

Intelligence and Income

Nurfatima Jandarova¹ Aldo Rustichini²

June 6, 2024

40th Summer Seminar of Finnish Economists 2024

¹Center of Excellence in Tax Systems Research, Tampere University

²Department of Economics, University of Minnesota

- Cognitive and noncognitive skills and family background are important determinants of education
- Policies can change the relative importance of these factors (Ichino, Rustichini, and Zanella 2022)
- Growing evidence on the importance of genes (Rustichini et al. 2023)

- A model of educational choice
 - individual skills, family background, labour market returns
 - general characterisation of regression function
- Estimate in the UK survey data
 - robustness to functional forms
 - subsample: genotype \rightarrow polygenic scores

Determinants of education: Heckman, Stixrud, and Urzua (2006), Almlund et al. (2011), Björklund and Salvanes (2011), Ichino, Rustichini, and Zanella (2022)

Consider all characteristics (skills and family background) together + polygenic analysis

Characterise link between regression equation and model solution

Genes and environment in education: Rustichini et al. (2023)

Nonlinear effort choice

UK Household Longitudinal Study (2009-)

Working sample: 22 881 individuals

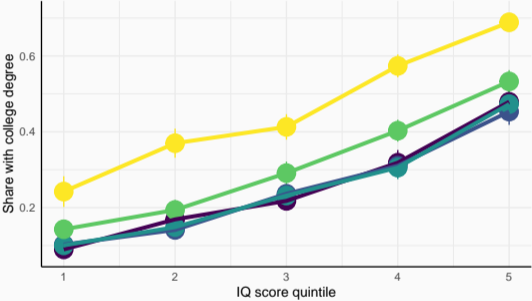
- college: ever had HE degree as highest qualification
- predicted discounted present value of earnings [Profiles](#)
- cognitive test scores [CFA](#), Big 5 personality scores [PCA](#)
- fam advantage score: parents' education and employment status [PCA](#)

METADAC

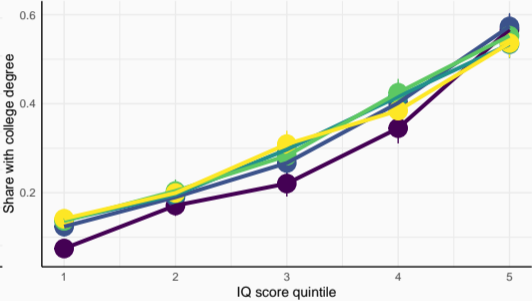
Genotyped subsample: 3 413 individuals [Table](#)

- polygenic score (PGS) of fluid intelligence (Savage et al. 2018) [Illustration](#) [Correlation](#)

College and individual characteristics



Fam score quintile 1 2 3 4 5



Big5 score quintile 1 2 3 4 5

Probability of college

	Born in 1950-64		Born in 1965-79		Born in 1980-94	
	OLS	Logit ME	OLS	Logit ME	OLS	Logit ME
IQ score	0.137*** (0.005)	0.152*** (0.007)	0.151*** (0.005)	0.151*** (0.005)	0.130*** (0.006)	0.130*** (0.006)
Fam score	0.050*** (0.005)	0.054*** (0.008)	0.077*** (0.006)	0.077*** (0.006)	0.074*** (0.006)	0.074*** (0.006)
Big5 score	0.006 (0.006)	0.015** (0.006)	0.010* (0.006)	0.010* (0.006)	0.014* (0.006)	0.014* (0.006)
Obs.	9 539	9 539	10 586	10 586	8 409	8 409

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

SEM of college and wages

	Born in 1950-64		Born in 1965-79		Born in 1980-94	
	Pred. wage	College	Pred. wage	College	Pred. wage	College
College	0.820*** (0.020)		0.804*** (0.016)		0.651*** (0.015)	
IQ score	0.082*** (0.009)	1.018*** (0.046)	0.077*** (0.007)	0.917*** (0.039)	0.045*** (0.007)	0.788*** (0.044)
Fam score		0.382*** (0.051)		0.492*** (0.040)		0.453*** (0.041)
Big5 score		0.082* (0.037)		0.081** (0.033)		0.128*** (0.037)
Indirect effect	0.126*** (0.006)		0.136*** (0.005)		0.096*** (0.005)	
Total effect	0.208*** (0.009)		0.213*** (0.007)		0.140*** (0.007)	
Obs.	9 496	9 496	10 488	10 488	8 382	8 382

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Model

Individuals described by $z \in Z = \Theta \times X \times Y$

- intelligence $\theta \in \Theta$
- family advantage score $x \in X$
- Big 5 personality score $y \in Y$

Human capital $H \equiv \{nc, c\}$ (no college vs college)

DPV of earnings $W(h, z, \delta) = \sum_{a=18}^{65} \delta^{a-18} W(h, z, a)$

Choose effort $e \in \mathbb{R}_+$ to acquire human capital $h = c$ given cost $\frac{c(e)}{\Gamma(z)}$

$$\max_e \pi(e) [W(c, z, \delta) - W(nc, z, \delta)] - \frac{c(e)}{\Gamma(z)}$$

Solution

Denote $A = (W(c, z, \delta) - W(nc, z, \delta)) \Gamma(z)$. Then, optimal effort solution

$$E^*(A; \pi) \equiv \arg \max_e \pi(e)A - e$$

Definition

Π is the set of functions $\pi : \mathbb{R}_+ \rightarrow [0, 1]$ that are strictly increasing, concave, continuous at 0, $\pi(0) = 0$, $\lim_{x \rightarrow \infty} \pi(x) = 1$.

Proposition

For $P(A)$ increasing in A and upper semi-continuous in ΔW , $\exists \pi \in \Pi$ such that

$$P(A) = \pi(E^*(A; \pi))$$

We consider four functional forms that can describe $P(A)$:

- Linear probability model $P(A) = A$
- Logit $P(A) = (1 + e^{-A})^{-1}$
- Logit power $P(A) = (1 + e^{-A})^{-\kappa}$, $\kappa \in \mathbb{R}_+$ Plot
- Cutoff power $P(A) = \min\{\max\{A, 0\}^\kappa, 1\}$, $\kappa \in \mathbb{R}_+$ Plot

Two-step estimation:

1. Given δ and κ , fit $P(A)$ to the observed college indicators.
2. Grid search $\hat{\delta}$ and $\hat{\kappa}$ that minimise sum of squared residuals. Plots

Results

	LPM	Logit	Logit power ¹	Cutoff power ¹
IQ score	0.118*** (0.005)	0.089*** (0.007)	0.104*** (0.006)	0.108*** (0.006)
Fam score	0.061*** (0.004)	0.054*** (0.004)	0.058*** (0.004)	0.058*** (0.004)
Big5 score	0.025*** (0.003)	0.038*** (0.004)	0.032*** (0.004)	0.031*** (0.004)
College premium, std	0.026*** (0.006)	0.060*** (0.009)	0.042*** (0.007)	0.040*** (0.007)
Obs.	31 571	31 571	31 571	31 571
δ	0.925	0.925	0.925	0.925
κ			2.90	1.20
College premium mean	36.95	36.95	36.95	36.95
College premium sd	9.77	9.77	9.77	9.77

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

¹ Bootstrapped standard errors

Results with polygenic scores

Simple logit without college premium

	LPM	Logit ME
IQ PGS	0.067*** (0.007)	0.066*** (0.007)
Fam score	0.094*** (0.008)	0.118*** (0.010)
Big5 score	0.019** (0.008)	0.019** (0.008)
Obs.	3 602	3 602

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Results with polygenic scores

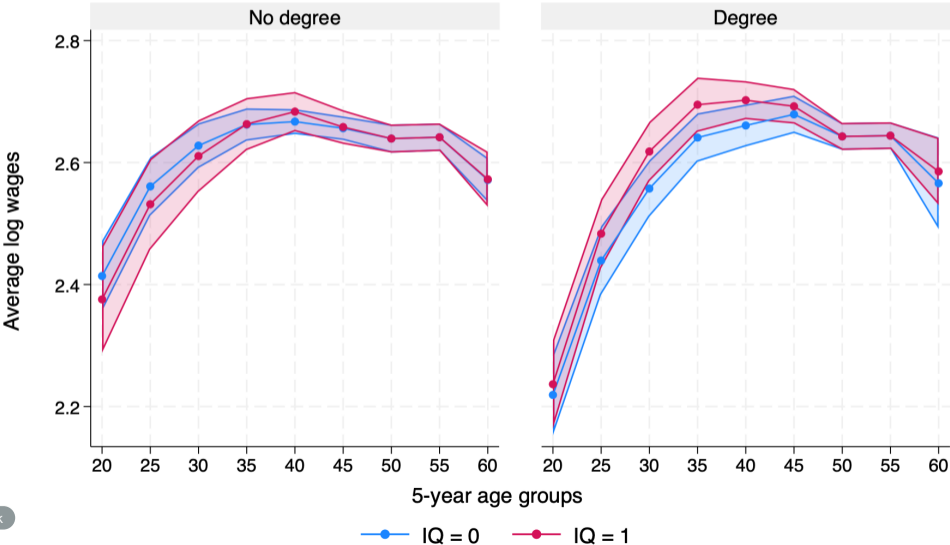
	LPM	Logit	Logit power	Cutoff power
IQ PGS	0.038*** (0.011)	0.034** (0.012)	0.036** (0.012)	0.037*** (0.011)
Fam score	0.085*** (0.008)	0.154*** (0.011)	0.146*** (0.011)	0.134*** (0.010)
Big5 score	0.109*** (0.031)	0.102** (0.034)	0.105** (0.034)	0.115*** (0.034)
College premium, std	0.007** (0.002)	0.006** (0.002)	0.006** (0.002)	0.007** (0.002)
Obs.	3 602	3 602	3 602	3 602
δ	0.925	0.925	0.925	0.925
κ			2.90	1.20
College premium mean	84.86	84.86	84.86	84.86
College premium sd	14.27	14.27	14.27	14.27

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

- Revisit role of intelligence, personality and family characteristics in education choice
- Conditions linking econometric specification to individual optimization
- Further analysis with polygenic scores

Appendices

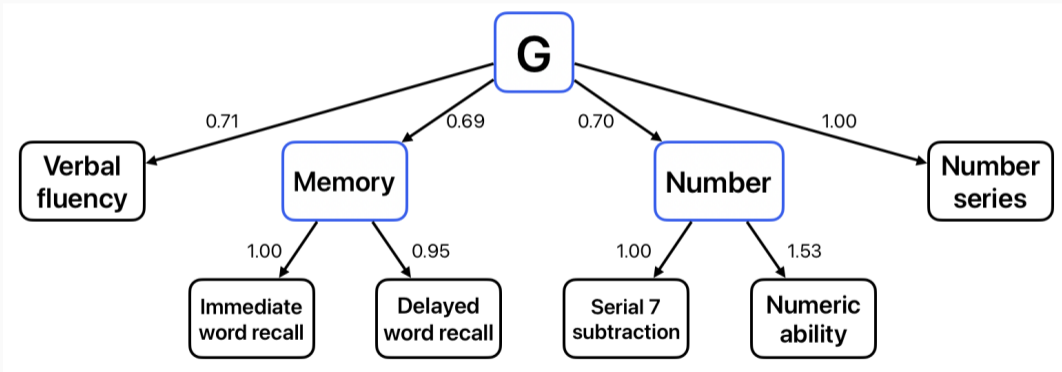
Predicted wage profile



Back

Intelligence score

Confirmatory factor analysis related to (Johnson and J.Bouchard 2005)



Big 5 personality score

Combine individual test scores using principal component analysis

Score	Loading
Agreeableness	0.4408
Conscientiousness	0.4970
Extraversion	0.4628
Neuroticism	-0.3751
Openness	0.4514

PC1 explains 36% of variation in the data

Family advantage score

Combine education of parents and their employment status using principal component analysis

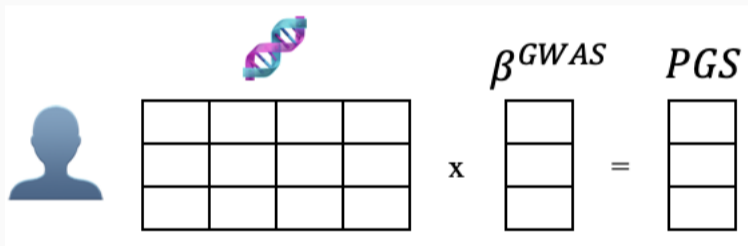
Variable	Loading (mother)	Loading (father)
Years of education	0.4020	0.4286
Work	0.2243	0.5527
Dead	-0.0889	-0.2958
Absent	-0.2042	-0.4023

PC1 explains 23% of variation in the data

METADAC vs full sample

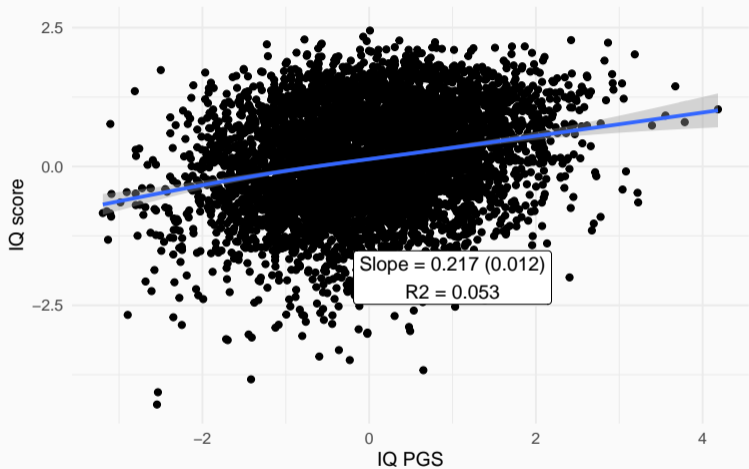
	Male	Age	White British	College	Mother worked	Mother's years of edu	Father worked	Father's years of edu
UKHLS								
Working sample	0.440 (0.496) 31 571	41.811 (13.907) 31 571	0.818 (0.386) 31 571	0.304 (0.460) 31 571	0.600 (0.490) 31 571	11.480 (2.830) 31 571	0.823 (0.382) 31 571	11.775 (3.464) 31 571
METADAC								
Full sample	0.428 (0.495) 7 281	46.345 (12.901) 7 236	0.973 (0.161) 7 281	0.264 (0.441) 7 251	0.666 (0.472) 5 248	11.259 (2.369) 7 281	0.887 (0.316) 5 256	11.298 (3.524) 7 281
Working sample	0.441 (0.497) 3 413	45.865 (10.890) 3 413	1.000 (0.000) 3 413	0.304 (0.460) 3 413	0.695 (0.460) 3 413	11.759 (2.420) 3 413	0.898 (0.303) 3 413	12.070 (3.423) 3 413

Polygenic score computation



Back

Polygenic score



Best EA4 R2 = 0.1328 (Okbay et al., 2022)

Proposition (continuously differentiable)

Redefine the effort choice problem as $e^*(\alpha; \pi) \equiv \arg \max_e \pi(e) - \alpha e$ where $\alpha = A^{-1}$.

Definition

The set of endogenous probabilities is the set \mathcal{Q} of multivalued functions $Q : \mathbb{R}_+ \rightarrow [0, 1]$ that are decreasing, closed valued, with $\lim_{\alpha \rightarrow 0} Q(\alpha) = 1$, $Q(\bar{\alpha}) = 0$ for some $\bar{\alpha} > 0$.

Proposition

For any function $Q \in \mathcal{Q}$ which is continuously differentiable strictly decreasing in the interval $[\underline{\alpha}, \bar{\alpha}]$, with $Q(\underline{\alpha}) = 0$, $Q(\bar{\alpha}) = 1$, there exists a continuously differentiable function $\pi \in \Pi$ such that for all $\alpha \in \mathbb{R}_+$, $Q(\alpha) = \pi(h(\alpha; \pi))$.

Proposition proof (continuously differentiable)

Note that $Q(\alpha) = 0$ for $\alpha > \bar{\alpha}$, so we may take the boundary condition

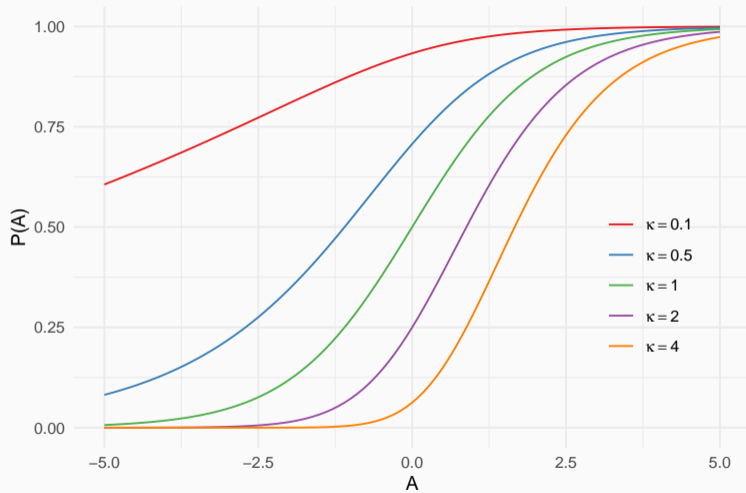
$$h(\bar{\alpha}) = 0$$

Consider the ordinary differential equation

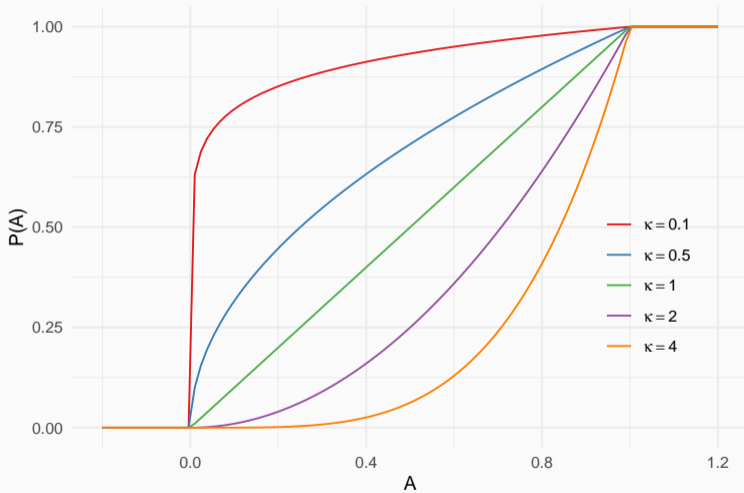
$$\frac{dh}{d\alpha} = \frac{Q'}{\alpha}, \quad \alpha > 0$$

We now define the function π as the solution of $\pi(h(\alpha)) = Q(\alpha)$. The function h satisfies the differential equation, which is the first order necessary and sufficient conditions for optimal effort choice problem, namely $\pi'(h(\alpha)) = \alpha$. Thus, our claim follows.

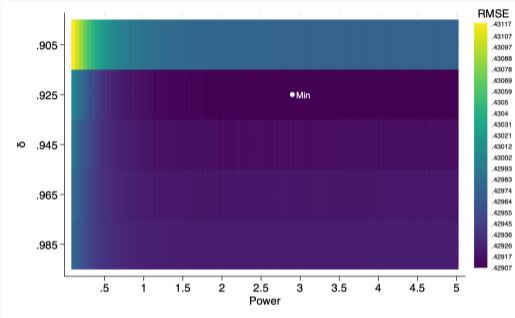
Logit power



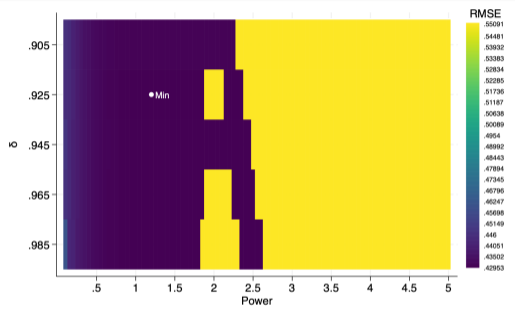
Cutoff power



Grid search



(a) Logit power



(b) Cutoff power

Figure 1: RMSE heatmap of grid search over δ and κ

- Almlund, Mathilde, Angela Lee Duckworth, James Heckman, and Tim Kautz. 2011. "Personality Psychology and Economics." In, 4:1–181. Elsevier. <https://doi.org/10.1016/B978-0-444-53444-6.00001-8>.
- Björklund, Anders, and Kjell G. Salvanes. 2011. "Education and Family Background." In, 201–47. Elsevier. <https://doi.org/10.1016/b978-0-444-53429-3.00003-x>.
- Heckman, James, Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *Journal of Labor Economics* 24 (3): 411–82.

- Ichino, Andrea, Aldo Rustichini, and Giulio Zanella. 2022. "College Education, Intelligence, and Disadvantage: Policy Lessons from the UK in 1960-2004." http://www.andreaichino.it/wp-content/uploads/IRZ_Sorting.pdf.
- Johnson, Wendy, and Thomas J. Bouchard. 2005. "The Structure of Human Intelligence: It Is Verbal, Perceptual, and Image Rotation (VPR), Not Fluid and Crystallized." *Intelligence* 33 (4): 393–416. <https://doi.org/https://doi.org/10.1016/j.intell.2004.12.002>.
- Rustichini, Aldo, William G. Iacono, James J. Lee, and Matt McGue. 2023. "Educational Attainment and Intergenerational Mobility: A Polygenic Score Analysis." *Journal of Political Economy* 131 (10): 2724–79. <https://doi.org/10.1086/724860>.

Savage, Jeanne E., Philip R. Jansen, Sven Stringer, Kyoko Watanabe, Julien Bryois, Christiaan A. de Leeuw, Mats Nagel, et al. 2018. “Genome-Wide Association Meta-Analysis in 269,867 Individuals Identifies New Genetic and Functional Links to Intelligence.” *Nature Genetics* 50 (7): 912–19.
<https://doi.org/10.1038/s41588-018-0152-6>.