ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 75/25

College, cognitive ability, and socioeconomic disadvantage: policy lessons from the UK in 1960-2004

Andrea Ichino, Aldo Rustichini, Giulio Zanella

www.rfberlin.com SEPTEMBER 2025

College, cognitive ability, and socioeconomic disadvantage: policy lessons from the UK in 1960-2004

Authors

Andrea Ichino, Aldo Rustichini, Giulio Zanella

Reference

JEL Codes: I23, I28, J24, O33

Keywords: college education, university, cognitive ability, disadvantage

Recommended Citation: Andrea Ichino, Aldo Rustichini, Giulio Zanella (2025): College, cognitive ability, and socioeconomic disadvantage: policy lessons from the UK in 1960-2004. RFBerlin Discussion Paper No. 75/25

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

College, cognitive ability, and socioeconomic disadvantage: policy lessons from the UK in 1960-2004*

Andrea Ichino† Aldo Rustichini‡ Giulio Zanella§

September 17, 2025
Forthcoming in *The Review of Economic Studies*

Abstract

University access has significantly expanded in OECD countries, and further growth figures prominently in political agendas. We study possible consequences of historical and future expansions in a stochastic, general equilibrium Roy model where tertiary educational attainment is determined by cognitive *ability* and socioeconomic *disadvantage*. In our analysis, individual productivity depends not only on education but also directly on cognitive ability. The expansion of university access in the UK that started in the 1960s provides an ideal case study to draw lessons for the future. We find that this expansion led to the selection into college of progressively less talented students from advantaged backgrounds. Appropriate counterfactual policies existed that would have achieved the dual goal of increasing college graduates' cognitive ability while improving tertiary education opportunities for the disadvantaged.

JEL Classification: I23, I28, J24, O33

Keywords: college education, university, cognitive ability, disadvantage,

*We are grateful to Josh Angrist, Sascha Becker, Richard Blundell, Russell Cooper, Tom Crossley, David Green, Daniela Iorio, Wenchao Jin, Lars Kirkebøen, David Levine, Lance Lochner, Tommaso Porzio, Guido Tabellini, Michela Tincani and seminar participants at Bayreuth, Bocconi, Bologna, Bristol, Cardiff, CEPR, CESifo, EALE, EIEF, Konstanz, Leibniz Institute Essen, Louvain-La-Neuve, Mannheim, CSEF Naples, NBER, PSE, SOLE, Stockholm School of Economics, UCL, and York for useful comments. Stav Federman, Nurfatima Jandarova and Johanna Reuter provided outstanding research assistance.

†European University Institute and University of Bologna; andrea.ichino@eui.eu
†University of Minnesota; aldo.rustichini@gmail.com
‡University of Bologna; giulio.zanella@gmail.com

1 Introduction

Enrollment in tertiary education has increased by a factor of 3.4 in OECD countries since 1970 (UNESCO Statistics, 2024)[1] and further growth remains a key political priority. For instance, the European Union's goal for 2030 is that "The share of 25-34 year-olds with tertiary educational attainment should be at least 45%" (EU Council, 2021). Using UK data that span four decades, this paper studies the consequences of such historical and planned expansion processes on the selection of college students in terms of pre-college *cognitive ability* and *socioeconomic disadvantage* (for brevity, hereafter *ability* and *disadvantage*). We use a general equilibrium Roy (1951) model where these traits affect the graduation probability and where their correlation determines how technological change and higher education policy alter incentives to pursue a college degree. In our analysis, productivity and wages depend directly on cognitive ability for a given educational attainment. The model is used to study how actual policies shaped the evolution of students' sorting into college, the college premium dynamics, and to simulate counterfactual policies. In the UK, the share of 17-30 year-olds in higher education rose from 5% in 1960 to 43% in 2007 (Chowdry et al., 2013),[2] an increase observed previously in the US (Goldin and Katz, 2008) and subsequently in other OECD countries (Schofer and Meyer, 2005; Meyer and Schofer, 2007). The UK expansion is thus an ideal case study. We illustrate its nature and consequences, drawing lessons to judge the ambitious targets currently set in Europe and elsewhere.

This expansion originates in the Robbins (1963) Report, which claimed the existence of large "reserves of untapped *ability* [that] may be greatest in the *poorer* sections of the community" (p. 53, our italics) and thus recommended that "all young persons qualified by ability and attainment to pursue a full-time course in higher education should have the opportunity to do so." (p. 49). According to the Report, "fears that expansion would lead to a lowering of the average ability of students in higher education [were] unfounded." (p. 53). These claims are typically made also in other countries by advocates of tertiary education expansions, but have not been adequately investigated for lack of data sets containing a pre-college cognitive ability measure. An upside of our data is that they enable the construction of such a measure, in addition to pre-college socioeconomic status (SES).

[1]This factor was about 1.6 in the US, 3.7 in Japan, 3.8 in France, 4.1 in Italy, and 5.5 in the UK. Enrollment is to any tertiary education program, of students who have successfully completed secondary education.

[2]Similar evidence can be found in Blackburn and Jarman (1993), Boliver (2011), Blanden and Machin (2004), Walker and Zhu (2008), Major and Machin (2018) and Blundell et al. (2022).

We find that: (i) on average, the ability of students selected into college declined by about 13% of a standard deviation between the 1960s and the 1990s; the average ability of high school graduates without a college degree also declined, indicating that students who earned a college degree in the 1990s (and who would not have in the 1960s) had higher ability than high school graduates of the 1960s, but lower ability than the average college graduate of the same period;[3] (ii) the college premium in terms of the discounted present value of earnings remained flat across cohorts despite *education-biased* technological change (EBTC);[4] (iii) the college premium declined at the top of the cognitive ability distribution while increasing at lower ability levels — a novel pattern within education groups that we bring to light, reflecting a phenomenon that we label *ability-biased* technological change (ABTC);[5] (iv) these facts reflect the interaction between technological change and an increase in the number of college graduates achieved by reducing non-tuition costs and by lowering qualification barriers at entry, without sufficient attention to cognitive ability; (v) although "untapped ability" did exist, the policy that prevailed was unfit to draw this ability into universities and favored primarily low-ability students from high-SES families;[6] (vi) policies explicitly designed to select students of sufficiently high ability (possibly combined with aid to the disadvantaged) were feasible and would have reconciled the goals of increasing college graduates' ability and tertiary education opportunities for the underprivileged.

Choosing a social welfare function to guide tertiary education expansions necessarily

[3] Walker and Zhu (2008) and Blundell et al. (2022) consider this hypothesis in their analysis of the evolution of the wage gap between college and high school graduates over years. They cannot test it because their data source (the UK LFS) does not contain an ability measure. Carneiro and Lee (2011) study the increase in college enrollment in the US in 1960-2000 and present evidence consistent with the possibility that the expansion drew into college marginal students of lower quality than average college students.

[4] This finding agrees with Blundell et al.'s 2022 evidence on the UK college premium in terms of hourly wages. Bianchi (2020) studies a large expansion of access to STEM majors enacted in Italy in the early 1960s and finds a negative impact on STEM graduates' wages. Our and their results are in sharp contrast with evidence for the US, where the college premium increased after the 1970s (see, among others, Katz and Murphy, 1992, Fortin, 2006, Goldin and Katz, 2008 and Autor et al., 2020). We comment in Section 5.2.3 on the US-UK difference in light of our results.

[5] This finding is consistent with Levy and Murnane's (1996) and Autor et al.'s (2003) analyses. Among equally educated workers performing a certain task, low-ability individuals may benefit from computerization more than individuals whose ability is already high. More recent evidence is discussed in Section 5.2.1.

[6] These findings agree with Blanden and Machin (2004), Machin (2007), Sutton Trust (2018), Boliver (2013) and Major and Machin (2018), who show that the expansion of UK higher education since the 1960s predominantly benefited children from high-income families. They also agree with Campbell et al. (2019) and Cooper and Liu (2019), who find evidence of a mismatch between ability and educational attainment in the UK and other OECD countries, respectively. We leave in the background other consequences of higher education expansions, such as over-education (Freeman, 1976), i.e., the mismatch between educational attainment and occupation. Cervantes and Cooper (2022) study both margins of mismatch in OECD countries.

involves value judgments into which we do not delve. What matters for our purposes is that a large class of such functions has the property that among two otherwise identical students, the one with higher cognitive ability should have a higher priority in college access. In our model, possible reasons for this property, which we support empirically, are that ability and higher education are complements in determining individual productivity and that, even if returns to college remain positive, a lower pre-college ability is associated (*ceteris paribus*) with a higher study effort cost.[7] In any case, preserving a high ability of students selected into college does not necessarily have to come at the cost of increasing inequality, as confirmed by our counterfactual policy simulations.[8]

Our framework is a general equilibrium model that extends the partial equilibrium setting of Katz and Murphy (1992) and Autor et al. (2020) to an active labor supply side that makes human capital investment decisions. The labor demand side has the standard feature that competitive firms produce output by combining college and high school graduates, thus affecting the college premium. An important novelty of our framework is that workers' productivity, and thus wages, depend on both their education and their cognitive ability. Technological change affects the productivity of college and high school graduates in ways that may differ as a function of their ability. It may favor either higher educational attainment (*education-biased*) or higher cognitive ability (*ability-biased*), or both. As a result, it may activate forces that increase the demand for college graduates or for more able students independently of any change in higher education policy.

The labor supply side is also novel. In our model, obtaining a college degree depends on two factors: an individual's cognitive ability and socioeconomic disadvantage as determined by family background. Ability and disadvantage affect the cost of study effort that an individual must exert to attain a college degree, thereby altering the graduation probability. The government can choose parameters that link ability and disadvantage to study effort cost, thus expanding university access for different groups in the population. The combination of ability and disadvantage with these effort cost parameters generates alternative combinations

[7]Other reasons may be considered in a richer model. For example, universities have a double role in society: providing higher education but also supporting basic research at an advanced level in all fields, a task that is facilitated by higher cognitive ability. Thus, the consequences of a decline in the average ability of graduates are going to be far-reaching, particularly if there is a reluctance to allow tertiary education institutions of higher quality to be more selective in their acceptance.

[8]Surprisingly, these concerns are absent in the cited EU Council (2021), which sets the goal of at least 45% of graduates in the EU by 2030. It is not even clear how this specific threshold has been chosen.

of these two traits that mark the boundary between higher and lower graduation probability regions, a stochastic generalization of the classical Roy (1951) model. A higher education policy is a way to alter such boundaries, thus affecting who is selected into college in terms of ability and disadvantage.

Given a policy, the evolution of graduates' characteristics depends on the correlation between ability and disadvantage at the time of selection into college. The reforms advocated by the Robbins Report in the past and by supporters of tertiary education expansions at present (e.g. EU Council, 2021) are motivated by the belief that we live in stratified societies where university access is hampered by SES disadvantage more than by lack of ability. In this case, if the correlation in question is positive, even expanding tertiary education opportunities for the disadvantaged may increase the fraction of graduates without reducing their average ability. We find that the UK society was indeed stratified, but was characterized by a *negative* correlation between pre-college ability and disadvantage, which is a possibility in the light of the economics of skill formation (e.g., Cunha et al., 2006 and Heckman and Mosso, 2014). The key lesson from the UK experience is that, in such a context, an expansion of tertiary education that aims at equality of opportunities cannot disregard cognitive ability.

The rest of the paper proceeds as follows. Section 2 presents the theoretical model. Section 3 describes the data, Section 4 illustrates the key facts, Section 5 estimates the model and uses it for counterfactual quantitative analysis. Section 6 concludes.

2 Model

2.1 Setup

Workers. There is a unit mass population of economic agents who are fully employed at equilibrium. Each individual is characterized by a given pair $(\theta, \lambda) \in \mathbf{\Theta} \times \mathbf{\Lambda} \subset \mathbb{R}_+ \times \mathbb{R}_+$.[9] Θ denotes cognitive *ability* and its support $\mathbf{\Theta}$ is ordered by the order on the real numbers; Λ summarizes non-cognitive *disadvantage*, i.e., a set of socioeconomic factors that increase study effort cost, and its support $\mathbf{\Lambda}$ is similarly ordered.[10] The joint distribution of ability and disadvantage is denoted by $\mu \in \Delta(\mathbf{\Theta} \times \mathbf{\Lambda})$. While disadvantage is publicly

[9] We use boldface to denote a set, capitals to denote random variables, and lower case to denote generic variables and realizations of random variables.

[10] We frame the analysis in terms of *disadvantage*, instead of *advantage*, because the policy debate on tertiary education typically emphasizes the need to help students who are impaired by a *less favorable* SES.

observable, employers see only a coarse measure of cognitive ability, namely the *ability group* $j = 1, 2, \cdots, J$ to which an individual belongs. The set of possible ability values of individuals in group j is $\mathbf{\Theta}_j$, so that J is the cardinality of the partition of the ability distribution support into the ability groups that are relevant to employers: $\mathbf{\Theta}_1 \cup \mathbf{\Theta}_2 \cup \cdots \cup \mathbf{\Theta}_J = \mathbf{\Theta}$.

Each individual is also characterized by an endogenous human capital level $k \in \mathbf{K}$. Given our focus on higher education, we follow the literature in considering only two levels, and so $\mathbf{K} = \{0, 1\} \equiv \{\text{high school, college}\}$. Educational attainment k is determined by an allocation function π that describes the probability on human capital obtained by an individual, for given cognitive skills and study effort level. The set of effort levels $\mathbf{S}$ is the positive real line. We assume that the human capital level obtained with college, once achieved, cannot be lost, so the only possible transition in human capital is from 0 to 1. In sum, we let $\pi : \mathbf{S} \times \mathbf{\Theta} \rightarrow [0, 1]$, where $\pi(s, \theta)$ is the probability of attaining a college degree for an individual whose ability is θ and who exerts effort s. For an individual of type (θ, λ), preferences are defined over lifetime consumption and leisure, and are represented by

$$u(c, s; \theta, \lambda) = \frac{c^{1-\sigma_c} - 1}{1 - \sigma_c} + \Omega(\lambda) \frac{(1 - \Gamma(\theta)s)^{1-\sigma_s} - 1}{1 - \sigma_s}, \tag{1}$$

where c denotes consumption and $\sigma_c > 0$ and $\sigma_s > 0$ are parameters; the functions $\Omega \geq 0$ and $\Gamma \geq 0$ are *effort cost shifts*, hence non-negative. They are influenced by the policy maker and depend on disadvantage and ability. Absent policy interventions, it is reasonable to assume $\frac{d\Gamma(\theta)}{d\theta} < 0$ (the marginal disutility of study effort decreases with ability) and $\frac{d\Omega(\lambda)}{d\lambda} > 0$ (the same marginal disutility increases with disadvantage).[11] Thus, *ceteris paribus*, college is less costly, in terms of leisure utility, for higher ability and higher SES students. However, for efficiency or equity reasons, higher education policy can alter the opportunity cost of study effort selectively on the basis of Θ and Λ. Examples are provided in Section 2.3.

Let $s^*(\theta, \lambda)$ be the optimal study effort of a (θ, λ)-type individual. Since the model is static, consumption is equal to earnings, which in turn depend on an individual's human capital and ability group. That is, the wage of an individual whose educational attainment is k and whose ability is $\theta \in \mathbf{\Theta}_j$ is given by $w(k, j)$. Given a vector of wages $\mathbf{w}_j \equiv [w(0, j), w(1, j)]$

[11] Effort cost shifts $\Omega(\lambda)$ and $\Gamma(\theta)$ affect utility asymmetrically because we hypothesize that ability improves the effectiveness of study effort directly, while disadvantage affects only the leisure utility at the given effort. E.g., Dillon and Smith (2020) find that student ability improves college completion, while Bailey and Dynarski (2011) and Hoxby and Avery (2013) find that student socioeconomic disadvantage hinders college enrollment.

for each ability group j, an individual solves:[12]

$$\max_{s\geq 0}\left(\pi(s,\theta)\Delta U(\mathbf{w}_j)+\Omega(\lambda)\frac{(1-\Gamma(\theta)s)^{1-\sigma_s}-1}{1-\sigma_s}\right), \tag{2}$$

where we denote

$$\Delta U(\mathbf{w}_j)\equiv\frac{w(1,j)^{1-\sigma_c}}{1-\sigma_c}-\frac{w(0,j)^{1-\sigma_c}}{1-\sigma_c}. \tag{3}$$

It is convenient to specify a simple multiplicative relation, with a cutoff, that is:

$$\pi(s,\theta)=\Pi(\theta s), \tag{4}$$

where $\Pi(\cdot)\equiv\min(\max(\cdot,0),1)$ is the cutoff function. The resulting probability of attaining a college education vs a high school degree is a piece-wise linear probability model. Thus, for a given level of effort, a higher-ability individual is more likely to attain a tertiary degree. In this setup, optimal study effort is unique and is given by:

$$s^*(\theta,\lambda)=\min\left(\max\left(\frac{1}{\Gamma(\theta)}\left(1-\left(\frac{\Omega(\lambda)\Gamma(\theta)}{\theta\Delta U(\mathbf{w}_j)}\right)^{\frac{1}{\sigma_s}}\right),0\right),1\right). \tag{5}$$

Although one can choose any positive effort level, it is not optimal to choose $s>\frac{1}{\Gamma(\theta)}$, because effort is costly and the college graduation probability would not change. From equation (5), given a utility gap $\Delta U(\mathbf{w}_j)$, this probability for an individual of type (θ,λ) is

$$\pi(\theta,\lambda)=\Pi\left(\frac{\theta}{\Gamma(\theta)}\left(1-\left(\frac{\Omega(\lambda)\Gamma(\theta)}{\theta\Delta U(\mathbf{w}_j)}\right)^{1/\sigma_s}\right)\right). \tag{6}$$

Let $x(k)$ denote the size of the population with educational attainment $K=k$, and denote by $x(k,j)$ the size of the population in ability group j. Then the aggregate supply vectors $\mathbf{x}^S\equiv[x^S(0),\ x^S(1)]$ and $\mathbf{x}^S_j\equiv[x^S(0,j),\ x^S(1,j)]$, with $\sum_j\mathbf{x}^S_j=\mathbf{x}^S$, are composed by:

$$x^S(1)=\int_{\boldsymbol{\Theta}\times\boldsymbol{\Lambda}}\pi(s^*(\theta,\lambda),\theta)d\mu(\theta,\lambda); \tag{7}$$

$$x^S(0)=1-x^S(1); \tag{8}$$

[12]Thus, we are implicitly assuming that family background (subsumed into λ) affects productive ability, and thereby wages, only indirectly via education choices for given cognitive ability. We present in the Online Appendix to Section 2 a more general version of the model that relaxes this assumption at the cost of using coarser measures of ability and disadvantage to have a sufficiently large sample size into ability-by-disadvantage cells, and to limit the number of additional parameters that need to be estimated. The Online Appendix to Section 5.2 brings this more general model to the data and shows that our conclusions, particularly those concerning our evaluation of the UK expansion policy, are essentially unchanged – although parameters are somewhat less precisely estimated

$$x^S(1,j) = \int_{\mathbf{\Theta}\times\mathbf{\Lambda}} \pi(s^*(\theta,\lambda),\theta)\mathbb{I}[\theta \in \mathbf{\Theta}_j]d\mu(\theta,\lambda); \tag{9}$$

$$x^S(0,j) = \int_{\mathbf{\Theta}\times\mathbf{\Lambda}} (1-\pi(s^*(\theta,\lambda),\theta))\mathbb{I}[\theta \in \mathbf{\Theta}_j]d\mu(\theta,\lambda) = f_j - x^S(1,j), \tag{10}$$

where $\mathbb{I}[\cdot]$ is the indicator function and $f_j \in (0,1)$ is the population in skill group j. The latter satisfy $\sum_j f_j = 1$.

Firms. A representative firm has a technology that maps a vector of labor allocation into output. This technology is of the CES type, and gives the following pro-capita output:

$$Q(\mathbf{x}) \equiv A^\rho \left(\sum_k \sum_j a(k,j)x(k,j)^\rho\right)^{\frac{1}{\rho}}, \tag{11}$$

where A is Total Factor Productivity (TFP). We have already taken advantage of the homogeneity of degree one to normalize $\sum_k \sum_j x^S(k,j) = 1$ (hence the term A^ρ), and also $\sum_k \sum_j a(k,j) = 1$. We assume $\rho \leq 1$, where $\rho \equiv \frac{\varsigma-1}{\varsigma}$, for ς the elasticity of substitution between *high school* and *college* labor inputs across cognitive ability groups.

The firm is competitive and solves, for any wage vector $\mathbf{w}_j$ taken as given, the problem: $\max_{\mathbf{x}\in\mathbb{R}^{2J}_+} \left(Q(\mathbf{x}) - \sum_j \mathbf{w}_j\mathbf{x}_j\right)$, where $\mathbf{x}_j = [x(0,j), x(1,j)]$ and $\mathbf{w}_j\mathbf{x}_j$ is the inner product. While unconditional and conditional aggregate labor supplies are constrained at equilibrium by equations (7)-(8) and (A-1)-(10) to add up to 1 and to f_j, respectively, the competitive firm ignores this constraint. The first-order conditions for an interior solution are:

$$w(k,j) = A^\rho a(k,j)x(k,j)^{\rho-1}Q(\mathbf{x})^{1-\rho}, \quad k \in \{0,1\},\ j = 1,\cdots,J, \tag{12}$$

and so labor demand by educational attainment and ability group, $x^D(k,j)$, satisfies

$$\frac{w(1,j)}{w(0,j)} = \frac{a(1,j)}{a(0,j)}\left(\frac{x^D(1,j)}{x^D(0,j)}\right)^{\rho-1} \quad \Leftrightarrow \quad r_j = \alpha_j(\xi_j^D)^{\rho-1}, \tag{13}$$

where $\alpha_j \equiv \frac{a(1,j)}{a(0,j)}$ denotes the technological productivity ratio within ability group j, while $r_j \equiv \frac{w(1,j)}{w(0,j)}$ and $\xi_j^D \equiv \frac{x^D(1,j)}{x^D(0,j)}$ are the college-to-high school wage and labor demand ratios within ability group j, respectively. In this model, technological change is represented by a change in A, or in any $a(k,j)$, or in ρ. Importantly, it can be skill-biased with respect to two skill dimensions, a novel aspect of our analysis that is formalized in the following definition:

Definition 1 *For a pair* (k, j)*, a change from* $a(k, j)$ *to* $a'(k, j)$ *is called* relative progress *if* $a'(k, j) \geq a(k, j)$*. A technological change is:*

- education-biased *if* $\alpha'_j = \frac{a'(1,j)}{a'(0,j)} \geq \frac{a(1,j)}{a(0,j)} = \alpha_j$ *for all* j *(i.e., it favors college graduates).*
- ability-biased *if* $\frac{a'(k,i)}{a'(k,j)} \geq \frac{a(k,i)}{a(k,j)}$ *for all* k *and* $i > j$*, (i.e., it favors higher cognitive ability).*

2.2 Equilibrium

Definition 2 *An equilibrium in an economy described by parameters* $(\Omega, \Gamma, \sigma_c, \sigma_s, A, \{\alpha_j\}, \rho)$ *is a vector* $[\{\boldsymbol{w}_j^*\}, s^*, \{\boldsymbol{x}_j^*\}]$ *such that: (i) individuals choose effort to maximize utility, and the aggregate labor supplies* $\boldsymbol{x}_j^S$ *are determined by equations (A-1)-(10); (ii) the firm chooses labor to maximize profits; (iii) the labor and goods market clear:* $\boldsymbol{x}_j^S = \boldsymbol{x}_j^D = \boldsymbol{x}_j^*$ *for each skill group* j*, and* $Q(\boldsymbol{x}^D) = \sum_k \sum_j w^*(k, j) x^S(k, j)$*.*

Using equation (12) to write wages at equilibrium as a function of the labor allocation, quantity $\Delta U(\mathbf{w}_j)$ defined in equation (3) can be written as

$$\Delta U(\mathbf{w}_j(\mathbf{x}^*)) = \frac{\left(q(\mathbf{x}^*)a(1,j)x^*(1,j)^{\rho-1}\right)^{1-\sigma_c} - \left(q(\mathbf{x}^*)a(0,j)x^*(0,j)^{\rho-1}\right)^{1-\sigma_c}}{1-\sigma_c}, \qquad (14)$$

where $q(\mathbf{x}) \equiv A^\rho Q(\mathbf{x})^{1-\rho}$. Thus, an equilibrium labor allocation vector $\mathbf{x}^*$ is fully characterized by the set of equations in the skilled population $x(1, j)$ for each skill group j,

$$x(1,j) = \int_{\boldsymbol{\Theta}\times\boldsymbol{\Lambda}} \Pi\left(\frac{\theta}{\Gamma(\theta)}\left(1 - \left(\frac{\Omega(\lambda)\Gamma(\theta)}{\theta\Delta U(\mathbf{w}_j(\mathbf{f}-\mathbf{x}(1), \mathbf{x}(1))}\right)^{\frac{1}{\sigma_s}}\right)\right)\mathbb{I}[\theta \in \boldsymbol{\Theta}_j]d\mu(\theta,\lambda), \qquad (15)$$

for $j = 1, \cdots, J$, where $\mathbf{f} \equiv [f_1, f_2, \cdots, f_J]$ is the distribution of ability groups in the population, and where we use the fact that at equilibrium, given $\mathbf{f}$, the wage vector for a given skill group j is a function of the collection $\mathbf{x}(1) \equiv [x(1, 1), x(1, 2), \cdots, x(1, J)]$ of labor allocations across *all* skill groups. Note that the equilibrium ratio $\xi_j = \frac{x(1,j)}{x(0,j)}$ coincides with the odds of college graduation for ability group j.

2.3 Higher education policy

In order to define higher education policy formally and in a tractable way, we follow the macroeconomic literature and set $\sigma_c = \sigma_s = 1$.[13] Under this assumption,

[13] For example, Prescott (2004) and Greenwood et al. (2017) set $\sigma_c = \sigma_s = 1$; Olivetti (2006), Guner et al. (2011), and Bick and Fuchs-Schündeln (2018) set $\sigma_c = 1$.

$$\Delta U(\mathbf{w}_j) = \ln w(1,j) - \ln w(0,j) \equiv \Delta \ln w_j. \tag{16}$$

Next, we specify the effort cost shifts as linear functions:[14]

$$\Omega(\lambda) = \delta + \beta\lambda \qquad \text{and} \qquad \Gamma(\theta) = \gamma + \tau\theta, \tag{17}$$

where the four parameters are controlled by the government, either actively (i.e., a purposeful stimulation of college attendance by students with certain characteristics) or passively (i.e., a mere accommodation of changes in students' demand for higher education driven by other factors). Therefore, in what follows, we refer to them as "policy parameters", and a higher education policy is a quadruple $G = (\delta, \beta, \gamma, \tau)$.[15] Combining equations (6) and (16)-(17), the equilibrium probability of college graduation for a student of type (θ, λ) at policy G is

$$\pi^*(\theta, \lambda; G) = \Pi\left(\frac{\theta}{\gamma + \tau\theta} - \frac{\beta}{\Delta \ln w_j(G)}\lambda - \frac{\delta}{\Delta \ln w_j(G)}\right). \tag{18}$$

For example, a government wishing to stimulate college attendance by disadvantaged students can offer means-tested grants, which in the model would be represented by a reduction of β. On the contrary, an active policy that increases β is the design of complex financial aid, tuition, and enrollment systems that disadvantaged households can hardly navigate. A public investment program to build new universities in response to an increased demand for college education across all families can be represented as a passive policy that allows δ to decrease. A government can also build new universities in the absence of such increased demand; this active policy would aim at increasing the probability of graduation of students living in the affected areas independently of their ability or family background, which in the model would again correspond to a reduction of δ. As for the remaining parameters, a policy that grants scholarships based on ability, or that ranks college applicants according to this same measure would correspond in the model to a reduction of τ. Beyond these examples, Section 4.1 provides historical evidence of expansive policies that took place in the UK since the 1960s, like the construction of universities and polytechnics (a decrease of δ or β) and the introduction of less stringent admission criteria (an increase of τ).

[14] Linearity allows for tractability while not limiting in any important way the types of higher education policies that we can analyze, as shown in Section 2.4.

[15] Recall that the logic of the problem requires that we consider values of G ensuring $\Omega(\lambda) \geq 0$ and $\Gamma(\theta) \geq 0$ for the values of θ and λ in the range of the economy. This restriction is imposed throughout the analysis.

A central question when expanding higher education policy is precisely the one that was addressed in the Robbins Report, namely whether an increase in college participation is possible that would select into college the "reserves of untapped ability" from the "poorer sections of the community". To answer this question, we define two paradigmatic expansion policies. Consider a status quo policy G and let G' denote a new policy.

Definition 3 *Policy G' induces an ability-enhancing expansion if and only if*

(2.1) $\mathbb{E}(\Theta|G', K = 0) \leq \mathbb{E}(\Theta|G, K = 0)$ *and*

(2.2) $\mathbb{E}(\Theta|G', K = 1) \geq \mathbb{E}(\Theta|G, K = 1)$.

Policy G' induces a disadvantage-mitigating expansion if and only if

(2.3) $\mathbb{E}(\Lambda|G', K = 0) \leq \mathbb{E}(\Lambda|G, K = 0)$ *and*

(2.4) $\mathbb{E}(\Lambda|G', K = 1) \geq \mathbb{E}(\Lambda|G, K = 1)$.

This definition corresponds to the notion of positive selection in the Roy (1951) model and its applications in economics (Willis and Rosen, 1979; Borjas, 1987) and econometrics (Heckman, 1979). It is an *outcome-based* definition in line with the Robbins Report's desiderata, and is more demanding than the typical *procedural-based* definitions of meritocratic or progressive expansions, which essentially consist of selecting the highest-ability and highest-disadvantage applicants, respectively. However, if the "reserves of untapped ability" are insufficient (in terms of quality and size), even a meritocratic expansion would "lead to a lowering of the ability" of college graduates (i.e., on average, worse engineers, medical doctors, etc.). According to Robbins, the "fears" that any expansion would fail to be ability-enhancing were "unfounded" for the UK, and one of our empirical goals is to verify this claim. A similar reasoning holds symmetrically for a progressive expansion.

To better understand the more demanding requirement of our definition, consider students who graduate from college under G' but not under G. The expansion induced by G' is ability-enhancing if these students' average ability satisfies two conditions: (i) it is not lower than the one of those who under G' still attain only a high school degree, so that the average ability of high school graduates weakly decreases; (ii) it is not lower than the one of those who attain a college degree under G, so that the average ability of college graduates weakly increases. And similarly for progressivity, replacing ability with disadvantage. While meritocracy and progressivity require only conditions (2.1) and (2.3), respectively, since study effort cost is lower for higher-ability students, it is desirable for society to expand college access in a way that also satisfies condition (2.2). In addition, an average higher ability of college graduates may lead to higher output. As for progressivity, inasmuch as society prefers more equal outcomes the expansion should also satisfy condition (2.4).

Note that, because of general equilibrium effects, the intended aim of a policy may not align with its actual outcome. Thus, the consequences of different expansion policies for mean conditional ability and disadvantage, $\mathbb{E}(\Theta|K)$ and $\mathbb{E}(\Lambda|K)$, are cumbersome to characterize analytically, and we illustrate them via numerical simulations of the model's equilibrium.

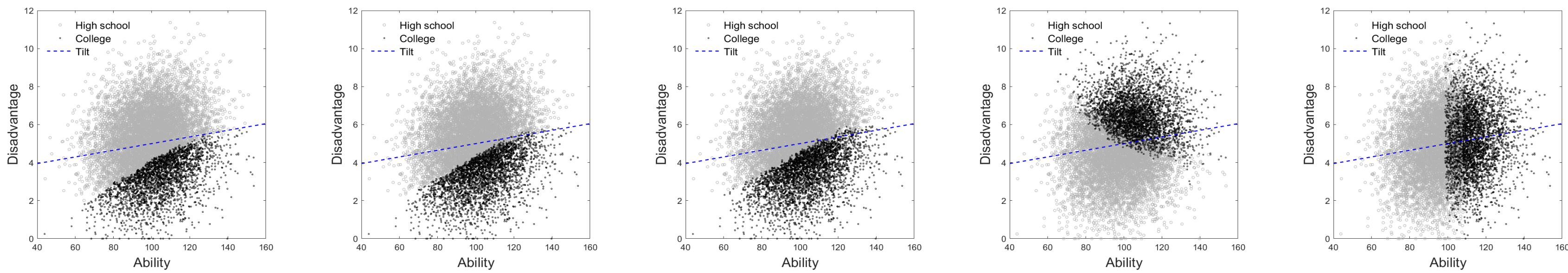
2.4 Numerical simulation

We assume that cognitive ability and SES disadvantage, $[\Theta\ \Lambda]$, are joint normal with mean $[m_\Theta\ m_\Lambda]$, standard deviation $[\sigma_\Theta\ \sigma_\Lambda]$, and correlation η. The effects of higher education policies of different types on the ability and disadvantage distribution in the college population depend on the "tilt" of the joint density $\mu(\Theta, \Lambda)$, which is given by the slope $\eta\frac{\sigma_\Lambda}{\sigma_\Theta}$ of the population regression of Λ on Θ. In our stochastic Roy model, selection into college is determined by *isoprobability curves* (loci of θ and λ combinations such that the probability of obtaining a college degree is constant) rather than by a deterministic cutoff locus.

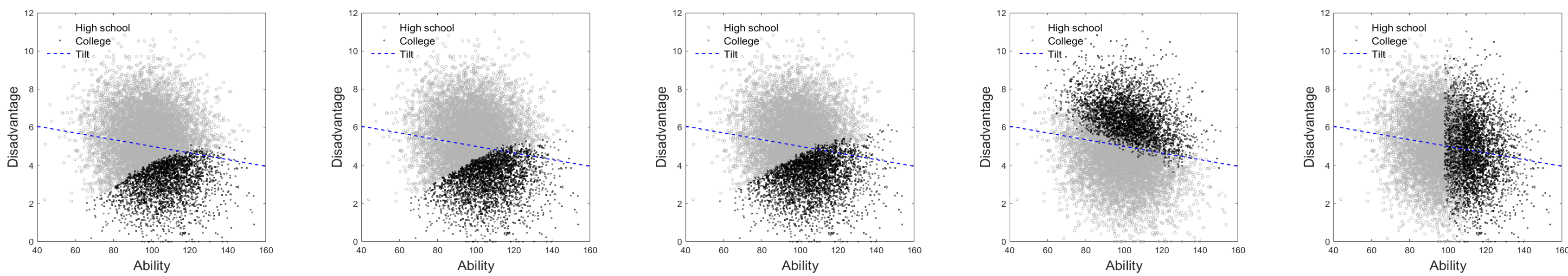
We simulate in Figure 1 a population of 10,000 individuals partitioned into $J = 100$ ability groups defined by the percentiles of the ability distribution. The top row represents Society 1, where $\eta = corr(\Theta, \Lambda) > 0$. The bottom rows illustrate Society 2, where $\eta < 0$. The plots show individuals of type (θ, λ) and their allocation to college or high school attainment – depending on whether the college graduation probability $\pi^*(\theta, \lambda; G)$ is above or below a random threshold – at equilibrium wages. The dashed line is the tilt of $\mu(\Theta, \Lambda)$. Productivity is increasing in ability, and so individuals in higher ability groups earn higher wages for given education. Parameters are chosen so that the status quo features college graduation odds similar to those observed in the UK during the 1960s, i.e., $\xi = 0.4$, and η matches the absolute value of $corr(\Theta, \Lambda)$ in the UK, i.e., $\eta = 0.2$ in Society 1 and $\eta = -0.2$ in Society 2. The latter is the empirical case during the period that we study (see Section 4.1). In the status quo of either society (column 1 of Figure 1), college graduates have, on average, higher ability and are less disadvantaged than high school graduates.

Figure 1: Status quo in Society 1 ($\eta > 0$, top) or Society 2 ($\eta < 0$, bottom), effects of education-biased technological change (EBTC), and effects of three expansion policies

Status quo ($\xi = 0.4$)	EBTC ($\xi = 0.5$)	IE policy ($\xi = 0.6$)	λ-target policy ($\xi = 0.6$)	θ-target policy ($\xi = 0.6$)



$\mathbb{E}(\Theta\|K=1) = 105.0$	$\mathbb{E}(\Theta\|K=1) = 104.5$	$\mathbb{E}(\Theta\|K=1) = 104.2$	$\mathbb{E}(\Theta\|K=1) = 107.3$	$\mathbb{E}(\Theta\|K=1) = 114.4$
$\mathbb{E}(\Theta\|K=0) = 98.0$	$\mathbb{E}(\Theta\|K=0) = 97.8$	$\mathbb{E}(\Theta\|K=0) = 97.5$	$\mathbb{E}(\Theta\|K=0) = 95.7$	$\mathbb{E}(\Theta\|K=0) = 91.4$
$\mathbb{E}(\Lambda\|K=1) = 3.14$	$\mathbb{E}(\Lambda\|K=1) = 3.28$	$\mathbb{E}(\Lambda\|K=1) = 3.40$	$\mathbb{E}(\Lambda\|K=1) = 6.59$	$\mathbb{E}(\Lambda\|K=1) = 5.26$
$\mathbb{E}(\Lambda\|K=0) = 5.75$	$\mathbb{E}(\Lambda\|K=0) = 5.86$	$\mathbb{E}(\Lambda\|K=0) = 5.96$	$\mathbb{E}(\Lambda\|K=0) = 4.04$	$\mathbb{E}(\Lambda\|K=0) = 4.84$



$\mathbb{E}(\Theta\|K=1) = 108.0$	$\mathbb{E}(\Theta\|K=1) = 107.4$	$\mathbb{E}(\Theta\|K=1) = 106.6$	$\mathbb{E}(\Theta\|K=1) = 102.8$	$\mathbb{E}(\Theta\|K=1) = 114.3$
$\mathbb{E}(\Theta\|K=0) = 96.8$	$\mathbb{E}(\Theta\|K=0) = 96.4$	$\mathbb{E}(\Theta\|K=0) = 96.0$	$\mathbb{E}(\Theta\|K=0) = 98.4$	$\mathbb{E}(\Theta\|K=0) = 91.4$
$\mathbb{E}(\Lambda\|K=1) = 3.06$	$\mathbb{E}(\Lambda\|K=1) = 3.20$	$\mathbb{E}(\Lambda\|K=1) = 3.33$	$\mathbb{E}(\Lambda\|K=1) = 6.61$	$\mathbb{E}(\Lambda\|K=1) = 4.76$
$\mathbb{E}(\Lambda\|K=0) = 5.78$	$\mathbb{E}(\Lambda\|K=0) = 5.89$	$\mathbb{E}(\Lambda\|K=0) = 6.00$	$\mathbb{E}(\Lambda\|K=0) = 4.03$	$\mathbb{E}(\Lambda\|K=0) = 5.14$

Notes: The scatter plots represent the joint distribution of ability and disadvantage in two societies. The dashed line is the tilt of the distribution, whose slope is $\eta\frac{\sigma_\Lambda}{\sigma_\Theta}$, where $\eta = \text{corr}(\Theta, \Lambda)$. This tilt is positive ($\eta = 0.2$) in Society 1 and negative ($\eta = -0.2$) in Society 2. The black dots are individuals who are selected into college at equilibrium. The tables below each plot describe the average ability and disadvantage conditional on educational attainment. In both societies, we simulate a population of 10,000 individuals with type (θ, λ) drawn from a jointly normal distribution ($m_\Theta = 100$; $\sigma_\Theta = 15$; $m_\Lambda = 5$; $\sigma_\Lambda = 1.75$). The population is partitioned into $J = 100$ ability groups defined by the percentiles of the ability distribution. In the first column (status quo), the policy parameters are such that the odds of college graduation are $\xi = 0.4$ at equilibrium, similar to the UK population during the 1960s. In this status quo, $\gamma = 3.42$ ($\gamma = 4.416$ in Society 2), $\tau = 0$, $\delta = 2$, $\beta = 2.5$. The technology parameters are $\alpha_j = j + 0.001(j-1)$, where $j = 1, \cdots, 100$ indexes ability groups, and $\rho = 0.584$ (Card and Lemieux, 2001, UK estimate). In the second column (education-biased technological change, EBTC), each α_j is increased by 0.35 in Society 1 (0.30 in Society 2) so as to induce higher equilibrium odds of college graduation, $\xi = 0.5$. For each policy experiment in the remaining columns, the parameters are set so as to increase the odds of college graduation to $\xi = 0.6$. The wage ratio adjusts to equilibrium in each ability group. Indiscriminate expansion (IE) policy: $\delta = 1.26$ ($\delta = 1.19$ in Society 2). λ-target expansion policy: $\beta = -0.3$ ($\beta = -0.31$ in Society 2), $\gamma = 17.775$ ($\gamma = 16.437$ in Society 2). θ-target expansion policy: $\tau = -0.3$, $\beta = 0$, $\gamma = 50.15$ ($\gamma = 49.692$ in Society 2), $\delta = 2.197$ ($\delta = 2.202$ in Society 2).

The crucial difference between the two societies is that, because of the different η, the pool of high-ability but disadvantaged students who are excluded from higher education is larger in Society 1 than in Society 2.[16] Therefore, Society 1 is the most favorable case for a government that wishes to expand access to college without reducing the quality of graduates or without favoring high SES students. Starting from this status quo, column 2 of Figure 1 illustrates the equilibrium effects of education-biased technological change (EBTC) that increases ξ to 0.5, absent any policy change. Such technological shock induces higher study effort across the board because it increases the college premium (by about 2% on average across ability groups). Since status quo college attainment is less frequent among low-ability, high-disadvantage students, EBTC has greater potential to expand college enrollment in this group; thus it reduces average ability and increases average disadvantage in the college population of either society.

The remaining columns describe the effects of policies that increase ξ to 0.6, like the UK ones in the 1990s. First, we consider a policy that decreases the intercept δ in effort cost shift $\Omega(\Lambda) = \delta + \beta\Lambda$. This policy may appear to favor college graduation independently of one's ability and disadvantage. Thus, in its intentions, it is an indiscriminate expansion (IE). But this conclusion ignores the effect of the policy on the college wage premium, which must decline due to the higher supply of graduates; such drop (about 0.6%-0.7% on average across ability groups) may offset the policy and ultimately *reduce* the average ability of individuals selected into college. This may be the case even in Society 1, where it should be easier to expand tertiary education without reducing the ability of college graduates: in the first row, this ability declines from 104.5 in column 2 to 104.2 in column 3. However, the incidence of low-SES college graduates increases relative to the status quo with EBTC (from 3.28 to 3.40). *A fortiori*, also in Society 2 an intended IE policy reduces the average ability of graduates with respect to the same status quo, (from 107.4 to 106.6).

Consider next, in column 4, a policy that intends to expand tertiary education opportunities for the disadvantaged (labeled as the λ-target policy) by decreasing the slope β of effort cost shift $\Omega(\Lambda) = \delta+\beta\Lambda$ while increasing the intercept γ of effort cost shift $\Gamma(\Theta) = \gamma+\tau\Theta$. In our simulation, such a policy takes a strong form because β turns from positive to negative, so that a low SES (i.e., a large realization of Λ) becomes an advantage in college access. The

[16] For example, in the north-eastern region of the scatter plot defined by the top terciles of the ability and disadvantage marginal distributions, the population fractions that are excluded from higher education in the status quo are 13% in Society 1 and 9% in Society 2.

policy effect on graduates' average ability is positive and large in Society 1 (from 104.5 to 107.3) but *negative* in Society 2 (from 107.4 to 102.8). Thus, this policy is ability-enhancing only in Society 1, where the correlation η between ability and disadvantage is negative.

This dilemma is resolved by the policy illustrated in the last column, which intends to expand tertiary education opportunities for the most able (labeled as the θ-target policy) by adjusting effort cost shifts $\Omega(\Lambda)$ and $\Gamma(\Theta)$ to $\tau < 0$ (a cost shift in favor of high-ability students) and $\beta = 0$ (so that one's SES becomes irrelevant), while γ and δ both increase to obtain the desired odds of college graduation ξ. The result is that the probability of obtaining a college degree increases much more at higher levels of ability. College graduates' average ability increases to almost 114.5 in both societies while their average disadvantage increases to 5.26 in the first one and to 4.76 in the second one. Thus, in either society, this policy is not only ability-enhancing but also disadvantage-mitigating because it draws sufficiently high-talent and high-disadvantage students into college.

3 Data

3.1 Data sources

Our main data source is Understanding Society (University of Essex, Institute for Social and Economic Research, 2023), henceforth USoc, a representative longitudinal survey of UK households. Wave 3 (2011-2013, $N = 49,692$) contains cognitive ability information. We restrict this sample to individuals: (i) with non-zero cross-sectional weights (38,223); (ii) white born in the UK (31,132); (iii) with non-missing education information (31,072); (iv) born between 1940 and 1984 (23,288). The left panel of Table 1 reports descriptive statistics for this sample. The central panel of the table reports the same statistics for the sub-sample of 22,175 subjects who satisfy these selection criteria and also have cognitive ability information. A comparison suggests that the ability measure is missing quasi-at-random for 1,113 subjects. In what follows, we use this representative sample of the UK population when we need statistics for this population. Since our analysis compares college and high school graduates, we further restrict the sample to the 17,890 subjects who, in addition to satisfying the selection criteria (i)-(iv), have the required high school credentials to apply for college admission. Table 1's right panel reports the descriptive statistics for this sample.

Table 1: The UK Understanding Society sample

	White UK 1940-1984					White UK 1940-1984 with non-missing intelligence score					White UK 1940-1984 with non-missing intelligence score and at least a high school degree				
	N	mean	sd	min	max	*N*	mean	sd	min	max	*N*	mean	sd	min	max
Individual characteristics															
Age	23,288	49.40	12.32	24	72	22,175	49.25	12.29	24	72	17,890	47.53	12.06	24	72
Female	23,288	0.52	0.50	0	1	22,175	0.52	0.50	0	1	17,890	0.51	0.50	0	1
Any tertiary degree	23,288	0.24	0.43	0	1	22,175	0.25	0.43	0	1	17,890	0.31	0.46	0	1
Age left school	22,896	16.26	1.11	7	21	21,794	16.29	1.12	7	21	17,536	16.52	1.09	7	21
Born in England	22,990	0.81	0.39	0	1	21,892	0.81	0.39	0	1	17,654	0.81	0.39	0	1
Health status	23,287	2.57	1.11	1	5	22,174	2.54	1.10	1	5	17,889	2.44	1.06	1	5
Number of marriages	20,475	1.01	0.61	0	4	19,469	1.01	0.61	0	4	15,472	0.98	0.59	0	4
N. of children < 18	23,288	0.36	0.81	0	8	22,175	0.36	0.81	0	8	17,890	0.40	0.84	0	8
Belongs to a religion	22,051	0.48	0.50	0	1	20,986	0.48	0.50	0	1	16,870	0.47	0.50	0	1
Real monthly income	23,288	2.00	1.71	-8	26	22,175	2.03	1.73	-8	26	17,890	2.22	1.82	-8	26
Family characteristics at age 14-16															
Father's yrs school	19,207	11.93	2.81	0	18	18,353	11.98	2.82	0	18	15,035	12.33	2.84	0	18
Mother's yrs school	19,846	11.47	2.44	0	18	18,950	11.51	2.44	0	18	15,458	11.84	2.47	0	18
Father employed	22,905	0.88	0.32	0	1	21,818	0.89	0.32	0	1	17,614	0.90	0.30	0	1
Mother employed	23,020	0.62	0.48	0	1	21,930	0.63	0.48	0	1	17,692	0.65	0.48	0	1

Notes: We start from the third wave (2011-2013) of the UK Understanding Society survey (USoc). This wave contains information on respondents' cognitive ability and consists of 49,692 observations. We first apply four selection criteria: we keep observations with non-zero cross-sectional response weights (38,223); we restrict to white respondents born in the UK (31,132); we keep observations with non-missing education information (31,072); and, we restrict the sample to individuals who were born between 1940 and 1984 (23,288). The left panel of the table reports descriptive statistics for this sample. The central panel reports the same descriptive statistics for the sub-sample consisting of 22,175 individuals with non-missing ability test scores. The similarity of the statistics in the two panels suggests that information on ability is missing quasi-at random. We will use this representative sample of the UK population whenever we need statistics for such a population. However, to compare college with high school graduates and estimate the parameters of the model, we further restrict the sample to the 17,890 subjects who, in addition to satisfying the above selection criteria, have the high school credentials that are necessary to apply for college admission. The right panel of Table 1 reports the descriptive statistics for this selected sample, which differ as expected from those of the entire population. Real monthly income is expressed in thousands of real GBP (2015 prices).

We leverage Usoc's panel structure to predict the Discounted Present Value of lifetime earnings (DPV), using the methodology proposed by Lagakos et al. (2018). Their identifying assumption is that wage profiles are flat in the final years of a worker's career, a "common prediction of theories of life cycle wage growth" (p. 799). We take these final years to be between age 51 and 60. Thanks to this assumption, observed wage growth for workers in a given cohort between 51 and 60 only reflects aggregate time effects. Once these are obtained, it is easy to estimate cohort and age effects for different cohorts.[17] We then use the estimated coefficients (including individual fixed effects and disregarding year effects) and the USoc sample with cognitive information (central panel of Table 1) to predict the wage of each subject at all ages between 18 and 65 and the resulting DPV, using an annual discount rate of 3% and full-time work. This DPV is taken as a proxy for model variable $w(k, j)$.

We also use two ancillary data sources described in the Online Appendix to Section 3.1: the 1970 British Cohort Study (BCS70) to strengthen our analysis of ability; and the University Statistical Record (USR) to describe how the UK expansion was enacted.

3.2 College graduation rate and college cohorts

Before 1992, students had two higher education options in the UK: traditional universities and polytechnics. As illustrated in Pratt (1997), Willet (2017) and Jandarova and Reuter (2021), these two types of institutions differed in many ways, e.g., funding, target populations, organization, subjects, and admission criteria.[18]

Following the Robbins Report, which had recommended the unification of UK higher education in consideration of the similarities between universities and polytechnics, the *Further and Higher Education Act* of 1992 allowed polytechnics to obtain university status. In line with the literature on the evolution of the wage gap between college and high school graduates in the UK (for example: Machin and McNally, 2007; Walker and Zhu, 2008; Blundell et al., 2022), in the present paper a "college graduate" is defined as a person who obtained

[17] We apply this method by pooling USoc waves 1-11, where individuals in different cohorts are observed during the period 2009-2019 (we discard waves 12 and 13 to avoid years affected by Covid 19). Specifically, we regress hourly wages on individual fixed effects and: year dummies, dummies for age between 18-50 years and between 61-65 (one for each age), so that the effects of dummies for ages between 51-60 are constrained to zero; interactions between the gender dummy and age dummies; interactions between education dummies and age; and interactions between age, gender, and education dummies. Note that individual fixed effect absorb all invariant individual characteristics that affect selection into college, including ability and disadvantage.

[18] Professionally-oriented public colleges (e.g., teacher training and nursing colleges) offered another option. This group was relatively small and so we consider it as part of "polytechnics".

a tertiary education degree of any kind. Combining together graduates from traditional universities and polytechnics is not a limitation given that we are studying the expansion of the UK higher education system and that ending the "binary divide" between universities and polytechnics was in fact part of this policy. The comparison group of "high school graduates" comprises the remaining individuals in our final USoc sample who had the high school credentials required to apply for college admission but who did not obtain a tertiary education degree.

To facilitate the interpretation of our results in relation to historical information on policy and technology trends, we aggregate individuals into "college cohorts" using 15-year windows to maintain a sufficiently large within-cohort sample size. These cohorts are groups of individuals in actual (for college graduates) or potential (for high school graduates) college attendance age. For such an age, we use the year of birth plus 20 as a label. The three 15-year windows are: 1960-1974 for individuals born between 1940 and 1954 (4,716 individuals in the final sample with at least a high school degree), 1975-1989 for those born between 1955 and 1969 (6,922 individuals from this same group), and 1990-2004 for subjects born between 1970 and 1984 (6,252 individuals from that same group).[19] We assume absence of first-order substitutability between college graduates across these cohorts (and similarly for high-school graduates), which is reasonable in the light of evidence that the time of entry in the labor market has long-term consequences on wages and employment along the life cycle.[20] At the same time, recall that our model implies, within cohorts, perfect substitutability between equally educated workers of comparable ability, and some substitutability across education and ability groups according to the value of parameter ρ.

3.3 Cognitive ability

In Wave 3, USoc respondents older than 16 were eligible for a cognitive ability test composed of the following six sub-tests: immediate word recall, delayed word recall, subtraction, number series, verbal ability, and numeric ability.[21] We observe, for each individual, the fraction of correct answers and whether help was received during the test – either specific help in answering a question or generic material aid during the test – resulting into 14 cognitive ability

[19] Labeling these groups as "college cohorts" avoids possible confusion with birth cohorts.

[20] See, among others, Kahn (2010), Oreopoulos et al. (2012), Schwandt and von Wachter (2019), von Wachter (2020) and Jandarova (2022).

[21] See McFall (2013) for a detailed description of these cognitive tests.

variables: six sub-test scores and eight dummies for whether help was received. Following the psychometric literature (Fawns-Ritchie and Deary, 2020), we use Principal Component Analysis (PCA), in the representative sample of the UK population with cognitive ability information described in the central panel of Table 1, to extract from these variables a measure of general cognitive ability.[22] The First Principal Component (FPC), which is the empirical counterpart of the ability construct Θ in the model, has an eigenvalue of 2.55 and explains 18.2% of the data variability. The corresponding eigenvector features positive values for the fractions of correct answers, negative values for 6 of the 8 help dummies, and positive but near-zero values for the remaining two help dummies (see Table A-1 in the Online Appendix to Section 3.3 for additional details). We, therefore, conclude that this FPC summarizes the cognitive ability of USoc respondents in a satisfactory way.

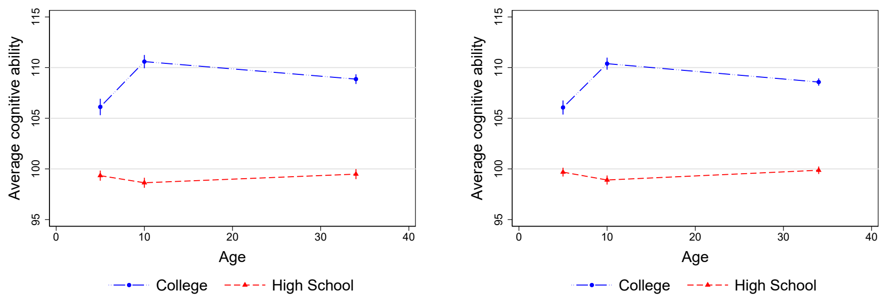
This ability score is taken to be a cardinal measure of the underlying cognitive ability construct, so any monotonic linear transformation of this measure is admissible and we must pick one. It is convenient to choose one such that variable Θ has mean 100 and standard deviation 15, so as to facilitate the comparison with the widely used IQ measure. This choice implies that we can identify γ and τ as policy parameters determining the cost of effort *relative to that scale*, as is evident in equation (18). Since we are interested in policy *changes*, the particular scale that we choose is irrelevant. The distribution of the resulting ability measure in USoc is illustrated in Figure A-1 of the Online Appendix to Section 3.3.

The USoc cognitive tests are administered *after* potential or actual college attendance, and so may be endogenous to tertiary education. Using evidence based on the BCS70 we show, in line with the literature, that this is not a concern. This data set is unique in that it contains scores in verbal and mathematical tests of the same individuals at ages 5, 10, and 34. Using PCA again to construct a normalized cognitive ability measure (see Tables A-2 to A-4 in the Online Appendix to Section 3.3 for additional details), we report in Figure 2 the average ability at these three ages of high school graduates who eventually obtained or did not obtain a college degree. The panels correspond to different ways of dealing with missing test scores at a particular age, as explained in the figure's note. In all panels, while the ability gap between the two groups increases until age 10, between age 10 and 34 the ability of college graduates does not increase, neither relative to that

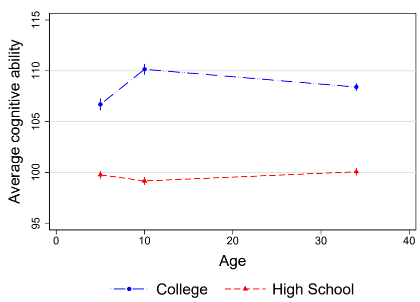
[22]Figure A-3 in the Online Appendix to Section 3.3 shows that an alternative cognitive ability measure produced by factor analysis has a correlation of 0.99 with the PCA measure.

Figure 2: The effect of higher education on cognitive ability

(A) No imputation, same subjects (B) No imputation, unequal sample sizes



(C) With imputation, equal sample sizes



(D) Sample sizes

Age	Panel A	Panel B	Panel C
5	4,386	5,748	7,369
10	4,386	5,308	7,369
34	4,386	7,270	7,369

The figure reports cognitive ability measures constructed from the 1970 British Cohort Study (BCS70, see the Online Appendix to Section 3.1). We use the second (1975), third (1980), and seventh (2004) waves, where participants were 5, 10, and 34 years old, respectively, and completed cognitive tests assessing verbal and mathematical skills. Principal Components Analysis (PCA) is used to create a cognitive ability measure at each age, standardized to have mean 100 and standard deviation 15. Panel (A) is based on the 4,386 participants who completed the cognitive tests at *all* of the three ages. Panel (B) is based on those who completed the cognitive tests at *any* of the three ages; therefore, in this panel, sample sizes differ across ages, as described in Panel (D) of the table. In Panel (C), the missing test scores at each age are imputed (see details in the Online Appendix to Section 3.3) and therefore sample size is identical for all ages. In panels (A), (B), and (C) and for each age, the circles mark the cognitive ability of college graduates, while the triangles mark the corresponding ability of subjects with high school certification that would be sufficient for enrollment in tertiary education. The table in panel (D) presents the sample sizes used to create panels (A), (B), and (C), which exclude high school dropouts.

of high school graduates nor in absolute terms.[23] These findings are consistent with the literature: while Brinch and Galloway (2012) provide evidence that pre-tertiary education may affect cognitive ability, Kremen et al. (2019) and Arum and Roksa (2011) show that this is not the case for college.[24] Ollikainen et al. (2023), using a regression discontinuity design, provide credible evidence that even the type of schooling (college track vs vocational track) has no relevant effect on cognitive skills. Moreover, consistent with evidence that general cognitive ability is unlikely to be malleable beyond infancy (Heckman and Mosso (2014); Protzko (2015), Ritchie et al. (2015)) this paper shows that any effect of schooling on specific cognitive skills is not mediated by general cognitive ability, which instead seems

[23] We test formally this claim in the Online Appendix to Section 3.3.

[24] Conditioning on cognitive ability measured in adolescence, Clouston et al. (2012) find that higher education is correlated with ability measured during midlife, but this evidence cannot be regarded as causal.

to be largely unaffected by education.[25]

We emphasize that our ability measure can therefore be interpreted as a *pre-college* measure of cognitive skills and is the outcome of both nature and nurture. As such, it also reflects socioeconomic disadvantage, as discussed in more detail below. Yet, the evidence from the BCS70 reported in Figure 2 and the economic literature on cognitive development (as summarized, for example, by Heckman and Mosso, 2014) suggest that our ability measure captures skills that are fixed relatively early in life.

There are also two well-known and widely discussed issues in measuring cognitive ability, which has been shown to vary over time for a given age and over age for a given cohort. The first variation is known as the "Flynn effect" because Flynn (1987) measured an apparent improvement in IQ scores in 14 nations during the 20th century (an effect that reversed itself in recent years). The second has been documented by Salthouse (2012, 2019), who observed that different types of cognitive skills evolve in different ways during the life cycle. By analogy, we label this as the "Salthouse effect". Since we want a measure of ability that does not reflect cohort age, the Salthouse effect must be removed by normalizing the USoc ability measure within birth years.[26] This comes at the cost of removing also the Flynn effect, which is less of a concern because this finding is more controversial. For example, using high-quality data from Norway that enable a within-family analysis of IQ, Bratsberg and Rogeberg (2018) argue that the Flynn effect and its reversal in recent years are explained by environmental factors. Consistent with these findings, a 13-year-long assessment by Dworak et al. (2023) in a large US sample indicates that the Flynn effect and its reversal do not generalize across age or education groups, casting doubts on whether these phenomena are genuine.

3.4 Socioeconomic disadvantage

A higher value of disadvantage Λ in our model means lower SES, a non-cognitive factor that reduces the probability of college graduation. For given cognitive ability, students from low-income, low-education, or single-parent families are less likely to enroll and complete college (Bailey and Dynarski, 2011; Hoxby and Avery, 2013). We measure such disadvantage in the

[25] This conclusion is consistent with Ritchie and Tucker-Drob (2018), since "the vast majority of the studies in [their] meta-analysis considered specific tests and not a latent g factor" (p. 1367).

[26] The comparison between the left and the right panel of Figure A-2 in the Online Appendix to Section 3.3 illustrates the effect of this normalization on the USoc measure.

USoc sample described in the central panel of Table 1, by aggregating via PCA eight relevant variables: parents' years of schooling and six dummies (referring retrospectively to when the respondent was 14) for whether either parent was employed, whether the respondent was living with only one parent, and whether either parent was deceased. The FPC explains 22% of the variability in these eight variables. The corresponding eigenvector contains negative values for whether either parent was absent or dead and positive values for the other variables (see Table A-6 in the Online Appendix to Section 3.4). We therefore conclude that this FPC summarizes a socioeconomic *advantage*. Since we want a measure of disadvantage, we simply invert the sign of this FPC. We also shift the support of the FPC distribution so that disadvantage has a minimum of zero.[27]

Like for cognitive ability, we take the disadvantage measure produced by PCA as a cardinal measure of the underlying concept and any monotonic linear transformation is admissible. Since there is no scale in the psychometric tradition for variable Λ, we use the translation of the PCA measure described above. As is again evident in equation (18), the scale of parameter β adapts to this particular scale, which is immaterial because we want to estimate *changes* in β across cohorts. Figure A-4 in the Online Appendix to Section 3.4 shows the distribution of our disadvantage measure. Given the discrete nature of the socioeconomic variables at our disposal and the lack of information on family income when the respondent was 14, the resulting FPC offers a stratified measure of disadvantage. This measure is relatively coarse but arguably captures long-term determinants of a young person's cost of study effort.

Finally, note that while we normalize ability within birth year, we do not do the same for disadvantage. The reason is that socioeconomic standards (e.g., parental education) have certainly improved in the UK during the period that we study, which is part of the reason why the demand for college education has increased. We therefore do not want to remove by construction the effects of this force from our empirical analysis.

[27]Figure A-5 in the Online Appendix to Section 3.4 shows that an alternative disadvantage measure produced by factor analysis has a correlation of 0.97 with measures obtained with PCA. A recent literature summarized by Corazzini et al. (2021) argues that non-cognitive personality traits are also a relevant determinant of tertiary education attainment, while Deming (2017) highlights their growing importance in the labor market. We show in the Online Appendix to Section 5.2 that augmenting our ability or disadvantage measures with the "Big Five" personality traits does not change our conclusions.

4 Key empirical facts associated with the expansion

The size of the UK tertiary education expansion was documented in Section 1. Using the full USoc sample that is representative of the UK population (central panel of Table 1), college graduates increased from 18.2% to 32.6% between the 1960-1974 and the 1990-2004 college cohorts. In our analysis sample, restricted to subjects with at least a high school degree (right panel of Table 1), these statistics are 27.9% and 35.2%, respectively. In this section, we document key facts associated with this expansion.

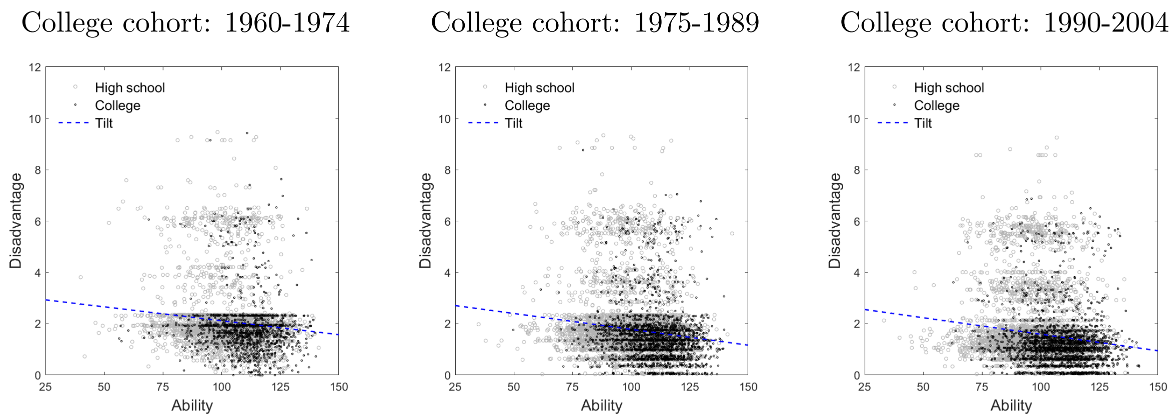
Figure 3 shows the empirical joint distribution of ability and disadvantage by cohort and who, among students with a high school degree, attains (black dots) or not (gray circles) a college degree. Like for the simulated joint distributions in Figure 1, the dashed line is the tilt, whose slope is $\eta\frac{\sigma_\Lambda}{\sigma_\Theta}$, where $\eta = \text{corr}(\Theta, \Lambda)$. This correlation is negative in all college cohorts, and is statistically and quantitatively stable over time: approximately -0.14. Thus, during the period that we study, the UK resembles Society 2 of Section 2.4. In light of the economics of skill formation (e.g., Cunha et al., 2006, Heckman and Mosso, 2014), this is not surprising. High-ability parents have better education and higher income; thus, they transmit to their children higher ability via genes and better nurturing, as well as higher SES. As such, η does *not* reflect a policy-invariant relation between pre-college cognitive ability and SES disadvantage. It merely tells us that, during 1960-2004, the UK was in a particular social equilibrium that matches Society 2 of Figure 1.[28]

In light of our model, we are particularly interested in three sets of facts that are key to inferring the features and impact of the UK tertiary education expansion: (1) the evolution of mean cognitive ability and mean SES disadvantage of students who graduate or not from college; (2) the way the expansion was enacted; and (3) the evolution of the college premium. When appropriate, we will illustrate these facts using groups defined by the terciles of the ability and disadvantage distributions in the overall UK population.[29] Note that this way we depart from the numerical simulation of Section 2.4, where we consider $J = 100$ ability

[28] For example, early childhood interventions can drive this correlation toward zero. However, late interventions are less, if at all, effective than early ones (see Cunha and Heckman, 2007), while being costly for the reasons that we also illustrate in this study. Thus, it would be a mistake to use university expansions to correct for the lack of adequate early education policies.

[29] That is, we compute terciles using the USoc sample with cognitive information (central panel of Table 1), which also includes subjects without a high school degree. Then we assign individuals to ability group 1 if their ability is below the first tercile, to group 2 if their ability is between the first and second terciles, and to group 3 if their ability is above the second tercile. And similarly for disadvantage groups. Thus, in our final sample of students with at least a high school degree these groups are not equally sized in a given cohort.

Figure 3: Empirical joint distribution of ability and disadvantage by education



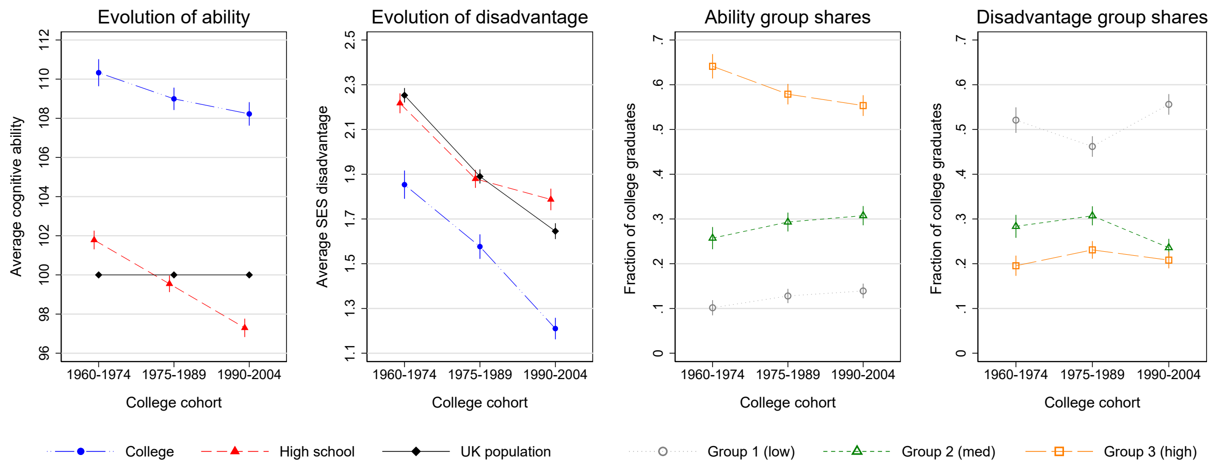
Notes: The figure shows the empirical counterpart of the scatter plots in Figure 1, by college cohort. Each point is an individual in our sample. Black dots indicate college graduates while gray circles denote high school graduates. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1). The dashed line is the tilt of the distribution, whose slope is $\eta \frac{\sigma_\Lambda}{\sigma_\Theta}$, where $\eta = \text{corr}(\Theta, \Lambda)$.

groups, by setting $J = 3$ for either ability or disadvantage. This limit is dictated by the number of parameters that we can estimate in our structural analysis of Section 5.

4.1 Students' cognitive ability and SES disadvantage

The first panel of Figure 4, shows that the average ability of college graduates declines by two points (13% of a standard deviation), from 110.3 in the 1960-1974 cohort to 108.2 in the 1990-2004 cohort. Note that 100 is the value to which the average ability of the entire UK population is normalized in all cohorts (see Section 3.3). Thus, during this period, college graduates became more similar to the population in terms of average ability. During the same period, also the average ability of high school graduates without tertiary education declined by almost five points (about 30% of a standard deviation), from 101.8 to 97.3. It follows that the UK tertiary education expansion was not ability-enhancing in the sense of Definition 3: although the expansion favored students whose ability was higher than the ability of those previously excluded from college, these students were not sufficiently able relative to the existing pool of college graduates. This possibility was conjectured by Walker and Zhu (2008) and Blundell et al. (2022), but they could not verify it because their data lacks cognitive information. Consistent with the decline of graduates' average ability, the third panel of Figure 4 shows that the share of highest-ability students declines across cohorts, while the shares of less able students increase.

Figure 4: Evolution of ability and disadvantage by education, and group shares



Notes: The first panel shows the evolution of average cognitive ability among students with a college degree and among students with a high school degree. This panel also plots mean ability in the UK population that is normalized by construction to 100 in all cohorts (see Section 3.3). Similarly, the second panel shows the evolution of average SES disadvantage among students with a college degree and among students with a high school degree. This panel also plots mean SES disadvantage in the UK population (continuous line, 22,175 subjects described in Table 1's central panel), a statistic that, contrary to ability, is not normalized to be constant at a specific value (see Section 3.4). The third and fourth panels show how the shares of three ability groups and three disadvantage groups (defined by the terciles of the respective distributions in the UK population) in the college population evolved across cohorts. Vertical lines are 95% C.I. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1), unless otherwise specified.

As for SES disadvantage, the second panel of Figure 4 reports both the evolution of the mean by educational attainment and the evolution of the same variable in the entire UK population. In contrast to cognitive ability, for which the UK population average is normalized to 100, average disadvantage is not constrained to be constant in this panel (see Section 3.4). In fact, it exhibits a declining trend that reflects the improving socioeconomic status of the UK population during the period that we study.[30] In both 1960-1974 and 1975-1989 college cohorts, high school graduates are very similar to the overall UK population in terms of disadvantage, which declines by about 15%. College graduates are evidently more advantaged than high school graduates in both college cohorts, but they experience a similar decline of disadvantage. Thus, initially, the expansion of tertiary education affected the average background of college and high school graduates only marginally, relative to the UK population. The outcome of the sorting process departs more substantially from mere population changes for the 1990-2004 cohort: compared to the 1975-1989 cohort, average disadvantage declined by 12.9% in the UK population, by 23.2% among college graduates,

[30]The standard deviation does not change and is about 1.3 in all cohorts. A declining mean and a constant standard deviation imply an increasing coefficient of variation, i.e., widening relative inequality in terms of disadvantage, as shown by, e.g., Machin (1996) and Office for National Statistics (2021).

and only by 4.9% among high school graduates. These numbers suggest that the more recent stage of the expansion brought into college high school graduates who were relatively advantaged in the group of those previously excluded, and also more advantaged than the average high school graduate who was previously admitted to college. Thus, the UK tertiary education expansion has not been disadvantage-mitigating either, again in the sense of Definition 3. The disadvantage group shares reported in the fourth panel of the figure support this claim from a different angle: the incidence of the most disadvantaged students in the college population is the lowest in all cohorts, and barely changes between the 1960s and the 1990s. At the other end of the spectrum, the share of the most advantaged students is always the highest and actually increases over this period.

To bring this descriptive analysis one step further, we next combine the two dimensions represented in the third and fourth panels of Figure 4. The resulting joint distribution of ability and disadvantage groups in the college population across the three cohorts is reported in Table 2, where we denote ability and disadvantage groups by θ_i and λ_j, for $i, j = \{1, 2, 3\}$.

Table 2: Joint distribution of ability and disadvantage groups in the college population

	Cohort 1960-1974			Cohort 1975-1989			Cohort 1990-2004		
	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3
λ_1	0.050	0.118	0.348	0.043	0.121	0.301	0.068	0.162	0.325
λ_2	0.032	0.082	0.168	0.039	0.097	0.169	0.026	0.081	0.126
λ_3	0.016	0.057	0.128	0.040	0.074	0.116	0.040	0.068	0.104

Notes: for each college cohort, a cell in this table reports the share of students in the college population who belong to the three ability and disadvantaged groups defined by the terciles of the corresponding distributions in the UK population: $\Pr(\theta_i = i, \lambda_j = j | K = 1, \text{cohort})$ for groups $i, j = \{1, 2, 3\}$. For each cohort, the entries add up to 1. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1).

The share of college graduates with the highest ability (θ_3) and the highest disadvantage (λ_3) declined from 12.8% to 10.4% between the first and the last cohort. These individuals correspond to the black dots in the northeastern portion of Figure 3. On the contrary, the analogous share with the lowest ability (θ_1) and the lowest disadvantage (λ_1) in the college population, corresponding to the black dots in the southwestern portion of Figure 3, increased from 5% to almost 7%. These simple statistics indicate that students who were both relatively more able and relatively more disadvantaged actually ended up with reduced opportunities to attain a college degree with respect to their lower-ability and lower-disadvantage counterparts.

This fact can also be appreciated by estimating the following linear model for the probability of college graduation as a function of ability, disadvantage, and their interaction:

$$k = b_0 + b_1\theta + b_2\lambda + b_3\theta\lambda + \epsilon. \tag{19}$$

As reported in Table 3, the interaction coefficient b_3 is negative in both the first and the last cohort, and decreases substantially over time. This means that during the UK expansion, cognitive ability boosted individual graduation probability more for advantaged students than for disadvantaged ones, and this penalization for the high-ability but disadvantaged students was more severe in 1990–2004 than in 1960–1974.

Table 3: Probability of college graduation as a linear function of ability and disadvantage

College cohort	b_0	b_1	b_2	b_3	N
1960-1974	−0.801	1.111	0.071	−0.105	4,716
	(0.093)	(0.092)	(0.035)	(0.035)	
1990-2004	−0.977	1.406	0.115	−0.175	6,252
	(0.062)	(0.062)	(0.030)	(0.031)	

Notes: The table reports OLS estimates of the coefficients of equation (19). Robust standard errors in parentheses. Sample: USoc, 10,968 white respondents who belong to cohorts 1960-1974 and 1990-2004, with non-missing education and ability information and with at least a high school degree that permits college enrollment.

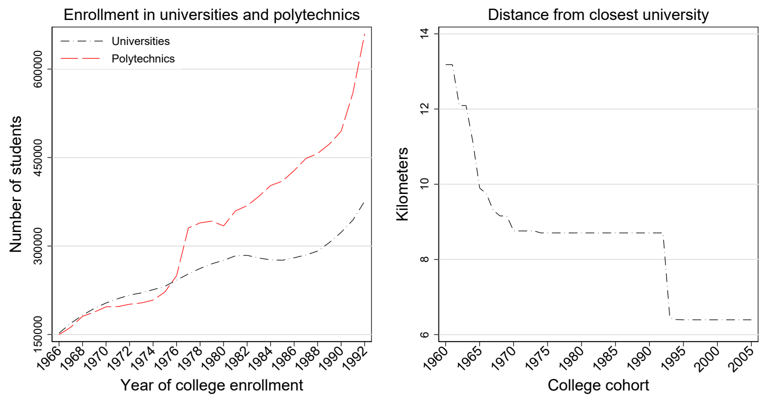
4.2 How was the expansion enacted?

The UK higher education expansion was enacted with several policy interventions: by ending the divide between traditional universities and polytechnics, by increasing the number of academic institutions without much consideration of quality, and by reducing ability requirements at entry. A first piece of evidence supporting these claims is provided in the left panel of Figure 5, which uses Pratt's (1997) data. Between 1966 and 1992, the stock of students enrolled in universities more than doubled (from 152,227 to 376,074). This increase is smaller than for Polytechnics (from 149,720 to 659,790) but is still substantial. While the growth of Polytechnics is a consequence of an explicit expansion policy,[31] the higher number of students in traditional universities is the result of more subtle policy changes. For example, the right panel of Figure 5 shows that graduation costs were reduced by increasing

[31] According to Pratt (1997), about thirty Polytechnics were created between 1966 and 1973. In 1988, the Education Reform Act reduced funding per student granted to this type of institution, inducing them to expand students' enrolment in order to keep constant the total amount of available resources.

the number of academic institutions and by bringing them closer to potential students.[32] In the USoc sample, the *average* distance from the closest university dropped by about 6km between 1960 and 2005. In this period most of these institutions had no fees, and so mobility costs were an important component of college costs.[33] However, such an indiscriminate expansion would not necessarily favor disadvantaged students. Without policies targeting these students specifically, only affluent families are well positioned to take advantage of increased college slots (e.g., Bailey and Dynarski, 2011; Hoxby and Avery, 2013).

Figure 5: Higher education enrollment and distance to the closest college



Notes: The left panel uses data in Table 3.3 of Pratt (1997) to produce a modified version of Figure 3.2 in this same book. The modification is that we aggregate "Polytechnichs" and "Other colleges". For the right panel, we use a list of all Royal Charters granted in the UK ever since the 13th century (the list can be found at https://privycouncil.independent.gov.uk/royal-charters/list-of-charters-granted/), and we selected entries corresponding to universities and colleges. Each entry has a legal address, which we use as a location point to count the number of universities active over time in each area. For each year we then count, how many active universities were located in each county. If this step returns a positive number, we set the distance to zero; if it returns a zero, we compute the distance (in km) to the nearest university from the county boundary. For each year, the figure plots the average distance over counties from the closest university.

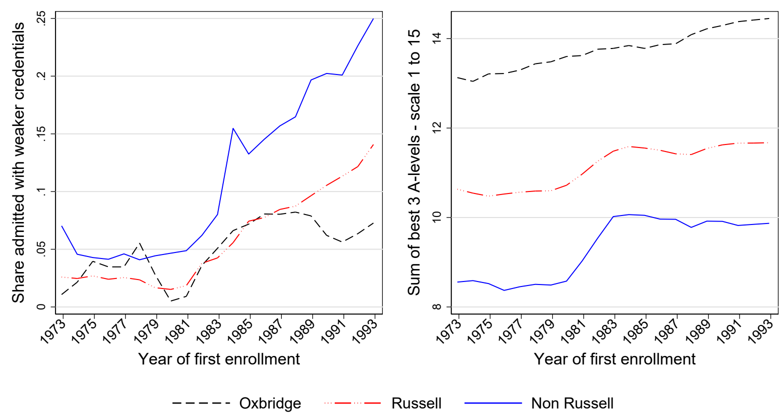
A second piece of evidence comes from USR, which provides information on the evolution of enrollment and entry criteria. The left panel of Figure 6 shows the fraction of students admitted with weaker high school credentials to three groups of UK universities: Oxbridge, the Russell group, and the remaining, less prestigious institutions. In all groups, this fraction increased between 1973 and 1993. The increase is particularly evident in the residual group, but it is visible also for the Russell group and even for Oxbridge. The USR documentation

[32]According to Blundell et al. (2022), more than twenty new universities were created in the 1960s. See also the evidence in Blackburn and Jarman (1993).

[33]As summarized by Willet (2017), the 1962 Education Act waved tuition and introduced maintenance grants for all UK students. In the 1980s, such grants were tied to family income in order to provide stronger support for more disadvantaged students. Only in 1998, with the Teaching and Higher Education Act, fees of 1,000 GBP per year were introduced. Only after the period that we study, with the 2004 Higher Education Act, fees raised to 3,000 GBP per year and then again to 9,000 GBP following the 2010 Independent Review of Higher Education Funding and Student Finance (the "Browne Review").

explains that this is an indicator of less demanding admission criteria because it refers to two main categories of students: those who had less than 3 A-Level scores (i.e., the regular minimum requirement for admission) and those admitted on the basis of HNC/HND/ONC/OND qualifications, which have a more vocational or technical nature.

Figure 6: Criteria for admission to a university



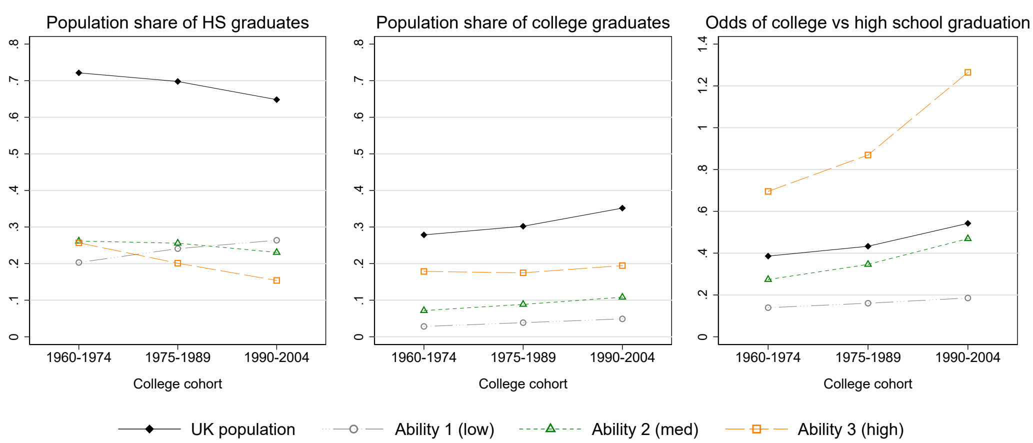
Notes: The left panel shows the fraction of students admitted with weaker high school credentials to three groups of UK universities: Oxbridge, the Russell group, and remaining institutions. The right panel reports the average sum of the best 3 A-Level scores for students admitted to the three groups of universities during the period covered by USR data. Source: USR.

The right panel reports instead the average sum of the best 3 A-Level scores for students admitted to the three groups during the period covered by USR data. As expected, in all years students admitted at Oxbridge have higher best A-level scores than students admitted at the Russell group, who in turn dominate students in the remaining institutions. What is more striking is that in all the three groups this indicator increases significantly over the period of observation. This increase has two possible interpretations. First, there was grade inflation in high schools that eventually facilitated college admission. A second alternative interpretation is that universities became more selective in admitting students but at the same time high school students improved over time their performance in A-Level exams. We are unable to establish which scenario is the correct one. However, if universities had become more selective, the average cognitive ability of their graduates would have increased. As shown in Section 4.1, this is not the case.

Another important policy change took place in 1988, when the GCSEs replaced the CSEs and O-Levels as the exams that UK students take at age 16. According to Blundell et al. (2022), this "reform led to an increase in educational attainment at the secondary level and hence an increase in the proportion of the young with sufficient academic credentials for potential admission to universities". This reform may have contributed as well to the decline of high school graduates' cognitive ability documented in the top left panel of Figure 4.

Figure 7 shows the consequences of these changes on the shares of high school (left panel) and college (central panel) graduates by ability, model variable $x(k, j)$, and the associated odds of college graduation (right panel), model variable ξ_j. The decline of the share of high school graduates in the population is driven by a drop in the intermediate and top-ability groups. Hence the corresponding increase in college graduates' shares for these groups. As for the bottom ability group, the shares of both high school and college graduates increase. The reason is that fewer of these students drop out of high school and more of them acquire the credentials to apply for college. As a result, in all ability groups the odds of college graduation increase. However, even if this increase is particularly pronounced in the top ability group, it is not sufficient to make the expansion ability-enhancing according to Definition 3.

Figure 7: Population shares of HS and college graduates by ability and associated odds

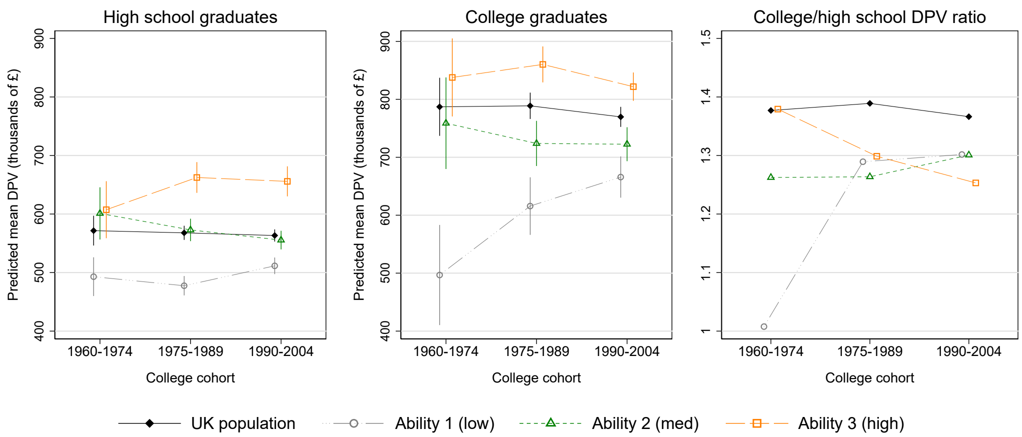


Notes: The left and central panels display the evolution across college cohorts of the population shares, $x(k, j)$, of college and high school (HS) graduates, respectively, in the UK population and in the three ability groups defined by the terciles of the cognitive ability distribution. The right panel describes the implied evolution of the odds of college graduation by ability group, ξ_j. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1).

4.3 DPV of lifetime earnings and the college premium

Figure 8 illustrates instead the consequences of the UK expansion on the DPV of lifetime earnings, model variable $w(k, j)$, and the associated college premium, model variable r_j. In the left panel, the flat profile of the pooled DPV of high school graduates hides a polarization of the underlying levels by ability group. It decreases for the intermediate group, while it increases for both the bottom and top groups between the first and the last cohorts. In the central panel, the DPV of college graduates exhibits a slight decline between these two

Figure 8: DPV of lifetime earnings by ability and associated college premium



Notes: The left and central panels display the evolution across college cohorts of the DPV of lifetime earnings of college and high school graduates, respectively, in the UK population and in the three ability groups defined by the terciles of the cognitive ability distribution. The DPVs are estimated as explained in Section 3.1. The right panel describes the implied evolution of the respective college premia. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that permits college enrollment (see the right panel of Table 1).

cohorts, which results from a remarkable increase for the least able that is more than offset by a decline for the two more able groups. As a result, the college premium in terms of DPV is essentially flat (right panel). This finding is in line with the evidence for the UK provided by Blundell et al. (2022) based on real hourly wages in the UK Labor Force Survey.

Both our and their findings are in sharp contrast with the evidence for the US, where the college premium has increased since the 1980s (Katz and Murphy, 1992; Fortin, 2006; Goldin and Katz, 2008; Autor et al., 2020). Such stark difference requires an explanation: while in the US technology won the "race" with education (thus causing the college premium to rise), this has not happened in the UK. Moreover, the novel evidence of Figure 8 suggests that the outcome of the "race" in the UK is an increase for the least able students and a decrease for their more able peers. A discussion of how our analysis contributes to understanding these patterns requires the estimation of the model's parameters, to which we turn next.

5 Structural analysis of the UK university expansion

5.1 Identification and estimation

The analysis with only three ability groups that we adopted in Section 4 will facilitate our structural estimation. In this case, there are five technology parameters in the model: three technological productivity ratios, α_j, one for each ability group $j \in \{1, 2, 3\}$; parameter ρ that is equal to one minus the inverse of the elasticity of substitution between high school and college graduates in production; and TFP, A. Equation (16) implies that TFP does not affect the consumption utility gap $\Delta U(\mathbf{w}_j)$, in equation (3), and so it can be ignored. Still, it is clear that without further assumptions, the α_j's and ρ are not separately identified in our model. We, therefore, calibrate ρ at the value of 0.584 that is implied by the preferred estimate for the UK of Card and Lemieux (2001, Tab. IV, column 6), which they obtained using data from the UK General Household Survey for the period 1974-1996 and by adopting the same definition of college and high school graduates that we use.

On the contrary, the policy parameters that govern the shifts in effort cost $\Omega(\Lambda) = \delta + \beta\Lambda$ and $\Gamma(\Theta) = \gamma + \tau\Theta$ are identified – a fact that follows immediately from equation (18) – and so we are left with parameters $(\gamma, \tau, \delta, \beta, \alpha_1, \alpha_2, \alpha_3)$ to estimate. Given $\rho = 0.584$, the empirical counterpart of the joint distribution $\mu(\Theta, \Lambda)$, and a target set of empirical moments, we estimate these parameters by Minimum Distance (MD) in each college cohort. Specifically, for each point in the discretized parameter space (the "grid"), we solve numerically for the equilibrium supply of college graduates in each ability group, $x(1, j)$, by finding for each group the unique fixed point of the following equation, which combines equations (14) and (15),

$$x(1,j) = \sum_{\lambda,\theta} \omega(\theta,\lambda)\Pi\left(\frac{\theta}{\gamma+\tau\theta} - \frac{\delta+\beta\lambda}{\ln\alpha + (\rho-1)(\ln x(1,j) - \ln(f_j - x(1,j)))}\right)\mathbb{I}[\theta \in \mathbf{\Theta}_j]\mu(\theta,\lambda), \tag{20}$$

where $\omega(\theta, \lambda)$ denote USoc cross-sectional response weights that add up 1. Given the equilibrium odds of college graduation in each ability group, $\xi_j = \frac{x(1,j)}{f_j - x(1,j)}$, we obtain the equilibrium college-to-high school wage ratios, r_j.

The equilibrium graduation probabilities are then used to classify each individual in the sample as a college graduate if that individual's probability is above an individual-specific random threshold. Finally, we pick the parameters that minimize the distance between ten informative theoretical moments and their empirical analogs: the odds of college graduation and the college premium in each ability group (six moments) and the average ability and disadvantage of college and high school graduates (four moments). These ten moments are the most informative for estimating our four policy parameters and the three technological productivity ratios because in our model it is precisely the change in higher education policy or technology that alters the labor market equilibrium and the sorting process into college.

Additional untargeted moments are set aside to check how well we match facts not used in our MD estimation. In consideration of the importance of ability and disadvantage in our analysis, we select for these purposes the 25^{th} and 75^{th} percentiles of the conditional (to educational attainment K) distributions of Θ and Λ, i.e., eight moments.

All of the empirical moments are estimated using the final USoc sample (right panel of Table 1), applying the appropriate sampling weights. Denoting by

$$\widehat{T} = \left[\widehat{\xi}_1 \ \widehat{\xi}_2 \ \widehat{\xi}_3 \ \widehat{r}_1 \ \widehat{r}_2 \ \widehat{r}_3 \ \widehat{\mathbb{E}}(\Theta|K=1) \ \widehat{\mathbb{E}}(\Theta|K=0) \ \widehat{\mathbb{E}}(\Lambda|K=1) \ \widehat{\mathbb{E}}(\Lambda|K=0)\right]$$

the target vector of empirical quantities, and by

$$T(\gamma,\tau,\delta,\beta,\alpha_1,\alpha_2,\alpha_3;\rho) = [\xi_1 \ \xi_2 \ \xi_3 \ r_1 \ r_2 \ r_3 \ \mathbb{E}(\Theta|K=1) \ \mathbb{E}(\Theta|K=0) \ \mathbb{E}(\Lambda|K=1) \ \mathbb{E}(\Lambda|K=0)]$$

its theoretical counterpart at equilibrium, the criterion function is

$$\mathcal{C}(\gamma,\tau,\delta,\beta,\alpha_1,\alpha_2,\alpha_3;\rho) = (T(\gamma,\tau,\delta,\beta,\alpha_1,\alpha_2,\alpha_3;\rho)-\widehat{T})\Upsilon W \Upsilon (T(\gamma,\tau,\delta,\beta,\alpha_1,\alpha_2,\alpha_3;\rho)-\widehat{T})',$$

where Υ is a diagonal matrix whose elements are the inverse of the elements of $\widehat{T}$, and W is a weighting matrix. Thus, the criterion function is a weighted sum of *percentage* squared deviations of the theoretical moments from the empirical ones. We set $W = I$, and we find the min of $\mathcal{C}(\cdot;\rho)$ over the grid. To produce standard errors, we repeat this MD estimation 1,000 times in samples obtained from random draws with replacement. The bootstrap s.e. are given by the standard deviation of each parameter's estimate across the replications.

A crucial question about our identification is whether the criterion function $\mathcal{C}(\cdot;\rho)$ attains a *global* minimum at the estimated parameters. The choice of starting values is important in this respect, and we address it by obtaining an initial set of estimates for policy parameters $G = (\gamma,\tau,\delta,\beta)$ by Nonlinear Least Squares (NLS) from the supply-side equation (18), after replacing $\Delta \ln w_j(G)$ with its empirical analog, $\ln \widehat{w}(1,j) - \ln \widehat{w}(0,j)$. As for the technological productivity ratios $(\alpha_1,\alpha_2,\alpha_3)$, we obtain starting values by fitting the demand-side equation (13) in isolation for each college cohort and ability group. Specifically, for each j and cohort, we solve equation (13) for α_j and then we plug into the resulting equation the empirical values of the odds of college graduation, ξ_j, and the wage ratio, r_j, to obtain a numerical value for α_j. In other words, the starting values of the policy and technology parameters are partial equilibrium estimates. More computational details and checks against the presence of local minima are provided in the Online Appendix to Section 5.1.

5.2 Parameter estimates

MD estimates of the technology and policy parameters are reported in panels [A] and [B], respectively, of Table 4. The Online Appendix to Section 5.2 shows that the ten targeted moments and the eight untargeted moments are matched quite satisfactorily. We next explain what these estimates tell us about the nature of technological change and higher education policy in the UK during the period that we study.

Table 4: Minimum-distance estimates of technology and policy parameters

	[A] Technology				[B] Policy		
	1960-1974	1975-1989	1990-2004		1960-1974	1975-1989	1990-2004
α_1	0.490	0.580	0.550	γ	9.890	3.940	4.080
	(0.028)	(0.018)	(0.005)		(0.040)	(0.022)	(0.013)
α_2	0.730	0.800	0.950	τ	-6.030	-1.820	-2.050
	(0.008)	(0.008)	(0.006)		(0.040)	(0.020)	(0.017)
α_3	1.180	1.160	1.340	δ	0.001	0.037	0.019
	(0.044)	(0.011)	(0.016)		(0.001)	(0.002)	(0.005)
ρ	0.584	0.584	0.584	β	0.003	0.006	0.005
N	4,716	6,922	6,252		(0.001)	(0.001)	(0.001)

Notes: The table reports Minimum-Distance (MD) estimates of technology and policy parameters, obtained under the assumption that $\rho = 0.584$, i.e., the value estimated for the UK by Card and Lemieux (2001). The MD criterion function is given by equation (21), and the weighting matrix is the identity matrix. Standard errors in parentheses are the standard deviation of point estimates over 1,000 bootstrap samples. The Online Appendix to Section 5.1 provides more computational details. College cohorts in the columns are defined by the period of actual or potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree (right panel of Table 1).

5.2.1 Understanding technological change

The estimated productivity ratios, α_j, increase across ability groups within each college cohort. Thus, the technology at any given point in time is such that college graduates' productivity advantage over high school graduates of comparable ability increases with ability.[34] These ratios also increase between the first and last college cohorts for each ability group j. With reference to Definition 1, this means *education-biased* technological change (EBTC), as established by a large literature (Katz and Murphy, 1992; Katz and Autor, 1999).[35]

[34]Some estimates of α_j are below 1. Recall that $\alpha_j = \frac{a(1,j)}{a(0,j)}$. The fact that $\alpha_j < 1$ does *not* imply that high school graduates are more productive than college graduates because marginal productivity depends on $a(k, j)$ but also, inversely, on the fraction of the workforce with educational attainment k.

[35]As explained in Section 3.1, the Lagakos et al. (2018) procedure to infer DPVs of lifetime earnings for individuals born in different cohorts using wage data observed in 2009-2019, allows us to estimate equilibria

Exploiting the two-dimensional notion of skills in our data, we can also assess the presence of *ability-biased* technological change (ABTC). To this end, we use equations (12)-(13) to back out the $a(k, j)$ parameters that appear in the production function (11). These parameters are reported in Table 5. We emphasize that they are normalized to add up to 1, and so any change across cohorts must be interpreted as *relative* technological progress or regress. The table shows that for any combination of ability levels $i > j$ and any education level k, the ratio $\frac{a(k,i)}{a(k,j)}$ declines over time (except for $\frac{a(0,3)}{a(0,2)}$, which is flat). According to Definition 1, technological change was *biased against ability.* As an illustration of this fact, the top-left panel of Figure 9 shows the declining evolution across cohorts of the ratio $\frac{a(k,3)}{a(k,1)}$ for college and high school graduates.

Table 5: Implied $a(k, j)$ parameters of the production function

College cohort	$a(0,1)$	$a(0,2)$	$a(0,3)$	$a(1,1)$	$a(1,2)$	$a(1,3)$	$\sum a(k,j)$
1960-74	0.148	0.201	0.201	0.071	0.145	0.234	1.000
1975-89	0.152	0.187	0.196	0.088	0.150	0.227	1.000
1990-04	0.167	0.172	0.171	0.092	0.163	0.235	1.000

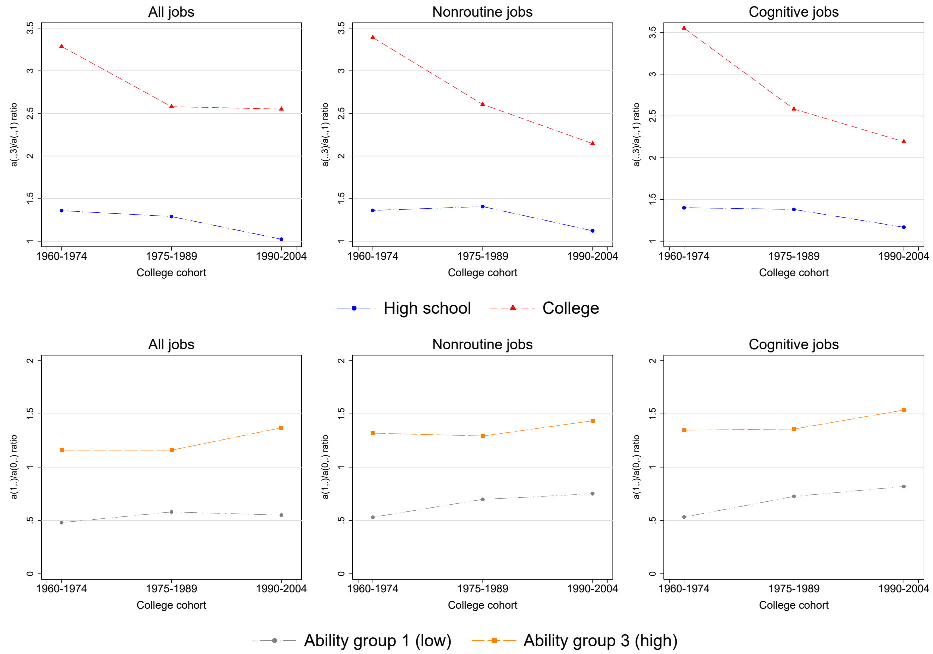
Notes: The table reports numerical values for the $a(k, j)$ parameters that feature in the production function (11), as implied by our estimates and the data. Given the $\alpha_j = a(1, j)/a(0, j)$ ratios reported in Table 4, we calculate $a(k, i)/a(k, j)$ for given k and $i > j$ by taking ratios of equation (12) in analogy to equation (13). Normalizing $a(0, 1) = 1$ within each cohort allows us to back out all the remaining $a(k, j)$. Dividing these values by the within-cohort total produced the numerical values in this table.

This finding sheds new light on the hypothesis, pioneered by Levy and Murnane (1996) and Autor et al. (2003), that computers and offshoring have progressively replaced routine workers in advanced economies, while at the same time increasing workers' productivity in non-routine and cognitive jobs. A large literature has studied the implications of this hypothesis for job and wage polarization, providing evidence for several advanced economies, including the UK.[36] An implication of this hypothesis is that college graduates' productivity increased relative to high school graduates, i.e., EBTC, as we also find in our data. However, it does not follow that the same should be true, within education groups, for workers with higher vs lower cognitive ability, i.e., that ABTC should operate in favor of more able workers. As pointed out by Levy and Murnane (1996), we can think of computers as increasing

for different cohorts that reflect the respective differences in technology, labor supply and labor demand by education and cognitive ability. Equivalent, alternative procedures are used in the literature; see, for example, the supplementary material in Blundell et al. (2022).

[36] See, among others, Goos and Manning (2007), Goos et al. (2009), Acemoglu and Autor (2011), Autor and Dorn (2013), Michaels et al. (2014).

Figure 9: Education and cognitive ability biases of technological change



Notes: Top row: evolution across college cohorts of the $a(k,3)/a(k,1)$ ratio for college ($k=1$) and high school ($k=0$) graduates, for all jobs (left), nonroutine jobs (center), and cognitive jobs (right). Bottom row: analogous evolution of the $a(1,j)/a(0,j)$ ratio (i.e., parameter α_j) for the bottom ($j=1$) and top ($j=3$) ability groups defined by the terciles of the cognitive ability distribution. Sample: 12,879 subjects with cognitive ability information, at least a high school degree, and an occupation.

individual ability, complementing cognitive skills in tasks that require them. Thus, among equally educated workers performing a certain task, low-ability individuals may benefit from computerization more than individuals whose ability is already high and who therefore are less in need of aid from a computer. For example, consider two architects, one in ability group 1 and the other in group 3. The advent of Computer-Aided Design may have helped the less-able architect more than the higher-ability one. Several recent papers provide evidence that is consistent with ABTC operating in favor of less able workers. Tafti (2024) demonstrates that the adoption of robots in surgery improves the relative performance of low-ability surgeons by reducing the performance advantage of high-ability ones. Brynjolfsson et al. (2025) find that providing customer-support agents at a business-process software firm with a generative AI-based conversational assistants increases substantially the productivity of low-experience, low-skilled agents, with minimal (and possibly negative) impact on the rest. Similarly, Hoffmann et al. (2024) show that access to a generative AI code completion tool for software developers increases the productivity of less able individuals.

The declining evolution of the $\frac{a(k,i)}{a(k,j)}$ ratios, for $i > j$, shown in Table 5 is consistent with the phenomena that these papers detect, and more generally with decreasing returns to cognitive ability (Castex and Dechter, 2014; Deming, 2017) and to unobservable skills within education levels (Lochner et al., 2025). That is, *within each education group*, the combined effect of computers replacing workers in routine tasks and enhancing workers' productivity in nonroutine and cognitive tasks is an increase in the productivity of low-ability individuals relative to their high-ability counterparts.

To corroborate this interpretation, we map workers' occupations in our Usoc sample (3-digit ISCO-88) to the five task types proposed by Acemoglu and Autor (2011): non-routine cognitive analytical, nonroutine cognitive interpersonal, routine cognitive, routine manual, and nonroutine manual physical.[37] The top central and right panels of Figure 9 illustrate, as an example, the declining evolution of the $\frac{a(k,3)}{a(k,1)}$ ratio in non-routine and cognitive jobs. It is particularly in these two (overlapping) types of jobs that we expect computers to have favored low-ability workers for given education, so the dynamics in the figure support our claim. Additional supporting evidence is provided in the bottom panels of Figure 9, which – again as an illustration – show that the $\frac{a(1,j)}{a(0,j)}$ ratios, i.e., parameters α_j, increased for both nonroutine and cognitive jobs in both ability groups 1 and 3. This fact confirms that technological change was education-biased for both nonroutine and cognitive jobs.[38]

This interpretation of our estimates of the technology parameters enables explaining the wage levels and college premia dynamics in Figure 8. In light of equation (12), the DPV level of the lowest-ability college graduates increased because, in the UK, these are the workers who most benefited from the introduction of computers in the workplace, while their share did not increase enough to offset this productivity advantage (see Figure 7). On the contrary, the DPV level of the highest-ability college graduates remains essentially flat as a result of the almost flat evolutions of their productivity and their share in the population. A similar interpretation applies to the evolution of wage levels and college premia for the remaining types of workers. The flat evolution of the overall college premium in the UK population is then a consequence of these opposing trends.

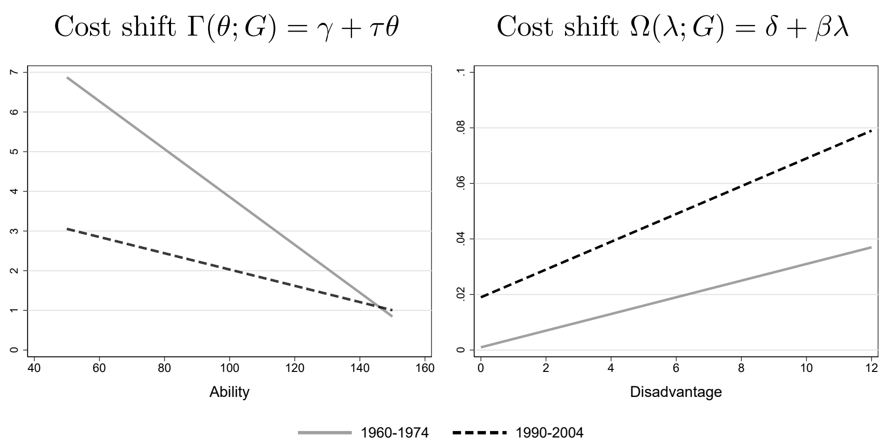
[37] We implement this classification using the code provided by Lewandoski (2020). Note that we only observe these occupations at the time when wave 3 Usoc subjects were interviewed in 2011-2013. Therefore, we have no occupation information for 5,011 workers (about 3,000 of whom belong to the 1960-1974 cohort, and the rest are evenly divided between the other two cohorts) who were not employed (plausibly retired) at the time of the interview. Moreover, for those who were instead employed, if they had changed occupation at some point due to technological change, we only observe their occupation after displacement.

[38] This evidence is only suggestive, given the limitations of the USoc occupation information. Moreover, while our model extends Katz and Murphy (1992) by allowing for different cognitive abilities within education levels, we cannot extend it also to allow for different tasks, as for example in Acemoglu and Autor (2011).

5.2.2 Understanding the UK higher education policy

The descriptive evidence in Section 4 has shown that even if the share of high-ability college graduates increased (Figure 7), the UK tertiary education expansion resulted in a declining average cognitive ability of college graduates (Figure 4) and therefore was not ability-enhancing in the sense of Definition 3. Our structural estimates confirm this conclusion. As shown in Table 4, we estimate a substantial drop in γ and an increase in τ, δ, and β between the first and the last cohorts. The economic meaning of these changes is illustrated in Figure 10. In the left panel, the continuous line indicates that effort cost shift $\Gamma(\theta) = \gamma + \tau\theta$ was declining with ability in the 1960-1974 cohort, as conjectured in Section 2. The tertiary education policies implemented in the UK produced the change described by the dashed line: a large reduction in the cost of attending college that was more pronounced for lower-ability students. This drop was sufficiently pronounced to induce a falling average ability of graduates, confirming that the college expansion was not ability-enhancing. In the right panel, the continuous line indicates that the effort cost shift $\Omega(\lambda) = \delta + \beta\lambda$ was increasing with disadvantage in the 1960-1974 cohort, as we also conjectured in Section 2. The dashed line shows that the expansion policy actually resulted in an *increased* cost of attending college that was more pronounced for more disadvantaged students. Such an increase induced a falling average disadvantage of graduates relative to the UK population. Thus, the college expansion was not disadvantage-mitigating either.

Figure 10: Estimated study effort cost shifts



Notes: Study effort cost shifts $\Gamma(\cdot)$ and $\Omega(\cdot)$ implied by the structural estimates of Table 4, for the 1960-1974 (continuous line) and 1990-2004 (dashed line) college cohorts, as a function of cognitive ability (left) or socioeconomic disadvantage (right).

5.2.3 Understanding the UK-US difference in college premium dynamics

Our analysis also contributes to explaining why the evolution of the college premium in the UK was so different from the US. An explanation was first proposed by Blundell et al. (2022) under the assumption that the UK is a technological follower of the US in terms of technology adoption and firm organization. The flat college premium in the UK would then reflect a switch from a centralized to a decentralized organizational mode by firms, which neutralized the effects of technological changes and of increasing college attendance.

In the light of our results, we suggest that the different types of higher education expansion that took place in the two countries may have also played a role. Technological change during the period that we study appears to be *education-biased* but also *biased against cognitive ability* within education groups, due to computerization (Figure 9). If this was the case also in the US, then, given that the vast majority of American colleges have become *less* selective since at least the 1970s,[39] the college wage premium may have increased in the US and not in the UK because the tertiary education expansion that took place in the US possibly selected into college an even larger fraction of low-ability students whose productivity was actually favored by workplace computerization.

5.3 Counterfactual expansion policies

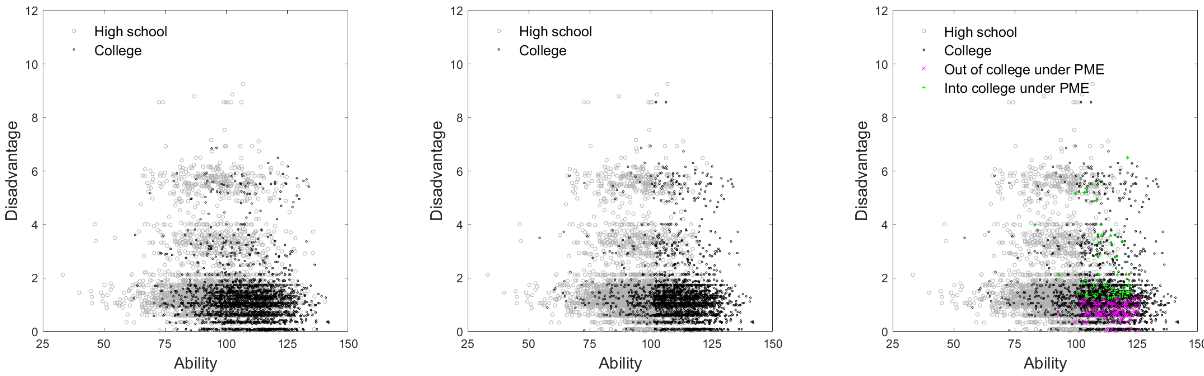
Alternative policies could have reached the "reserves of untapped ability [...] in the poorer sections of the community" (Robbins, 1963, p. 53) that were a central concern in the Robbins Report but that were disregarded by the UK expansion. We next describe two of these counterfactual policies under the constraint that they deliver an increase of the aggregate odds of college graduation similar to the observed one (i.e., from $\xi = 0.39$ to $\xi = 0.54$).

The first is a policy that expands tertiary education opportunities for the most able, similar to the last column (θ-target) of simulated expansions in Figure 1. This goal is pursued by setting τ to a more negative value than in the observed policy, but only for students above average ability ($\tau = 0$ for the others), while reducing β to zero. For example, financial aid for university studies only to students who perform sufficiently well in a cognitive ability test. The second counterfactual policy expands tertiary education opportunities for

[39] This fact is documented by Hoxby (2009), who shows that in the US "[o]nly the top 10 percent of colleges are substantially more selective now than they were in 1962. Moreover, at least 50 percent of colleges are substantially less selective now than they were in 1962." (p. 95).

Figure 11: Effects of actual and counterfactual expansion policies on the 1990-2004 cohort

Actual policy	Counterfactual θ-target policy	Counterfactual (θ, λ)-target policy



Actual policy	Counterfactual θ-target policy	Counterfactual (θ, λ)-target policy
$\xi = 0.54$; (pooled)	$\xi = 0.54$ (pooled)	$\xi = 0.53$ (pooled)
$\xi_1 = 0.19;\ r_1 = 1.30$	$\xi_1 = 0.11;\ r_1 = 1.40$	$\xi_1 = 0.11;\ r_1 = 1.40$
$\xi_2 = 0.47;\ r_2 = 1.30$	$\xi_2 = 0.33;\ r_2 = 1.50$	$\xi_2 = 0.33;\ r_2 = 1.51$
$\xi_3 = 1.26;\ r_3 = 1.25$	$\xi_3 = 1.96;\ r_3 = 1.01$	$\xi_3 = 1.86;\ r_3 = 1.03$
$\mathbb{E}(\Theta \mid K=1) = 108.2$	$\mathbb{E}(\Theta \mid K=1) = 111.7$	$\mathbb{E}(\Theta \mid K=1) = 111.6$
$\mathbb{E}(\Theta \mid K=0) = 97.3$	$\mathbb{E}(\Theta \mid K=0) = 95.5$	$\mathbb{E}(\Theta \mid K=0) = 95.7$
$\mathbb{E}(\Lambda \mid K=1) = 1.21$	$\mathbb{E}(\Lambda \mid K=1) = 1.39$	$\mathbb{E}(\Lambda \mid K=1) = 1.46$
$\mathbb{E}(\Lambda \mid K=0) = 1.79$	$\mathbb{E}(\Lambda \mid K=0) = 1.69$	$\mathbb{E}(\Lambda \mid K=0) = 1.65$

Notes: Effects of two counterfactual expansion policies that would have achieved odds of college graduation similar to the observed $\xi = 0.54$, reporting for comparison the effect of the actual policy (left panel). In the central panel, a θ-target expansion – relative to the 1960-1974 status quo – decreases τ to -6.3 for students above average ability (while setting $\tau = 0$ for those below average), and sets $\beta = 0$, $\gamma = 9.1$, and $\delta = 0$. In the right panel, a (θ, λ)-target expansion that, relative to the θ-target policy, decreases γ only for students whose disadvantage is above the median and, for these students, in proportion to their ability: $\gamma = 9.3 - \mathbb{I}[\lambda > \text{med}(\Lambda)](0.235 + 0.135\theta)$. The density in the north-eastern portion of the cloud is not sufficient for the (θ, λ)-target policy to expand the odds of college graduation beyond 0.53 without violating the constraint of positive cost shifts $\Omega(\cdot) \geq 0$ and $\Gamma(\cdot) \geq 0$. Students marked with a "×" graduate from college under the counterfactual θ-target policy but not under the counterfactual (θ, λ)-target policy; for students marked with a "+" the opposite happens. In the scatter plots, each point is an individual in the 1990-2004 college cohort in our sample. Sample: USoc, 17,890 white respondents born in the UK in 1940-1984 with non-missing education and ability information and with at least a high school degree that qualifies for college enrollment (see the right panel of Table 1).

the most able, with additional aid to the disadvantaged, and so we label it as the (θ, λ)-target policy. To reach this goal, the θ-target expansion is modified by reducing parameter γ for students whose disadvantage is above the median and, for these students, in proportion to their ability. An example is the PACE policy implemented in Chile (Tincani et al., 2022).[40]

Figure 11 illustrates the outcomes of the actual policy (left panel) and of these two counterfactual expansion modes (central and right panels) on the selection of high school graduates into tertiary education. Under the θ-target policy (central panel), isoprobability

[40] As these authors report, "PACE targets disadvantaged schools, and it offers students who graduate in the top 15 percent of their high school guaranteed admissions to colleges participating in the centralized admission system, eliminating the entrance exam score requirement. [...] Students in PACE high schools are considerably disadvantaged: they have 10th-grade standardized test scores that are 1.5 standard deviations below those of regular college entrants and 0.49 standard deviations below the OECD average."

curves of college completion (see Section 2.4) become nearly vertical. In this counterfactual scenario, relative to the observed status quo in the left panel of Figure 3, we observe a decline in the fraction of college graduates in the south-western portion of the population cloud (low ability, low disadvantage) and an increase in the north-eastern portion (high ability, high disadvantage). As a consequence, relative to the status quo of the 1960s, the average cognitive ability of college graduates *increases* by 1.4 percentage points (111.7 vs 110.3) instead of the actual two-point drop. Their counterfactual average disadvantage is still lower than in the 1960s (1.39 vs 1.85) – which partly reflects the declining average disadvantage of the UK population during this period depicted in the bottom left panel of Figure 4 – but higher than the actual average disadvantage of graduates in the 1990s (1.21). Therefore, this counterfactual θ-target policy is both ability-enhancing and more disadvantage-mitigating than the policy that was actually implemented.

The alternative (θ, λ)-target policy (right panel) reinforces the aid to high-ability, high-disadvantage students. In this panel, we mark with a "×" students who, in the counterfactual, graduate from college under the θ-target policy but not under the (θ, λ)-target policy, and with a "+" students for whom the opposite holds. Clearly, the (θ, λ)-target policy raises graduation barriers for lower-ability, more advantaged students while correspondingly lowering such barriers for higher-ability, more disadvantaged ones, on average (recall that our experiments keep the total number of graduates constant to the observed level). Relative to the counterfactual θ-target policy, the (θ, λ)-target policy cannot expand the odds of college graduation beyond 0.53 without violating the constraint of positive cost shifts $\Omega(\cdot) \geq 0$ and $\Gamma(\cdot) \geq 0$, because the density in the north-eastearn portion of the (θ, λ) distribution is insufficient. However, it selects into college students whose average ability is about the same as under the θ-target policy, and thus more than three points above the actual average for the 1990-2004 cohort (111.6 vs 108.2).At the same time, the (θ, λ)-target policy favors more disadvantaged students substantially more than the θ-target policy (1.46 vs 1.39). This is the outcome that was envisioned by Robbins (1963), and is the opposite of what actually happened in the UK. To further appreciate this fact, in comparison to the θ-target benchmark, note that in the right panel of Figure 11, the 121 "×" students who do not graduate from college under the (θ, λ)-target policy relative to the θ-target policy have an average ability of 114.5 and an average disadvantage of 0.79, while the 89 "+" students who replace them in college have an average ability of 112.5 and an average disadvantage of 2.25.

6 Conclusions

We have introduced into the analysis of higher education policy the systematic consideration of individuals' cognitive ability, in addition to more conventional measures of socioeconomic disadvantage. In our framework, technological change is possible in two different dimensions: the change can be education-biased but also ability-biased. This perspective provides new insights into the interpretation of the flat evolution of the college wage premium in the UK between the 1960s and the 2000s. The notion of ability as a scarce resource to be allocated across different education levels is an important consideration in the rich analysis of the Robbins (1963) Report. Such consideration and analysis are instead conspicuously absent in the current debate, and notably in the document that states the European Union's goal for 2030 of increasing to 45% the share of 25-34 year-old EU residents with tertiary educational attainment (EU Council, 2021). This target is set without mention of the cognitive skills that students selected into college ought to have. While such neglect may be inspired by equity aspirations, the key lesson conveyed by our analysis is that it is possible to reconcile graduates' quality with better college opportunities for the disadvantaged in the design of tertiary education expansions that are both ability-enhancing and disadvantage-mitigating.

References

Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," Elsevier, vol. 4 of *Handbook of Labor Economics*, 1043–1171.

Arum, R. and J. Roksa (2011): "Limited learning on college campuses," *Society*, 48, 203–207.

Autor, D., C. Goldin, and L. F. Katz (2020): "Extending the Race between Education and Technology," *AEA Papers and Proceedings*, 110, 347–51.

Autor, D., F. Levy, and R. J. Murnane (2003): "The skill content of recent technological change: an empirical exploration," *The Quarterly Journal of Economics*, 118, 1279–1333.

Autor, D. H. and D. Dorn (2013): "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, 103, 1553–97.

Bailey, M. J. and S. M. Dynarski (2011): "Inequality in Postsecondary Education," in *Whither Opportunity*, ed. by G. J. Duncan and R. J. Murnane, Russell Sage, 117–132.

Bianchi, N. (2020): "The indirect effects of educational expansions: Evidence from a large enrollment increase in university majors," *Journal of Labor Economics*, 38, 767–804.

Bick, A. and N. Fuchs-Schündeln (2018): "Taxation and labour supply of married couples across countries: A macroeconomic analysis," *The Review of Economic Studies*, 85, 1543–1576.

Blackburn, R. M. and J. Jarman (1993): "Changing Inequalities in Access to British Universities," *Oxford Review of Education*, 19, 197–215.

Blanden, J. and S. Machin (2004): "Educational Inequality and the Expansion of UK Higher Education," *Scottish Journal of Political Economy*, 51, 230–249.

Blundell, R., D. A. Green, and W. Jin (2022): "The UK as a Technological Follower: Higher Education Expansion, Technological Adoption, and the Labour Market," *The Review of Economic Studies (forthcoming)*.

Boliver, V. (2011): "Expansion, differentiation, and the persistence of social class inequalities in British higher education," *Higher Education*, 61, 229–242.

——— (2013): "How fair is access to more prestigious UK universities?" *The British Journal of Sociology*, 64, 344–364.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *The American Economic Review*, 77, 531–553.

Bratsberg, B. and O. Rogeberg (2018): "Flynn effect and its reversal are both environmentally caused," *Proceedings of the National Academy of Sciences*, 115, 6674–6678.

Brinch, C. N. and T. A. Galloway (2012): "Schooling in adolescence raises IQ scores," *Proceedings of the National Academy of Sciences*, 109, 425–430.

Brynjolfsson, E., D. Li, and L. Raymond (2025): "Generative AI at Work," *The Quarterly Journal of Economics*, 140, 889–942.

Campbell, S., L. Macmillan, G. Wyness, A. Bryson, L. Stokes, and D. Wilkinson (2019): "Mismatch in higher education: prevalence, drivers and outcomes," .

Card, D. and T. Lemieux (2001): "Can falling supply explain the rising return to college for younger men? A cohort-based analysis," *The Quarterly Journal of Economics*, 116, 705–746.

Carneiro, P. and S. Lee (2011): "Trends in quality-adjusted skill premia in the United States, 1960-2000," *American Economic Review*, 101, 2309–2049.

Castex, G. and E. K. Dechter (2014): "The Changing Roles of Education and Ability in Wage Determination," *Journal of Labor Economics*, 32, 685–710.

Cervantes, C. V. and R. Cooper (2022): "Labor market implications of education mismatch," *European Economic Review*, 148, 104179.

Chowdry, H., C. Crawford, L. Dearden, A. Goodman, and A. Vignoles (2013): "Widening participation in higher education: analysis using linked administrative data," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176, 431–457.

Clouston, S. A., D. Kuh, P. Herd, J. Elliott, M. Richards, and S. M. Hofer (2012): "Benefits of educational attainment on adult fluid cognition: international evidence from three birth cohorts," *International Journal of Epidemiology*, 41, 1729–1736.

Cooper, R. and H. Liu (2019): "Mismatch in human capital accumulation," *International Economic Review*, 60, 1291–1328.

Corazzini, L., S. D'Arrigo, E. Millemaci, and P. Navarra (2021): "The influence of personality traits on university performance: Evidence from Italian freshmen students," *PLOS ONE*, 16, 1–20.

Cunha, F. and J. Heckman (2007): "The technology of skill formation," *American economic review*, 97, 31–47.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006): "Interpreting the evidence on life cycle skill formation," *Handbook of the Economics of Education*, 1, 697–812.

Deming, D. J. (2017): "The Growing Importance of Social Skills in the Labor Market," *The Quarterly Journal of Economics*, 132, 1593–1640.

Dillon, E. W. and J. A. Smith (2020): "The consequences of academic match between students and colleges," *Journal of Human Resources*, 55, 767–808.

Dworak, E. M., W. Revelle, and D. M. Condon (2023): "Looking for Flynn effects in a recent online U.S. adult sample: Examining shifts within the SAPA Project," *Intelligence*, 98, 101734.

EU Council (2021): "Council Resolution on a strategic framework for European cooperation in education and training towards the European Education Area and beyond (2021-2030)," Resolution 2021/C 66/01, Official Journal of the European Union, 64.

Fawns-Ritchie, C. and I. J. Deary (2020): "Reliability and validity of the UK Biobank cognitive tests," *PLOS ONE*, 15, 1–24.

Flynn, J. R. (1987): "Massive IQ gains in 14 nations: What IQ tests really measure," *Psychological Bulletin*, 101, 171–191.

Fortin, N. M. (2006): "Higher-Education Policies and the College Wage Premium: Cross-State Evidence from the 1990s," *American Economic Review*, 96, 959–987.

Freeman, R. (1976): *The overeducated american*, Academic press.

Goldin, C. and L. F. Katz (2008): *The Race between Education and Technology*, Cambridge, MA: Harvard University Press.

Goos, M. and A. Manning (2007): "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, 89, 118–133.

Goos, M., A. Manning, and A. Salomons (2009): "Job Polarization in Europe," *The American Economic Review*, 99, 58–63.

Greenwood, J., N. Guner, and G. Vandenbroucke (2017): "Family Economics Writ Large," *Journal of Economic Literature*, 55, 1346–1434.

Guner, N., R. Kaygusuz, and G. Ventura (2011): "Taxation and Household Labour Supply," *The Review of Economic Studies*, 79, 1113–1149.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Heckman, J. J. and S. Mosso (2014): "The economics of human development and social mobility," *Annual Review of Economics*, 6, 689–733.

Hoffmann, M., S. Boysel, F. Nagle, S. Peng, and K. Xu (2024): "Generative AI and the Nature of Work," Tech. rep., CESifo Working Paper 11479.

Hoxby, C. and C. Avery (2013): "The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students," Tech. rep., Brookings Papers on Economic Activity.

Hoxby, C. M. (2009): "The changing selectivity of American colleges," *Journal of Economic perspectives*, 23, 95–118.

Jandarova, N. (2022): "From bad to worse: long-term effects of recession in adolescence," Ph.D. thesis, European University Institute.

Jandarova, N. and J. Reuter (2021): "Multiple Imputation of University Degree Attainment," *Chapter of Jandarova's Ph.D. dissertation at the European University Institute.*

Kahn, L. B. (2010): "The long-term labor market consequences of graduating from college in a bad economy," *Labour Economics*, 17, 303–316.

Katz, L. F. and D. H. Autor (1999): *Changes in the Wage Structure and Earnings Inequality*, 1463–1555.

Katz, L. F. and K. M. Murphy (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *The Quarterly Journal of Economics*, 107, 35–78.

Kremen, W. S., A. Beck, J. A. Elman, D. E. Gustavson, C. A. Reynolds, X. M. Tu, M. E. Sanderson-Cimino, M. S. Panizzon, E. Vuoksimaa, R. Toomey, et al. (2019): "Influence of young adult cognitive ability and additional education on later-life cognition," *Proceedings of the National Academy of Sciences*, 116, 2021–2026.

Lagakos, D., B. Moll, T. Porzio, N. Qian, and T. Schoellman (2018): "Life Cycle Wage Growth across Countries," *Journal of Political Economy*, 126, 797–849.

Levy, F. and R. J. Murnane (1996): "With What Skills Are Computers a Complement?" *The American Economic Review*, 86, 258–262.

Lewandoski, P. (2020): "Ageing of routine jobs in Europe," *Economic Systems*, 44.

Lochner, L., Y. Park, and Y. Shin (2025): "The Evolution of Unobserved Skill Returns in the U.S.: A New Approach Using Panel Data," Working Paper 33407, National Bureau of Economic Research.

Machin, S. (1996): "Wage Inequality in the UK," *Oxford Review of Economic Policy*, 12, 47–64.

——— (2007): "Education expansion and intergenerational mobility in Britain," MIT Press, Schools and the equal opportunity problem.

Machin, S. and S. McNally (2007): "Tertiary education systems and labour markets," *Tertiary Review*, 1–57.

Major, E. and S. Machin (2018): *Social Mobility and its Enemies*, Pelican Books.

McFall, S. (2013): "Understanding Society:UK Household Longitudinal study: Cognitive Ability Measures," Institute for Social and Economic Research, University of Essex.

Meyer, J. W. and E. Schofer (2007): "The University in Europe and the World: Twentieth Century Expansion," Transcript verlag, Towards a Multyversity. Universities between Global Trends and National Traditions.

Michaels, G., A. Natraj, and J. V. Reenen (2014): "Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years," *The Review of Economics and Statistics*, 96, 60–77.

Office for National Statistics (2021): "Household income inequality, UK: financial year ending 2020," Statistical Bullettin, 21 January.

Olivetti, C. (2006): "Changes in women's hours of market work: The role of returns to experience," *Review of Economic Dynamics*, 9, 557–587.

Ollikainen, J.-P., T. Pekkarinen, R. Uusitalo, and H. Virtanen (2023): "Effect of secondary education on cognitive and non-cognitive skills," .

Oreopoulos, P., T. von Wachter, and A. Heisz (2012): "The Short- and Long-Term Career Effects of Graduating in a Recession," *American Economic Journal: Applied Economics*, 4, 1–29.

Pratt, J. (1997): *The Polytechnic Experiment: 1965-1992*, Bristol: Society for Research into Higher Education, Open University Press.

Prescott, E. C. (2004): "Why do Americans Work so Much More than Europeans?" Working Paper 10316, National Bureau of Economic Research.

Protzko, J. (2015): "The environment in raising early intelligence: A meta-analysis of the fadeout effect," *Intelligence*, 53, 202–210.

Ritchie, S. J., T. C. Bates, and I. J. Deary (2015): "Is education associated with improvements in general cognitive ability, or in specific skills?" *Developmental psychology*, 51, 573.

Ritchie, S. J. and E. M. Tucker-Drob (2018): "How much does education improve intelligence? A meta-analysis," *Psychological Science*, 29, 1358–1369.

Robbins, L. (1963): "Higher Education: Report of the Committee Appointed by the Prime Minister Under the Chairmanship of Lord Robbins, 1961-63," .

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

Salthouse, T. (2012): "Consequences of age-related cognitive declines," *Annual review of psychology*, 63, 201–226.

Salthouse, T. A. (2019): "Trajectories of normal cognitive aging." *Psychology and aging*, 34, 17–24.

Schofer, E. and J. W. Meyer (2005): "The Worldwide Expansion of Higher Education in the Twentieth Century," *American Sociological Review*, 70, 898–920.

Schwandt, H. and T. von Wachter (2019): "Unlucky Cohorts: Estimating the Long-Term Effects of Entering the Labor Market in a Recession in Large Cross-Sectional Data Sets," *Journal of Labor Economics*, 37, S161–S198.

Sutton Trust (2018): "Increasing higher education participation amongst disadvantaged young people and schools in poor communities," National Council for Educational Excellence.

Tafti, E. A. (2024): "Technology, skills, and performance: the case of robots in surgery," Tech. rep., Working Paper.

Tincani, M. M., F. Kosse, and E. Miglino (2022): "The Effect of Preferential Admissions on the College Participation of Disadvantaged Students: The Role of Pre-College Choices," .

UNESCO Statistics (2024): "School enrollment, tertiary (% gross)," retrieved from World Bank Open Data, https://data.worldbank.org/indicator/SE.TER.ENRR, last accessed on Sept. 17, 2024.

University of Essex, Institute for Social and Economic Research (2023): "Understanding Society: Waves 1-13, 2009-2022 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 18th Edition." .

von Wachter, T. (2020): "The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources," *Journal of Economic Perspectives*, 34, 168–94.

Walker, I. and Y. Zhu (2008): "The College Wage Premium and the Expansion of Higher Education in the UK," *The Scandinavian Journal of Economics*, 110, 695–709.

Willet, D. (2017): *A University Education*, Oxford: Oxford University Press.

Willis, R. J. and S. Rosen (1979): "Education and Self-Selection," *Journal of Political Economy*, 87, S7–S36.