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

Andrea Ichino[†] Aldo Rustichini[‡] Giulio Zanella[§]

May 25, 2023

Abstract

University access has greatly expanded during the past decades 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*. The enlargement of university access enacted in the UK following the 1963 Robbins Report 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 and to a declining college wage premium across cohorts. Our structural estimates indicate that the implemented policy was unfit to reach high-ability, disadvantaged individuals as Robbins had instead advocated. We show that counterfactual meritocratic selection policies would have attained that goal and so would have also been progressive.

JEL Classification: I23, I28, J24, O33

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

[*]We are grateful to Sascha Becker, Richard Blundell, Russell Cooper, Tom Crossley, David Green, Daniela Iorio, Wenchao Jin, Lars Kirkebøen, David Levine, Tommaso Porzio, and seminar participants at Bocconi, Bologna, Bristol, Cardiff, CESifo, EALE, EUI, Konstanz, Leibniz Institute Essen, Louvain-La-Neuve, Mannheim, NBER, 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 increased by a factor of 3.4 in OECD countries since 1970 (UNESCO Statistics, 2021)[1] and further expansion figures prominently in political agendas. For example, 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%" (Council of the EU, 2021). We investigate 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* (ability and disadvantage, henceforth). In our stochastic, general equilibrium Roy (1951) model, these traits determine the graduation probability, and their correlation is crucial to understand how technological progress and higher education policy alter incentives to pursue tertiary education. The model is used to study how actual policies shaped the evolution of students' sorting into college in terms of these traits and to simulate counterfactual policies, using UK data that span four decades of expansion. 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 experience offers an ideal case study. We illustrate its nature and consequences, drawing lessons to judge the ambitious targets currently set in Europe and elsewhere.

The UK 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 have not been adequately investigated for lack of data sets containing cognitive ability measures. An upside of our data is that we observe such measures, in addition to pre-college socioeconomic status (SES).

We find that: (i) on average, pre-college ability of graduates declined by about 13% of a SD between the 1960s and the 1990s; non-graduates' ability also declined, indicating

[1] This factor was about 1.9 in the US, 3.7 in France and in Japan, 3.9 in Italy, and 4.5 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), Riddell et al. (2013), Major and Machin (2018) and Blundell et al. (2022).

that students who attained a college degree in the 1990s (and who would have not in the 1960s) had higher ability than non-graduates of the 1960s, yet lower ability than the average graduate of the same period;[3] (ii) the reason is a non-meritocratic increase in the number of graduates, achieved by reducing non-tuition costs and by lowering qualification barriers at entry; (iii) the wage gap between college graduates and non-graduates declined progressively across cohorts;[4] like the increase in the supply of graduates, this pattern may mimic with a lag the college premium decline of the 1970s in the US, which was only later followed by an increase (e.g., Katz and Murphy, 1992, Fortin, 2006, Goldin and Katz, 2008 and Autor et al., 2020); (iv) although "untapped ability" did exist, the policy that prevailed was unfit to draw this ability into universities and ended up favoring primarily low-ability students from high-SES families;[5] (v) meritocratic policies based on the selection of high-ability students from any socioeconomic background (possibly with a subsidy to the study effort of more disadvantaged students) could have achieved the Robbins Report's progressive goals.

Although we eschew the difficult question of which social welfare function should be used to determine the decision to expand university access (a question that we postpone to future research), we claim that increasing college graduation opportunities for talented students can hardly be characterized as an undesirable outcome. In our model the reason is that, even if returns to college remain positive, a lower pre-college ability is associated (*ceteris paribus*) with a higher study effort cost, which implies a social welfare loss.[6] The Robbins Report

[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] Bianchi (2020) studies a large expansion of access to STEM majors enacted in Italy in the early 1960s and finds a similar impact on STEM graduates' wages. Our finding is instead in contrast with the weakly increasing wage gap over cohorts between college and high-school graduates in the UK reported by Blundell et al. (2022) in Figure 4 of their Online Appendix, which is puzzling given that we use the same methodology and Labour Force Survey (LFS) data to construct cohort wage ratios net of age effects. Our Online Appendix to Section 4.4 shows that the reason is the different group to which we compare college graduates: *all individuals without a college degree* instead of *high-school graduates* only. Section 3.2 explains why this is the appropriate comparison group to answer our research question.

[5] These results 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 mismatch between ability and educational attainment in the UK and other OECD countries, respectively. We leave in the background other consequences of higher education expansion 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.

[6] Other reasons may be considered in a richer model. For example, universities have a double role in

clearly mentions that preserving the high ability of students selected into college should be a condition for expanding tertiary education. Moreover, this condition does not necessarily have to come at the cost increasing inequality, as also suggested by the Robbins Report and confirmed by our counterfactual policy simulations.[7]

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 features: competitive firms produce output by combining graduate and non-graduate workers, thus affecting the wage gap. Skill-biased technical change increases the productivity of graduate workers and activates a force that increases the demand for college graduates independently of any change in higher education policy.

The labor supply side is more novel. In our model, obtaining a college degree is the outcome of two factors. One is simply the cognitive *ability* of the individual. The other is socioeconomic *disadvantage*, as determined by family background (e.g., parents' education, their presence in the household, and their employment status at the time a respondent was young).[8] In the model, ability and disadvantage affect the cost of study effort that a student must exert to attain a college degree, thereby altering an individual's graduation probability. The government can shape the parameters that link ability and disadvantage to the cost of study effort, thus expanding university access in different ways. To clarify the exposition, at the cost of some simplification, we adopt the following labels for three paradigmatic government interventions: a *Meritocratic expansion* (ME) policy favors more able students; a *Progressive expansion* (PE) policy favors more disadvantaged students; an *Indiscriminate expansion* (IE) policy enlarges university access independently of ability and disadvantage. The combination of ability and disadvantage with the effort cost parameters as shaped by policy generates *isoprobability curves* in the corresponding space (i.e., alternative combinations of ability and disadvantage such that the graduation probability is constant). These curves mark the boundary between higher and lower graduation probability regions,

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 reluctance to allow the tertiary education institutions of higher quality to be more selective in their acceptance.

[7]Surprisingly, these concerns are absent in Council of the EU (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.

[8]In the Online Appendix to Section 5.2 we extend the analysis to include the Big Five personality traits.

a stochastic generalization of the classical Roy (1951) model. A higher education policy is a way to change the position and slope of these curves.

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 were motivated by the belief that the UK was a stratified society where university access was facilitated more by high SES than by ability. In this case, if the correlation in question is positive, even an indiscriminate or progressive expansion may increase the fraction of graduates without reducing their average ability, as the Report claimed. Our evidence suggests that the UK society was indeed stratified, but was characterized by a *negative* correlation between pre-college ability and disadvantage – an unsurprising finding in light of what we know from 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 contexts, only a shift towards meritocratic policies aimed at increasing the graduation probability more strongly for higher ability students (possibly with a twist in favor of those sufficiently able but disadvantaged) could achieve the desiderata of the Robbins Report.

Our analysis has of course some limitations. One of them must be highlighted upfront so that the reader can calibrate expectations: to maintain tractability, we assume that higher ability reduces the effort cost of acquiring a college degree, and this is the reason why it is desirable, *ceteris paribus*, that the pre-college ability of college attendants is higher. However, we abstract from other implications of higher ability, like in particular the possibility of a direct effect on productivity for a given education level. Likewise, we abstract from the possible consumption value of a college education.

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

2 Model

We adopt a Becker-style human capital model in which education increases productivity. An innovation is the introduction of a study effort cost that depends on cognitive ability and on socioeconomic disadvantage, in a way that is affected by policy.

2.1 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 *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. The two variables are assumed to be publicly observable, and their joint distribution is denoted by $\mu \in \Delta(\mathbf{\Theta} \times \mathbf{\Lambda})$.

Each individual is also characterized by an endogenous human capital level $k \in \mathbf{K}$, where $\mathbf{K}$ is an ordered set of human capital levels. Given our focus on higher education, we consider only two levels, and so $\mathbf{K} = \{0, 1\}$ ($\equiv$ {no-college, college}).[10] 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 "college", once achieved, cannot be lost, so the only transition in human capital is from 0 to 1. In sum, we let $\pi : \mathbf{S} \times \mathbf{\Theta} \to [0, 1]$, 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 marginal disutility increases with disadvantage).[11] Under these assumptions, *ceteris paribus*,

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

[10] By adopting this definition, we deviate from the literature that studies the evolution of the wage gap between college and high school graduates, e.g., Machin and McNally (2007), Walker and Zhu (2008), and Blundell et al. (2022). The reason is that we are interested in evaluating whether the UK expansion was successful in drawing into college talented students who were previously likely to drop out of education at any lower level, not just at the high-school level.

[11] The two effort cost shifts $\Omega(\lambda)$ and $\Gamma(\theta)$ enter the utility function in an asymmetric way because we hypothesize that higher ability improves effectiveness of study effort directly, while disadvantage affects only the leisure utility at the given effort. For example, 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.

it is less costly in terms of leisure utility to admit to college 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 an individual's θ and λ. Examples are provided in Section 2.4.

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 *only* on an individual's human capital. That is, an individual's wage is given by $w(k)$.[12] Given a vector of wages $\mathbf{w} \equiv (w(0), w(1))$, an individual solves:

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

where we denote

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

It is convenient to specify

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

where $\Pi(\cdot) \equiv \min(\max(\cdot, 0), 1)$ is the cut-off function. The resulting probability of attaining college education 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. Under assumptions (2), (3) and (4), the optimal effort is unique and given by:

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

Note that although an individual can choose any positive effort level, it is never optimal to choose any $s > \frac{1}{\Gamma(\theta)}$, because effort is costly and the probability of college would not change. From equation (5), given a utility gap $\Delta U(\mathbf{w})$, the probability of college graduation for an individual of type (θ, λ) is

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

[12] As mentioned in the Introduction, this is an assumption that we make for tractability. However, note that even if the productivity of college graduates does not depend on ability, in this framework it is still desirable – *ceteris paribus* – to select high-ability students in college because their study effort cost is lower. A more general model where additional factors (for example cognitive ability, background, or gender) may affect wages directly is a task for future research.

Let $x(k)$ denote the population fraction with educational attainment k. The aggregate supply vector $\mathbf{x}^S \equiv [x^S(0)\ x^S(1)]$ is composed by

$$x^S(1) = \int_{\mathbf{\Theta}\times\mathbf{\Lambda}} \pi(s^*(\theta,\lambda),\theta)d\mu(\theta,\lambda); \qquad x^S(0) = 1 - x^S(1). \tag{7}$$

2.2 Firms

A representative firm has a technology that maps a vector of labor allocation into quantity of output produced. This technology is of the CES type; for every $\mathbf{x} \in \mathbb{R}^2_+$,

$$Q(\mathbf{x}) \equiv A\left(\sum_{k\in\{0,1\}} a(k)x(k)^\rho\right)^{\frac{1}{\rho}}, \tag{8}$$

where A is total factor productivity (TFP), the product of population size and the additional factor that allows us to normalize $\sum_k x^S(k) = 1$ *and* also $\sum_k a(k) = 1$. We assume $\rho \leq 1$, where $\rho \equiv \frac{\varsigma-1}{\varsigma}$, for ς the elasticity of substitution between *no-college* and *college* labor inputs.

The firm is competitive and solves, for any wage vector $\mathbf{w}$ taken as given, the following problem: $\max_{\mathbf{x}\in\mathbb{R}^2_+}(Q(\mathbf{x}) - \mathbf{w}\mathbf{x})$, where $\mathbf{w}\mathbf{x}$ is the inner product. Note that while aggregate labor supply is constrained at equilibrium by equation (7) to add up to 1, the competitive firm ignores this constraint. The first-order conditions for an interior solution are:

$$w(k) = A^\rho a(k)x(k)^{\rho-1}Q(\mathbf{x})^{1-\rho}, \quad k \in \{0,1\}, \tag{9}$$

and so labor demand by educational attainment, $x^D(k)$ for $k = 0, 1$ satisfies

$$\frac{w(1)}{w(0)} = \frac{a(1)}{a(0)}\left(\frac{x^D(1)}{x^D(0)}\right)^{\rho-1} \quad \Leftrightarrow \quad r = \alpha(\xi^D)^{\rho-1}, \tag{10}$$

where $\alpha \equiv \frac{a(1)}{a(0)}$ is the technological skill ratio and $r \equiv \frac{w(1)}{w(0)}$ and $\xi^D \equiv \frac{x^D(1)}{x^D(0)}$ are the college to no-college wage and labor demand ratios, respectively. In this model, technical change is represented by any change in A, α, or ρ. A change from $a(k)$ to $a'(k)$ is called *progress* if for all k, $a'(k) \geq a(k)$. A progress *favors college graduates* (i.e., is *skill-biased*) if $\alpha' \geq \alpha$.

2.3 Equilibrium

Existence of an equilibrium as defined below (and uniqueness in the special case that we consider in the empirical analysis), is discussed in Online Appendix to Section 2.3.

Definition 1 *An equilibrium in an economy described by the parameters* $(\Omega, \Gamma, \sigma_c, \sigma_s, A, \alpha, \rho)$ *is a vector* $(\boldsymbol{w}^*, s^*, \boldsymbol{x}^*)$ *such that: (i) individuals choose effort to maximize utility, and the aggregate labor supply* $\boldsymbol{x}^S$ *is determined by equation (7); (ii) the firm chooses labor to maximize profits; (iii) the labor and goods market clear:* $\boldsymbol{x}^S = \boldsymbol{x}^D = \boldsymbol{x}^*$ *and* $Q(\boldsymbol{x}^D) = \sum_k w^*(k) x^S(k)$.

In our structural estimation we use a characterization of the equilibrium labor allocation that provides a convenient computational algorithm. Using equation (9) to write wages at equilibrium as a function of the labor allocation, ΔU can be written as

$$\Delta U(\mathbf{w}(\mathbf{x}^*)) = \frac{(q(\mathbf{x}^*)a(1)x^*(1)^{\rho-1})^{1-\sigma_c} - (q(\mathbf{x}^*)a(0)x^*(0)^{\rho-1})^{1-\sigma_c}}{1-\sigma_c}, \tag{11}$$

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 following equation in the skilled labor fraction $x(1)$,

$$x(1) = \int_{\mathbf{\Theta}\times\mathbf{\Lambda}} \Pi\left(\frac{\theta}{\Gamma(\theta)}\left(1 - \left(\frac{\Omega(\lambda)\Gamma(\theta)}{\theta\Delta U(\mathbf{w}(1-x(1), x(1))}\right)^{\frac{1}{\sigma_s}}\right)\right) d\mu(\theta, \lambda), \tag{12}$$

where we use the fact that at equilibrium the wage vector is a function of the pair $(1 - x(1), x(1))$ of labor allocation from equation (9).

2.4 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,

$$\Delta U(\mathbf{w}) = \ln w(1) - \ln w(0) \equiv \Delta \ln w. \tag{13}$$

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

$$\Omega(\lambda) = \delta + \beta\lambda \qquad \text{and} \qquad \Gamma(\theta) = \gamma + \tau\theta \tag{14}$$

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

[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 $\sigma_c = 1$.

[14] As shown below, linearity allows for tractability while not limiting in any important way the types of higher education policies that we can analyze.

factors). Therefore, in what follows we refer to them as to "policy parameters", and a higher education policy is a quadruple $G = (\delta, \beta, \gamma, \tau)$.[15] Combining equations (6) and (13)-(14), the equilibrium probability of college graduation for a student of type (θ, λ) at policy G is

$$\pi^*(\theta, \lambda; G) = \Pi\left(\frac{\theta}{\gamma + \tau\theta} - \frac{\beta}{\Delta \ln w(G)}\lambda - \frac{\delta}{\Delta \ln w(G)}\right). \tag{15}$$

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 policies that took place after the Robbins (1963) Report, 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).

Given a status quo policy G, it is convenient to classify the possible interventions into three abstract categories of expansionary higher education policies G'.

Definition 2 *Let $\Delta\pi^*(\theta, \lambda) = \pi^*(\theta, \lambda; G') - \pi^*(\theta, \lambda; G)$.*

1. *An Indiscriminate Expansion (IE) policy is a G' that induces a function $\Delta\pi^*(\theta, \lambda) > 0$ and equal for all θ and λ.*

2. *A Progressive Expansion (PE) policy is a G' that induces a function $\Delta\pi^*(\theta, \lambda) > 0$ and increasing in λ for all θ.*

3. *A Meritocratic Expansion (ME) policy is a G' that induces a function $\Delta\pi^*(\theta, \lambda) > 0$ and increasing in θ for all λ.*

[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.

We emphasize that these categories are intended to provide a benchmark for the evaluation of real policies that do not necessarily match these requirements exactly. Note also that the intended aim of a policy is not necessarily the same as the actual outcome of the policy, once general equilibrium effects are considered. We return on this point below.

A central question when expanding university access is precisely the one addressed in the Robbins Report: namely, whether an increase in college participation is possible that would put unexploited ability to good use. To answer this question we need an operational definition of the notion of *untapped ability* that is at center stage in the Robbins Report. A possible definition is that untapped ability exists if there are two individuals i and j with $\theta_i > \theta_j$ and $k_i < k_j$ (i.e., i has higher ability than j but j achieves a college degree while i does not). However, because individuals cannot be forced to go to college, we need a policy-relevant definition of *reachable ability*. To this end, we denote by $\xi(G)$ the college to no-college labor ratio at equilibrium under policy G.

Definition 3 *Reachable ability at a policy G exists if there exists a different policy G' such that* $: \mathbb{E}(\Theta|K = 1; G') \geq \mathbb{E}(\Theta|K = 1; G)$ *and* $\xi(G') > \xi(G)$.

That is, there are skills that can be put to good use via higher education if there exists a policy such that, at equilibrium, the fraction of population with a university degree is higher and mean ability of college graduates is not smaller. As remarked in the Introduction, the Robbins Report claimed the existence of reachable ability by ruling out "that expansion would lead to a lowering of the average ability of students in higher education." (p. 53).

In order to establish the conditions for the existence of reachable ability, we assume (for mere illustrative purposes) that pre-college ability and 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 type on the ability distribution in the college population depends on the relation between the slopes of two functions that link Θ and Λ.

The first is the tilt of the joint density $\mu(\Theta, \Lambda)$. Its inclination can be conveniently characterized by the slope $\eta\frac{\sigma_\Lambda}{\sigma_\Theta}$ of the population linear regression of Λ on Θ,

$$\Lambda - m_\Lambda = \eta\frac{\sigma_\Lambda}{\sigma_\Theta}(\Theta - m_\Theta). \tag{16}$$

The second slope is that of the *isoprobability curves* of obtaining a college degree, i.e., the locus of (θ, λ) combinations such that the probability of graduating is constant. Using equation (15), it is immediate that this slope is given by

$$\frac{\partial \Lambda}{\partial \Theta}(\theta, \lambda) = \frac{\gamma \Delta \ln w(G)}{(\gamma + \tau\theta)^2 \beta}. \tag{17}$$

When $\tau = 0$, isoprobability curves are straight lines. The comparison between the two slopes is crucial in the following analysis, so it is convenient to label the difference:

$$\psi(\theta, G) \equiv \frac{\gamma \Delta \ln w(G)}{(\gamma + \tau\theta)^2 \beta} - \eta \frac{\sigma_\Lambda}{\sigma_\Theta}. \tag{18}$$

The comparative statics of interest is how different expansive higher education policies alter mean conditional ability and disadvantage, $\mathbb{E}(\Theta|K)$ and $\mathbb{E}(\Lambda|K)$, of students selected or not selected into college. Such policies induce equilibrium responses with effects that vary across regions of the policy space G. Thus, their consequences are cumbersome to characterize analytically. Leaving a theoretical illustration to the Online Appendix to Section 2.4, here we resort to numerical simulations of the model's equilibrium.

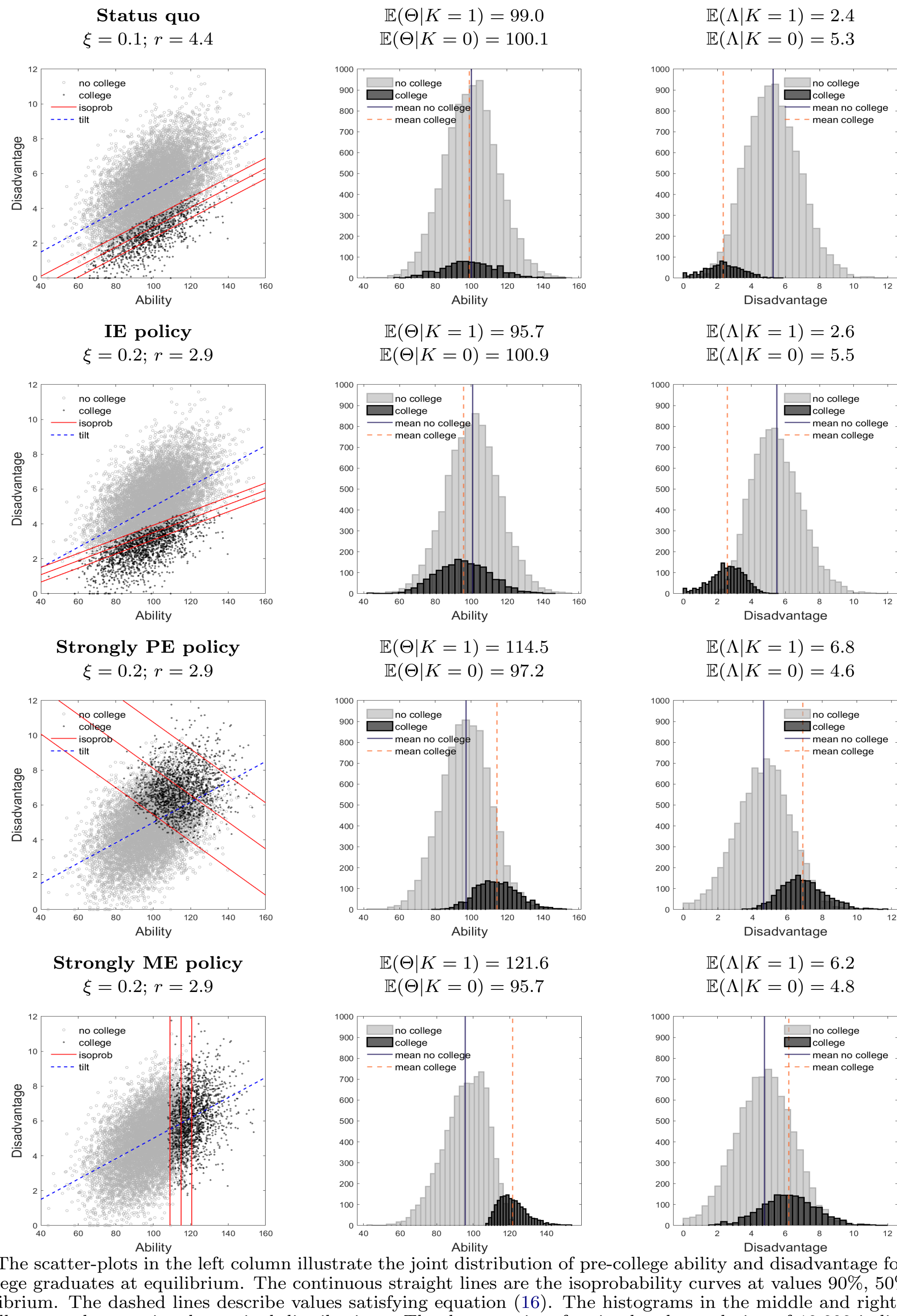
2.5 Numerical simulation

Using simulated data, Figures 1 and 2 describe the effects of the three higher education policies of Definition 2 in two paradigmatic societies. In Figure 1 (Society 1), $\eta > 0$ (i.e., pre-college ability Θ and disadvantage Λ are positively correlated), but $\psi(\cdot, G) < 0$ (i.e., isoprobability lines are flatter than the line describing the tilt of the joint distribution $\mu(\Theta, \Lambda)$). Figure 2 (Society 2) features instead $\eta < 0$, in which case it is necessarily $\psi(\cdot, G) \geq 0$.[16]

The top rows illustrate the role of ψ and η in determining the conditional distribution of Θ and Λ at equilibrium. The scatter plots on the left represent individuals of type (θ, λ) in the population and their allocation to college and no-college attainment at equilibrium wages. The dashed line graphs equation (16), which measures the tilt of $\mu(\Theta, \Lambda)$. The three continuous lines are isoprobability curves associated with graduation probability of 0.9 (bottom), 0.5 (middle), and 0.1 (top). For each isoprobability curve, an individual above or below the line has a college graduation probability $\pi^*(\theta, \lambda; G)$ smaller or larger than the probability associated with that curve, respectively. Each individual is assigned to college or no-college attainment if $\pi^*(\theta, \lambda; G)$ is above or below a random threshold. In the status quo, it is $\tau = 0$ and so these curves are straight lines. A policy change from G to G'

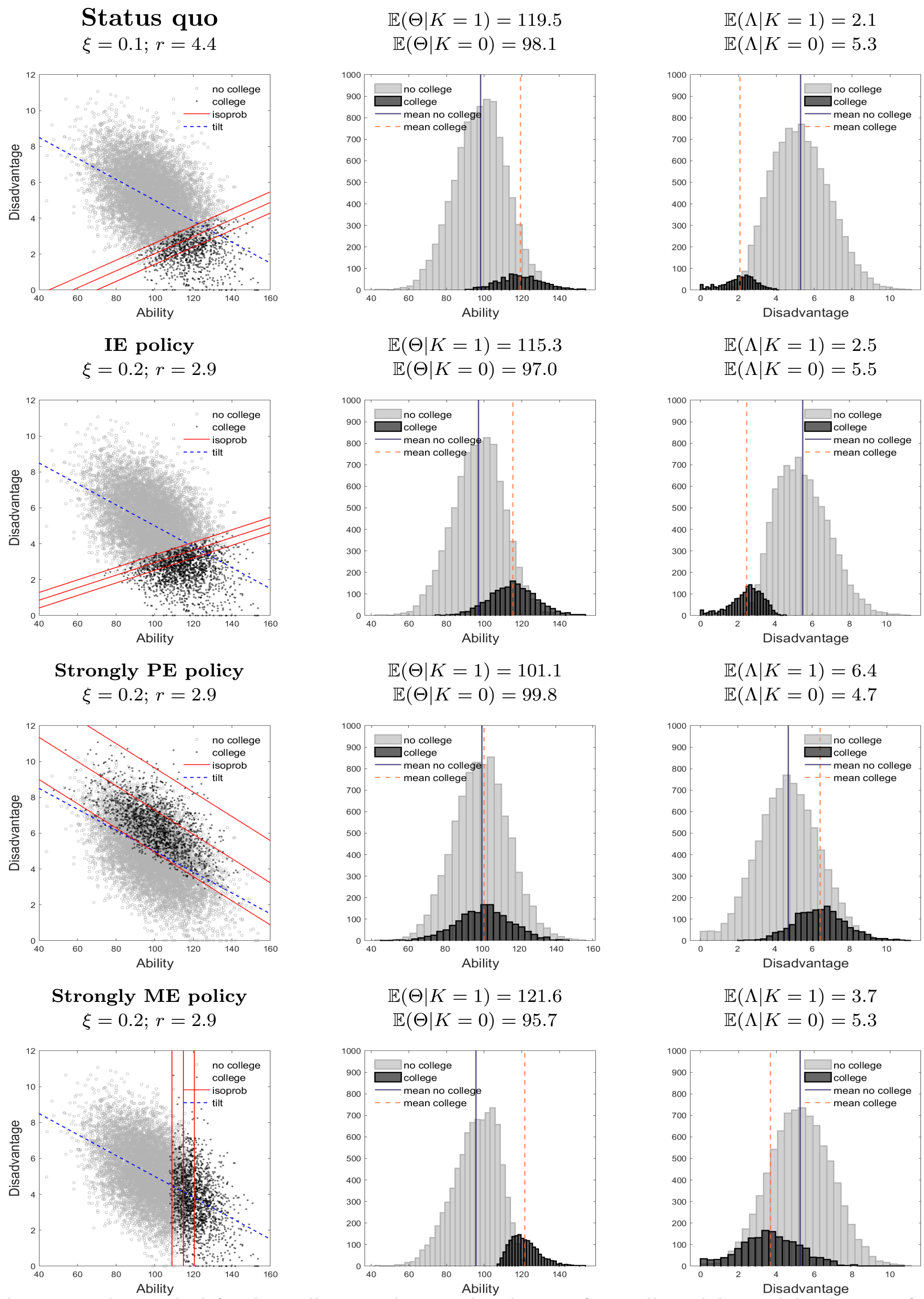
[16]For completeness, the third possible society characterized by $\eta > 0$ and $\psi(\cdot, G) \geq 0$ is considered in Figure A-1 of the Online Appendix to Section 2.5.

Figure 1: Status quo in Society 1 ($\eta > 0, \psi < 0$) and effects of three expansion policies: Indiscriminate Expansion (IE), Progressive Expansion (PE), Meritocratic Expansion (ME).



Notes: The scatter-plots in the left column illustrate the joint distribution of pre-college ability and disadvantage for no-college and college graduates at equilibrium. The continuous straight lines are the isoprobability curves at values 90%, 50% and 10%, at equilibrium. The dashed lines describe values satisfying equation (16). The histograms in the middle and right columns of panels illustrate the associated marginal distributions. The data consist of a simulated population of 10,000 individuals with type (θ, λ) drawn (for mere illustrative purposes) from a jointly normal distribution ($m_\Theta = 100$; $\sigma_\Theta = 15$; $m_\Lambda = 5$; $\sigma_\Lambda = 1.75$; $\mathrm{corr}(\Theta, \Lambda) = \eta = 0.5$). In the first row (status quo), the policy parameters are set to generate $\xi = 0.1$: $\gamma = 26.1$, $\tau = 0$ (so isoprobability curves are straight lines), $\delta = 2$, $\beta = 1$. The technology parameters are $\alpha = 1.1$ and $\rho = 0.4$. For each policy experiment in the other rows, the parameters are set so as to double the college to no-college labor ratio. The wage ratio adjusts to equilibrium. IE policy: $\delta = 0$. Strongly PE policy: $\beta = -0.16$, $\gamma = 86$. Strongly ME policy: $\tau = -8$, $\beta = 10^{-6}$, $\gamma = 30.1$, $\delta = 5.3$.

Figure 2: Status quo in Society 2 ($\eta < 0, \psi < 0$) and effects of three expansion policies: Indiscriminate Expansion (IE), Progressive Expansion (PE), Meritocratic Expansion (ME).



Notes: The scatter-plots in the left column illustrate the joint distribution of pre-college ability and disadvantage for no-college and college graduates at equilibrium. The continuous straight lines are the isoprobability curves at values 90%, 50% and 10%, at equilibrium. The dashed lines describe values satisfying equation (16). The histograms in the middle and right columns of panels illustrate the associated marginal distributions. The data consist of a simulated population of 10,000 individuals with type (θ, λ) drawn (for mere illustrative purposes) from a jointly normal distribution ($m_\Theta = 100$; $\sigma_\Theta = 15$; $m_\Lambda = 5$; $\sigma_\Lambda = 1.75$; $\text{corr}(\Theta, \Lambda) = \eta = -0.5$). In the first row (status quo), the policy parameters are set to generate $\xi = 0.1$: $\gamma = 31$, $\tau = 0$ (so isoprobability curves are straight lines), $\delta = 2$, $\beta = 1$. The technology parameters are $\alpha = 1.1$ and $\rho = 0.4$. For each policy experiment in the other rows, the parameters are set so as to double the college to no-college labor ratio. The wage ratio adjusts to equilibrium. IE policy: $\delta = 0$. Strongly PE policy: $\beta = -0.18$, $\gamma = 87$. Strongly ME policy: $\tau = -8$, $\beta = 10^{-6}$, $\gamma = 30.1$, $\delta = 5.3$.

changes the slope of isoprobability curves, which is given by equation (17), or their vertical intercept, which for some probability level π is given by $-\frac{\Delta \ln w(G)}{\beta}\left(\pi + \frac{\delta}{\Delta \ln w(G)}\right)$, or both. The histograms are the resulting conditional distributions of ability and disadvantage.

In Society 1, many high-ability students with a disadvantaged background are excluded from higher education (north-eastern region of the scatter plot in the top row of Figure 1), hence the paradoxical outcome that the population in college has on average lower ability than the population outside college. In this society, university access is easier for students from affluent families even if they are not very talented. As shown below, this is the most favorable case for a government that wishes to expand access without reducing the quality of graduates: in Society 1 it is relatively easy to reduce access barriers and draw talented students from any socioeconomic background into college.

In Society 2, graduates have on average higher ability than non-graduates and have a relatively advantaged background. The pool of high-ability students with a disadvantaged background that are excluded from higher education is evidently smaller than in Society 1. There is still reachable ability in Society 2 but less than in Society 1. Section 4.5 shows that this is the case that best characterizes the UK in the years that we study.

The remaining rows of Figures 1 and 2 illustrate policy effects. Starting from a college to no-college graduation rate of $\xi = 0.1$, we simulate three policy changes that increase this rate to $\xi = 0.2$. First, an *intended* indiscriminate expansion (IE) policy, which decreases the intercept δ in effort cost shift $\Omega(\Lambda) = \delta + \beta\Lambda$ (equation 14). By decreasing δ, this policy may appear to shift isoprobability curves upward without affecting their slope and thus, since $\eta > 0$, allow high-ability and low-SES students to access college. But this conclusion ignores the effect of the policy on the college wage premium, which would decline due to the higher supply of graduates; this may offset the policy change by flattening isoprobability curves (see equation 17) and ultimately *reduce* the average ability of individuals selected into college. Parameters are chosen to demonstrate that this may be the case even in Society 1 (where students not attaining higher education have on average higher ability than those who do): although the government intends to shift isoprobability lines up as an easy way of reaching the many talented students outside college, the drop in the wage ratio from $r = 4.4$ to $r = 2.9$ reduces the slope of the lines and the policy ends up favoring primarily the not-so-talented students from relatively advantaged families in the southwestern portion of the scatter plot. The incidence of low-SES graduates increases only marginally relative to the status quo.

A fortiori, also in Society 2 an intended IE policy reduces mean ability of graduates while not affecting their average background much. Note that in this society – which is the empirically relevant one in our case study – the expansion decreases mean ability both conditional on having attained a college degree *and* conditional on not having attained it. This means that in Society 2 the indiscriminate expansion draws into college students who have higher pre-college ability than the average non-graduate, yet lower ability than the average graduate. As shown in Section 4.2, this is a pattern that we find in the data and that, as remarked in the Introduction, is not surprising (e.g., Cunha et al., 2006).

Consider next an intended progressive expansion (PE) policy that decreases the slope β of effort cost shift $\Omega(\Lambda) = \delta + \beta\Lambda$ while increasing the intercept γ of effort cost shift $\Gamma(\Theta) = \gamma + \tau\Theta$. This reform aims at decreasing the importance of a student's SES relative to ability in determining the graduation probability. In Figures 1 and 2, it takes a strong form because β turns from positive to negative, so that a low SES (large λ) becomes an advantage in college access, as indicated by the fact that isoprobability curves become negatively sloped. This policy induces a large increase in the incidence of disadvantaged graduates in both societies. However, its effect on their average ability is positive and large in Society 1 but *negative* in Society 2. Expanding university access without decreasing graduates' ability is not easy when the correlation η between pre-college ability and disadvantage is negative.

This dilemma is resolved by the strongly meritocratic (ME) policy illustrated in the bottom row of the two figures. Here the parameters of the effort cost shifts $\Omega(\Lambda)$ and $\Gamma(\Theta)$ are adjusted to make $\tau < 0$ (so there is a cost shift in favor of high-ability students), while β approaches zero (so that one's SES becomes irrelevant) and γ and δ both increase to obtain the desired college to no-college rate ξ. The result is that isoprobability curves become nearly vertical. This strongly ME policy raises the incidence of high-ability *and* high-disadvantage individuals in the college population of the two societies. Such a policy not only increases the average ability of students in higher education; it is also an egalitarian one, in the sense that it draws into college talented students with a disadvantaged background. In Society 2, this is the *only* one among the three classes of expansion strategies that achieves these goals. In Society 1 they can be obtained with a wider range of policies.[17]

[17] Using the code in our replication package, the reader can verify that these conclusions are fairly general.

3 Data

We next describe our data sources and the measurement of the four variables that are at center stage in the model: college attainment, ability, disadvantage, and earnings.

3.1 Data sources

Our main data source is Understanding Society (USoc), a representative longitudinal survey of UK households. Wave 3 (2011-2013) is composed of, 49,692 individuals and contains cognitive ability information. We restrict this sample to individuals: (i) with non-zero cross-sectional response weights (38,223); (ii) white born in the UK (31,132), so as to work with homogeneous cohorts; (iii) with non-missing education information (31,072); (iv) born between 1940 and 1984 (23,288). Table 1 reports descriptive statistics. Since 1,113 observations have missing ability information ability, the table compares all individuals with those with cognitive test scores to show that the ability measure is missing quasi at random. Our final USoc sample consists of 22,175 individuals with non-missing ability measure. Two ancillary data sets (described in the Online Appendix to Section 3.1) are used to strengthen our analysis of pre-college cognitive ability: the UK Biobank (UKB) and the 1970 British Cohort Study (BCS70). Moreover, in order to provide evidence on how the UK expansion was enacted, we use data from the University Statistical Record (USR), which contain administrative information on the universe of students enrolled at UK universities between 1972 and 1993. Finally, following Blundell et al. (2022), we use the UK Labour Force Survey (LFS) for the analysis of the evolution of the wage gap between college graduates and non-graduates. These data sources are also described in the Online Appendix to Section 3.1.

3.2 College graduation rate and college cohorts

Before 1992, students wishing to pursue higher education in the UK had two options: enrolling in a traditional university or attending a polytechnic.[18] 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. Following the Robbins Report, which had recommended the unification of UK higher edu-

[18]There was also the option of attending professionally-oriented public colleges, such as teacher training and nursing colleges. This group was relatively small and so we consider it as part of "polytechnics".

Table 1: The UK Understanding Society sample

	White UK born in 1940-1984					White UK born in 1940-1984 with non-missing ability scores				
	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
Female	23,288	0.52	0.50	0	1	22,175	0.52	0.50	0	1
Any tertiary degree	23,288	0.24	0.43	0	1	22,175	0.25	0.43	0	1
Age left school	22,896	16.26	1.11	7	21	21,794	16.29	1.12	7	21
Age left FT edu	11,450	22.08	6.18	15	67	11,146	22.10	6.17	15	67
Born in England	22,990	0.81	0.39	0	1	21,892	0.81	0.39	0	1
Health status	23,287	2.57	1.11	1	5	22,174	2.54	1.10	1	5
Number of marriages	20,475	1.01	0.61	0	4	19,469	1.01	0.61	0	4
N. of children < 18	23,288	0.36	0.81	0	8	22,175	0.36	0.81	0	8
Religious belonging	22,051	0.48	0.50	0	1	20,986	0.48	0.50	0	1
Real monthly income	23,288	2.00	1.71	-8	26	22,175	2.03	1.73	-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
Mother's yrs school	19,846	11.47	2.44	0	18	18,950	11.51	2.44	0	18
Father employed	22,905	0.88	0.32	0	1	21,818	0.89	0.32	0	1
Mother employed	23,020	0.62	0.48	0	1	21,930	0.63	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 apply four selection criteria: first, we keep observations with non-zero cross-sectional response weights (38,223); second, we restrict to white respondents born in the UK (31,132); third, we keep observations with non-missing education information (31,072); finally, 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 right panel reports the same descriptive statistics for our final USoc 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. Real monthly income is expressed in thousands of 2015 GBP.

cation 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 a higher education degree of any kind. This 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.

We instead depart from this literature in the definition of the comparison group (see also Section 2.1). We are interested in evaluating whether the UK expansion was successful in drawing into college those talented students who were previously likely to drop out of education at *any* lower level, not just at the high school level. Therefore, our comparison group is composed by individuals with any educational attainment below a tertiary degree in the population that we study.[19]

To facilitate the interpretation of our results in relation to historical information on policy and technology trends, we aggregate individuals into "college cohorts". These are groups of individuals in actual (for graduates) or potential (for non-graduates) college attendance age. For such age, we use as a label the year of birth plus 20. The large sample size available in the UKB allows us to construct college cohorts using 5-year windows. For the smaller USoc sample that we use for inference, we construct three 15-year periods in order to increase sample size and thereby statistical power. These three periods are: 1960-1974 for individuals born between 1940 and 1954 (7,103 individuals in the final sample), 1975-1989 for those born between 1955 and 1969 (8,329 individuals), and 1990-2004 for subjects born between 1970 and 1984 (6,743 individuals). Labeling these groups as "college cohorts" avoids possible confusion with birth cohorts. In light of evidence suggesting that the time of entry in the labor market has long-term consequences on wages and employment along the life cycle,[20] it is reasonable to assume the absence of first-order substitutability between college graduates across these cohorts, and similarly for non-college graduates.

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 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) to extract from these variables a measure of pre-college general cognitive

[19] To quantify the difference in the definition of the comparison groups, out of the 936,135 observations in our LFS sample, 146,565 (15.7% of the total) are not high school graduates according to the definition of Blundell et al. (2022) and so are not included in their comparison group, while they are included in ours.

[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.

ability. 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-3 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.

The UKB provides instead a Fluid Intelligence Score (FIS), the sum of the correct answers to 13 cognitive questions: numeric addition, identification of largest number, word interpolation, positional arithmetic, family relationship calculation, conditional arithmetic, synonim, chained arithmetic, concept interpolation, arithmetic sequence recognition, antonym, square sequence recognition, and subset inclusion logic.[22] There is no reason to aggregate the subtest scores in a way different from the one adopted by the UKB, and so we use FIS as the ability measure in this data set. Fawns-Ritchie and Deary (2020) validate the presence of a general cognitive ability component in FIS and conclude that "despite the brief and non-standard nature of the UK Biobank cognitive assessment, a measure of general cognitive ability can be created using these tests" (p. 19).

These ability scores are taken to be cardinal measures of the underlying cognitive ability construct, so any monotonic linear transformation (MLT) of these measures is admissible and we must pick one. It is convenient to choose a MLT such that variable Θ has mean 100 and standard deviation 15, so as to make 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* of the ability measure, as is evident in equation (15). Since we are interested in policy *changes*, the particular scale that we choose is irrelevant. The distribution of the resulting ability measures in USoc and UKB are illustrated, respectively, in the left and right panels of Figure A-2 in the Online Appendix to Section 3.3.

The USoc and UKB cognitive tests are administered *after* potential or actual college attendance and so may be endogenous to tertiary education. In line with the literature, the BCS70 indicates 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 age 5, 10, and 34. Using again PCA to construct a normalized cognitive ability measure (see Tables A-4 to A-6 in the Online

[22]See the UK Biobank data show case for a detailed description of these cognitive tests.

Appendix to Section 3.3 for additional details), we report in Figure 3 the average ability at these three ages of subjects who eventually obtained (circles, top line) or did not obtain (triangles, bottom line) 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 in the two groups varies before age 10, there is no significant variation afterwards. In particular, college graduates do not experience an increase in cognitive ability relative to non-graduates. 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.[23] 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) show that any effect of schooling on specific cognitive skills is not mediated by the general cognitive ability, which instead seems to be largely unaffected by education.[24]

We emphasize that our ability measures can therefore be interpreted as *pre-college* measures of cognitive skills and are the outcome of both nature and nurture. As such, they also reflect socioeconomic disadvantage, as discussed in more detail below. Yet, the evidence from the BCS70 reported in Figure 3 and the economic literature on cognitive development (as summarzied, for example, by Heckman and Mosso, 2014) suggest that our ability measures capture skills that are fixed relatively early in life.

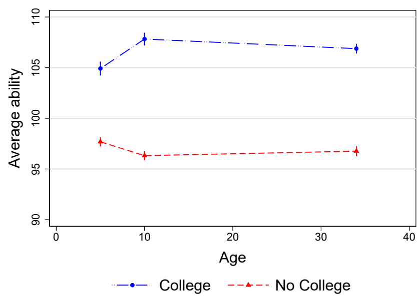
Like all variables in econometric analysis, our cognitive ability indicators contain measurement error. For example, it is has been argued that cognitive ability varies over time for a given age and over age for a given cohort. The first variation is known as the "Flynn effect" because Flynn (1987) measured 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 the average age of a cohort,

[23]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.

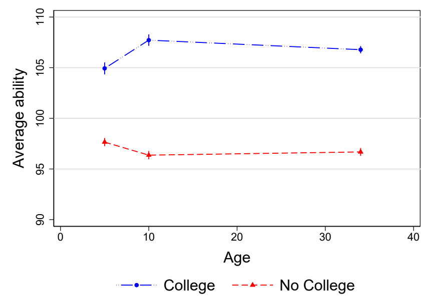
[24]This conclusion is consistent with Ritchie and Tucker-Drob (2018), since "the vast majority of the studies in [their] meta-analysis considered specific tests and not a latent g factor" (p. 1367).

Figure 3: The effect of higher education on cognitive ability

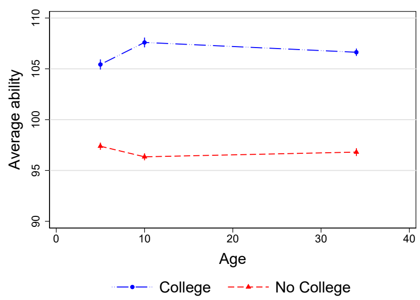
(A) No imputation, same subjects



(B) No imputation, unequal sample sizes



(C) With imputation, equal sample sizes



(D) Sample sizes

Age	Panel A	Panel B	Panel C
5	5,657	7,479	9,665
10	5,657	6,927	9,665
34	5,657	9,448	9,665

Notes: 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 SD 15. Panel (A) is based on the 5,657 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 the table of Panel (D). 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 non-graduates.

the Salthouse effect must be removed by normalizing both the Usoc and the UKB ability measures within birth years.[25] This comes at the cost of removing also the Flynn effect, which is less of a concern because this finding is more questionable. 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. The within-birth year normalization implies that the policy parameters γ and τ that we will estimate incorporate any residual measurement error.

[25] The comparison between Figure A-3 and Figure A-4 in the Online Appendix to Section 3.3 illustrates the effect of this normalization on the two measures.

3.4 Socioeconomic disadvantage

A large realization λ of disadvantage in our model represents low SES, a non-cognitive factor that reduces the probability of college graduation. For example, 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 USoc 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-7 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.[26]

Like for cognitive ability, we take the disadvantage measure produced by PCA as a cardinal measure of the underlying concept and any MLT is admissible. Since there is no scale in the psychometric tradition for variable Λ, we simply use the translation of the PCA measure that we have described above. As is again evident in equation (15), the scale of parameter β adapts to this particular scale, which is immaterial because we want to estimate *changes* in β across cohorts. However, like for γ and τ in the case of ability, our estimates of policy parameters δ and β will contain some measurement error because we cannot include in the PCA all the factors that are relevant determinants of variable Λ. Figure A-5 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 unavailability 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 more long-term determinant 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

[26] A more recent literature summarized by Corazzini et al. (2021) argues that non-cognitive personality traits are also a relevant determinant of tertiary education attainment. We show in the Online Appendix to Section 5.2 that augmenting our ability or disadvantage measures with the "Big Five" personality traits (which are available in USoc for the cohorts that we study) does not change our conclusions.

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, contrary to the removal of the Salthouse effect which is instead desirable, as discussed in Section 3.3.

4 Key empirical facts

In this section we document four key facts that the model is required to reproduce: increasing fraction of college graduates (with an illustration of the main policies that implemented such expansion); decreasing average ability of both college and non-college graduates; decreasing average disadvantage, for the entire population and by graduation status; and declining wage ratio between college graduates and non-graduates. We also document a fifth key fact that is relevant for interpreting the consequences of the UK expansion: negative correlation between pre-college cognitive ability and socioeconomic disadvantage; this means that, during the period that we consider, the UK resembles Society 2 of Figure 2.

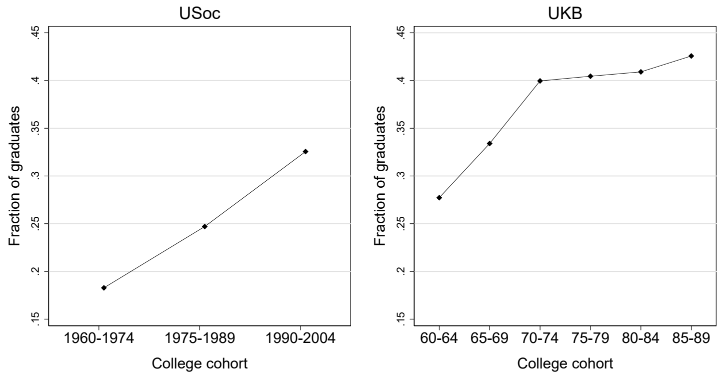
4.1 The fraction of college graduates increased steeply

Figure 4 shows that in the USoc sample (left panel) the fraction of graduates increased from about 17% in college cohort 1960-1974 to about 32% in cohort 1990-2004. A similar trend is observed in our UKB data sample (right panel). Since UKB respondents are on average more educated than the UK population, their college graduation rate is higher than in USoc; yet we observe a similar increase in college graduation rates: from about 28% to about 43%.[27]

This expansion was enacted in a mostly non-meritocratic way, by ending the binary divide between traditional universities and polytechnics, by increasing the number of academic institutions, 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. The stock of students enrolled in universities more than doubled (from 152,227 to 376,074) between 1966 and 1992. This increase is smaller than for Polytechnics (where numbers more than quadrupled in the same period, from 149,720 to 659,790), but is still substantial. While

[27] Figure 4 plots model variable $x(1)$, i.e., the fraction of graduates. The structural analysis is in terms of ξ, i.e., the college to no-college labor ratio. There is a 1:1 mapping between the two because $\xi = \frac{x(1)}{1-x(1)}$.

Figure 4: Fraction of college graduates by college cohort



Notes: The left panel displays the fraction of graduates – model variable $x(1)$ – in three Usoc college cohorts (sample: 22,175 white respondents born in the UK between 1940 and 1984, with non-missing education and ability score; see Table 1). The right panel displays the same variable in six UKB college cohorts (sample: 212,624 white respondents born in the UK between 1940 and 1969, with non-missing education and ability score; see Table A-1).

the growth of Polytechnics is a consequence of an explicit expansion policy,[28] 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 the number of academic institutions and by bringing them closer to potential students.[29] In the USoc sample, the *average* distance from the closest university dropped by about 6 km 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.[30] 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).

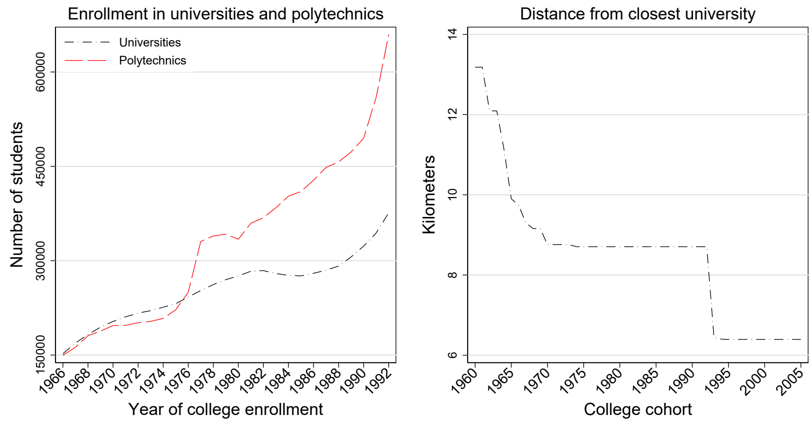
A second piece of evidence comes from USR, which provides information on the evolution of enrollment and entry criteria. The left panel of Figure 6 shows the fraction of students admitted without A-levels to three groups of UK universities: Oxbridge, the Russell group, and the remaining, less prestigious institutions. In all groups, the fraction of students admit-

[28]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.

[29]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).

[30]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. And 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").

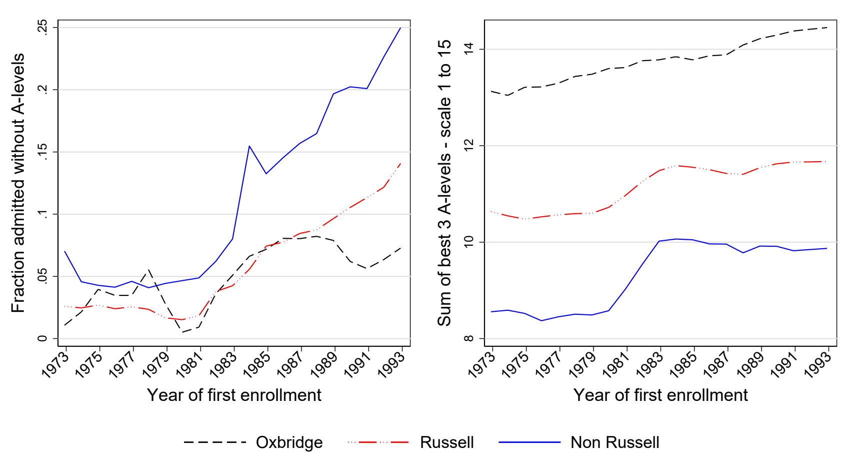
Figure 5: Higher education enrollment and distance to the closest college



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.

ted without A-levels 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 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 mininum 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 displays the fraction of students admitted without A-level scores to three groups of UK universities: Oxbridge, the Russell group, and remaining institutions. The right panel reports instead 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.

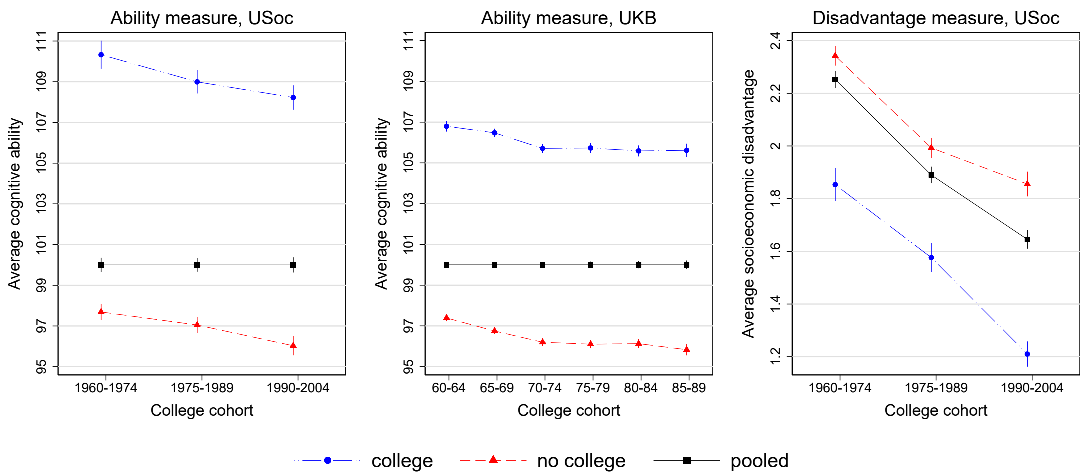
The right panel reports instead the average sum of the best 3 A-Level scores for students admitted to the three groups during the period covered by USR data. As expected, in all years students admitted at Oxbridge have higher best A-level scores than students admitted at the Russell group, which in turn 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 so as to facilitate college admission. Second, universities became more selective in admitting students or 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 next, this is not the case (see Figure 7).

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".

4.2 Graduates' average ability declined

The left and middle panels of Figure 7 show the evolution of average cognitive ability in our samples across the different college cohorts, by college graduation status. In USoc, the average ability of the population is constant by construction (see Section 3.3), at a value of 100. However, for college graduates (left panel) it declined by about two points (13% of a standard deviation), from 110.3 in the 1960-1974 college cohort to 108.2 in the 1990–2004 cohort. Interestingly, during the same period, also the average ability of non-graduates declined by about two points, from 97.7 to 96.0. Similar dynamics are observed in the UKB sample, where graduates' average FIS (middle panel) declined from 106.8 in the 1960-1964 cohort to 105.6 in the 1985-1989 cohort; for non-graduates the decline was from 97.4 to 95.8. The declining average ability of both graduates and non-graduates suggests that the expansion of higher education that was enacted in the UK brought into college students who had higher ability than average in the group of those previously excluded, yet lower ability than the average student who was previously admitted to college, as conjectured by Walker and Zhu (2008) and Blundell et al. (2022).[31]

Figure 7: Dynamics of ability and disadvantage measures by college attainment



Notes: The left and right panels display, respectively, the dynamics of average ability and average disadvantage in the population and by college graduation status across the three USoc college cohorts (sample: 22,175 white respondents born in the UK between 1940 and 1984, with non-missing education and cognitive test scores; see Table 1). The middle panel shows the dynamics of average FIS in the population and by college graduation status across the UKB college cohorts (sample: 212,624 white respondents born in the UK between 1940 and 1969, with non-missing education and cognitive test scores, see Table A-1).

4.3 Graduates' average disadvantage declined

The right panel of Figure 7 shows the evolution of average socioeconomic disadvantage in the USoc sample, for the entire population and by college attainment status. This variable is not constrained to be constant on average in the population (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 consider.[32] Between the 1960-1974 and the 1975-1990 college cohorts, the decline was about 15% both among college graduates and non-graduates. This means that, initially, the expansionary higher education policy affected the average background of college and non-college students only marginally.

The outcome of the sorting process departs more substantially from mere population changes for the 1990-2004 college cohort: relative to the 1975-1990 cohort, average disadvantage declined by 12.9% in the population, 23.2% among college graduates, and 6.9%

[31] Additional evidence supporting this interpretation is offered in Table A-9 of the Online Appendix to Section 5.2, which shows that the bottom percentiles of the ability distribution of graduates declined significantly while the top remained almost unchanged. The opposite happened for non-graduates.

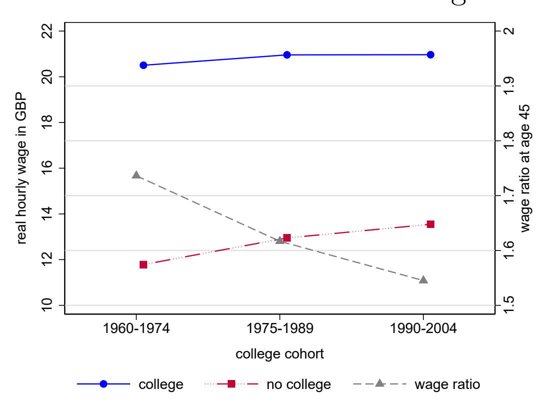
[32] The standard deviation does not change and is about 1.3 in all cohorts. A declining mean and a constant SD imply a declining coefficient of variation, i.e., widening relative inequality along the disadvantage dimension, in line with evidence in, e.g., Machin (1996) and Office for National Statistics (2021).

among non-graduates. These figures suggest that the more recent stage of the expansion process brought students into college who were relatively advantaged in the group of those previously excluded, and also more advantaged than the average student who was previously admitted to college. This fact is in line with existing evidence that the enlargement of higher education in the UK has predominantly benefited children from high-income families (e.g., Blanden and Machin, 2004; Machin, 2007; Sutton Trust, 2018; Boliver, 2013; and Major and Machin, 2018, among others).

4.4 The college to no-college wage ratio declined

The evolution of the college to no-college wage ratio over cohorts is illustrated in Figure 8, using the LFS data described in Online Appendix Table A-2. Since the college cohorts that we study are observed over different age ranges in the 1993–2019 period covered by our LFS sample, we adopt the methodology of Blundell et al. (2022) to remove age effects. Specifically, we aggregate the data in cells defined by the combination of college cohort and age. Using these cells as observations, we regress the average real hourly wage of college graduates on dummies for each age and for each cohort. The three circles connected by a continuous line in Figure 8 represent the average real hourly wage of college graduates at age 45 in each cohort, net of age effects. The three squares connected by a dotted–dashed line represent the analogous average real hourly wage of non-graduates. Finally, we compute the college to no-college wage ratio in each cell and we regress it on dummies for each age and cohort. The triangles connected by a dashed line describe the evolution of this wage ratio.

Figure 8: Evolution over cohorts of the wage ratio at age 45



Notes: Evolution of real hourly wages at age 45 for college graduates and non-graduates, and of the corresponding ratio; left scale: real monetary values obtained using the 2022 edition of the OECD GDP deflator; right scale: ratio between the two wage levels. Sample: 936,135 LFS respondents surveyed between 1993 and 2019, born in 1940-1984 (see Table A-2).

The real hourly wage at age 45 of students who obtained a college degree between 1990 and 2004 increased by about 0.5 GBP with respect to those who graduated thirty years earlier (from 20.5 to 21.0 GBP). For non-graduates, the real hourly wage increased instead by more than 1.5 GBP (from 11.8 to 13.5 GBP). As a result, the wage ratio declined by about 11 percent, from 1.74 to 1.55. This finding contrasts with the evidence of a weakly increasing wage gap between *college and high-school graduates* reported in the literature for UK post-WW2 cohorts, particularly by Blundell et al. (2022).[33] The Online Appendix to Section 4.4 shows that this discrepancy is essentially due to our different definition of the comparison groups (college graduates vs non-graduates; see Section 3.2 for the rationale of this choice in our analysis). In fact there is no contrast between our evidence and the literature when we compare the Blundell et al. (2022) groups (college graduates vs high-school graduates), given that we use their same data and methodology to remove age effects. Other differences, namely the range of the LFS sample (1993–2019 instead of 1993–2016), the use of mean instead of median wages by education group, the use of age dummies instead of age polynomials in the regressions to remove age effects, or the focus on only three cohorts of 15 birth years between 1940 and 1984 instead of eight cohorts of 5 birth years between 1950 and 1989 are less, if at all, relevant.[34]

4.5 Ability and disadvantage are negatively correlated

Our measures of pre-college cognitive ability (Θ) and socioeconomic disadvantage (Λ) are *negatively* correlated in the USoc sample. This correlation (labeled η in Section 2) is reported in Table 2 for the three college cohorts, alongside its standard error. It is about -0.15, and varies over time by a statistically insignificant amount. Thus, during 1960-2004 the UK is represented by Society 2 of Figure 2.

In light of the economics of skill formation (e.g., Cunha et al., 2006, Heckman and Mosso, 2014), this is unsurprising. 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 socioeconomic status. Since early childhood interventions can in principle drive this correlation toward zero, η does not reflect a deep relation between pre-college cognitive ability

[33] Other studies include Blanden and Machin (2004), O'Leary and Sloane (2005), Walker and Zhu (2008), Devereux and Fan (2011), and Chowdry et al. (2013).

[34] A national minimum wage was introduced in the UK only in 1998, and so this legislated wage increase for low-skilled workers is unlikely to be responsible for the evidence displayed in Figure 8.

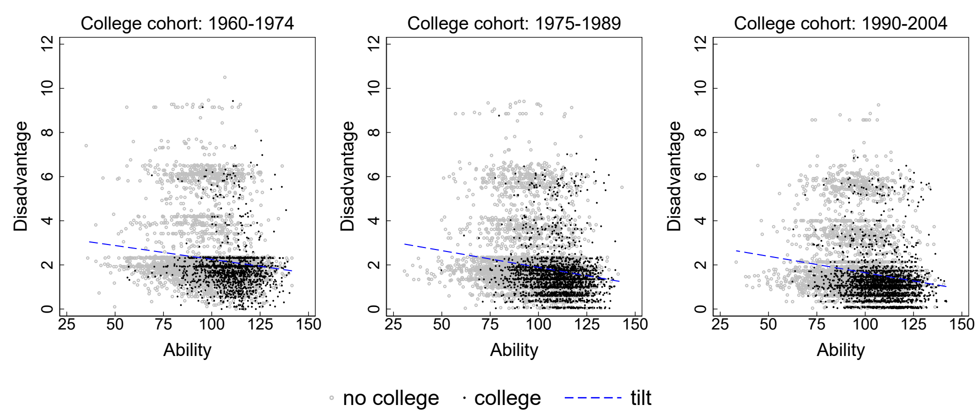
Table 2: Correlation between cognitive ability and socioeconomic disadvantage in USoc

	College cohort		
	1960-1974	1975-1989	1990-2004
$\eta = \text{Corr}(\Theta, \Lambda)$	-0.143	-0.173	-0.164
	(0.013)	(0.012)	(0.013)
N	7,103	8,329	6,743

Notes: The table reports the correlation between our measures of ability (Θ) and disadvantage (Λ). The standard error is produced via the delta method. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1).

and socioeconomic disadvantage. However, late interventions are less, if at all, effective than early ones (e.g. Cunha and Heckman, 2007), while being costly for the reasons discussed in this study. Thus, it would be a mistake to use university expansions to correct for the lack of adequate early education policies. Figure 9 displays the empirical counterpart of the scatter plot in the top row of Figure 2. The negative correlation between ability and disadvantage is reflected in the negative tilt of the underlying distribution (dashed line).

Figure 9: Empirical joint distribution of ability and disadvantage in the three college cohorts



Notes: The figure displays the empirical counterpart of the scatter plot in the top row of Figure 2, for the three USoc college cohorts. Each point is an individual in our sample. The dashed line represents the tilt of the underlying joint distribution of ability and disadvantage (see equation 16). Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1).

5 The UK higher education policy and its consequences

Next, we estimate the model presented in Section 2 to infer the policy that prevailed in the UK after the Robbins Report and to contrast it with the one that should have been implemented to achieve the Report's goals, drawing lessons for future expansions.

5.1 Identification and estimation

There are three technology parameters in the model: α (the technological skill ratio), ρ (one minus the inverse of the elasticity of substitution between school and college labor inputs), and A (TFP). Equation (13) implies that TFP does not affect the consumption utility gap $\Delta U(\mathbf{w})$, equation (3), and so parameter A can be ignored. Still, it is clear that without further assumptions, α and ρ are not separately identified in our model – only the locus given by equation (10) is identified. Using US data, Katz and Murphy (1992) and more recently Autor, Goldin, and Katz (2020) produce estimates of ρ around 0.3 and 0.4, respectively, in a partial equilibrium model where technology follows a linear trend. The implicit assumption is that the relative supply of college graduates is exogenous, and specifically does not respond to unobserved (to the econometrician) wage shocks originating from the demand side. Such an assumption cannot hold in our general equilibrium framework. However, since we do not need to identify the technology parameters separately, we can estimate the model for three alternative values of ρ in the range that is typically found in the literature: 0.3, 0.4, and 0.5. It follows that technical change is represented only by changes in α. An increase in α represents skill-biased technical change (see Section 2.2).

On the contrary, the policy parameters that control the effort cost shifts $\Omega(\Lambda) = \delta + \beta\Lambda$ and $\Gamma(\Theta) = \gamma + \tau\Theta$ (equation 14), are identified – a fact that follows immediately from equation (15) – and so we are left with parameters $(\gamma, \tau, \delta, \beta, \alpha)$ to estimate. For a given value of ρ, 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 graduates, $x(1)$, by finding the unique fixed point of the following equation, which combines equations (11) and (12),

$$x(1) = \sum_{\theta,\lambda} \omega(\theta,\lambda)\Pi\left(\frac{\theta}{\gamma+\tau\theta} - \frac{\delta+\beta\lambda}{\ln\alpha + (\rho-1)(\ln x(1) - \ln(1-x(1)))}\right)\mu(\theta,\lambda), \tag{19}$$

where $\omega(\theta, \lambda)$ denote USoc cross-sectional response weights, adding up to 1. Given the equilibrium college to no-college workforce ratio $\xi = \frac{x(1)}{1-x(1)}$, we obtain the equilibrium college to no-college wage ratio r. The equilibrium individual 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 six informative theoretical moments and their empirical analogs: the college to no-college workforce ratio, the college to no-college wage ratio, and the average ability and disadvantage of graduates and non-graduates. These six moments are the most informative for estimating our four policy parameters and the residual technology parameter because in our model it is precisely the change in higher education policy or technical progress that alter 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 the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the conditional (to educational attainment K) distributions of Θ and Λ, i.e., eight moments.

All of the moments are estimated using USoc, except for the college to no-college wage ratio which is based on the LFS, applying the appropriate weights in all cases. Denoting by

$$\widehat{T} = \left[\widehat{\xi} \;\; \widehat{r} \;\; \widehat{\mathbb{E}}(\Theta|K=1) \;\; \widehat{\mathbb{E}}(\Theta|K=0) \;\; \widehat{\mathbb{E}}(\Lambda|K=1) \;\; \widehat{\mathbb{E}}(\Lambda|K=0)\right] \tag{20}$$

the target vector of empirical quantities and by

$$T(\gamma,\tau,\delta,\beta,\alpha;\rho) = [\xi \;\; r \;\; \mathbb{E}(\Theta|K=1) \;\; \mathbb{E}(\Theta|K=0) \;\; \mathbb{E}(\Lambda|K=1) \;\; \mathbb{E}(\Lambda|K=0)] \tag{21}$$

its theoretical counterpart at equilibrium, the criterion function is

$$J(\gamma,\tau,\delta,\beta,\alpha;\rho) = (T(\gamma,\tau,\delta,\beta,\alpha;\rho) - \widehat{T})\Upsilon W \Upsilon (T(\gamma,\tau,\delta,\beta,\alpha;\rho) - \widehat{T})', \tag{22}$$

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 $J(\cdot;\rho)$ over the grid. To produce standard errors, we repeat this MD estimation 10,000 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 $J(\cdot;\rho)$ 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 higher education policy parameters $G = (\gamma,\tau,\delta,\beta)$ by Nonlinear Least Squares (NLS) from the supply-side equation (15), after replacing $\Delta \ln w(G)$ with its empirical analog, $\ln \widehat{w}(1) - \ln \widehat{w}(0)$. More details are provided in the Online Appendix to Section 5.1.

5.2 Estimates and policy anatomy

Our MD estimates of the structural parameters are reported in panel [A] of Table 3 for $\rho = 0.4$, which is the value estimated for the US by Autor, Goldin, and Katz (2020). The Online Appendix to Section 5.2 reports estimates for $\rho = 0.3$ (Katz and Murphy, 1992) and $\rho = 0.5$. The remaining panels of Table 3 compare the model-predicted values of the six targets to the empirical values computed from the data. The six target moments are matched remarkably well. Table A-9 in the Online Appendix to Section 5.2 shows that the eight untargeted moments are also well matched. The MD estimates of the policy parameters are very close to the NLS estimates, which are reported in Table A-8 of the Online Appendix to Section 5.1. This is unsurprising given that we search for a minimum around these values, but is also reassuring in consideration of (i) the different objective functions that the two estimators optimize; and (ii) the fact that our MD estimator involves five, not necessarily independent parameters while the NLS estimator involves four parameters only.

We estimate a significant increase of about 41% in technology parameter α between the 1960-1974 and 1990-2004 college cohorts, which indicates skill-biased technological progress.[35] As for the policy parameters, we find a substantial drop in γ and δ, and an increase in τ and β during this period. The economic meaning of these changes is illustrated in Figure 10. In the left panel, the continuous line indicates that, as expected, effort cost shift $\Gamma(\theta) = \gamma + \tau\theta$ was declining with ability in the 1960-1974 cohort. 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. Thus, the college expansion was non-meritocratic. In the right panel, the continuous line indicates that, again as expected, the effort cost shift $\Omega(\lambda) = \delta + \beta\lambda$ was increasing with disadvantage in the 1960-1974 cohort. The dashed line shows that the expansion policy actually resulted in an *increased* cost of attending college that was more pronounced for less advantaged students. Thus, the college expansion was also non-progressive. These conclusions are consistent with the descriptive evidence presented in Section 4.1.

The consequences for the sorting process into higher education are illustrated in Figure 11, which shows empirical isoprobability curves (constructed using the estimates in Table 3)

[35] Our estimates of α are all below 1, which implies $a(0) > a(1)$ in the production function, equation (8). This inequality does *not* imply that non-graduates are more productive than graduates because marginal productivity depends on $a(k)$ but also, inversely, on the fraction of the workforce in education group k.

Table 3: Minimum-distance estimates of model parameters for $\rho = 0.4$

[A] Parameter estimates

	College cohort 1960-1974	1975-1989	1990-2004
γ	6.169	5.366	3.549
	(0.027)	(0.024)	(0.019)
τ	−3.456	−2.953	−1.777
	(0.032)	(0.026)	(0.023)
δ	0.073	0.065	0.071
	(0.003)	(0.003)	(0.002)
β	0.021	0.015	0.028
	(0.003)	(0.002)	(0.003)
α	0.706	0.829	0.998
	(0.013)	(0.013)	(0.016)
N	7,103	8,329	6,743

[B] Labor market targets

	College cohort 1960-1974	1975-1989	1990-2004
1. College to no-college workforce ratio, ξ			
model	0.224	0.328	0.483
	(0.007)	(0.009)	(0.013)
data	0.224	0.328	0.483
	(0.007)	(0.009)	(0.014)
2. College to no-college earnings ratio, r			
model	1.735	1.617	1.545
	(0.007)	(0.005)	(0.006)
data	1.736	1.617	1.545
	(n/a)	(n/a)	(n/a)

[C] Ability targets

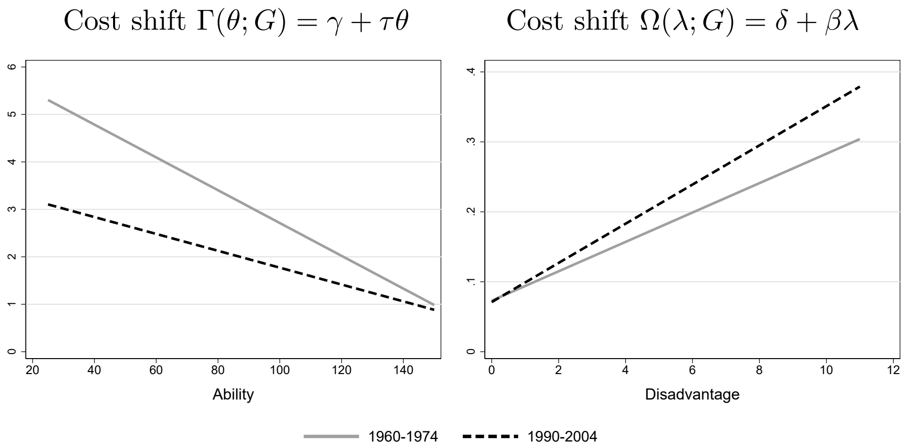
	College cohort 1960-1974	1975-1989	1990-2004
3. Graduates' ability, $\mathbb{E}(\Theta\|K=1)$			
model	110.3	108.9	108.2
	(0.5)	(0.5)	(0.5)
data	110.3	109.0	108.2
	(0.4)	(0.3)	(0.3)
4. Non-graduates' ability, $\mathbb{E}(\Theta\|K=0)$			
model	97.7	97.1	96.0
	(0.2)	(0.2)	(0.3)
data	97.7	97.0	96.0
	(0.2)	(0.2)	(0.2)

[D] Disadvantage targets

	College cohort 1960-1974	1975-1989	1990-2004
5. Graduates' disadvantage, $\mathbb{E}(\Lambda\|K=1)$			
model	1.86	1.57	1.22
	(0.03)	(0.03)	(0.03)
data	1.85	1.58	1.21
	(0.03)	(0.03)	(0.02)
6. Non-graduates' disadvantage, $\mathbb{E}(\Lambda\|K=0)$			
model	2.34	1.99	1.85
	(0.02)	(0.02)	(0.02)
data	2.34	1.99	1.86
	(0.02)	(0.02)	(0.02)

Notes: The table reports the mean and standard deviation of minimum-distance (MD) estimates of model parameters over 10,000 bootstrap samples, setting $\rho = 0.4$, and of model-predicted vs empirical values of the six targets. The MD criterion function is given by equation (22), and the weighting matrix is the identity matrix. The Online Appendix to Section 5.1 provides more computational details. The ability score is expressed in units in the table but in hundreds units in the estimation, so as to reduce the order of magnitude of the estimated γ and τ. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1).

Figure 10: Estimated study effort cost shifts

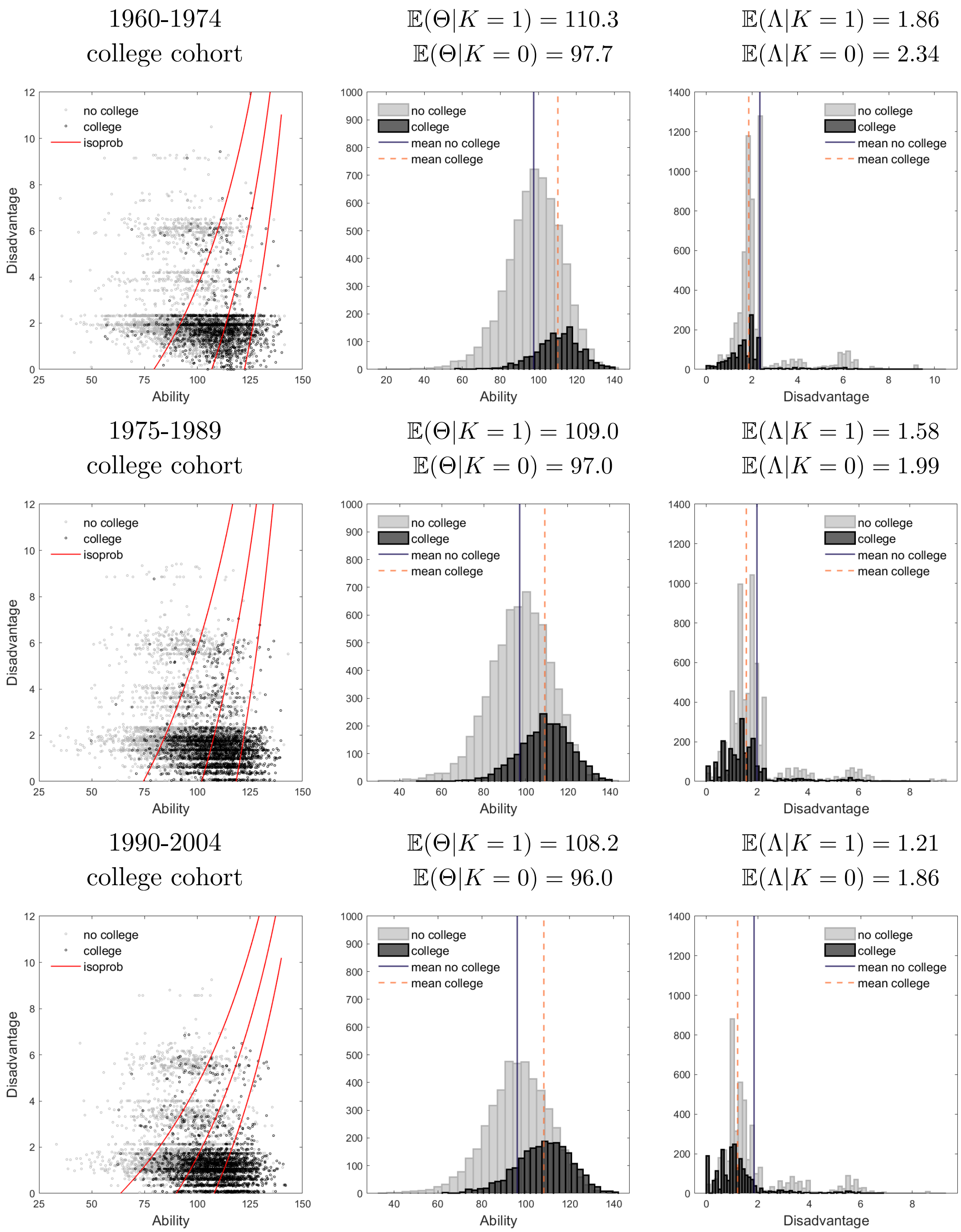


Notes: The figure shows the study effort cost shifts $\Gamma(\cdot)$ and $\Omega(\cdot)$ implied by the structural estimates of Table 3, for the 1960-1974 (continuous line) or 1990-2004 (dashed line) college cohorts, as a function of cognitive ability (left panel) or socioeconomic disadvantage (right panel).

alongside the empirical joint distribution of ability and disadvantage (left column) and the empirical conditional distributions of ability and disadvantage in the three college cohorts (middle and right columns). The isoprobability curves represented in the figure are those associated with probabilities of graduating from college equal to 0.1, 0.3, and 0.5. Moving from the 1960–1974 (top row) to the 1990–2004 (bottom row) college cohorts, we observe a clockwise rotation of the isoprobability curves, which results from a higher vertical intercept $-\frac{\Delta \ln w(G)}{\beta}\left(\pi + \frac{\delta}{\Delta \ln w(G)}\right)$ and a reduced slope $\frac{\gamma \Delta \ln w(G)}{(\gamma+\tau\theta)^2 \beta}$. This change becomes particularly evident in the comparison between the 1975–1989 and the 1990–2004 cohorts. Considering the estimates in panel [A] of Table 3, the main drivers of the increase in the intercept is the increase of β in the effort cost shift $\Omega(\lambda)$. The ensuing increase in the number of graduates reduced the wage ratio $\Delta \ln w(G)$, which further contributes to increasing the intercept. As for the slope, its decline is the result of the reduced wage gap in the numerator and the larger β in the denominator that were not compensated by a sufficiently large decline of τ in the effort cost shift $\Gamma(\theta)$.

Therefore, the scatter plots of Figure 11 indicate that the tertiary education expansion implemented in the UK during the period that we study brought into college a large number of less disadvantaged and lower ability students (i.e., individuals with low λ and low θ). The more disadvantaged and more able students in the northeastern portion of the scatter plot, actually ended up with *reduced* opportunities to access higher education.

Figure 11: Status quo in the UK and effects of actual expansion policies



Notes: The figure displays the empirical counterpart of the scatter plots and histograms in Figure 2, for the three USoc college cohorts. In the scatter plots, each point is an individual in our sample. The continuous lines are the 10%, 30%, and 50% isoprobability lines (see equation 15). Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1).

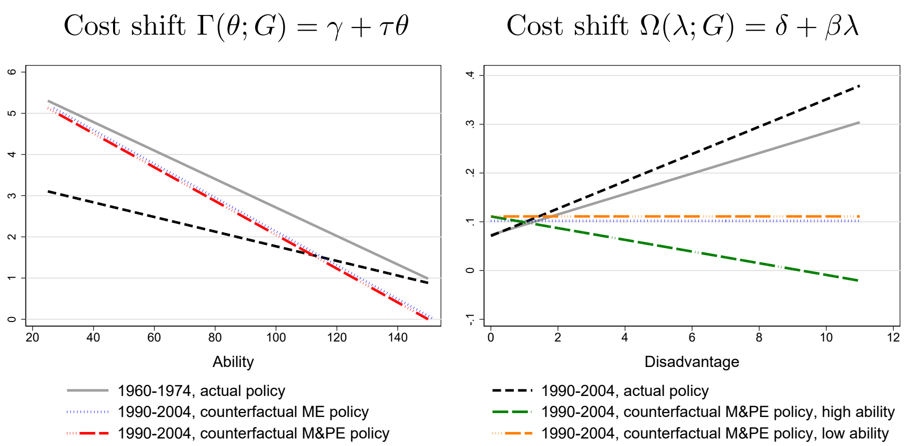
5.3 Counterfactual expansion policies

Alternative policies with a stronger meritocratic component, instead, could have reached the "reserves of untapped ability [...] in the poorer sections of the community" (p. 53) that were a central concern in the Robbins Report, even allowing for an increase of the college to no-college workforce ratio equal to the observed one (from $\xi = 0.22$ to $\xi = 0.48$). This claim is illustrated by two counterfactual policy experiments. The first one is a strongly ME policy that would have reduced τ towards larger, negative values and reduced β to essentially zero, such as conditioning college admission to passing a cognitive ability test. The second one is a variant of this counterfactual strongly ME policy that reduces parameter δ in proportion to a student's disadvantage, provided that the student is above average in terms of ability, which we label as the "meritocratic *and* progressive expansion (M&PE) policy". An example of M&PE policy is offering college admission to disadvantaged students who perform sufficiently well in a cognitive ability test, similar to the PACE policy in Chile (Tincani et al., 2022).

Figure 12 reproduces Figure 10 with added lines that describe the implications of these two counterfactual policies. The left panel shows that ME and M&PE would have had the same effect on the cost shift $\Gamma(\theta; G) = \gamma + \tau\theta$, described by the switch from the continuous line to the overlapping dotted line and two-dash-and-dotted line. The result is a cost reduction that is more pronounced for higher ability students. In the right panel, the effort cost shift $\Omega(\lambda; G) = \delta + \beta\lambda$ becomes flat for the counterfactual ME policy (dotted line) because this policy makes effort cost independent of disadvantage. The same happens for low-ability students in the case of the ME&P policy (three-dash-and-dotted line), for the same reason. On the contrary, under this ME&P policy, $\Omega(\lambda; G)$ declines with disadvantage for high-ability students (four-dash-and-dotted line), indicating that for them disadvantage becomes an advantage.

Figure 13 illustrates the consequences for the sorting process into university education. In the case of the ME policy experiment (top row), isoprobability curves would have become nearly vertical. Relative to the observed status quo (1960-1974 college cohort, top row of Figure 11), the average cognitive ability of college graduates would have *increased* by about one point, instead of decreasing by two points. Their average disadvantage would have been still lower than in the 1960s – which partly reflects the declining average disadvantage of the British population during this period depicted in the right panel of Figure 7 – but higher than the actual average in the 1990s. Therefore, this counterfactual strongly ME

Figure 12: Actual and counterfactual study effort cost shifts



Notes: The figure reproduces Figure 10, adding lines that describe the implications of the two counterfactual policy experiments described in the text. Specifically, the figure shows the study effort cost shifts $\Gamma(\cdot)$ and $\Omega(\cdot)$ implied by different actual (for the 1960-1974 or 1990-2004 college cohorts) or counterfactual (for the 1990-2004 college cohorts) higher education policies, as a function of ability (left panel) or disadvantage (right panel). The shifts implied by the actual policies are computed using the estimated policy parameters. The shifts implied by the two counterfactual strongly meritocratic expansion (ME) and meritocratic&progressive expansion (M&PE) policies are computed using the policy parameters underlying the counterfactual experiments illustrated in Figure 13.

policy would have brought into college more high-talent disadvantaged students and fewer low-talent advantaged ones than the policy that was actually implemented.[36]

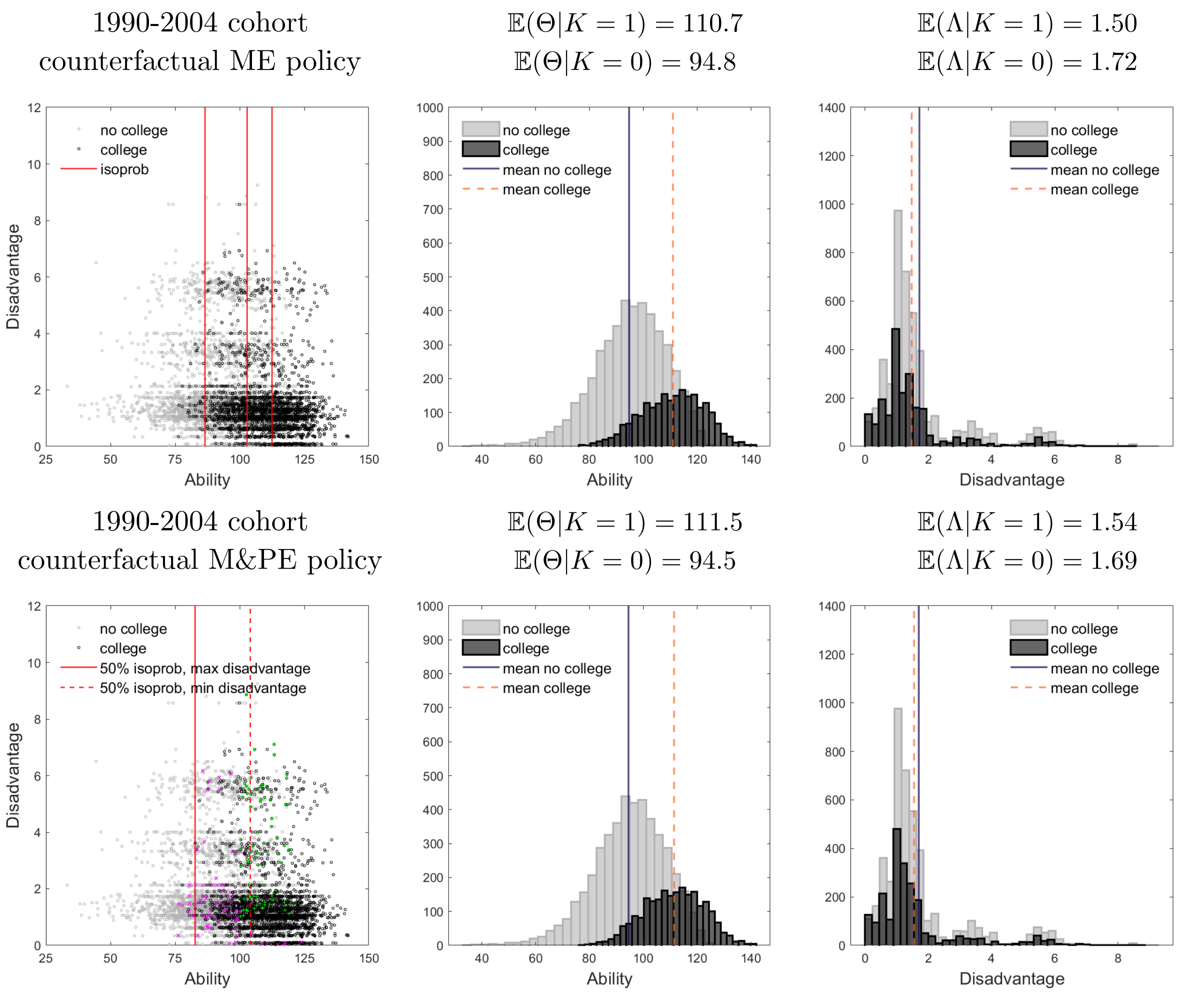
In the case of the ME&P policy experiment (bottom row of Figure 13), isoprobability curves become individual-specific and shift leftward as disadvantage increases, as indicated by the two 50% isoprobability lines represented in the figure and associated with the least (dashed line) and most (continuous line) disadvantaged students, respectively, in our sample. In the figure, we mark with a "×" students who would have graduated from college under the counterfactual ME policy but not under the counterfactual M&PE policy, and with a "+" those students for whom the opposite holds. Clearly, this M&PE policy raises enrollment and graduation barriers for lower-ability, more advantaged students while correspondingly lowering such barriers for higher-ability, more disadvantaged ones, on average.[37]

[36] Note that this counterfactual meritocratic expansion of tertiary education would have been more effective if complemented by secondary education policies aimed at improving the attainment of talented teenagers from disadvantaged families. As pointed out by Chowdry et al. (2013), "poor achievement in secondary schools is more important in explaining lower [Tertiary Education] participation rates among pupils from low socio-economic backgrounds than barriers arising at the point of entry to [Tertiary Education]" (p. 431).

[37] The 80 "×" students in the bottom row of Figure 13 have an average ability of 91 and an average disadvantage of 1.62, while the 62 "+" students in the bottom row of Figure 13 have an average ability of 109.4 and an average disadvantage of 3.04. The size of the two groups is different due to the stochastic nature of the model.

Relative to the counterfactual ME policy, M&PE selects into college students whose average ability is about one point higher (more than three points above the actual average for the 1990-2004 college cohort, and two points above the 1960-1974 cohort level) and whose average disadvantage is also (marginally) higher. This is the outcome that was envisioned by Robbins (1963), which is the exact opposite of what happened in the UK since the 1960s.

Figure 13: Effects of two counterfactual expansion policies in the UK



Notes: The figure displays the effects on the USoc college cohort 1990-2004 of two counterfactual expansion policies that would have achieved the observed graduate-to-school workforce ratio $\xi = 0.48$. First (top row), a strongly Meritocratic Expansion (ME) policy that – relative to the 1960-1974 status quo – decreases τ to -4.1 and β to 10^{-6}, adjusting γ and δ to values of 6.15 and 0.102, respectively. This policy turns isoprobability curves into essentially vertical lines. Second (bottom row), a ME policy with a Progressive Expansion (PE) component that, relative to the ME policy in the top row, sets δ to 0.111 for students whose ability is below average and to $0.111 - 0.12\lambda$ for the above-average ones. This policy makes isoprobability curves individual-specific, shifting them to the left for more disadvantaged students. Students marked with a "×" graduate from college under the counterfactual ME policy but not under the counterfactual M&PE policy; for students marked with a "+" the opposite happens. The policy that was actually implemented is represented in the bottom panel of Figure 11. In the scatter plots, each point is an individual in the 1990-2004 college cohort in our sample, and the lines are the 10%, 30%, and 50% isoprobability lines (see equation 15) for the top row, and the 50% isoprobability lines of the most disadvantaged (continuous) and least disadvantaged (dashed) students in the sample for the bottom row. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1).

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. The notion of ability as a scarce resource to be allocated across different education levels was 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 (Council of the EU, 2021). This target is set with no mention of the cognitive skills that students selected into college should have. While such neglect may be inspired by equity considerations, the key lesson conveyed by our analysis of the UK experience is that appropriate meritocratic policy options can actually reconcile graduates' quality with better college opportunities for the disadvantaged.

References

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.

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.

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.

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

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

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.

Council of the EU (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.

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.

Devereux, P. J. and W. Fan (2011): "Earnings returns to the British education expansion," Ucd centre for economic research working paper series.

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

Dodgeon, B., J. Elliott, M. Hancock, J. Johnson, S. Parsons, A. Peters, and P. Shepherd (2020): "1970 British Cohort Study Age 34 Sweep User Guide," UCL Centre for Longitudinal Studies.

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.

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.

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

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.

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 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.

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.

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

O'Leary, N. and P. Sloane (2005): "The Return to a University Education in Great Britain," *National Institute Economic Review*, 193.

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.

Riddell, S., L. Hunter Blackburn, and S. Minty (2013): "Widening Access to Higher Education: Scotland in UK Comparative Perspective," Tech. rep., Centre for Research in Education Inclusion and Diversity, University of Edinburgh.

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.

Shaked, M. and J. G. Shanthikumar (2007): *Stochastic orders*, Springer Science.

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

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 (2021): "School enrollment, tertiary (% gross)," retrieved from World Bank Open Data, https://data.worldbank.org/indicator/SE.TER.ENRR.

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.

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

ONLINE APPENDIX

Andrea Ichino, Aldo Rustichini, Giulio Zanella

May 25, 2023

Appendix to Section 2.3

We focus on vectors of labor allocation that satisfy the necessary equilibrium condition (10) and the constraint in (7). The following observation is convenient to establish existence of the equilibrium (and uniqueness in the special case that we consider in the empirical analysis).

Lemma 1 *For every $r > 0$ there is a unique pair $\boldsymbol{x}(r) = (x(0,r), x(1,r))$ and a corresponding pair $\boldsymbol{w}(r) = (w(0,r), w(1,r))$ such that*

$$x(0,r) + x(1,r) = 1, \quad \frac{Q_{x(1)}(\boldsymbol{x}(r))}{Q_{x(0)}(\boldsymbol{x}(r))} = r; \tag{A-1}$$

$$\forall k : w(k,r) = A^{\rho} Q(\boldsymbol{x}(r))^{1-\rho} a(k) x(k,r)^{\rho-1}. \tag{A-2}$$

Any pair $(\boldsymbol{x}^*, \boldsymbol{w}^*)$ which is part of an equilibrium is of the form in equations (A-1) and (A-2) for some value of the wage ratio r.

Proof. For given technological skill ratio α and equilibrium graduate-to-school labor ratio $\xi(r)$, we obtain from equation (9) that

$$\xi(r) = \left(\frac{\alpha}{r}\right)^{\frac{1}{1-\rho}}, \tag{A-3}$$

and we obtain from the aggregate constraint on labor supply that

$$x(1,r) = \frac{\xi(r)}{1+\xi(r)}; \quad x(0,r) = \frac{1}{1+\xi(r)}. \tag{A-4}$$

Replacing (A-4) into (9) and using (A-3) yields the skilled and unskilled wages as functions of r, at the demand of the firm. In particular:

$$w(0,r) = Aa(0)^{\frac{1}{\rho}} \left(1 + \alpha^{\frac{1}{1-\rho}} r^{\frac{\rho}{\rho-1}}\right)^{\frac{1-\rho}{\rho}}. \tag{A-5}$$

We can replace equation (A-5) into (3) to express the difference in utility of consumption between a graduate and a non-graduate worker as

$$\Delta U(\mathbf{w}) = w(0,r)^{1-\sigma_c}\left(\frac{r^{1-\sigma_c}-1}{1-\sigma_c}\right). \tag{A-6}$$

Using equations (A-3) and (A-4), we see that the demand for skilled labor is a decreasing function of r. If we consider the supply of skilled labor, it is clear from equation (5) that effort, and thus supply of skilled labor, is increasing in ΔU, for every pair of individuals characteristics (θ, λ), and any value of σ_s. We can consider ΔU as a function of r, using the expression in (A-6), so that the supply of skilled labor is in turn a function of r, written as $\Delta U(r)$. When $\sigma_c = 1$ it is immediate that $\Delta U(r) = \ln r$, increasing in r. Thus the equilibrium exists and is unique in this case. ■

Appendix to Section 2.4

The behavior of mean ability in the population of college graduates and so the existence of reachable ability depends on the sign of the derivative of the function D defined as:

$$D(\theta) \equiv \theta\left(\frac{1}{\gamma+\tau\theta} - \frac{\beta\eta\sigma_\Lambda}{\Delta\ln w(G)\sigma_\Theta}\right). \tag{A-7}$$

To appreciate the role played by this function, consider the simple case in which $\tau = 0$. In this case, the expression in parentheses on the RHS of equation (A-7) is independent of Θ and (when $\beta > 0$) is positive or negative depending on whether the slope of the isoprobability curves (which in this case is constant and is given by $\frac{\Delta\ln w(G)}{\gamma\beta}$) is larger or smaller than the tilt of the density (which is given by $\eta\frac{\sigma_\Lambda}{\sigma_\Theta}$). When larger, D is an increasing function, and then part (ii) of the following proposition states that the probability of attaining a college degree, conditional on ability θ, is increasing in θ.

Proposition 1 *At the equilibrium:*

(i) The probability of attaining a college degree conditional on θ is

$$P(K=1|\theta) = \mathbb{E}_{\phi_\epsilon}\pi\left(D(\theta) - \frac{\beta}{\Delta\ln w(G)}\epsilon + \frac{\beta\eta\sigma_\Lambda}{\Delta\ln w(G)\sigma_\Theta}m_\Theta - \frac{\delta}{\Delta\ln w(G)}\right), \tag{A-8}$$

with ϵ a normal random variable, independent of Θ; its density ϕ_ϵ has parameters $(m_\epsilon, \sigma_\epsilon^2) = (m_\Lambda, (1-\eta^2)\sigma_\Lambda^2)$.

(ii) *If π is any increasing function from $\mathbb{R}$ to $[0,1]$, then for any θ_1, θ_2:*

$$P(K=1|\theta_2) \geq P(K=1|\theta_1) \textit{ if and only if } D(\theta_2) \geq D(\theta_1).$$

(iii) *If π is increasing and D is increasing over $\mathbf{\Theta}$, for any increasing function g on $\mathbf{\Theta}$:*

$$\mathbb{E}(g|K=1) \geq \mathbb{E}(g|K=0), \tag{A-9}$$

with a strict inequality if g is strictly increasing. In particular, when g is the identity function, equation (A-9) states that the mean ability among college graduates is higher than among school graduates.

(iv) *The mean ability of college graduates has the selection equation form in (A-12).*

Proof. For part (i), consider the linear transform of Λ that is normal, uncorrelated with, and hence independent, from θ:

$$\epsilon \equiv \Lambda - \eta \frac{\sigma_\Lambda}{\sigma_\Theta}(\Theta - m_{\Theta)}. \tag{A-10}$$

Let $\phi_{\Theta\epsilon}$ denote the density of the joint distribution of (Θ, ϵ), and denote by ϕ_Θ and ϕ_ϵ its marginal densities. ϕ_ϵ is a normal density with parameters $(m_\epsilon, \sigma_\epsilon^2) = (m_\Lambda, (1-\eta^2)\sigma_\Lambda^2)$. Expressing the probability of obtaining a college degree as a function of (θ, ϵ) yields (A-8).

Part (ii) follows from equation (A-8) and the assumption that π is increasing.

We now consider Part (iii). First recall the definition of likelihood ratio order (e.g., Shaked and Shanthikumar, 2007, definition 1.C.1):

Definition 2 *Given two densities f_1 and f_0 on Θ, we say that f_1 is larger than f_0 in the likelihood ratio order if the function $\theta \to \frac{f_1(\theta)}{f_0(\theta)}$ is increasing.*

Take f_i in Definition 2 to be $P(\cdot|K=i)$ for $i=0,1$. To verify the condition in this definition, we consider

$$\begin{aligned} \frac{P(\theta|K=1)}{P(\theta|K=0)} &= \frac{P(K=0)}{P(K=1)}\frac{P(\theta, K=1)}{P(\theta, K=0)} = \frac{P(K=0)}{P(K=1)}\frac{P(K=1|\theta)}{P(K=0|\theta)} \\ &= \frac{P(K=0)}{P(K=1)}\left(\frac{P(K=1|\theta)}{1-P(K=1|\theta)}\right). \end{aligned}$$

Therefore, by part (ii) above, if $D(\theta)$ is increasing then function $\theta \to P(K=1|\theta)$ is increasing and conditional probability $P(\cdot|K=1)$ is larger than $P(\cdot|K=0)$ in the likelihood ratio order, by definition of this order. The conclusion then follows from the fact that the likelihood ratio order implies the stochastic order (Shaked and Shanthikumar (2007), Theorem 1.C.1), and from well-known properties of the stochastic order.

To establish part (iv), define

$$(P\phi)(\theta) \equiv \frac{\phi(\theta; m_\Theta, \sigma_\theta^2)P(K=1|\theta)}{\int_{\mathbb{R}} \phi(\tau; m_\Theta, \sigma_\theta^2)P(K=1|\tau)d\tau} \tag{A-11}$$

and the moment-generating function $M_{P\phi}(t) \equiv \int_{\mathbb{R}}(P\phi)(\theta)e^{t\theta}d\theta$. The mean ability in the population in college is given by $\mathbb{E}(\Theta|K=1) = \frac{d}{dt}M_{f\phi}(t)|_{t=0}$, which we can compute:

$$\mathbb{E}(\Theta|K=1) = m_\Theta + \sigma_\Theta^2 \int_{\mathbb{R}} \left(\frac{\phi(z;0,1)P'(m_\Theta + \sigma_\Theta z|K=1)}{(\int_{\mathbb{R}} \phi(x;0,1)P(m_\Theta + \sigma_\Theta x|K=1)dx)} dz \right). \tag{A-12}$$

■

Proposition 1 holds for any increasing function π. Thus, this result is an extension of the standard selection problem in the Roy (1951) model, when selection is determined by a function of Θ between 0 and 1 described in (A-8) rather than by passing a threshold. In fact equation (A-12) is a general form of the standard selection equation, which is its special case when the function P is the indicator function of a half line. Following the same steps, a symmetric result can be derived that characterizes $\mathbb{E}(\Lambda|K=1)$.

Appendix to Section 2.5

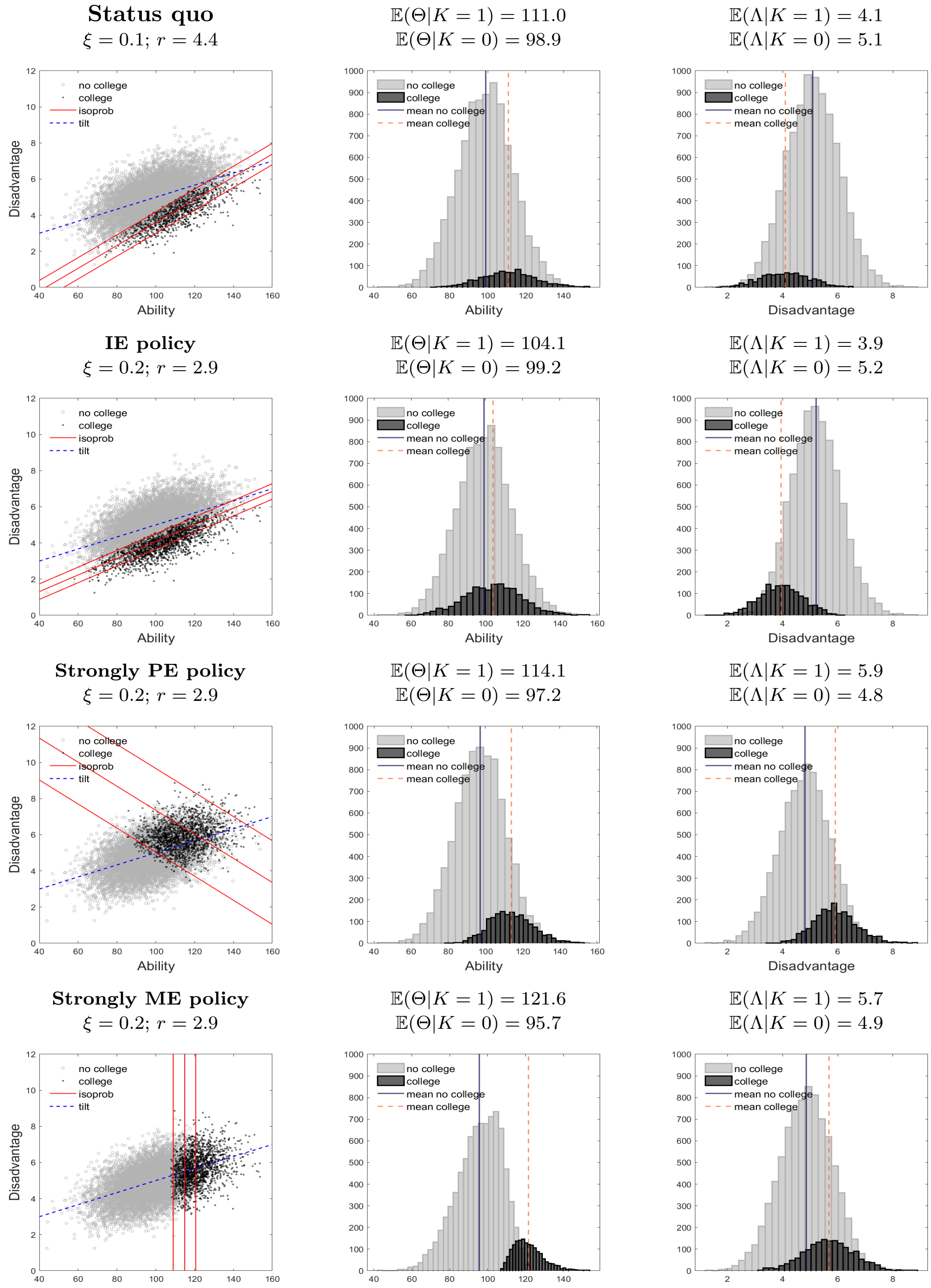
Figure A-1 describes a third type of society, characterized by $\eta > 0$ (like in Society 1) and $\psi(\cdot, G) > 0$ (like in Society 2), and the effects that alternative expansion policies would have in this society. In contrast to Society 1 described in Section 2.5 of the paper, the average ability of the population in college is now higher than outside college. There are still high ability students from disadvantaged families who are outside college, but fewer than in Society 1. This fact can be appreciated by drawing in each status-quo scatter plot (first row of the figure) horizontal and vertical lines at, say, $\Lambda = 8$ and $\Theta = 120$, and noting how many observations fall in the resulting northeastern quadrants. Thus there is reachable ability also in this third society but, again, less than in Society 1. This third is type society is meant to illustrate precisely this fact, i.e., that a positive correlation η between ability Θ and disadvantage Λ is necessary but not sufficient for large "reserves of untapped ability [...] in the poorer sections of the community" (Robbins, 1963).

As the figure illustrates, in this third society the IE policy, implemented by decreasing the intercept δ in effort cost shift $\Omega(\Lambda)$ as described in second row of the figure, reduces the mean ability of graduates while not affecting their average social background much. As for Society 1, the reason is the equilibrium response of the wage ratio, which drops from $r = 4.4$ to $r = 2.9$, with a flattening effects on the isoprobability lines.

The progressive expansion (PE) policy turns a disadvantaged socioeconomic background into an advantage in college access by inverting the sign of slope β in effort cost shift $\Omega(\Lambda) = \delta + \beta\Lambda$ (i.e., the isoprobability lines become negatively sloped as described in the third row of the figure). Like in Society 1 or 2, this policy induces a large increase in the incidence of graduates with a disadvantaged background also in this third society. However, in this case the effect on their average ability is not as large as in Society 1 although it remains positive.

Finally, the effect of the strongly meritocratic expansion (ME) policy mix, illustrated in the bottom row of the figure, changes the policy parameters in effort cost shifts $\Omega(\Lambda) = \delta+\beta\Lambda$ and in effort cost shift $\Gamma(\Theta) = \gamma+\tau\Theta$ (Eq. 14) so that the isoprobability lines become vertical. In this case, only students whose ability is above a certain threshold experience an increase in graduation probability. Like in Society 1 or 2, this strongly ME raises the incidence of high-ability *and* high-disadvantage individuals.

Figure A-1: Status quo in Society 3 ($\eta > 0, \psi > 0$) and effects of three expansion policies: Indiscriminate Expansion (IE), Progressive Expansion (PE), Meritocratic Expansion (ME).



Notes: The scatter-plots in the left column illustrate the joint distribution of ability and disadvantage for school and college graduates at equilibrium. The continuous straight lines are the isoprobability curves at values 90%, 50% and 10%, at equilibrium. The dashed lines describe values satisfying equation (16). The histograms in the middle and right columns of panels illustrate the associated marginal distributions. The data consist of a simulated population of 10,000 individuals with type (θ, λ) drawn (for mere illustrative purposes) from a jointly normal distribution ($m_\Theta = 100$; $\sigma_\Theta = 15$; $m_\Lambda = 5$; $\sigma_\Lambda = 1$; $\text{corr}(\Theta, \Lambda) = \eta = 0.5$). In the first row (status quo), the policy parameters are set to generate $\xi = 0.1$: $\gamma = 31$, $\tau = 0$ (so isoprobability curves are straight lines), $\delta = 2$, $\beta = 1$. The technology parameters are $\alpha = 1.1$ and $\rho = 0.4$. For each policy experiment in the other rows, the parameters are set so as to double the college to no-college labor ratio. The wage ratio adjusts to equilibrium. IE policy: $\delta = 0$. Strongly PE policy: $\beta = -0.18$, $\gamma = 86$. Strongly ME policy: $\tau = -8$, $\beta = 10^{-6}$, $\gamma = 30.1$, $\delta = 5.3$.

Appendix to Section 3.1

UK Biobank

In order to corroborate some of the evidence produced using USoc, we also use data from the UK Biobank (UKB). Sample size is considerably larger than USoc, but the UKB is not a random sample of the UK population because subjects are adult volunteers who are older and more educated than average. Like USoc, the UKB contains information on educational attainment and ability. Starting from 502,365 UKB subjects who did not later (as of 4/25/2023) withdraw from the survey, we retain white respondents born in the UK (434,086) between 1940 and 1969 (417,205), with non-missing information on education (411,644). Information on cognitive ability is missing for 199,034 observations. Descriptive statistics are reported in Table A-1 for the four variables that can be directly compared with USoc. This table suggests that also in the UKB the ability measure is missing quasi at random. Our final UKB sample consists of 212,624 observations with non-missing ability score.

Table A-1: The UK Biobank sample

	White UK born in 1940-1984					White UK born in 1940-1984 with non-missing ability score				
	N	mean	sd	min	max	N	mean	sd	min	max
Age	411,644	56.41	7.78	39	70	212,624	56.56	7.78	39	70
Female	411,644	0.54	0.50	0	1	212,624	0.54	0.50	0	1
Any tertiary degree	411,644	0.31	0.46	0	1	212,624	0.36	0.48	0	1
Age left school	279,275	16.64	2.21	0	35	134,521	16.82	2.22	0	35

Notes: Starting from 502,365 UK Biobank subjects who did not later (as of 4/25/2023) withdraw from the survey, we retain white respondents born in the UK (434,086) between 1940 and 1969 (417,205), with non-missing information on education (411,644). Information on cognitive ability is missing for 199,020 of these observations. The left panel of the table reports descriptive statistics for the four UKB variables that can be directly compared with USoc. The right panel reports the same statistics for our final UKB sample consisting of 212,624 observations with non-missing ability score. The similarity of the statistics in the two panels suggests that information on ability is missing quasi at random.

The 1970 British Cohort Study

The 1970 British Cohort Study (BCS70) is a longitudinal study following the lives of 17,198 individuals born in England, Scotland, Wales, and Northern Ireland in a single week of April 1970. The BCS70 is representative of the British population at each respective age at which participants are observed. After the initial birth survey, conducted in 1970, data were collected in 10 additional surveys of the same participants. The last data collection was carried out in 2021, when the cohort members were 51 years old. Each survey monitored the cohort members' health, education, and economic outcomes. The data can be downloaded using the UK Data Service at this URL `https://ukdataservice.ac.uk`. For additional details see (Dodgeon et al., 2020).

We use the second wave (1975), third wave (1980), and seventh wave (2004), where cohort members were 5, 10, and 34 years old, respectively. The respective usable sample sizes for these waves are 2,699, 3,723, and 5,585. In these waves, cohort members completed tests assessing their verbal and logical/mathematical skills.

University Statistical Record

The University Statistical Record (USR) contains administrative information on the universe of students enrolled at UK universities between 1972 and 1993. USR was initiated following the Robbins Report, which had stressed the need for better data for the proper design of higher education policies. It was subsequently discontinued and replaced by the Higher Education Statistics Agency (HESA) in 1993. Unfortunately, pre-1993 USR information was not merged into HESA. Out of the initial 8,103,977 person/year records of students enrolled in a UK higher education institution in this period, we keep the 6,889,425 records of white individuals born in the UK, so as to match the final USoc and UKB samples. These records correspond, after some minor data cleaning, to 1,523,192 students born during 1948-1976.

Labour Force Survey

Our UK Labour Force Survey (LFS) sample is 1993:Q1–2019:Q4. The LFS is a quarterly survey of about 100,000 adults who, after applying the appropriate weights, are representative of the UK population in terms of individual characteristics and earnings. Respondents are asked about earnings during the first and fifth quarters in the survey. We discard those with missing information on age, gender, education, earnings, and hours worked, missing or zero weights for earnings and personal characteristics, or a foreign educational attainment.

Real hourly wages are constructed for each respondent as the ratio between the weekly wage in the main job and the actual weekly hours. Nominal values are deflated using the 2022

Table A-2: The UK Labour Force sample (1993–2019)

	UK employees born in 1940-1984				
	N	mean	sd	min	max
Age	936,135	41.35	11.41	16	79
Female	936,135	0.50	0.50	0	1
Any tertiary degree	936,135	0.25	0.43	0	1
Real hourly wage, college graduates	213,632	19.53	12.87	0.80	149.00
Real hourly wage, non-graduates	722,503	11.96	8.61	0.80	148.89

Notes: Starting from the UK Labour Force Survey (LFS) 1993:Q1–2019:Q4, we keep only the first and fifth quarters for each respondent, i.e., the instances that contain earnings. Respondents with missing information on age, gender, education, weakly earnings, and weekly hours worked, or with missing or zero weights for earnings and personal characteristics, or with a foreign education attainment are discarded. We also drop the top and bottom 0.1% outliers of the real wage distribution. Nominal values are deflated using the 2022 edition of the OECD GDP deflator. Yearly observations range between 25,000 and 50,000.

edition of the OECD GDP deflator (base year: 2015). While Blundell et al. (2022) study *median* wages by education group, the relevant variable in our model is the *average* wage, at a given age, of college graduates and non-graduates. The Online Appendix to Section 4.4 shows that the use of mean wages instead of median wages is essentially irrelevant. To neutralize the effect of outliers on average wages, we also drop the top and bottom 0.1% of the real wage distribution.

Our final sample are 936,135 subjects observed during at least one year between 1993 and 2019, with observations in each year ranging between about 25,000 and 50,000. Table A-2 presents descriptive statistics for the relevant variables. Section 4.4 of the paper explains how we use this information to measure the evolution of the college to no-college wage ratio over cohorts.

Appendix to Section 3.3

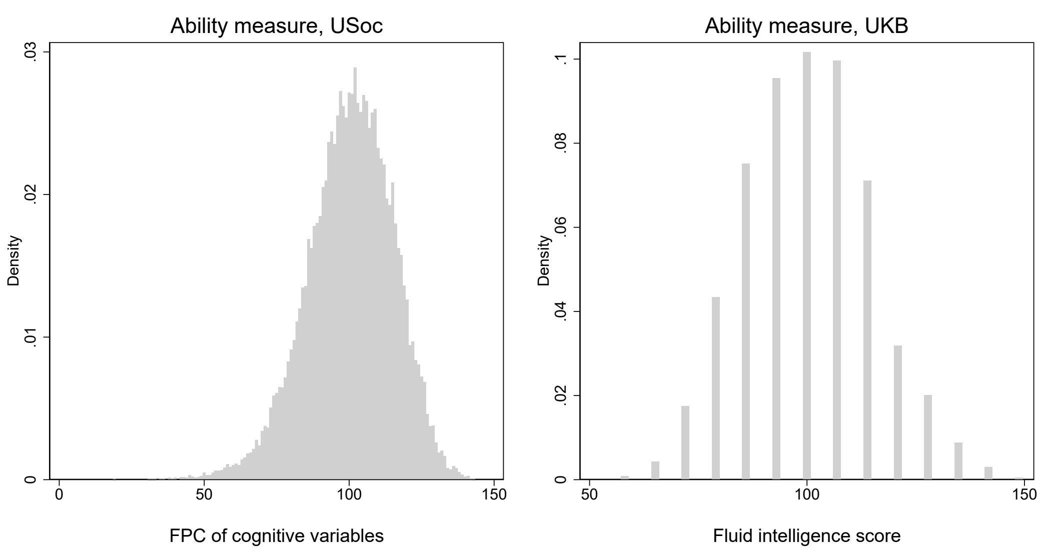
USoc and UK Biobank

Table A-3: Eigenvector of the PCA of cognitive ability measures in USoc

Immediate word recall	0.457	Help in immediate word recall	–0.011
Delayed word recall	0.449	Help in delayed word recall	0.004
Correct subtractions	0.318	Help in substractions test	–0.050
Number series	0.413	Help in number series test	–0.034
Verbal ability	0.365	Help in verbal ability test	–0.015
Numeric ability	0.423	Help in numeric ability test	–0.004
		Material aid in recall test	0.011
		Material aid in subtraction test	–0.040

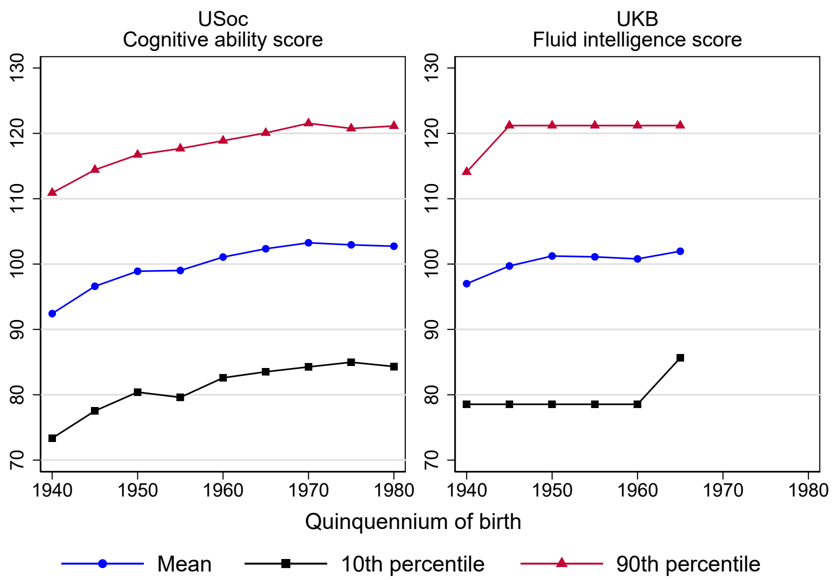
Notes: The table reports the eigenvector of the Principal Components Analysis of the 14 cognitive ability measures contained in Usoc. The First Principal Components (FPC) is the measure of ability that we use in our analysis. It has an eigenvalue of 2.55 and explains 18.2% of the data variability. The left panel of the table displays the positive values of the eigenvector terms for the fractions of correct answers in the 6 cognitive questions. The right panel, shows instead that the eigenvector values are negative for 6 out of 8 help dummies. For the the two remaining help dummies the values are positive but close to zero.

Figure A-2: Distribution of cognitive ability



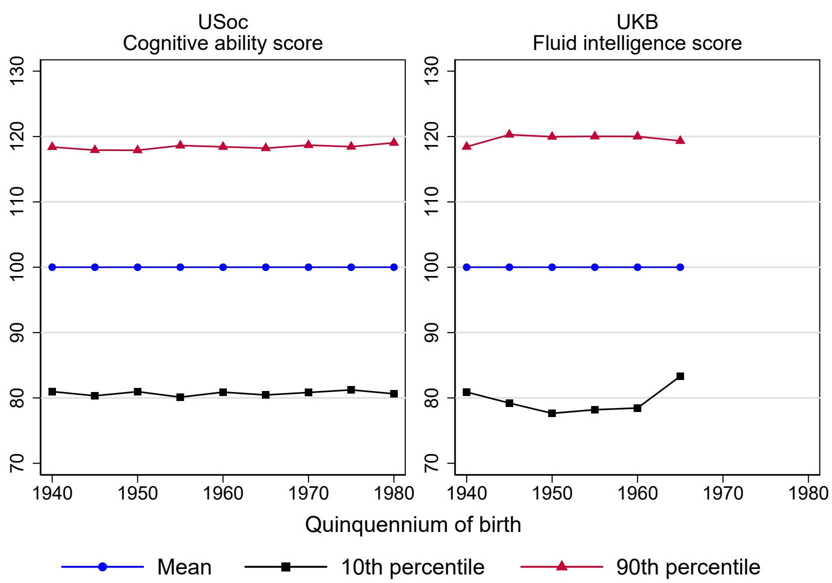
Notes: The figure illustrates the empirical distribution of our measures of ability. In the left panel, the USoc measure is the FPC of 14 cognitive ability variables. In the right panel, the UKB measure is the Fluid Intelligence Score, resulting from the sum of correct answers in 13 cognitive questions.

Figure A-3: Evolution of cognitive ability scores standardized over all birth years



Notes: The left panel of the figure displays the mean, the 10th and the 90th percentiles of the average ability score standardized over all birth years in our USoc sample of 22,175 white respondents born in the UK between 1940 and 1984, with non-missing education and ability score (see Table 1). The right panel displays the same statistics for the Average ability score (FIS) in our UKB sample of 212,659 white respondents born in the UK between 1940 and 1969, with non-missing education and ability score (see Table A-1).

Figure A-4: Evolution of ability scores standardized in each birth year



Notes: The left panel of the figure displays the mean, the 10th and the 90th percentiles of the average ability score standardized within each birth year in our USoc sample of 22,175 white respondents born in the UK between 1940 and 1984, with non-missing education and ability score (see Table 1). The right panel displays the same statistics for the Average ability score (FIS) in our UKB sample of 212,659 white respondents born in the UK between 1940 and 1969, with non-missing education and ability score (see Table A-1).

The 1970 British Cohort Study

PCA in the BCS70

Table A-4: Eigenvector of the PCA of cognitive ability measures in BCS70 age 5

Schonell Reading Test	0.220
Human Figure Drawing 1	0.516
Human Figure Drawing 2	0.523
English Picture Vocabulary Test	0.353
Complete a Profile	0.318
Copying Designs	0.431
Dummy for all tests missing	0.008

Notes: The table reports the eigenvector of the Principal Components Analysis of the 6 cognitive ability measures contained in BCS70 sweep 2 (age 5). The First Principal Components (FPC) is the measure of ability that we use in our analysis. It has an eigenvalue of 2.55 and explains 36.5% of the data variability.

Table A-5: Eigenvector of the PCA of cognitive ability measures in BCS70 age 10

Pictorial Language Comprehension	0.346
Spelling Dictation Task	0.411
Friendly Maths Test	0.486
Edinburgh Reading Test	0.496
British Ability Scales (BAS)	0.479
Dummy for all tests missing	-0.001

Notes: The table reports the eigenvector of the Principal Components Analysis of the 5 cognitive ability measures contained in BCS70 sweep 3 (age 10). The First Principal Components (FPC) is the measure of ability that we use in our analysis. It has an eigenvalue of 2.811 and explains 46.85% of the data variability.

Table A-6: Eigenvector of the PCA of cognitive ability measures in BCS70 age 34

Literacy Skills	0.707
Numeracy Skills	0.707
Dummy for all tests missing	-0.029

Notes: The table reports the eigenvector of the Principal Components Analysis of the 2 cognitive ability measures contained in BCS70 sweep 7 (age 34). The First Principal Components (FPC) is the measure of ability that we use in our analysis. It has an eigenvalue of 1.66 and explains 55.8% of the data variability.

Imputation in the BCS70

Not all BCS70 subjects have all cognitive scores at all ages and there are several codes for missing information. This appendix explains in greater detail the different imputation strategies that we follow in the panels of Figure 3 of the main text. In panel (A) or (B) no imputation of such missing information is attempted. In panel (C), the missing answers were coded as follows. First, "Not stated", "no answer", and "more than one answer" were treated as incorrect answers, and "Not applicable" and "not scorable" were coded as missing. Then, the missing answers were imputed as follows:

- If an observation had some missing answers in a specific test, but not all answers were missing, the missing answers were coded to be incorrect.
- If all answers for a test were missing, but the score for at least one test of the same age was available, the score of the missing test was set to be equal to the child's average score from the other tests.
- If all answers to all the tests were missing, the child's score in each test was set to the average score of the other children on the test, conditional on college attainment.

While creating a general cognitive score using the PCA, a dummy for having all missing in all tests was used.

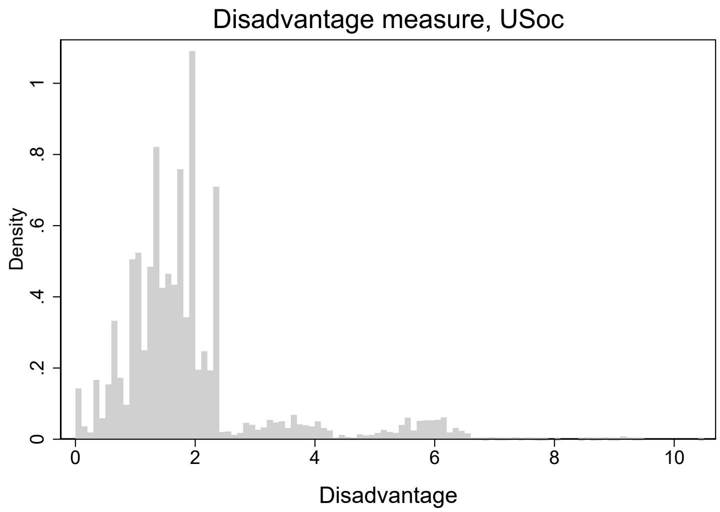
Appendix to Section 3.4

Table A-7: Eigenvector of the PCA of socioeconomic factors generating advantage in college enrollment and graduation in USoc

Father education	0.277
Mother education	0.290
Mother work	0.191
Mother dead	–0.125
Mother absent	–0.224
Father work	0.617
Father dead	–0.416
Father absent	–0.428

Notes: The table reports the eigenvector of the Principal Components Analysis of the 8 socioeconomic background variables in Usoc (referring retrospectively to when the respondent was 14 years of age) on which we base our measure of socioeconomic disadvantage. The First Principal Component (FPC) has an eigenvalue of 1.76 and explains 22% of the data variability. The table displays negative values for the variables that, as expected, reduce the FPC and increase disadvantage: whether either parent was dead or absent when the respondent was 14 years of age.

Figure A-5: Distribution of socioeconomic disadvantage



Notes: The figure illustrates the empirical distribution of our measure of disadvantage The measure is the the FPC of 8 socioeconomic variables at age 14, rescaled so that the minimum is zero.

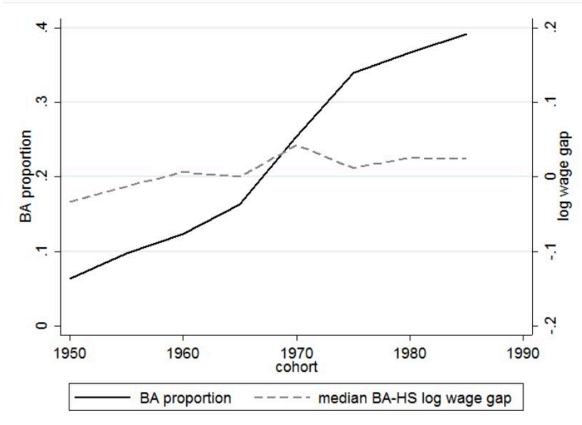
Appendix to Section 4.4

The evolution of the wage gap between college and high school graduates in the UK has been studied by Blanden and Machin (2004), O'Leary and Sloane (2005), Walker and Zhu (2008), Devereux and Fan (2011), Chowdry et al. (2013), and Blundell et al. (2022), among others. Our finding in Section 4.4 of a declining wage ratio between college graduates and less educated individuals over consecutive cohorts is apparently in contrast with the evidence of a weakly increasing gap reported in this literature, particularly by Blundell, Green, and Jin (2022) – BGJ, henceforth. In this appendix we show that the discrepancy is essentially due to the different definition of the comparison groups: college graduates vs high school graduates in BGJ, college graduates vs non-graduates in our paper.[A-1] We have explained in Section 3.2 the rationale of this choice in our analysis. In fact there is no contrast when we use BGJ's comparison groups, given that we use their same data source (LFS) and their methodology to remove age effects. Other differences between the two studies are less relevant.

Figure A-6 reproduces the right panel of Figure 4 in Appendix 5 of BGJ. The dashed line describes the evolution of the college-to-high school wage gap across 5-year birth cohorts,

[A-1] Specifically, we compare college graduates defined as individuals who have obtained a university degree or any other tertiary education diploma to all other subjects with a lower educational attainment. They compare college graduates (defined in the same way) to high-school graduates only (i.e., individuals with a secondary or some tertiary education below a university degree level, where the bottom line of secondary education is Grade C in the GCSE exam).

Figure A-6: College-to-high school wage ratio across birth cohorts in Blundell et al. (2022)



Notes: This figure a screenshot of the right panel of Figure 4 in Section 5 of the Online Appendix of Blundell et al. (2022). Their note to this panel reads: "We aggregate LFS data 1992-2016 up to the level of 5-year-birth-cohorts and age, where age is restricted to 20-59. We look at cohorts 1950-1985 only, so that each cohort appears many years in the data. ... For the right sub-figure, we regress the BA proportion" (proportion of college graduates) "on cohort dummies and an age polynomial of order 5. For the BA proportion, the cohort effects are scaled to the observed proportion for 1965 cohort at 30 year old. For the wage gap, the cohort effects are normalized to 0 for the 1965 cohort."

from 1950-54 until 1985-89. The methodology followed by the authors is to aggregate observations in cells defined by these 5-years cohorts and age in years. The difference between the log of median wages by education group in each cell is then regressed on cohort dummies and on a fifth-order polynomial in age. The dashed line plots the coefficients of the cohort dummies from this regression, normalized to zero in 1965, and suggests a weakly *increasing* pattern of the wage gap across successive cohorts. According to BGJ, the wage ratio was about 4% higher for the 1985-89 birth cohort relative to the 1950-54 cohort.

The top-left panel of Figure A-7 reproduces Figure 8 of the main tex, which shows a *decreasing* wage ratio across the three college cohorts that we consider, in contrast with BGJ. A first possible reason of this discrepancy is the fact that (for the reasons explained in Section 3.1) we compare *mean* wages by education group while BGJ compare *median* wages. The top-right panel of Figure A-7 shows that if we use median wages while sticking to all our other specifications, the wage ratio between college graduates and non-graduates exhibits a similar decrease, so using the mean or the median is actually irrelevant for the dynamics of the wage ratio across consecutive cohorts.

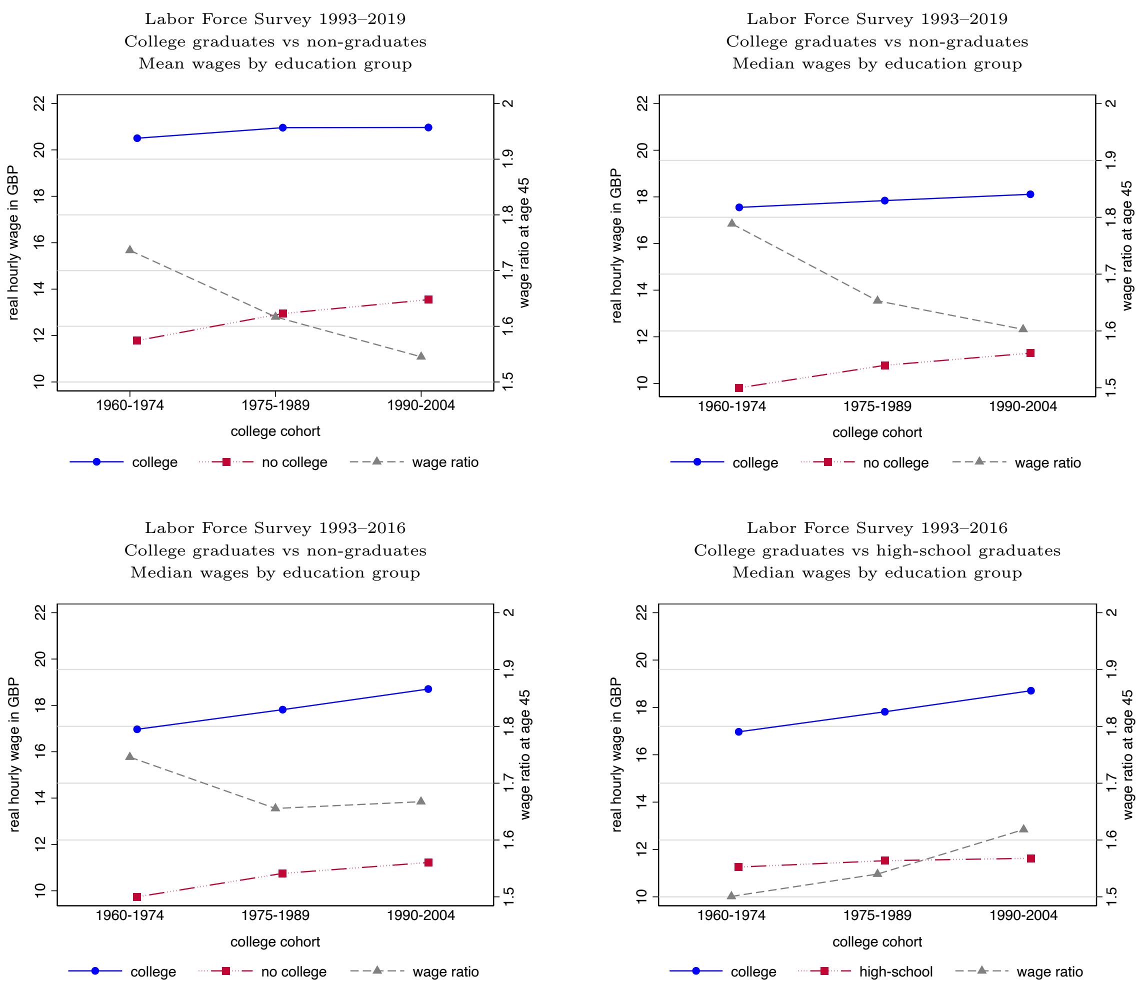
The bottom-left panel of Figure A-7, in addition to using medians, makes another step towards the BGJ specification by using 1993-2016 LFS data instead of 1993-2019. All our other specifications are preserved. As expected, this change affects mainly the wage ratio of the most recent cohort, which is now observed for fewer years, and results in a slightly flatter pattern. Finally, in the right-bottom pattern we change the comparison groups to those used by BGJ, while still using the median and the 1993-2016 sample: college graduates vs high-school graduates (see footnote A-1 for the exact definitions). Now the discrepancy between BGJ and us disappears: the wage ratio over consecutive cohorts increases as in BGJ, from 1.50 to 1.62, i.e., by about 8 percent.[A-2]

We conclude from this analysis that the discrepancy between the decreasing wage ratio that we find and the weakly increasing wage ratio found by BGJ is essentially due to the different educational groups considered in the two papers. As explained in Section 3.2, our broader definition of non-college graduates is justified by the fact that we study policies aimed at expanding university access so as to bring into higher education untapped ability from *any* less educated group that was previously excluded from college, not only from the pool of high school graduates. To quantify the difference in the definition of the comparison groups, out of the 936,135 observations in our LFS sample, 146,565 (15.7% of the total) are not high school graduates according to the definition of BGJ and so are not included in their comparison group, while they are included in ours.

We conclude from this analysis that the discrepancy between the decreasing wage ratio that we find and the weakly increasing wage ratio found by BGJ is essentially due to the different educational groups considered in the two papers. As explained in Section 3.2, our broader definition of non-college graduates is justified by the fact that we study policies aimed at expanding university access so as to bring into higher education untapped ability from *any* less educated group that was previously excluded from college, not only from the pool of high school graduates. To quantify the difference in the definition of the comparison groups, out of the 936,135 observations in our LFS sample, 146,565 (15.7% of the total) are not high school graduates according to the definition of BGJ and so are not included in their comparison group, while they are included in ours.

[A-2] We do not report here analogous figures showing that the remaining differences are practically irrelevant, namely: the consideration of only three cohorts, each one spanning 15 birth years between 1940 and 1984 (instead of BGJ's eight birth cohorts of 5 years from 1950 until 1989) and the use of age dummies (instead of BGJ's age polynomials) in the regressions that removes age effects.

Figure A-7: Wage levels and ratios for different specifications



Notes: The top-left panel reproduces Fig. 8 in the main text, based on our specifications and definitions. The top-right panel shows how this figure changes when median wages by education are used instead of mean wages. The bottom-left panel is like the top-right, except that the 1993-2016 LFS sample is used (like in BGJ) instead of the 1993-2019 LFS sample. Finally, the bottom-right panel shows how the figure in the bottom-left panel changes when education groups are defined as in BGJ: college graduates vs high school graduates, instead of college graduates vs non-graduates as in the three other panels (see footnote A-1 for the exact definitions).

Appendix to Section 5.1

It is plausible that the criterion function,

$$J(\gamma,\tau,\delta,\beta,\alpha;\rho) = (T(\gamma,\tau,\delta,\beta,\alpha;\rho) - \widehat{T})\Upsilon W \Upsilon (T(\gamma,\tau,\delta,\beta,\alpha;\rho) - \widehat{T})',$$

has local minima, and given that the grid is finite, the "wrong" starting point for the search process may yield estimates that correspond to one of them. This is particularly worrisome as there is no reference scale for policy parameters γ, τ, δ, and β – while for technology parameter α a natural reference point is 1, i.e., $a(1) = a(0)$ – and so one does not know where the grid should be centered in $\mathbb{R}^5$ in order not to get stuck into a local minimum.

We solve this problem by noting that a researcher not interested in disentangling the impact of higher education policy $G = (\gamma,\tau,\delta,\beta)$ from changing technology and socioeconomic characteristics or not interested in using the model for equilibrium policy analysis, can obtain a partial set of estimates by Nonlinear Least Squares (NLS) from the supply-side equation (15), after replacing $\Delta \ln w(G)$ with its empirical analog, $\ln \widehat{w}(1) - \ln \widehat{w}(0)$. The NLS estimates provide a guess that should be close to the actual policy parameters, i.e., the "right" starting values, even if it ignores the equilibrium effects of higher education policy. Such initial estimates are reported in Table A-8.

Our MD estimates and standard errors are then obtained as follows. Starting from the initial NLS estimates of the policy parameters and considering the reference value $\alpha = 1$ for the technology parameter, we set up a grid to locate the global minimum of criterion function $J(\gamma,\tau,\delta,\beta,\alpha;\rho)$, for a given value of ρ. Anchoring the grid search process to the NLS estimates of $G = (\gamma,\tau,\delta,\beta)$ and the natural reference value for α increases our confidence that the MD algorithm – which instead takes into account that $\Delta \ln w(G)$ depends on the parameters to be estimated – does not end up at a local minimum.

In order to mitigate the consequences of the curse of dimensionality, we design an algorithm that starts from a small grid composed by 40,500 points: 3 for each of the four policy parameters (the NLS estimate and two neighboring points, at distance 0.01 for γ and τ and distance 0.001 for δ and β) and 500 for α (from 0.01 to 5, in steps of 0.01). We then solve numerically for the model's equilibrium at each point of this grid by finding the unique fixed point of equation (19) for that particular combination of $(\gamma,\tau,\delta,\beta,\alpha)$, and we obtain a MD estimate by locating the minimum of $J(\cdot;\rho)$ over the grid. If this MD estimate hits a grid

Table A-8: Initial NLS estimates of policy parameters

	College cohort		
	1960-1974	1975-1989	1990-2004
γ	6.195	5.366	3.550
	(1.033)	(0.690)	(0.564)
τ	-3.415	-2.969	-1.756
	(0.772)	(0.511)	(0.375)
δ	0.075	0.071	0.071
	(0.018)	(0.014)	(0.019)
β	0.016	0.012	0.024
	(0.001)	(0.002)	(0.002)
N	7,103	8,329	6,743

Notes: The table reports Nonlinear Least Squares (NLS) estimates of parameters in equation (15), after replacing $\Delta \ln w(G)$ with its empirical value, i.e., $\ln w(1) - \ln w(0)$. The ability score is expressed in hundreds in the estimation, so as to reduce the order of magnitude of the estimated γ and τ. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

boundary (for example, if the estimate for γ is the minimum or the maximum in the vector of values for γ that is used to build the grid), then a point is added to enlarge that boundary and estimation is repeated with the expanded grid.

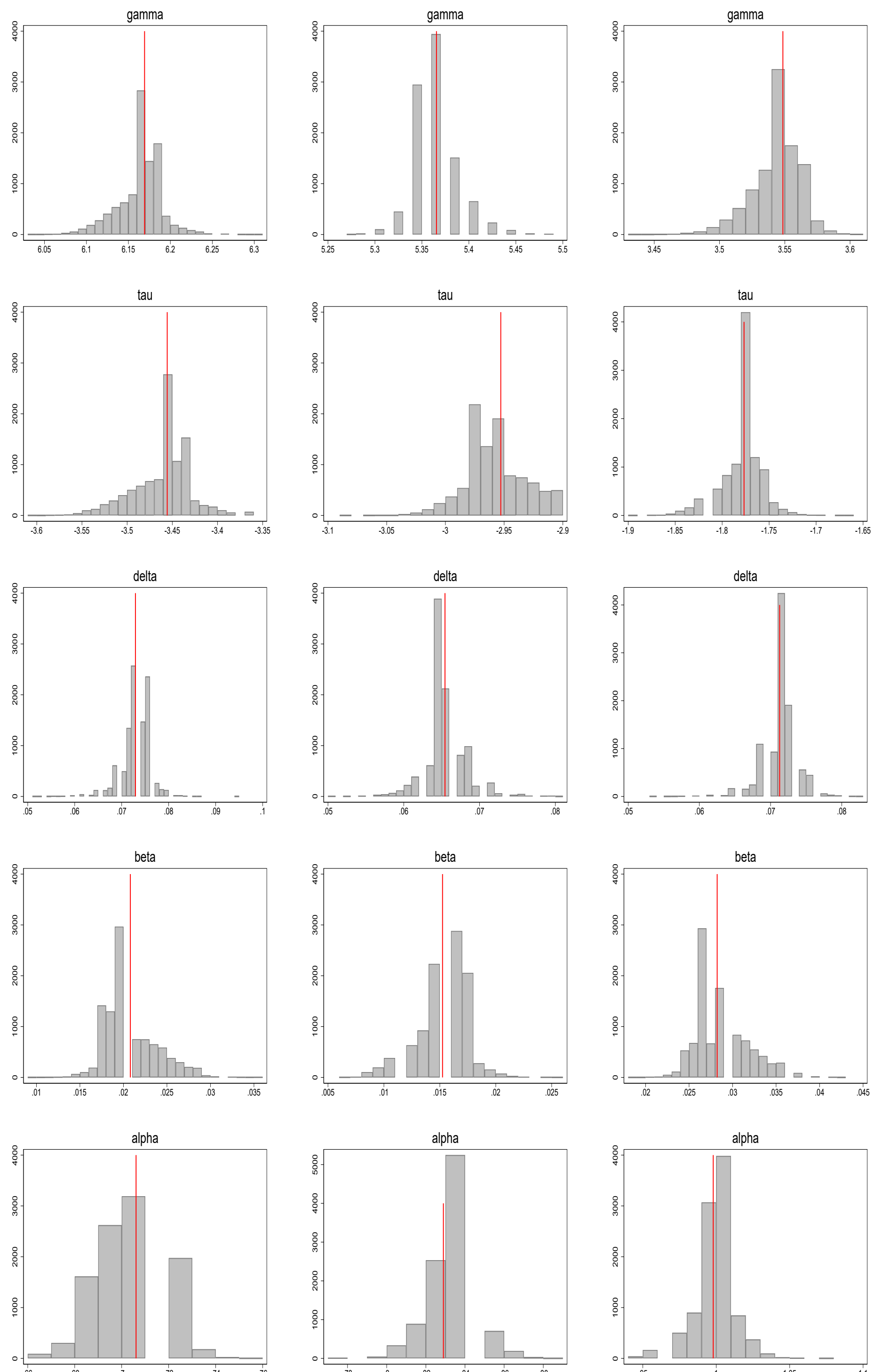
This process is iterated until the MD estimates are at an interior point of the grid. When calibrating $\rho = 0.4$, in the initial, actual sample ("one-shot" estimates) this occurs in final grids of: 96,000 points for college cohort 1960-1974 (the minimum value of the criterion function is min = 0.0000128); 720,000 points for cohort 1975-1989 (min = 0.0000213); and 96,000 points for cohort 1990-2004 (min = 0.0000140). Thus, the advantage of anchoring the MD starting values to the partial equilibrium NLS estimates is that we can greatly reduce the grid size. Despite this computational gain, we needed a further expedient in order to complete the 10,000 bootstrap replication in reasonable time for each cohort. The expedient is that the initial grid for each bootstrap sample consists of only $3^5 = 243$ points, resulting from vectors of 3 points for each parameter (the MD, one-shot estimates and 2 neighboring points); estimation is iterated according to the "no boundary estimates" rule described above, and repeated in the 10,000 bootstrap samples. The distribution of the resulting 10,000 bootstrap estimates of each parameter (conditional on $\rho = 0.4$) is illustrated in Figure A-8 for three three college cohorts. The vertical lines mark the averages that we report as our point estimates in Table 3 of the main text, and the standard deviations are our bootstrap standard errors in that table.

A crucial question is whether our MD algorithm produces estimates corresponding to a global minimum or not. To increase our confidence that it does, we inspect two- and three-dimensional sections of the criterion function over a much wider grid than the one employed in our computational algorithm. The two-dimensional sections are shown in Figure A-9 for each cohort and for $\rho = 0.4$. A panel plots the value of the log of the MD criterion as a function of a parameters, keeping the remaining 4 parameters fixed at the one-shot MD estimates. The global minimum as well as local minima are clearly visible in each panel. Note that despite the appearance of a cusp, the function is smooth around the minimum. This appearance is produced by the log scale, which is convenient but produces a large negative value at the minimum because it is very close to zero. The associated three-dimensional sections are shown in Figure A-10 for college cohort 1960-1974. Here we fix 3 parameters at the one-shot MD estimates and we plot the contour lines of the MD criterion as a function of the 10 possible combinations of the remaining 2 parameters. The minimum is marked by the intersection of the two dashed lines. It is again clear that the NLS estimates provide a guess that helps us locating the global minimum in the presence of several local minima. Our replication package can be used to produce the analogous figures for college cohorts 1975-1989 and 1990-2004 and to inspect the criterion function over any subset of $\mathbb{R}^5$.

Figure A-8: Distribution of MD estimates across 10,000 bootstrap samples, for $\rho = 0.4$

College cohort

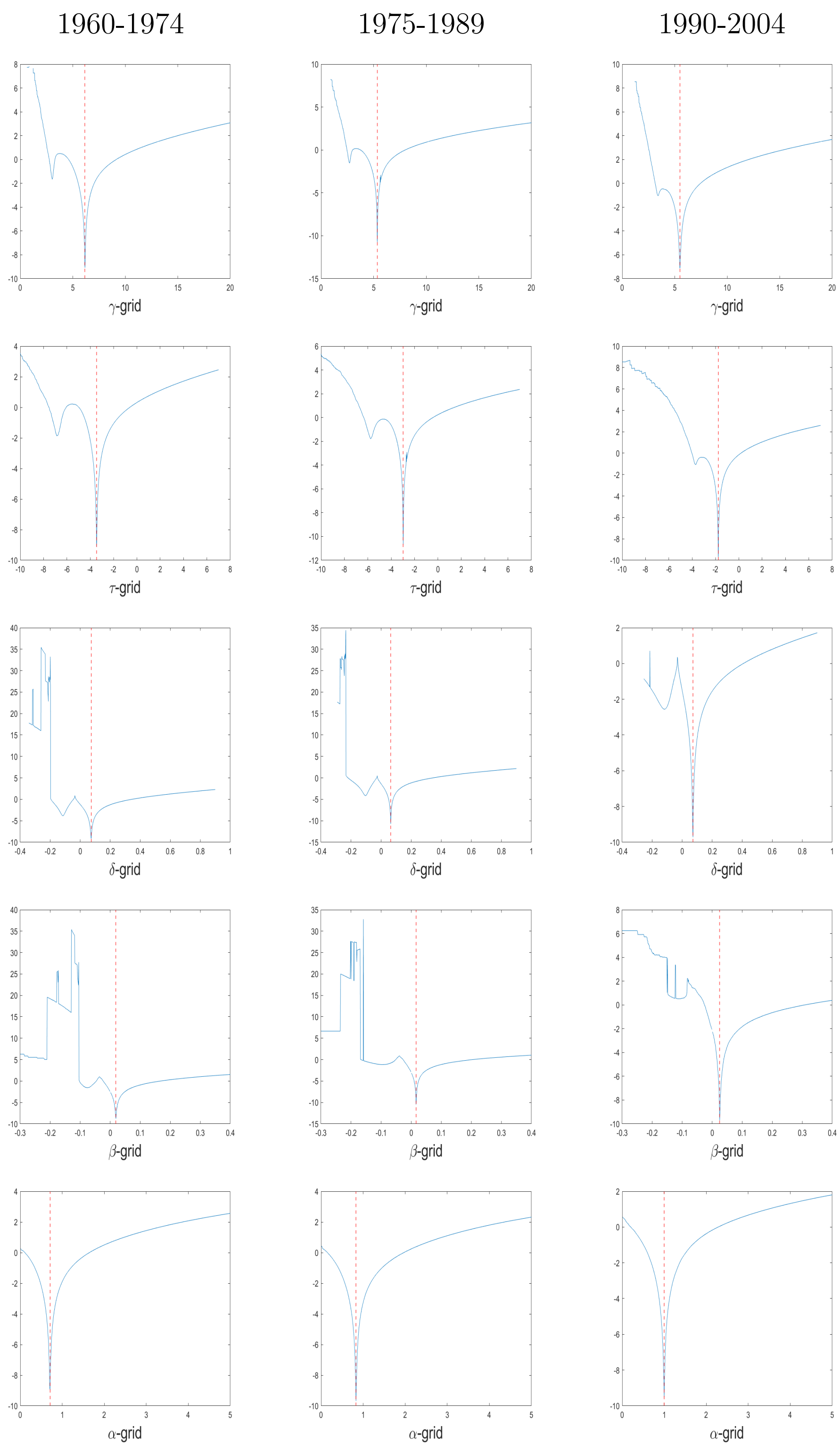
1960-1974 1975-1989 1990-2004



Notes: The figure illustrates the distribution of MD estimates of the five structural parameters of interest across 10,000 bootstrap samples, conditional on $\rho = 0.4$. The vertical line is the mean of the distribution. The point estimates and standard errors reported in Table 3 of the main text are the means and standard deviations, respectively, of these distributions.

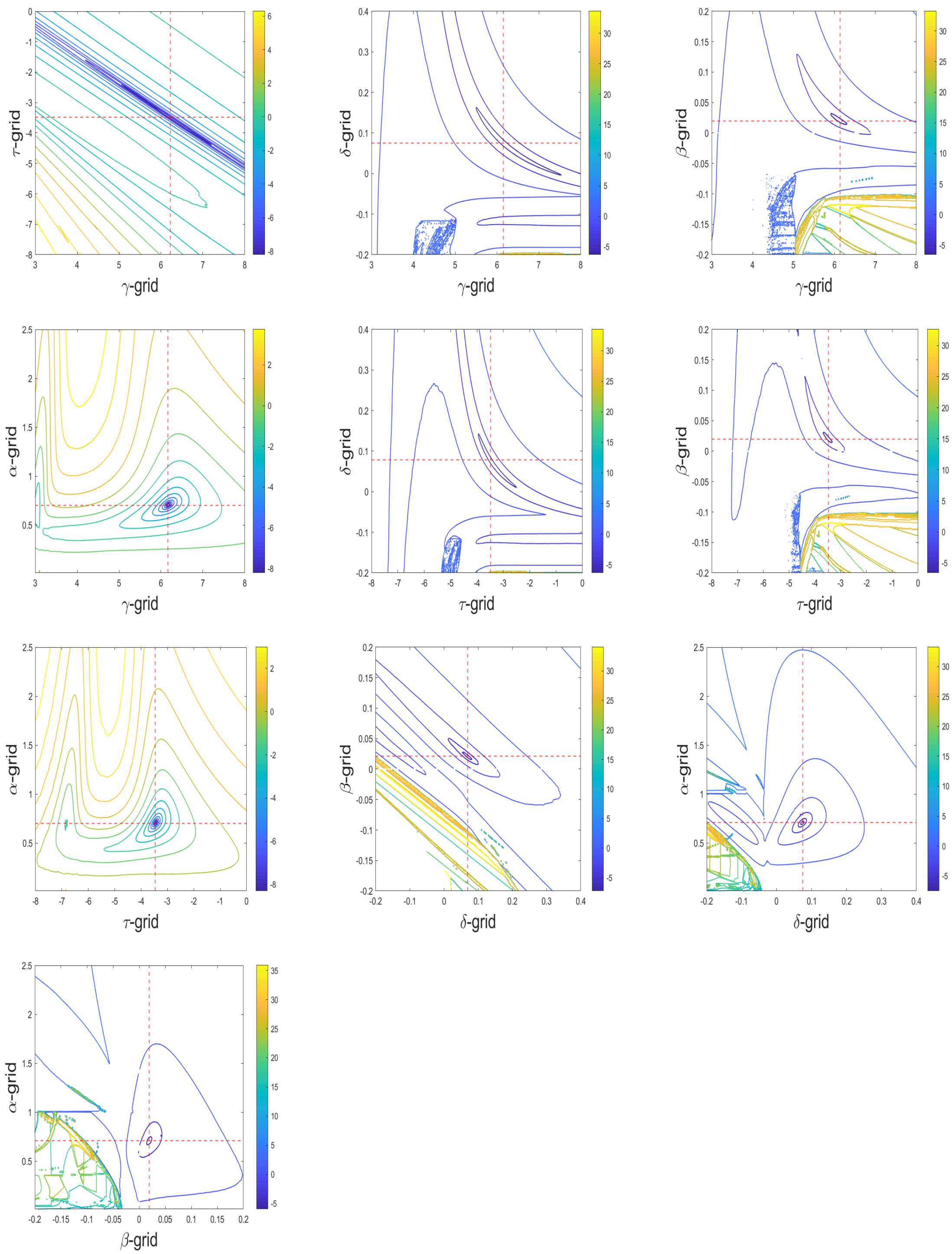
Figure A-9: 2D sections of criterion function for $\rho = 0.4$, log scale

College cohort



Notes: Each panel plots the value of the log of the MD criterion as a function of one parameter, keeping the remaining four parameters fixed at the MD estimates obtained in the actual (as opposed to bootstrap) sample. The dashed line marks the global minimum, which corresponds to our MD estimates. Local minima are clearly visible, and anchoring the grid to the initial NLS estimates of the policy parameters (see Table A-8) helps avoiding them. Note that despite the appearance of a cusp, the function is smooth around the global minimum, which takes a large negative value on the log scale because it is very close to zero.

Figure A-10: 3D sections of criterion function for $\rho = 0.4$ and cohort 1960-1974, log scale



Notes: Each panel plots the contour lines of the log of the MD criterion as a function of two parameters, keeping the remaining three parameters fixed at the MD estimates obtained in the actual (as opposed to bootstrap) sample. The possible $\binom{5}{2} = 10$ combinations are represented. The intersection of the two dashed lines marks the global minimum, which corresponds to our MD estimates. Local minima are clearly visible, and anchoring the grid to the initial NLS estimates of the policy parameters (see Table A-8) helps avoiding them.

Appendix to Section 5.2

Robustness and diagnostics of the MD estimates

This Appendix shows how well we match eight untargeted moments not used for estimation (namely, the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the conditional – to educational attainment K – distributions of Θ and Λ, Table A-9) and replicates the results in Table 3 of the main text and Table A-9 in this Appendix for the cases in which ρ is assumed to be equal to 0.3 (Table A-10 and Table A-11) or 0.5 (Table A-12 and Table A-13). Our replication package can be used to replicate also for these alternative values of ρ the visual analysis of the criterion function performed in the previous Appendix to Section 5.1.

Table A-9: Quality of match for eight untargeted moments at minimum-distance estimates of model parameters for $\rho = 0.4$

	Ability distribution			**Disadvantage distribution**		
	College cohort			College cohort		
	1960-1974	1975-1989	1990-2004	1960-1974	1975-1989	1990-2004
	Graduates' ability, 25th percentile			*Graduates' disadvantage, 25th percentile*		
model	102.3	100.9	99.7	1.51	1.03	0.75
	(0.5)	(0.4)	(0.4)	(0.03)	(0.02)	(0.03)
data	103.0	101.6	100.3	1.31	0.98	0.64
	(0.5)	(0.3)	(0.4)	(0.03)	(0.01)	(0.01)
	Graduates' ability, 75th percentile			*Graduates' disadvantage, 75th percentile*		
model	118.2	117.7	117.2	2.16	1.84	1.46
	(0.5)	(0.3)	(0.3)	(0.05)	(0.03)	(0.02)
data	117.7	117.6	117.0	2.04	1.79	1.46
	(0.4)	(0.3)	(0.3)	(0.05)	(0.02)	(0.01)
	Non-graduates' ability, 25th percentile			*Non-graduates' disadvantage, 25th percentile*		
model	89.2	88.3	87.5	1.70	1.34	1.01
	(0.3)	(0.3)	(0.3)	(0.02)	(0.00)	(0.02)
data	89.4	88.3	87.6	1.73	1.35	1.06
	(0.3)	(0.2)	(0.3)	(0.01)	(0.00)	(0.01)
	Non-graduates' ability, 75th percentile			*Non-graduates' disadvantage, 75th percentile*		
model	108.1	107.9	106.2	2.31	2.00	1.89
	(0.2)	(0.2)	(0.3)	(0.00)	(0.03)	(0.04)
data	107.9	107.6	106.2	2.31	2.00	1.83
	(0.2)	(0.2)	(0.3)	(0.00)	(0.03)	(0.05)

Notes: The table reports the mean and standard deviation over 10,000 bootstrap samples (at the respective minimum-distance estimates obtained setting $\rho = 0.4$) of model-predicted vs empirical values of eight untargeted moments targets. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

Table A-10: Minimum-distance estimates of model parameters for $\rho = 0.3$

[A] Parameter estimates

	College cohort 1960-1974	1975-1989	1990-2004
γ	6.183	5.356	3.552
	(0.026)	(0.022)	(0.019)
τ	−3.461	−2.973	−1.782
	(0.031)	(0.025)	(0.024)
δ	0.071	0.069	0.072
	(0.003)	(0.003)	(0.003)
β	0.021	0.015	0.028
	(0.003)	(0.002)	(0.003)
α	0.609	0.741	0.926
	(0.013)	(0.014)	(0.018)
N	7,103	8,329	6,743

[C] Ability targets

	College cohort 1960-1974	1975-1989	1990-2004
3. Graduates' ability, $\mathbb{E}(\Theta\|K=1)$			
model	110.3	109.0	108.2
	(0.5)	(0.4)	(0.5)
data	110.3	109.0	108.2
	(0.4)	(0.3)	(0.3)
4. Non-graduates' ability, $\mathbb{E}(\Theta\|K=0)$			
model	97.7	97.0	96.0
	(0.2)	(0.2)	(0.3)
data	97.7	97.0	96.0
	(0.2)	(0.2)	(0.2)

[B] Labor market targets

	College cohort 1960-1974	1975-1989	1990-2004
1. College to no-college workforce ratio, ξ			
model	0.224	0.328	0.482
	(0.007)	(0.009)	(0.013)
data	0.224	0.328	0.483
	(0.007)	(0.009)	(0.014)
2. College to no-college earnings ratio, r			
model	1.738	1.617	1.543
	(0.007)	(0.005)	(0.006)
data	1.736	1.617	1.546
	(n/a)	(n/a)	(n/a)

[D] Disadvantage targets

	College cohort 1960-1974	1975-1989	1990-2004
5. Graduates' disadvantage, $\mathbb{E}(\Lambda\|K=1)$			
model	1.86	1.57	1.22
	(0.03)	(0.03)	(0.03)
data	1.85	1.58	1.21
	(0.03)	(0.03)	(0.02)
6. Non-graduates' disadvantage, $\mathbb{E}(\Lambda\|K=0)$			
model	2.34	1.99	1.85
	(0.02)	(0.02)	(0.02)
data	2.34	1.99	1.86
	(0.02)	(0.02)	(0.02)

Notes: The table reports the mean and standard deviation of minimum-distance (MD) estimates of model parameters over 10,000 bootstrap samples, setting $\rho = 0.3$, and of model-predicted vs empirical values of the six targets. The MD criterion function is given by equation (), and the weighting matrix is the identity matrix. The Online Appendix to Section 5.1 provides more computational details. The ability score is expressed in units in the table but in hundreds in the estimation, so as to reduce the order of magnitude of the estimated γ and τ. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

Table A-11: Quality of match for eight untargeted moments at minimum-distance estimates of model parameters for $\rho = 0.3$

	Ability distribution			**Disadvantage distribution**		
	College cohort			College cohort		
	1960-1974	1975-1989	1990-2004	1960-1974	1975-1989	1990-2004
	Graduates' ability, 25th percentile			*Graduates' disadvantage, 25th percentile*		
model	102.2	101.1	99.8	1.51	1.03	0.75
	(0.5)	(0.4)	(0.4)	(0.03)	(0.02)	(0.03)
data	103.0	101.6	100.3	1.31	0.98	0.64
	(0.5)	(0.3)	(0.4)	(0.03)	(0.01)	(0.01)
	Graduates' ability, 75th percentile			*Graduates' disadvantage, 75th percentile*		
model	118.2	117.8	117.2	2.16	1.84	1.46
	(0.5)	(0.3)	(0.3)	(0.05)	(0.02)	(0.01)
data	117.7	117.6	117.0	2.04	1.79	1.46
	(0.4)	(0.3)	(0.3)	(0.05)	(0.02)	(0.01)
	Non-graduates' ability, 25th percentile			*Non-graduates' disadvantage, 25th percentile*		
model	89.2	88.3	87.5	1.70	1.34	1.01
	(0.3)	(0.3)	(0.3)	(0.02)	(0.00)	(0.02)
data	89.4	88.3	87.6	1.73	1.35	1.06
	(0.3)	(0.2)	(0.3)	(0.01)	(0.00)	(0.01)
	Non-graduates' ability, 75th percentile			*Non-graduates' disadvantage, 75th percentile*		
model	108.1	107.8	106.2	2.31	2.00	1.89
	(0.2)	(0.3)	(0.3)	(0.00)	(0.03)	(0.04)
data	107.9	107.6	106.2	2.31	2.00	1.83
	(0.2)	(0.2)	(0.3)	(0.00)	(0.03)	(0.05)

Notes: The table reports the mean and standard deviation over 10,000 bootstrap samples (at the respective minimum-distance estimates obtained setting $\rho = 0.3$) of model-predicted vs empirical values of eight untargeted moments targets. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

Table A-12: Minimum-distance estimates of model parameters for $\rho = 0.5$

[A] Parameter estimates

	College cohort 1960-1974	1975-1989	1990-2004
γ	6.171	5.347	3.542
	(0.026)	(0.022)	(0.018)
τ	-3.453	-2.961	-1.773
	(0.031)	(0.025)	(0.023)
δ	0.072	0.069	0.072
	(0.003)	(0.003)	(0.002)
β	0.021	0.015	0.028
	(0.003)	(0.002)	(0.003)
α	0.821	0.926	1.072
	(0.013)	(0.012)	(0.015)
N	7,103	8,329	6,743

[C] Ability targets

	College cohort 1960-1974	1975-1989	1990-2004
3. Graduates' ability, $\mathbb{E}(\Theta \mid K=1)$			
model	110.3	109.0	108.2
	(0.5)	(0.4)	(0.5)
data	110.3	109.0	108.2
	(0.4)	(0.3)	(0.3)
4. Non-graduates' ability, $\mathbb{E}(\Theta \mid K=0)$			
model	97.7	97.1	96.0
	(0.2)	(0.2)	(0.3)
data	97.7	97.0	96.0
	(0.2)	(0.2)	(0.2)

[B] Labor market targets

	College cohort 1960-1974	1975-1989	1990-2004
1. College to no-college workforce ratio, ξ			
model	0.224	0.328	0.482
	(0.007)	(0.009)	(0.013)
data	0.224	0.328	0.483
	(0.007)	(0.009)	(0.014)
2. College to no-college earnings ratio, r			
model	1.737	1.616	1.544
	(0.006)	(0.005)	(0.006)
data	1.736	1.617	1.545
	(n/a)	(n/a)	(n/a)

[D] Disadvantage targets

	College cohort 1960-1974	1975-1989	1990-2004
5. Graduates' disadvantage, $\mathbb{E}(\Lambda \mid K=1)$			
model	1.86	1.57	1.22
	(0.03)	(0.03)	(0.03)
data	1.85	1.58	1.21
	(0.03)	(0.03)	(0.02)
6. Non-graduates' disadvantage, $\mathbb{E}(\Lambda \mid K=0)$			
model	2.34	1.99	1.85
	(0.02)	(0.02)	(0.02)
data	2.34	1.99	1.86
	(0.02)	(0.02)	(0.02)

Notes: The table reports the mean and standard deviation of minimum-distance (MD) estimates of model parameters over 10,000 bootstrap samples, setting $\rho = 0.5$, and of model-predicted vs empirical values of the six targets. The MD criterion function is given by equation (), and the weighting matrix is the identity matrix. The Online Appendix to Section 5.1 provides more computational details. The ability score is expressed in units in the table but in hundreds in the estimation, so as to reduce the order of magnitude of the estimated γ and τ. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

Table A-13: Quality of match for eight untargeted moments at minimum-distance estimates of model parameters for $\rho = 0.5$

	Ability distribution			**Disadvantage distribution**		
	College cohort			College cohort		
	1960-1974	1975-1989	1990-2004	1960-1974	1975-1989	1990-2004
	Graduates' ability, 25th percentile			*Graduates' disadvantage, 25th percentile*		
model	102.2	101.0	99.7	1.51	1.03	0.75
	(0.5)	(0.4)	(0.4)	(0.03)	(0.02)	(0.03)
data	103.0	101.6	100.3	1.31	0.98	0.64
	(0.5)	(0.3)	(0.4)	(0.03)	(0.01)	(0.01)
	Graduates' ability, 75th percentile			*Graduates' disadvantage, 75th percentile*		
model	118.2	117.8	117.2	2.16	1.84	1.46
	(0.5)	(0.3)	(0.3)	(0.05)	(0.02)	(0.01)
data	117.7	117.6	117.0	2.04	1.79	1.46
	(0.4)	(0.3)	(0.3)	(0.05)	(0.02)	(0.01)
	Non-graduates' ability, 25th percentile			*Non-graduates' disadvantage, 25th percentile*		
model	89.2	88.3	87.5	1.70	1.34	1.01
	(0.3)	(0.3)	(0.3)	(0.02)	(0.00)	(0.01)
data	89.4	88.3	87.6	1.73	1.35	1.06
	(0.3)	(0.2)	(0.3)	(0.01)	(0.00)	(0.01)
	Non-graduates' ability, 75th percentile			*Non-graduates' disadvantage, 75th percentile*		
model	108.1	107.8	106.2	2.31	2.00	1.89
	(0.2)	(0.2)	(0.3)	(0.00)	(0.03)	(0.04)
data	107.9	107.6	106.2	2.31	2.00	1.83
	(0.2)	(0.2)	(0.3)	(0.00)	(0.03)	(0.05)

Notes: The table reports the mean and standard deviation over 10,000 bootstrap samples (at the respective minimum-distance estimates obtained setting $\rho = 0.5$) of model-predicted vs empirical values of eight untargeted moments targets. A college cohort is defined by the period of actual of potential college attendance, which is an individual's age plus 20. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper).

Robustness of our results to the inclusion of the "Big Five" personality traits

Using the "Big Five" (B5) personality traits to produce alternative measures of ability or disadvantage does not alter our conclusions. This is shown in Table A-14 and Figure A-11.

In the table, row 1 reproduces (for convenience) the correlations displayed in Table 2 between ability Θ an disdvantage Λ, as measured in the main text without using the B5. The remaining rows report this correlations when: the B5 traits are added to the socioeconomic indicators in the PCA for disadvantge (row 2), only the B5 variables contribute to the PCA for disadvantage (row 3), the B5 variables are added to the cognitive ability indicators in the PCA to produce a broader ability measure (row 4).

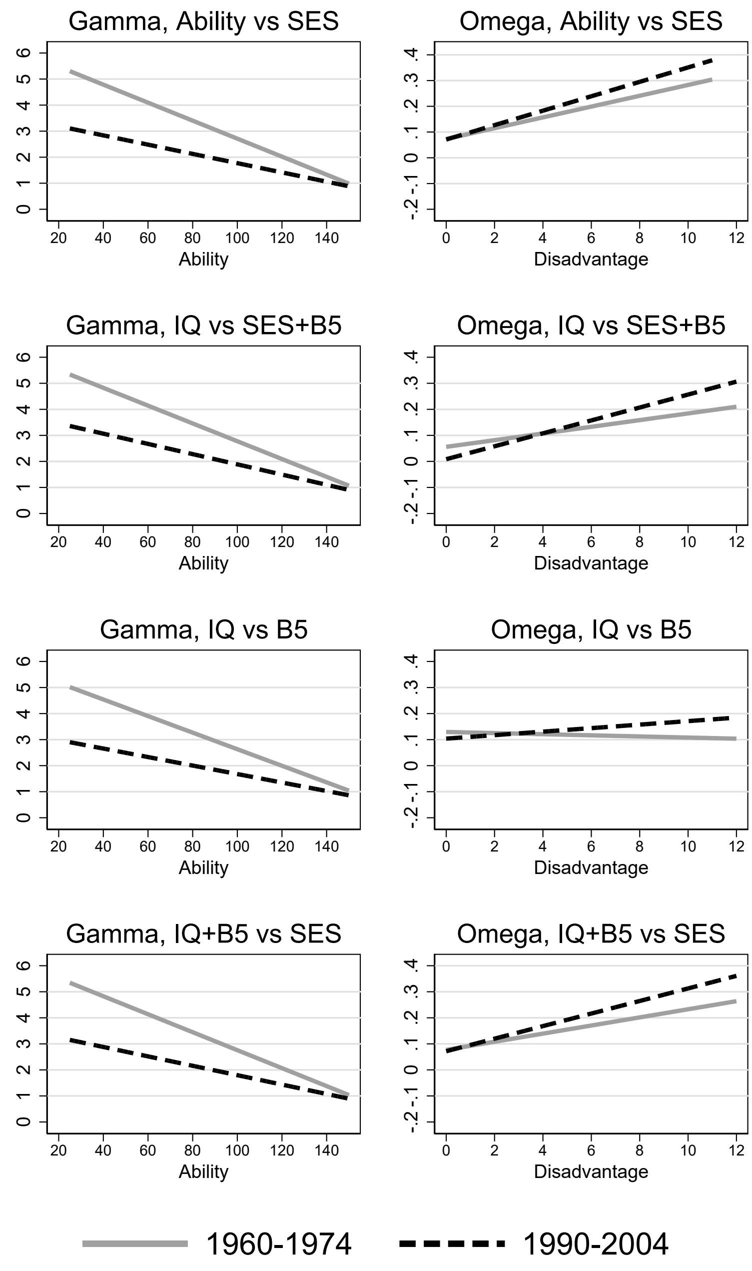
Figure A-11 follows the same exposition pattern to describe how the cost shifts $\Omega(\Lambda)$ and $\Gamma(\Theta)$ changed between the first and the final college cohorts, depending on whether and where the B5 variables are included. In all cases our central conclusion stands: the cost of study effort decreased more for low-ability students than for high-ability ones and increased more for high-disadvantage students than for low-disadvantage ones.

Table A-14: Correlation η between alternative measures of cognitive ability Θ and socioeconomic disadvantage Λ in USoc

	College cohort		
	1960-1974	1975-1989	1990-2004
Θ = FPC(cognitive variables); Λ = FPC(SES variables)	−0.143	−0.173	−0.164
	(0.013)	(0.012)	(0.013)
Θ = FPC(cognitive variables); Λ = FPC(SES & B5 variables)	−0.163	−0.198	−0.199
	(0.013)	(0.012)	(0.014)
Θ = FPC(cognitive variables); Λ = FPC(B5 variables)	−0.054	−0.081	−0.088
	(0.013)	(0.013)	(0.015)
Θ = FPC(cognitive & B5 variables); Λ = FPC(SES variables)	−0.143	−0.183	−0.168
	(0.013)	(0.012)	(0.013)
N	7,103	8,329	6,743

Notes: The table reports the correlation between alternative measures of ability (Θ) and disadvantage (Λ) that combine in different ways cognitive ability variables, socioeconomic status (SES) variables at the time the respondent was 14, and the "Big Five" (B5) personality variables available in USoc: in the first row, Θ is the FPC of cognitive variables and Λ is the FPC of SES variables (this first row is identical to Table 2 in the paper); in the second row, Θ is the FPC of cognitive variables and Λ is the FPC of SES and B5 variables; in the third row, Θ is the FPC of cognitive variables and Λ is the FPC of B5 variables; in the fourth row, Θ is the FPC of cognitive and B5 variables, and Λ is the FPC of SES variables. Standard errors are produced via the delta method. Cross-sectional response weights are applied. Sample: USoc, 22,175 white respondents born in the UK in 1940-1984 with non-missing education and ability information (see Table 1 in the paper in the paper).

Figure A-11: Estimated study effort cost shifts under alternative definitions of ability and disadvantage that include the "Big Five" (B5) personality traits



Notes: The figure shows the study effort cost shifts $\Gamma(\cdot)$ and $\Omega(\cdot)$ implied by NLS estimates of the policy parameters, for the 1960-1974 or 1990-2004 college cohorts, as a function of ability (left panel) or disadvantage (right panel), using alternative definitions of ability and disadvantage that combine in different ways cognitive ability variables, socioeconomic status (SES) variables at the time the respondent was 14, and the "Big Five" (B5) personality variables available in USoc, following the same order of Table A-14. To facilitate a comparison, the first row corresponds to the definition employed in the paper, i.e., ability is the FPC of cognitive variables and disadvantage is the FPC of socioeconomic variables. This first row is identical to Figure 10 in the paper. In the second row, ability is the FPC of cognitive variables and disadvantage is the FPC of SES and B5 variables; in the third row, ability is the FPC of cognitive variables and disadvantage is the FPC of B5 variables; in the fourth row, ability is the FPC of cognitive and B5 variables, and disadvantage is the FPC of SES variables.