475)	(0.0060)
ϕ	0.23					

numerical standard errors in parentheses

Estimates Period 2003 - 2011

	Medical & Life Sciences	STEM	Economics & Business	Arts & Humanities	Education	Other
μ_{Math}	-0.24	0.04	-0.84	-0.55	-0.27	-0.54
	(0.0422)	(0.0153)	(0.0975)	(0.0348)	(0.2776)	(0.0529)
μ_{Verbal}	-0.41	-0.37	-0.23	-0.51	-0.14	-0.24
	(0.0443)	(0.0375)	(0.0240)	(0.1737)	(0.0509)	(0.0117)
μ_{Inter}	-0.25	-0.72	-0.08	-0.13	-0.59	-0.38
	(0.0255)	(0.0129)	(0.0032)	(0.0029)	(0.0092)	(0.0137)
σ_{Math}	0.30	0.41	0.64	0.25	0.31	0.32
	(0.0163)	(0.0129)	(0.0068)	(0.0525)	(0.0079)	(0.0060)
σ_{Verbal}	0.29	0.32	0.36	0.29	0.19	0.38
	(0.0162)	(0.0059)	(0.0703)	(0.0271)	(0.0026)	(0.0139)
σ_{Inter}	0.38	0.36	0.21	0.35	0.45	0.39
	(0.0189)	(0.0206)	(0.0030)	(0.0046)	(0.0057)	(0.0066)
ϕ	0.23					

numerical standard errors in parentheses

Skills - Means

	Mathematical/Technical			Verbal/Organisatorial			Interpersonal		
Subject	1994 — 2002	2003 — 2011	Change	1994 — 2002	2003 — 2011	Change	1994 — 2002	2003 — 2011	Change
Medical & Life Sciences	0.89	0.82	-0.06	0.84	0.69	-0.14	0.79	0.84	0.05
STEM	0.91	1.13	0.22	0.68	0.72	0.04	0.70	0.52	-0.18
Economics & Business	0.54	0.53	-0.01	0.97	0.85	-0.12	0.80	0.95	0.14
Arts & Humanities	0.54	0.59	0.06	0.54	0.62	0.08	1.13	0.94	-0.20
Education	0.85	0.80	-0.05	0.70	0.88	0.18	0.99	0.61	-0.37
Other	0.60	0.62	0.02	0.80	0.85	0.05	0.89	0.74	-0.15
Overall Mean	0.72	0.76	0.04	0.75	0.74	-0.01	0.86	0.79	-0.07

Skills Distribution Aggregate

	Mathematical/Technical			Verbal/Organisatorial			Interpersonal		
	1994 — 2002	2003 — 2011	Change (%)	1994 — 2002	2003 — 2011	Change (%)	1994 — 2002	2003 — 2011	Change (%)
Gini Coefficient	0.297	0.263	- 11.55	0.240	0.189	-21.13	0.242	0.226	-6.59
	(0.001)	(0.001)	-	(0.000)	(0.001)	-	(0.001)	(0.000)	-
Upper Tail ($\frac{P90}{P50}$)	1.914	1.820	- 4.92	1.672	1.541	-7.87	1.696	1.596	-5.87
	(0.005)	(0.005)	-	(0.004)	(0.003)	-	(0.004)	(0.003)	-
Lower Tail ($\frac{P10}{P50}$)	0.442	0.558	26.19	0.533	0.653	22.59	0.536	0.538	0.39
	(0.001)	(0.002)	-	(0.002)	(0.001)	-	(0.002)	(0.002)	-
Within Subject	0.125	0.088	-29.71	0.079	0.054	-31.16	0.084	0.063	-24.68
	(0.001)	(0.001)	-	(0.001)	(0.000)	-	(0.001)	(0.000)	-
Between Subject	0.030	0.043	45.34	0.020	0.008	-60.07	0.016	0.022	33.33
	(0.000)	(0.001)	-	(0.000)	(0.000)	-	(0.000)	(0.000)	-

Data simulated from 100.000 representative draws. Bootstrapped standard errors in parentheses.

Skills - Distributions IQR

	Mathematical/Technical [p25 - p75]		Verbal/Organisatorial [p25 - p75]		Interpersonal [p25 - p75]	
Subject	1994 – 2002	2003 – 2011	1994 – 2002	2003 – 2011	1994 – 2002	2003 – 2011
Medical & Life Sciences	[0.62 - 1.07]	[0.64 - 0.96]	[0.66 - 0.97]	[0.54 - 0.81]	[0.50 - 0.95]	[0.61 - 1.00]
STEM	[0.65 - 1.10]	[0.79 - 1.37]	[0.52 - 0.81]	[0.55 - 0.86]	[0.45 - 0.86]	[0.38 - 0.62]
Economics & Business	[0.33 - 0.67]	[0.28 - 0.66]	[0.74 - 1.15]	[0.63 - 1.01]	[0.58 - 0.97]	[0.80 - 1.07]
Arts & Humanities	[0.32 - 0.67]	[0.49 - 0.68]	[0.33 - 0.67]	[0.49 - 0.73]	[0.92 - 1.31]	[0.69 - 1.12]
Education	[0.55 - 1.05]	[0.62 - 0.94]	[0.44 - 0.87]	[0.76 - 0.99]	[0.60 - 1.23]	[0.41 - 0.75]
Other	[0.32 - 0.75]	[0.47 - 0.73]	[0.54 - 0.98]	[0.61 - 1.02]	[0.63 - 1.07]	[0.53 - 0.89]

Conclusion - what have we learned?

- This paper has provided novel quantitative evidence on the changing distribution of graduate's skills in the UK. There are clear differences of skill endowments for students from different subject areas, but within subject differences due to ther potential factors are also sizeable if not more so.

- Overall, skill inequality has been falling so that graduates today are more similar than 20 years ago.
- Reduction in skill inequality has been driven by an increase in the lower tail.
- Within subject skill inequality is also falling, but there appears to be an uptick of between subject inequality with respect to mathematical and technical skills.

Derivation of marginal product

$$h_{io} = e^{\sum_K \lambda_{ok} s_{ik}} \qquad \text{(Workers HC)}$$

$$Y = F(H_1, ..., H_O) \qquad \text{(Aggregate Production Function)}$$

$$\frac{\partial Y}{\partial h_{io}} = \frac{\partial F}{\partial H_o}\frac{\partial H_o}{\partial h_{io}} = \frac{\partial F}{\partial H_o} e^{\sum_K \lambda_{ok} s_{ik}} \qquad \text{(Take derivative)}$$

$$\frac{\partial F}{\partial H_o} = \eta_o$$

$$w_{io} = \eta_o + \sum_K \lambda_{ok} s_{ik} \qquad \text{(Log Wage Equation)}$$

back

Derivation of Likelihood

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

(Rewrite unconditional probability)

$$\frac{Pr(o_i^*, w_i^{obs}|s_i)}{Pr(o_i^*|s_i)} = Pr(w_i^{obs}|s_i, o_i^*) \qquad \text{(Bayes Rule)}$$

$$Pr(o_i^*, w_i^{obs}|s_i) = Pr(o_i^*|s_i) * Pr(w_i^{obs}|s_i, o_i^*) \qquad \text{(rearange)}$$

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

(Integrate over skill distribution)

Derivation of Likelihood cont.

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

$$\Pr(w_i^{obs}|s_i, o_i^*) = \frac{e^{\left(-\frac{\nu_i^2}{2\phi^2}\right)}}{\sqrt{2\pi\phi^2}} \qquad \text{(Normal PDF)}$$

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

back

Estimation Strategy

1. Find $\Pr(o^*, w_i^{obs}|s_i;\theta)$ using the structure of the economic model.
2. Integrate over the unobserved skill distribution $f(s|\theta))$: $\Pr(o^*, w_i^{obs}|\theta)) = \int \Pr(o^*, w_i^{obs}|s_i;\theta))f(s|\theta))ds$.
3. Find vector of parameters $\hat{\theta}$ that maximizes the LogLikelihood function: $ll(\theta) = \sum_{i=1}^{N} \log(\Pr_i(o^*, w_i^{obs}|\theta)))$

back