The Great Gatsby goes to College: Tuition, Inequality and Intergenerational Mobility in the U.S.*

Damien Capelle
Princeton University
Job Market Paper
October 23, 2019 (check here for the most recent version).

Abstract

This paper studies the role of the higher education system, including government financial aid and transfers to colleges, in shaping income inequality and intergenerational mobility. I introduce a model of college choice with overlapping generations of heterogeneous households subject to a borrowing constraint and with heterogeneous colleges that maximize quality. First, I show that in response to the observed rise in the return to human capital the model yields predictions consistent with increases in five outcomes in the U.S. since 1980: income inequality, tuition, the dispersion of spending per-student across colleges, the exclusion of low-income students from top colleges, and the intergenerational elasticity of earnings (IGE). I quantify the model with rich micro-data from the U.S. About 6% of the observed increase in income inequality results from changes in how students and resources are allocated across colleges. Second, I use the model to run policy counterfactuals. If all students received the same higher education, the Gini coefficient and the IGE would decrease by up to 9% and 33%, respectively. Current government interventions—financial aid and transfers to colleges—decrease the Gini coefficient by 3% and the IGE by 12% compared to a laissez-faire policy. Need-blind admissions can be particularly useful at increasing mobility and correcting for the misallocation of students and resources, thereby increasing GDP, at the expense of slightly higher income inequality.

*I am deeply grateful to Roland Bénabou, Oleg Itskhoki, Nobuhiro Kiyotaki, Richard Rogerson for their comments, guidance, and support throughout all stages of this project. I also benefited greatly from conversations and input from Jonathan Heathcote, Gregor Jarosch, Adam Kapor, Adrien Matray, Jack Mountjoy, Chris Neilson, Alessandra Peter, Jonathan Payne, Esteban Rossi-Hansberg, Adrien Bilal, Mingyu Chen, Charly Porcher, Matthew Tauzer, along with participants in the Princeton Macroeconomics lunch. All errors are the sole responsibility of the author.
1 Introduction

To what extent does the higher education system in the U.S. accentuate or dampen income inequality and enhance intergenerational mobility? Access to college has traditionally been seen in the U.S. as one of the main pathways to upward mobility. However, access remains extremely selective and unequal, especially at top-quality universities. For example, Chetty et al. [2019] report that children whose parents are in the top 1% of the income distribution are 77 times more likely to attend an Ivy League college than those whose parents are in the bottom income quintile. What are the forces determining the sorting of students and financial resources across colleges of different quality? To what extent does parental income matter relative to the students abilities? How does this sorting in turn shapes inequality at the next generation and intergenerational mobility?

These questions regarding the contributions of colleges to income inequality and intergenerational mobility are all the more important as trends over the past forty years show that (a) the market returns to education and income inequality have increased [Autor et al., 2008, Piketty and Saez, 2003]; (b) the dispersion of expenditures per students across colleges has increased [Capelle, 2019c]; (c) the share of students from the lowest income quintile in top colleges has stagnated [Bailey and Dynarski, 2011, Chetty et al., 2017]; (d) tuition fees before financial aid have increased by a factor of four in real terms since 1980 (figure A4 in appendix); and (e) there is evidence of a slight increase in the intergenerational elasticity of income (hereafter IGE) corresponding to a decrease in intergenerational mobility [Black and Devereux, 2011, Urrutia and Tavares, 2016, Davis and Mazumder, 2017].

In this paper, I provide a framework to understand the interaction between the allocation of students and financial resources across heterogeneous colleges, income inequality and intergenerational mobility. In particular, I offer a unified explanation for the stylized facts (a) to (e) described in the previous paragraph and study how the higher education system—the endogenous response of colleges and government policies—accentuates or dampens income inequality and intergenerational (im)mobility and propagates the increase in income inequality.

The household side of the model builds on a large theoretical literature that formalizes how human capital transmission across generations perpetuates inequalities [Becker and Tomes, 1994, Benabou, 2002]. As in Benabou [2002], a continuum of heterogeneous households characterized by their human capital transmit, with some randomness, abilities to their children and make an educational investment choice subject to a borrowing constraint. After high school, households decide which college to send their child to. The market for higher education is populated by a continuum of heterogeneous colleges ordered by the quality they offer to their students. Households face an equilibrium tuition schedule that depends on
college quality, student ability and parental income. After college, each child becomes an adult, sells their human capital—a combination of their ability, college quality and some labor market shock—on a competitive labor market and gives birth to a child.

The novelty of my framework is to embed into this overlapping generation general equilibrium model an endogenously-determined distribution of heterogeneous colleges which are clubs [Buchanan, 1965]. Colleges seek to maximize the quality they provide to their students, subject to a social objective. They are clubs because their quality depends not only on the amount of educational resources spent per student but also on the average ability of the student body, what will be referred to as the “peer-effect”. Colleges have an incentive to attract high-ability students because of the peer-effect, as well as students from rich families who bring in additional financial resources to finance educational spending. The latter incentive is tempered by the assumed social objective. This microfoundation of the college sector is borrowed from a literature that estimates partial equilibrium models of higher education [Rothschild and White, 1995, Epple et al., 2017, Heathcote and Cai, 2016]. As in Heathcote and Cai [2016], colleges are price-takers and the tuition schedule clears each segment of the higher education market. Finally, I close the model with an educational sector that produces educational services and a government that implements non-linear merit and need-based financial aid to students and non-linear transfers to colleges.

The first contribution of the paper is to provide an analytical characterization of the equilibrium allocation and a unified explanation for the stylized facts (a) to (e): under weak conditions, an increase in the returns to education—a primitive of the model—is shown to lead to an increase in income inequality, an increase in the inequality of resources across colleges, a decrease in the share of low income students at top colleges, a decline in intergenerational mobility and an increase in tuition. Intuitively, the increase in the returns to education increases the dispersion of labor earnings for a given distribution of human capital, thereby increasing income inequality. This leads richer households to demand higher quality of higher education, incentivizing top colleges to raise tuition, increasing the dispersion of revenues and educational spending across colleges. This increased dispersion translates into more inequality in human capital at the following generation. Individuals from low-income background are priced out of top colleges, hence the stagnation of their shares and the decline in mobility. Higher education thus contributes to the gradual shift of the U.S. society to the right side of the Great Gatsby curve [Corak, 2013, Krueger, 2012].

I then estimate a more general version of the model that I use to quantify the above effects. I allow for some degree of intergenerational transfers of financial wealth and allow individuals

---

1In Benabou [2002], households buy a educational "good" traded at a constant unit price, independent on the households/students’ characteristics and there is no notion of quality ladder.
to not go to college. The model is estimated using several microdata sources: (i) the National Longitudinal Survey of Youth of 1997, a representative panel of high-schoolers, with very detailed information on parental background, the children’s abilities, their journey through the higher education system and their income in their early thirties; (ii) the NCES-NPSAS, a detailed student-level dataset on net tuition and financial aid; and (iii) the NCES-IPEDS, a panel of the universe of colleges. The estimation proceeds in two steps: first I estimate the closed-form model which I show is—under some assumptions—perfectly identified for the set of targeted moments and I then look around this set of parameters to estimate the more quantitative version. The closed-form expressions for the targeted moments in the first step make it very transparent which moments are important to pin down each parameter. They also make this first step computationally quick, allowing to estimate a large set of parameters.

The second contribution is to quantify the extent to which the higher education system shapes income inequality and intergenerational (im)mobility. I run four main policy counterfactuals. In the first one, I randomly allocate students and equalize spending across colleges. This leads to a decrease in the income Gini by up to 9.2% (4.3 p.p.) and a decrease in the IGE by up to 33%. This is a sizable effect. In the second one, I implement policies that neutralize the effect that parental income has on the sorting of students across colleges conditional on child ability, such as very progressive need-based aid or very progressive need-affirmative admissions policies. I find that this leads to a decrease in the income Gini by up to 3% (1.4 p.p.) and in the IGE by up to 20%. The third policy experiment derives the *laissez-faire* equilibrium by removing all government interventions in higher education. I find that the Gini coefficient would be 3% higher, and the IGE 10% higher. Moreover the most powerful dimension of current policy is the subsidy to colleges, which contributes to a reduction of inequality by 2%, followed by need-based financial aid, which decreases inequality by 1%. Merit-based financial aid has virtually no effect. I document elsewhere that government transfers to colleges have become significantly less redistributive over the past forty years [Capelle, 2019c]. If the degree of progressivity of the transfer schedule had remained at what it was at the beginning of the eighties, income inequality would be 1% lower. Finally, a conservative version of the 'College for All' proposal would have virtually no effect on inequality and average mobility, though it would benefit high-ability and poor children.

The third contribution is to assess the quantitative effects of a rise in the return to human capital and decompose the rise in inequality into a direct effect and the endogenous propagation through the higher education sector. I find that, given the estimated increase in the returns to education, the model generates an increase in the income Gini coefficient by 13 p.p., which corresponds to 130% of the empirical change, an increase in the expenditure per student Gini by 5 p.p. corresponding to 100% of the empirical change and an increase in
the IGE by 6%. In a counterfactual world in which the returns to human capital increase but do not propagate through the higher education sector, the increase in the Gini coefficient of income would have been 6% lower. There are two sources of amplification: the allocation of resources and the allocation of students quality across clubs. More than 100% of the total amplification is coming from the former. The latter actually dampens the increase in inequality because the positive assortative matching of students along the quality ladder of colleges worsens as relatively richer and less able children buy their way to top colleges.

The last set of contributions is normative. I first study a benchmark economy where a social planner decides upon the allocation of all goods, factors of production across sector and the sorting of students and resources across colleges. This first best allocation features perfect positive assortative matching of students by ability across colleges. In contrast, the decentralized equilibrium with incomplete financial markets displays too many rich and not so smart children in good colleges and too many poor but smart children in low quality colleges. I then show that the key source of inefficiencies in this economy is the borrowing constraint: smart and poor children lack access to liquidity to go to good colleges and rich but not very smart children over-accumulate human capital both because they can’t get direct financial transfers from their parents. Interestingly, in the extreme case where colleges are pure club and educational services do not matter, the allocation of students across colleges is the same as in the first best. Finally, in a second-best allocation, the government could restore the static optimal sorting of students with a mix of progressive need-based and merit-based tuition financial aid and transfers to colleges. But in general, it should not do so, because these transfers have implications for labor supply and intertemporal insurance. In a quantitative exercise, I find that the current level and progressivity of transfers to students and colleges are too low, even when the government’s aversion to inequality is arbitrary low. I also find that after a shock to returns to human capital, the government should raise the progressivity of transfers, although this result depends on the exact assumptions one is willing to make.

1.1 Literature Review

with a dominant role for early education. In these papers, in contrast with mine, there is a representative college whose decision—essentially tuition fees—is not microfounded and households have to make a binary decision—enrolling or not into college. The exception is [Hellier, 2017] who shows in a stylized framework that the French two-tier system is responsible for the low intergenerational mobility at the top of the income distribution.

In contrast, another stream of the literature models in detail and in static partial equilibrium the admission and tuition setting decisions of colleges and the rich heterogeneity in colleges and student types, but focuses mainly on the impact of financial aid policies [Epple et al., 2006, 2017, Fu, 2014] and affirmative action [Arcidiacono, 2005, Kapor, 2015, Ramos and Herskovic, 2017] on the sorting by itself, while my focus is the study of equilibrium inequality and mobility. More recently and more closely related, [Heathcote and Cai, 2016] shows that income inequality can fully account for the observed rise in average net tuition since 1990 in a static model where households choose a quality of college and price-taking colleges maximize the quality they deliver. My model has a flavor of theirs. We differ in that (i) there is a continuum of student ability instead of two types, which—maybe surprisingly—simplifies the analysis; (ii) higher education is not purely a consumption good but does matter for the accumulation of human capital. Incidentally, the equilibrium allocation is fully efficient in their model; and (iii) more importantly, my paper is dynamic and is, to the best of my knowledge, the first one to embed a sorting problem of heterogeneous students across heterogeneous clubs into an intergenerational setting, and still remains tractable and lends itself to the study of the role of higher education in the transmission of economic status over generation. One implication of this tractability is the possibility to work in an environment in which existence and uniqueness of the equilibrium—two issues that have plagued the theoretical and quantitative literature on clubs [Ellickson et al., 1999]—can be characterized.

Another related stream of the literature studies the macroeconomic determinants of tuition and the effects of financial aid. Lucca et al. [2015] stresses the role of the expansion of credit supply, Jones and Yang [2016] the importance of financial aid and [Gordon and Hedlund, 2017] the rising cost of universities input and professors—the Baumol’s disease, a mechanism our model accommodates—implied by the rise in the skill premium, to explain the rise in tuition. In this paper, I stress the role of the increase in the returns to education to explain the rise in the average and the dispersion of tuition fees, due to the increase in demand by households for higher quality of higher education, especially at the top of the distribution. This mechanism is akin to the revenue theory of cost by Bowen [1980], but applied to a framework with a ladder of colleges, whereby universities raise all the money they can through tuition fees and then spend it on projects that enhance quality. Martin et al. [2017] have estimated that this mechanism accounts for two third of the increase in
the average real tuition fees. Regarding financial aid policy, [Abbott et al., 2013] shows that current financial aid policy improves efficiency and increases GDP. Although we share the same deep source of inefficiency—a borrowing constraint—I focus on the misallocation of heterogeneous students across heterogeneous colleges rather than on the enrollment rate.


Finally, the last part of the paper focusing on the optimal higher education policies relates to two streams of the literature. The first one characterized the optimal education policy [Benabou, 1996, Krueger and Ludwig, 2013, Stantcheva, 2015]. I build on the non-linear income tax and transfer schedule proposed by Benabou [1996] and construct in the same spirit transfers to students based on their abilities and on their parents income—called respectively merit-based and need-based financial aid—as well as transfers to colleges. To my knowledge, this paper is the first to model government transfers policy in higher education in a rich and still tractable way. I also relate to the theoretical paper on clubs [Cole and Prescott, 1997, Caucutt, 2001, 2002] and the optimal solution to assignment problems when there is complementarity in the human capital production function and/or when there is a peer-effect [Durlauf and Seshadri, 2003a,b, Salle et al., 2008, Chade et al., 2017].

The rest of the paper is organized as follows. Section 2 presents the model and section 3 explains the closed-form equilibrium expressions. Section 4 derives the key analytical comparative statics: an increase in the return to human capital generates facts (a) to (e). Section 5 derives properties of the first best allocation, show the sources of inefficiencies in the decentralized equilibrium and characterizes the second-best allocation within the class of log-linear transfers schedules. Section 6 explains the estimation procedure. Section 7 derives the quantitative results regarding the role of the higher education sector for the amplification of inequality and the reproduction of economic status over generation. Section 8 concludes.
2 Human Capital Transmission with a Hierarchy of Colleges

The economy is populated by three types of agents: dynastic households, colleges and a government. At each generation, households imperfectly transmit human capital to their child and decide which college to send them to after high school. Colleges choose their pool of students as well as educational spending to maximize the quality they deliver, subject to a social objective. And the government sets a non-linear schedule of need-based and merit-based financial aid to students as well as transfers to colleges.

2.1 Households

There is a mass one of dynasties, indexed by $i \in [0, 1]$. Individuals live for two periods: one as a child and one as an adult. Each adult has one child. A generation $t$ household of dynasty $i$ is characterized by its level of human capital $h^i_t$ and the child’s human capital at the end of high school $h^i_{s,t}$. They choose consumption $c^i_t$, labor supply $\ell^i_t$ and college quality $q^i_t$ for their child to maximize their utility $U^i_t$. Dropping the generation and dynasty subscripts when no confusion results, the dynastic utility is

$$
\ln \mathcal{U}(h, h_s) = \max_{c, \ell, q} \left\{ (1 - \beta) \left[ \ln c - \ell \eta \right] + \beta E \left[ \ln \mathcal{U}(h', h'_s) \right] \right\}
$$

(1)

A child’s human capital at the end of high school is modeled as a log-linear combination of parents’ human capital $h$ and the birth shock $\xi_b$, capturing the randomness of the transmission process.

$$
h_s = (\xi_b h)^{\alpha_1}
$$

(2) 

A household’s lifetime after tax and transfers earnings denoted $y$ is a function of their level of human capital $h$, their supply of raw labor $\ell$ and the tax schedule. More details on this earning function and on the income tax schedule are given in section 2.3 and 2.4 respectively.

Households are subject to a lifetime budget constraint. Their after-tax income $y$ can be spent on consumption and on tuition fees, which depend on college quality $q$, household income $y$ and the child ability $h_s$. Normalizing the price of the final good to one, it is given by

$$
y = (1 + a_c) c + e(q, y, h_s)
$$

(3) 

Household Lifetime Budget Constraint
where \( a_c \) is the consumption tax rate.

This budget constraint implies that households face an intergenerational borrowing constraint, \textit{i.e.} the current adults cannot leave bequest or pass-on debt along to their offspring. This assumption draws on a large set of evidence that borrowing constraints do matter for college choices. Carneiro and Heckman [2002], Belley and Lochner [2007], Lochner and Monge-Naranjo [2011], Brown et al. [2012] and Johnson [2012] all find strong evidence of borrowing constraints in the data—though they differ in the exact number of students at the constraint. Lochner and Monge-Naranjo [2012] offers a review of the literature on borrowing constraint in education. Although this specification rules out net financial transfers across generation, the quantitative version I introduce later partially relaxes this assumption\(^2\).

The adulthood human capital of the child after college is a log-linear combination of its pre-college ability, the quality of the college they went to parental human capital and a labor market shock.\(^3\) It is given by

\[
h' = h_s q^{\alpha_2} h^{\alpha_3} \xi_y
\]

Child’s Post-College Human Capital \(^{(4)}\)

Although not crucial for the qualitative results, the parents’ human capital is included on top of the child’s ability and the college quality, because as we will see later on in the empirical section, I find a role for parental status in determining the labor market outcomes of children, conditional on abilities and college, through networks for example.

There are two sources of randomness in the accumulation process of human capital. The birth shock \( \xi_b \) is known before the college quality decision has to be made, while the labor market shock \( \xi_y \) is realized once the child enters the labor market. It is assumed that the birth and labor market shocks are i.i.d across generations and households and log-normally distributed.\(^4\)

\[
\begin{align*}
\ln \xi_b &\sim \text{i.i.d. } \mathcal{N}(\mu_b, \sigma_b^2) & \text{Birth Shock} \\
\ln \xi_y &\sim \text{i.i.d. } \mathcal{N}(\mu_y, \sigma_y^2) & \text{Labor Market Shock}
\end{align*}
\]

\(^{2}\)Ruling out net financial transfers across generations doesn’t prevent gross flows to exist within a lifetime. For example, children are allowed to borrow from their parents early in life and repay later. It doesn’t rule out student loans as long as they are exactly offset by a parental transfer of the same amount, \textit{i.e.} student debt is possible as it is paid by parents.

\(^{3}\)There is direct empirical evidence that the law of accumulation of human capital is characterized by complementarities across stages of education. Dillon and Smith [2018] finds evidence of complementarity between student ability and college quality for long-term earnings. See also Cunha et al. [2005] for a review of life-cycle skill formation.

\(^{4}\)This formulation of the household problem draws from and extends Benabou [2002]. The latter can be seen as the case where there is no birth shock, \( \sigma_b^2 = 0 \) and a constant unitary price for education \( e(q,y,h_s) = q \).
2.2 Colleges

Technology. There is a mass one of ex-ante identical colleges indexed by \( j \in [0, 1] \). They are all of the same size, \( c > 0 \). Given the relative mass of colleges and students, in equilibrium \( c = 1 \). A college is a technology that delivers to its students a quality that depends on educational spending per student \( I_j \) and the (geometric) average of student ability \( \theta_j \), which will be referred to as the "peer effect." Furthermore, I assume that quality depends negatively on the degree of dispersion of abilities and parental income within the college, \( \sigma_u^2 \), which is defined later. The production function of quality is given by

\[
\ln q_j = \ln I_j^\omega_1 \theta_j^{\omega_2} - \omega_1 \frac{\sigma_{u,j}^2}{2}
\]

where \( \omega_1, \omega_2 > 0 \).

College are clubs because who belongs to the college matters for the quality delivered to all members, through \( \theta \). The notion that colleges compete for the best students and seek to maximize \( \theta \) is supported by Hoxby [2018]. Whether it really matters for the value-added they deliver to their students and should enter the production function of quality is less well-established. Smith and Stange [2016], Sacerdote [2011] and Mehta et al. [2018] do find evidence of peer-effects, especially from roommates, for achievements while in college. Zimmerman [2019] finds evidence that the network and social capital [Coleman, 1988] built while in college matters for labor market outcomes.

I make two assumptions about the negative impact of student heterogeneity on quality. One is implicitly embedded in the peer-effect, \( \theta_j \), which is defined as the geometric average of student abilities and which therefore punishes heterogeneity relative to an arithmetic average:

\[
\ln \theta_j = E_{\phi(.)}[\ln(h_s)]
\]

where \( \phi(.) \) denotes the distribution of student types within the college. Secondly, through \( \sigma_{u,j} \), I explicitly assume that the more heterogeneous the class in terms of student ability and social background the more difficult it is for a college to deliver a given level of quality to its student. I define \( \sigma_{u,j}^2 \) as the within-college variance of a weighted average of (log) ability and parental background, \( \log h_s^{\omega_2} y^{-\frac{\omega_3}{\omega_1}} \):

\[
\sigma_{u,j}^2 = E_{\phi(.)} \left( \left( \log \left( \theta_j^{\omega_2} D_j^{-\frac{\omega_3}{\omega_1}} \right) - \log h_s^{\omega_2} y^{-\frac{\omega_3}{\omega_1}} \right)^2 \right)
\]

where \( D_j \) denotes the (geometric) average parental income within the college. Defining \( \sigma_{u,j}^2 \)
in this manner ensures tractability by making $I_j \times e^{-\frac{\sigma^2_{u,j}}{2}}$ a geometric average of tuition fees. The solution to this problem would therefore be the same if colleges maximized a weighted geometric average of tuition and student ability, a point I come back to when explaining how I embed this problem in a more quantitative version of the model.\footnote{From this perspective, the college’s problem has a flavor of Fu [2014], where colleges maximize a weighted average of average student ability and a quadratic function of net tuition.}

Educational spending $I_j$ is financed through the collection of tuition fees from all students. The static budget constraint of a college is

$$p_I I_j = E_{\phi(.)}[e_u(q, h_s, y)]$$

where $p_I$ denotes the price of educational services.

**Objective and Problem.** Each college seeks to maximize $\ln \mathcal{Y}_j$, the quality they deliver to their students ($\ln q_j$) net of a social objective. The social objective is a penalty increasing in the (geometric) average of parental incomes, $D_j$ and parametrized by $\omega_3$:

$$\ln \mathcal{Y}_j = \ln q_j - \omega_3 \ln D_j$$

Taking the tuition schedule $e(q, h_s, y)$ and the price of educational services $p_I$ as given, a college chooses the amount of educational services per student $I_j$ and the composition of the student body $\phi_j(h_s, y)$—a density over $(h_s, y)$ (which determines the average student ability $\theta_j$ and the average parental income of the students $D_j$)—to maximize the quality $q_j$ they deliver to their students. For simplicity, asymmetries of information are assumed away and clubs have perfect information about the type of applicants, $(h_s, y)$. Dropping the college subscript when no confusion results, the problem of a college is:

$$\max_{I, \theta, D, \phi(.)} \ln \mathcal{Y} = \ln q - \omega_3 \ln D \quad (8)$$

subject to:

$$\ln q = \ln I^{\omega_1} \theta^{\omega_2} - \omega_1 \frac{\sigma^2_u}{2} \quad \text{Production Function} \quad (9)$$

$$p_I I = E_{\phi(.)}[e_u(q, h_s, y)] \quad \text{Budget Constraint} \quad (10)$$

$$\ln \theta = E_{\phi(.)}[\ln(h_s)] \quad \text{Average Student Ability} \quad (11)$$

$$\ln D = E_{\phi(.)}[\ln(y)] \quad \text{Average Parental Income} \quad (12)$$

The formulation of the college problem follows the literature that studies the behavior of universities [Epple et al., 2017, Fu, 2014]. I depart from this literature by assuming that...
they behave competitively as in Heathcote and Cai [2016]. Unlike the latter paper, colleges have a social objective, which makes tuition an increasing function of parental income in equilibrium. Beyond the (quite recently) observed progressivity of tuition fees with income, there is also direct evidence for this social objective: the claimed and (growing) public effort of private and public universities to recruit low-income students.

**Entry and Positioning Game.** Before operating, *ex ante* identical colleges play a positioning game on the line of qualities: taking the position of all other colleges as given, they choose which quality to offer. The payoff for operating a given quality is given by (8) and is assumed to be \( V = 0 \) if the college is non-operating, *i.e.* of size zero. A Nash equilibrium of this positioning game is a mapping from the set of colleges \( j \in [0, 1] \) to the set of qualities such that given the positioning of all other colleges, no college wants to change its position.

This structure for entry ensures that all positive qualities are offered in equilibrium. The assumption that a non-operating college gets \( V = 0 \) sets the lower bound of the support of qualities offered in equilibrium, \( q = 0 \). The assumption that all colleges must be of size \( c = 1 > 0 \) ensures that colleges do not agglomerate at the highest quality level with each one of them operating with an infinitely small mass of students.\(^7,8\)

### 2.3 Technologies for the Final Good and Educational Services

Apart from the college sector, there are two other sectors in the economy: the consumption good sector and the educational services sector. The consumption good is produced by households who operate their own production function and who sell their output on a competitive market at a price normalized to 1. The market earnings function is

\[
y_m = Ah^\lambda \ell^\mu
\]

**Household Market Income**

\(^6\)This paper further departs from the literature by assuming that colleges maximize profits. In my paper, it can be shown that the dual problem of maximizing profit subject to a constraint on \( V \) leads to the same first order conditions. However it is not easy to interpret the nature of such a constraint. If \( \omega_3 = 0 \), this constraint becomes a quality constraint which is more easily interpretable.

\(^7\)A sufficient assumption is that there is a lower bound on the size of a college. With no lower bound, all colleges would prefer to operate at the highest quality level. Assuming such a maximum exists, let’s denote it \( \bar{q} \). (This argument is purely heuristic as the support for \( q \) in equilibrium will be \([0, +\infty)\).) For a mass \( f_{\bar{q}} \) of students applying at \( \bar{q} \) and a fixed size parameter \( c = 1 \), there exists a critical mass of colleges, \( m_{\bar{q}} = c \times f_{\bar{q}} \), such that an additional college wouldn’t find any student and would therefore get \( V = 0 \). Such a college would strictly prefer operating at a lower quality.

\(^8\)One can see these joint assumptions as the equivalent of the free-entry/non-profit condition in Heathcote and Cai [2016]. A key difference, however, is that contrary to a free-entry condition that equalizes payoff for all colleges, in a Nash equilibrium of the positioning game in my setup, colleges receive heterogeneous payoffs if (and only if) they offer different qualities. All colleges would like to be Harvard, but there is only one Harvard, etc...
where $A$ is an aggregate constant and $\lambda, \mu > 0$. The elasticity of income to human capital, $\lambda$ will be called the 'returns to human capital'. This parameter plays an important role in the rest of the paper. I argue in section 4 that an increase in $\lambda$ is able to rationalize the trends observed in higher education and explained in introduction.

Although very simple, this functional form is also the reduced-form expression of a more sophisticated production function with physical capital and/or the payoff to a household involved in an aggregate production process with some degree of substitution/complementarity across heterogeneous skills, as I show in appendix C.3.

Educational services are produced using the final good as input. Each unit of final good gives $A_I/A$ units of educational services:

$$y_I = \frac{A_I}{A} y_m$$

Colleges buy services from the educational sector at relative price $p_I$. In equilibrium, it will be the case that $p_I = 1/A_I$.

Given that higher education is intensive in its use of high skilled labor, it is of interest to extent the analysis to this case. This would make the price of educational services endogenous to economy-wide conditions. In particular it would imply a positive relationship between the returns to human capital, $\lambda$, and the price of education services $p_I$, through the increase in the relative wages of faculty. I provide a tractable generalization along these lines in appendix A.3.2.9

### 2.4 Government

The government implements four kind of taxes: two are specific to higher education (non-linear merit-based and need-based financial aid to college students and non-linear transfers to colleges) and two that are more standard (a linear consumption tax and a progressive income tax).

**Progressive Income Tax Schedule** The household labor income is subject to a progressive tax schedule with $a_y$ the average tax rate and $\tau_y$ its progressivity. The after-tax and 

---

9In this generalization, the production function of educational services is given by $y_I = A_I h^{\bar{\lambda}} \mu^\mu$ with $\bar{\lambda} \geq \lambda$, so that high human capital individuals have a comparative advantage in the educational sector. None of the key mechanisms emphasized in this paper depends on this generalization. And all key results are derived with the generalized version of the educational sector.
transfers lifetime earnings is given by

\[ y = (1 - a_y) y_m^{1-\tau_y} T_y \]

Household After-Tax Income

where \( T_y \) is a normalizing aggregate endogenous factor ensuring that \( a_y \) parametrizes the average income tax rate. This specification was introduced by Feldstein [1969], Persson [1983], Benabou [2000] and estimated more recently in Heathcote et al. [2017b]. The non-linear schedules for financial aid and the college subsidy are in the same spirit as this income tax schedule.

**Merit and Need-Based Financial Aid** Financial aid is allowed to be progressive with income and regressive with abilities:

\[ e(q, h_s, y) = T_e h_s^{\tau_m} y^{\tau_n} e_u(q, h_s, y) (1 + a_h) \]

where \( e(q, h_s, y) \) is the after financial net tuition faced by households, as specified in (3) and \( e_u(q, h_s, y) \) is the before financial aid price, commonly referred to as the sticker price. \( \tau_m \) is the rate of progressivity (or rather regressivity) of the merit-based subsidy, \( \tau_n \) is the rate of progressivity of the need-based subsidy and \( T_e \) ensures that \( a_h \) is the average financial aid to students.

**Transfers to Colleges** Financial transfers to colleges by states and the federal government are large and highly progressive, as is documented in a companion paper Capelle [2019c]. Colleges that spend less per student receive relatively more subsidies.\(^{10}\) This progressivity is closely related to the location of public and private colleges in the distribution of quality. Most papers modeling the higher education sector differentiate between public and private colleges. For example, in Epple et al. [2017], tuition fees at public universities are set exogenously while the private sector chooses them endogenously and public universities receive exogenous transfers from the government. In contrast, I do not specify any ex ante differences across colleges.\(^{11}\) In my model, the bottom and middle of the distribution of qualities, \textit{i.e.} the

\(^{10}\)In particular it doesn’t refer to the progressivity with respect to the average students parental income that populate these colleges. Even if students from richer families are more likely to be in high revenue colleges, it might be that overall these transfers are regressive since many children from low-income background do not enroll in college—a mechanism I abstract from in this version of the model but allow for in the quantitative version of the model presented in section 7.

\(^{11}\)There is little agreement in the literature about what really differentiates public colleges’ behaviors from non-profit private ones, especially in the more recent periods. Epple et al. [2017] and Heathcote and Cai [2016] assume that tuition fees at public colleges are fixed by States’ legislatures and that they receive public subsidies. But in most States, public colleges have increasing and now considerable autonomy in their tuition
colleges that receive relatively more transfers from the government, can be interpreted as public colleges. This way of modeling government transfers allows me to keep the model tractable while capturing most of the heterogeneity in government transfers along the quality distribution. Taking into account these transfers, the budget constraint of a college is:

\[ p_I I = T_u (1 + a_u) (E_{h_s,y}[e_u(q, h_s, y)])^{1 - \tau_u} \]

(16)

where \( \tau_u \) is the degree of progressivity of subsidies to universities and \( T_u \) ensures that \( a_u \) is the average amount of transfers per student received by colleges. The budget constraint presented in the college problem, (10), was the special case when \( \tau_u = 0 \).

**Government Budget Constraints and Definition of** \( T_u, T_y, T_e \) **There are two kinds of constraints. The first one is the aggregate budget constraint that states that revenues (income tax and consumption tax) must equal spending (transfers to colleges and students) at any period.**

\[
\int_0^1 a_y y(i) + a_c c(i) + e(i)di = \int_0^1 e(i)(1 + a_u)(1 + a_h)di
\]

(17)

The other three constraints, (19), (18) and (20), pin down \( T_u, T_y, T_e \) such that \( a_y, a_h, a_u \) parametrize respectively the average rate of income tax, financial aid and transfers to college.

\[
\int_0^1 y(i)^{1 - \tau_s} T_y(di) = \int_0^1 y(i)di
\]

(18)

\[
(1 + a_h) \int_0^1 e(i)di = \int_0^1 e_u(i)di
\]

(19)

\[
\int E_{h_s,y}[e_u(q, h_s, y)]f_q dq = \int T_u (E_{h_s,y}[e_u(q, h_s, y)])^{1 - \tau_u} f_q dq
\]

(20)

where \( f_q \) denotes the mass of students in colleges of quality \( q \).

### 2.5 Equilibrium

An equilibrium path is a sequence of tuition schedule, prices of educational services, household’s policy functions, colleges’ policy functions, a sorting rule, a distribution of human capital \{\( e_{\ell}(q, h_s, y), p_{I\ell}, c_{\ell}(h, \xi_b), \ell_{\ell}(h, \xi_b), \phi_{\ell}(q, y, h_s), I_{\ell}(q), q_{\ell}(j), q_{\ell}(h, \xi_b), f_{\ell}(h) \}_{\ell=0}^{\infty} \) such that

1. Given the sequence of prices, the household’s policy functions \( c_{\ell}(h, \xi_b), \ell_{\ell}(h, \xi_b) \) are solution to (1),

and hiring policies [Mc Guinness, 2011]. And even if the legislatures were omnipotent, it is not clear that their objective would be radically different than maximizing the quality delivered. For-profit colleges do display different behavior, but they make up a very small part of total enrollment [Deming et al., 2013].
2. Given the sequence of prices, the college’s policy functions $\phi_t(q, y, h_s), I_t(q)$ are solution to (8),

3. The allocation of colleges along the quality line $q_t(j)$ is a Nash equilibrium of the positioning game,

4. The demand for quality $q$ from students of type $(h_s, y)$ is matched by a supply for this type at that quality,

5. The final good market as well as the educational services market clear,

6. The evolution of the distribution of human capital, $f_t(h)$, is consistent with the inter-generational law of motion of human capital and the sorting rule, $q_t(h, \xi_b)$.

### 3 Properties of the Decentralized Equilibrium

I construct an equilibrium in which the distribution of human capital is log-normal. A necessary and sufficient condition for this distribution to remain log-normal over generations is for the tuition schedule to be a log-linear function of college quality $q$, student ability $h_s$ and parental income $y$. Given the assumptions laid out in the previous section, the unique tuition schedule compatible with the equilibrium conditions and colleges being in an interior solution is log-linear. These two restrictions—log-normality of human capital and interior solutions for colleges—ensure the tractability of the equilibrium expressions.\footnote{I cannot however rule out the existence of equilibria outside of this class.}

#### 3.1 Equilibrium Tuition Schedule

Consider a college that decides to supply quality $q$. It then has to choose the optimal combination of inputs—educational spending $I_j$ and the distribution of students’ quality that are consistent with $q$. Given the substitutability between educational resources and student ability, a college will trade off lower tuition for higher student ability. Similarly, because of its social objective, a college is willing to offer tuition discounts to students from low-income families. The first-order conditions with respect to the density over student types and to the level of spending in the college’s problem reflect these trade-offs. The following proposition gives the unique equilibrium tuition schedule that is compatible with all colleges being at an interior solution. It takes a log-linear form and, incidentally, implies that all colleges are indifferent between all student types.\footnote{Although it is natural to focus on interior solutions, I cannot rule out the existence of other equilibria where tuition fees deviate from this log-linear expression and some colleges are at corner solutions for some student types. The real world tuition schedule does display kinks and a log-linear tuition schedule should be
Proposition 3.1. Denoting by $\Sigma_h$ the equilibrium standard deviation of (log) human capital in the economy, the equilibrium before-financial-aid tuition schedule is given by

$$e_{u,t}(q, h_s, y) = \left( \frac{p_{I,t}}{(1 + a_{u,t})T_{u,t}} q^{1 + \frac{1}{c_1,t} h_s - \frac{c_2,t - c_3,t}{c_1,t}} \left( \frac{y}{\kappa_{2,t}} \right)^{c_3,t} \right)^{1 - \frac{1}{c_1,t}}$$

where $\varepsilon_{l,t} = \frac{\omega_l}{1 - \nu_l(\Sigma_{h,t})\omega_3} \forall l = 1, 2, 3$

with $\nu_l(\Sigma_t)$ the elasticity of mean parental income within a college to quality

$$D_t(q) = \kappa_{2,t} q^{\nu_l(\Sigma_{h,t})}$$

and all colleges are indifferent between all types.

The variables $\varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}$ are equilibrium objects that depend on current aggregate states, in particular the dispersion of human capital in the economy $\Sigma_h$, and the policy parameters. Mathematical expressions for $\nu_l$ and $\kappa_{2,t}$ are given in appendix A.4.5. $\kappa_{2,t}$ depends not only on current states but also on all future states through the labor supply decision $\ell_t$. The notation $\nu_l(\Sigma_{h,t})$ makes explicit that the elasticity of mean parental income to quality depends on the dispersion of human capital in the economy. As I show in appendix A.8, it is increasing in the latter. It also depends on the current policy parameters and $\lambda$ the returns to human capital. Note that when colleges have no social objective and only maximize quality, $\omega_3 = 0$, then $\varepsilon_l = \omega_l$ for all $l = 1, 2, 3$ and the $\omega$’s are independent of the state of the economy. I provide more intuition after the next paragraph.

Tuition fees are increasing in quality $q$, decreasing in student ability $h_s$ and increasing in parental income $y$ with respective elasticities of $\frac{1}{\varepsilon_{1}(1-\tau_u)}, \frac{\varepsilon_{2}}{\varepsilon_{1}(1-\tau_u)}, \frac{\varepsilon_{3}}{\varepsilon_{1}(1-\tau_u)}$. These elasticities are intuitive. Colleges of higher quality need to finance higher expenses, hence require higher tuition. If educational services are very important for the production of college quality, $\omega_1$ is high hence $\frac{1}{\varepsilon_{1}(1-\tau_u)}$ is low, tuition will not be very elastic to quality, because a small increase in revenues implies a large increase in quality. The elasticity $-\frac{\varepsilon_{2}}{\varepsilon_{1}(1-\tau_u)} = -\frac{\omega_2}{\omega_1(1-\tau_u)}$ captures the importance of the peer-effect relative to educational spending: if peers significantly matter, colleges have strong incentives to subsidize high ability students to attract them. The last elasticity $\frac{\varepsilon_{3}}{\varepsilon_{1}(1-\tau_u)} = \frac{\omega_3}{\omega_1(1-\tau_u)}$ captures the strength of the social objective: the larger $\omega_3$, the more progressive tuition fees are.\(^\text{14}\)

interpreted as a smooth approximation of reality. Although looking at a more general class of equilibria is potentially interesting, it is beyond the scope of this analysis and would defeat a key purpose of this paper, as all tractability would be lost.

\(^\text{14}\)The equilibrium tuition function turns out to be very similar to the one in Epple et al. [2017]. While the progressivity of tuition fees with parental income originates from market power in their framework, it comes from the social objective in this paper.
The correction by \( \frac{1}{1-\nu_t(\Sigma_{h,t})}\omega_3 \) that transforms \( \omega_1 \) into \( \varepsilon_1 \) reflects the cross-subsidization from high-income to low-income families implied by the social objective. Tuition fees for a family with a given income \( y \) increase with a lower elasticity with respect to quality when colleges have a social objective. This family becomes poorer and poorer relative to the within-college mean as one climbs the college quality ladder since parental income increases in equilibrium with quality. This effect is all the more pronounced as the social objective parameter \( \omega_3 \) and the equilibrium elasticity of parental income to college quality \( \nu \) are large.\(^{15}\)

### 3.2 Household Policy Functions

Given the equilibrium tuition schedule (21), households choose where to send their offspring. Since the tuition schedule is monotonic in \( q \), this decision amounts to choosing how much of their income to spend on higher education. Define the saving rate of a household of type \((h_s, y)\) going to college of quality \(q\):

\[
s_t(q, h_s, y) = \frac{e_t(q, h_s, y)}{y} = \frac{T_{e,t}h_s^{-\tau_{m,t}}y^{\tau_{m,t}}e_{u,t}(q, h_s, y)}{y(1 + a_{h,s})}
\]

where the second term decomposes tuition fees into financial aid and before-financial-aid tuition fees given respectively by (15) and (21).

An attractive feature of the class of models with unitary elasticity of intergenerational substitution, log-normal innovations and log-linear technologies is the possibility to obtain analytic expressions for the optimal saving rate and labor supply. The following proposition characterizes the solution to the F.O.Cs associated with the households’ problem.

**Proposition 3.2.** Defining \( U = \frac{\partial \ln u}{\partial \ln h} \), the elasticity of the value function to human capital, one has that, in equilibrium, for all \( i \), the households’ saving rate, labor supply and marginal

---

\(^{15}\)If inequality increases for exogenous reasons—as will be the case in our comparative statics with respect to the returns to education \( \lambda \)—the endogenous increase in \( \nu \) provides a partial mitigating force by making colleges willing to endogenously redistribute more across students, provided \( \omega_3 > 0 \).
value of human capital $U$ are given by:

$$s_t = \frac{\beta \alpha_2 \varepsilon_{1,t} (1 - \tau_{u,t}) U_{t+1}}{1 - \beta + \beta \alpha_2 \varepsilon_{1,t} (1 - \tau_{u,t}) U_{t+1}}$$  \hspace{1cm} (22)$$

$$\ell_t = \left[ (1 - \tau_{y,t}) \frac{\mu}{\eta} \left( 1 + \frac{\beta}{1 - \beta} \alpha_2 \varepsilon_{1,t} (1 - \tau_{u,t})(1 - \tau_{n,t}) - \varepsilon_{3,t} \right) U_{t+1} \right]^\frac{1}{\eta}$$  \hspace{1cm} (23)$$

with $U_t = (1 - \beta) \sum_{k=0}^{\infty} \beta^k (1 - \tau_{y,t+k}) \lambda_{t+k} \prod_{m=0}^{k-1} \alpha_{h,t+m} (\Sigma_{h,t+m})$  \hspace{1cm} (24)$$

and $\alpha_{h,t} = \alpha_1 + \alpha_3 + \alpha_2 [\varepsilon_{A,t} (\Sigma_{h,t}) + \varepsilon_{I,t} (\Sigma_{h,t})]$

The saving rate and labor supply are independent of the household type and depends positively on $U_{t+1}$ which is also common to all households. The latter depends positively on all current and future $\alpha_{h}$’s, which is the IGE. It thus depends on all current and future $\Sigma_{h}$’s, an endogenous state variable. The higher the future IGEs the more incentive the current generation has to invest in human capital and work. Importantly $V_t$—thus $s_t$—is also increasing in both the current and future returns to education—$\lambda$. It will play a key role in the dynamics of human capital afterwards. Notice also the presence of $(\varepsilon_{1,t} (1 - \tau_{u,t})(1 - \tau_{n,t}) - \varepsilon_{3,t})$ in the expression for labor supply: increasing the progressivity of financial aid or transfers to students disincentivizes households from working, and will play a role in the optimal tax analysis.

### 3.3 Equilibrium Sorting Rule

By combining the equilibrium tuition schedule and the equilibrium positioning of colleges on the quality line—the "supply side"—with the household spending rule—the "demand side"—one obtains the equilibrium sorting rule, a mapping from the set of household and student types into the set of qualities of higher education.

**Proposition 3.3.** In equilibrium, the sorting rule is given by

$$q_t = \left( \frac{s_t y_{t}^{1 - \tau_{n,t}} h_{s,t}^{\tau_{m,t}} (1 + a_{h,t})}{T_{e,t}} \right)^{\varepsilon_{1,t}(1 - \tau_{u,t})} \left( \frac{(1 + a_{u,t}) T_{u,t}}{p_{l,t}} \right)^{\varepsilon_{1,t}} h_{s,t}^{\varepsilon_{2,t}} \left( \frac{y_{t}}{k_{2,t}} \right)^{-\varepsilon_{3,t}}$$  \hspace{1cm} (25)$$

In the special case without government policy, it writes

$$q_t = y_{t}^{\varepsilon_{1,t} - \varepsilon_{3,t}} h_{s,t}^{\varepsilon_{2,t}} \left( \frac{s_t}{p_{l,t}} \right)^{\varepsilon_{1,t}} k_{2,t}^{\varepsilon_{3,t}}$$

Equation (25) tells us where a student from family background $y$ and ability $h_s$ goes to
college. The elasticity of quality to income and ability capture the strength of what I call, respectively, the income-sorting and ability-sorting channel:

\[
\frac{\varepsilon_{I,t}}{(1 - \tau_y)\lambda_t} = \varepsilon_{1,t}(1 - \tau_u)(1 - \tau_n) - \varepsilon_{3,t} \quad \text{Elasticity to Income}
\]

\[
\frac{\varepsilon_{A,t}}{\alpha_1} = \varepsilon_{2,t} + \tau_m(1 - \tau_u)\varepsilon_{1,t} \quad \text{Elasticity to Ability}
\]

The elasticity to income—associated with the income-sorting channel—captures the fact that richer households are able to buy a higher quality of college, not only because they are richer but also because colleges are ready to trade-off parental income for ability because they need financial resources. This mechanism is tempered by the fact that the government subsidies to colleges are progressive with slope \((1 - \tau_u)\) as well as by the progressive need-based financial aid, \(\tau_n\), and the social objective of colleges, \(\omega_3\). In theory, this elasticity could be negative, if the social objective parameter is large enough.

The elasticity with respect to ability—associated with the ability-sorting channel—captures the desire of colleges to attract high ability students because of the peer effect. It is amplified by the merit-based subsidy of the government, \(\tau_m\).

### 3.4 Government Budget, Educational Sector and Market Clearing

I outline here the implications of the remaining static equilibrium conditions. The aggregate government budget constraint (17) imposes, in all periods, a restriction on the path of the consumption tax rate \(a_{c,t}\) given an exogenous path of income tax \(a_{y,t}\), higher education subsidies \(a_{h,t}, a_{u,t}\) and endogenous saving rate \(s_t\). Analytic expressions for this constraint as well as for equations (18),(19) and (20) defining respectively \(T_{y,t}, T_{e,t}\) and \(T_{u,t}\) are derived in appendix A.3.1.

In the simple case considered here, the price of educational services is pinned down by the relative productivity parameter \(p_f = 1/A_I\). Market clearing then simply pins down the share of final good allocated to consumption and to the production of educational services. The more general and realistic case with a fully microfounded educational sector is described in detail in A.3.3.

### 3.5 Law of Motion of Human Capital

Having described the static equilibrium conditions, I now derive the law of motion for the distribution of human capital. Since the first two moments of this distribution are the only aggregate states, it also describes the dynamics of the aggregate economy. I start with the law of motion of human capital at the individual level.
Intergenerational Transmission of Status. Plugging the expression for the equilibrium sorting rule (25) into the law of accumulation of human capital (4) and gathering all terms in $\ln h$ (and dropping the generation subscript when no confusion results) gives the following intergenerational law of motion of human capital:

$$\ln h_{t+1} = \alpha_{h,t} \ln h_t + \ln \xi_y + (\alpha_1 + \alpha_2 \varepsilon_{A,t}) \ln \xi_b + \ln \kappa + \alpha_2 x_t$$

with $\alpha_{h,t}$ the intergenerational elasticity of human capital and $x_t$ an aggregate variable.\footnote{\textsuperscript{16}}

The IGE is a linear combination of the before, during and after college transmission of human capital. This paper really focuses on and opens the box of the transmission of economic status through college. The transmission during college decomposes itself into the two sub-channels introduced in the previous paragraph: the income-sorting channel that emphasizes the role that parental income and the ability-sorting channel that emphasizes the role that ability plays in the sorting of students across the ladder of college quality.

In the case without government intervention in higher education, $\tau_{u,t} = \tau_{n,t} = \tau_{m,t} = 0$, the income-sorting channel depends on the difference $\varepsilon_{1,t} - \varepsilon_{3,t}$ and the returns to human capital $\lambda$, while the ability-sorting channel depends only on the strength of the peer-effect $\varepsilon_{2,t}$

$$\alpha_{h,t} = \alpha_1 + \alpha_2 \varepsilon_{A,t}$$

If one additionally sets the peer effect and the social objective parameters to zero, $\omega_1 = 1, \omega_2 = \omega_3 = \tau_{u,t} = \tau_{n,t} = \tau_{m,t} = \alpha_3 = 0$, the IGE is the same as in Benabou [2002]\footnote{\textsuperscript{17}}

$$\alpha_{h,t} = \alpha_1 + \alpha_2 (1 - \tau_{y,t}) \lambda_t$$

Need-based financial aid and progressive subsidies to colleges mitigate the income sorting channel. Merit-based financial aid on the contrary reinforces the ability-sorting channel:

$$x_t = \varepsilon_{1,t} \ln (1 - \tau_a) + \ln (1 + a_{h,t}) + \ln (1 + a_{n,t}) \ln (1 - \tau_{u,t}) + (\varepsilon_{1,t} - \varepsilon_{3,t}) \ln (1 - \tau_{y,t}) + \ln (1 - a_{y,t}) \ln (1 + a_{y,t}) \ln (1 - \tau_{u,t}) + \ln (1 - \tau_{u,t}) \ln \kappa_{2,t}$$

\footnote{\textsuperscript{16} In section 3 of his paper, he separately derives the equilibrium allocation with progressive education subsidies, which is the case $\tau_y = 0$ and $\tau_n > 0$ in my model, \textit{i.e.} $\alpha_{h,t} = \alpha_1 + \alpha_2 (1 - \tau_{n,t}) \lambda_t$.}
Aggregate Law of Motion of Human Capital. Using the assumption of log-normality of both shocks, (5) and (6), if the economy starts from a log-normal distribution then human capital stays log-normally distributed along the equilibrium path:

**Proposition 3.4.** If \( \ln h_t \sim N(m_{h,t}, \Sigma^2_{h,t}) \) then

\[
\ln h_{t+1} \sim N(m_{h,t+1}, \Sigma^2_{h,t+1}) \tag{26}
\]

\[
m_{h,t+1} = \rho_t m_{h,t} + X_1(m_{h,t}, \{\Sigma_h\}_{s=t}^\infty) \quad \text{Mean of (log) Human Capital} \tag{27}
\]

\[
\Sigma^2_{h,t+1} = (\alpha_{h,t}(\Sigma_{h,t}))^2 \Sigma^2_{h,t} + X_2(t)(\Sigma_{h,t}) \quad \text{Variance of (log) Human Capital} \tag{28}
\]

where

\[
\rho_t = \alpha_1 + \alpha_3 + \alpha_1 \alpha_2 \omega_2 + \alpha_2 \omega_1 \lambda_t
\]

\[
X_1(t)(m_{h,t}, \{\Sigma_h\}_{s=t}^\infty) = -\frac{\sigma^2_y}{2} + \ln \kappa \
+ \frac{\tau_{u,t}}{1 - \tau_{u,t}} \left( \frac{\alpha_1(1 - \tau_{u,t})}{(1 - \nu_t(\Sigma_{h,t})t)} \right)^2 \left( \frac{\omega_1(1 - \tau_{u,t})}{\tau_{u,t} + \omega_2(1 - \tau_{n,t})\nu_t(\Sigma_{h,t})} \right)^2 \\
- \alpha_1 \left( \alpha_2 \left( \omega_2 + \omega_1(1 - \tau_{u,t})(\tau_{m,t}^2 \alpha_1) + 1 \right) \right) \frac{\sigma^2_y}{2}
\]

\[
+ \alpha_2 \omega_1 \left( \ln A^\mu_t(\{\Sigma_h\}_{s=t}^\infty)(1 - a_{y,t})s_t(\{\Sigma_h\}_{s=t}^\infty)(1 + a_{u,t})(1 + a_{h,t}) - \ln p_{I,t}(m_{h,t}, \Sigma_{h,t}) \right)
\]

\[
+ \alpha_2 \omega_1 \left( \ln A^\mu_t(\{\Sigma_h\}_{s=t}^\infty)(1 - a_{y,t})s_t(\{\Sigma_h\}_{s=t}^\infty)(1 + a_{u,t})(1 + a_{h,t}) - \ln p_{I,t}(m_{h,t}, \Sigma_{h,t}) \right)
\]

\[
X_2(t)(\Sigma_{h,t}) = \sigma^2_y + (\alpha_1[1 + \alpha_2(\varepsilon_{2,t}(\Sigma_{h,t}) + \tau_{m,t}(1 - \tau_{u,t})\omega_{1,t}(\Sigma_{h,t}))])^2 \sigma^2_b.
\]

In general the expression (27) is not a linear recursive formulation for the law of motion of \( m_h \) because \( s \) and \( \ell \) are forward looking variables that depend on all the future \( \Sigma_h \)'s via the \( \varepsilon \)'s. In contrast, the law of motion of \( \Sigma_h \), given by (28), is recursive—although in general not linear since both the autoregressive coefficient and the shifter depend on \( \Sigma_h \). The full system, (27) and (28), is therefore block-recursive which allows us to characterize the existence and uniqueness of the equilibrium path in section 3.7.

It is intuitive that the shifter in the law of motion of the mean of the distribution (27) is increasing in the saving rate \( s_t \), labor supply \( \ell_t \), subsidies and financial aid \( a_{u,t}, a_{h,t} \) but decreasing in the price of educational services \( p_{I,t} \). Regarding the law of motion of the variance (28), it is the mathematical expression of the Great Gatsby curve: the positive relationship
between the level of inequality $\Sigma_h$ and the strength of the intergenerational transmission of status, $\alpha_h$.

3.6 Distribution of Students along the Quality Ladder and Within-College Distribution of Students

Recall facts (b) and (c) noted in introduction: the dispersion of expenditures per student across colleges has increased and the share of low-income students at top colleges has stagnated. One can actually derive analytical expressions for the distribution of students across college qualities (and the implied distribution of expenditures) and for the within-college distributions of parental income and student ability. These closed-form solutions enable us to shed light on the forces that determine these two objects and will prove useful for the derivation of comparative statics in the next section. These three distributions are log-normal and their first and second moments depend on the aggregate states, directly and indirectly through the income-sorting and ability-sorting elasticities,

$$
\begin{align*}
\varepsilon_{I,t} &= (\varepsilon_{1,t} (1 - \tau_{u,t}) (1 - \tau_{n,t}) - \varepsilon_{3,t})(1 - \tau_{y,t}) \lambda_t \\
\varepsilon_{A,t} &= \alpha_1 (\varepsilon_{2,t} + \varepsilon_{1,t} (1 - \tau_{u,t}) \tau_{m,t}).
\end{align*}
$$

As the following proposition establishes, the dispersion of qualities is an increasing function of both of these variables. But the dispersion of parental income and and abilities within a college will be a function of their ratio. The former is increasing with the ratio $\varepsilon_A/\varepsilon_I$ while the latter is decreasing: the more students sort into colleges based on parental income, the less social diversity there is in a college and the more students sort into colleges based on abilities, the less abilities heterogeneity.

**Proposition 3.5.** 1. The distribution of college quality is given by

$$
\ln q \sim \mathcal{N} \left( \mu_{1,t} (m_{h,t}, \Sigma_{h,t}), \sigma_{1,t}^2 (\Sigma_{h,t}, \varepsilon_{I,t} (\Sigma_{h,t}), \varepsilon_{A,t} (\Sigma_{h,t})) \right)
$$

and $\sigma_{1,t}$ is increasing in $\varepsilon_{A,t}$, $\varepsilon_{I,t}$ and $\Sigma_{h,t}$.

2. Within a college of quality $q$, the distribution of parents’ (log) income is given by:

$$
\ln h|q \sim \mathcal{N} \left( \mu_{2,t} (q, m_{h,t}, \Sigma_{h,t}), \sigma_{2,t}^2 (\Sigma_{h,t}, \varepsilon_{I,t} (\Sigma_{h,t}), \varepsilon_{A,t} (\Sigma_{h,t})) \right)
$$

with $\mu_2$ increasing in $q$ and $\sigma_{2,t}$ increasing in $\varepsilon_{A,t}$ and $\Sigma_{h,t}$ but decreasing in $\varepsilon_{I,t}$.
3. And the distribution of students’ (log) abilities is given by
\[ \ln h_s | q \sim \mathcal{N}(\mu_3, \tau_m, \Sigma_{h,t}, \sigma^2_{3,t}(\Sigma_{h,t}, \varepsilon_{l,t}(\Sigma_{h,t}), \varepsilon_{A,t}(\Sigma_{h,t})) \]
with \( \mu_3 \) increasing in \( q \) and \( \sigma^2_{3,t} \) increasing in \( \varepsilon_{l,t} \) and \( \Sigma_{h,t} \) but decreasing in \( \varepsilon_{A,t} \).

3.7 Existence and Uniqueness of the Equilibrium Path

We are now ready to characterize the existence and uniqueness properties of the steady-state and macroeconomic equilibrium path, within the class of equilibria with log-normally distributed human capital. This is important since our goal is to derive comparative statics with respect to the returns to education \( \lambda \) and to policy parameters. I first gather the equations characterizing it.

Lemma 1 (Equations for Equilibrium Path). An equilibrium path is a sequence of exogenous variables \( (a_y,t, a_u,t, a_h,t, \tau_{y,t}, \tau_{n,t}, \tau_{m,t}, \tau_{u,t}, \lambda_t) \), initial aggregate states \( (m_{h,0}, \Sigma_{h,0}) \) and endogenous variables \( (s_t, \ell_t, U_t, m_{h,t}, \Sigma_{h,t}, h_t, p_{I,t}, a_{c,t}) \) such that \( (22) \)—\( (24), (27) \)—\( (43), (53) \) and \( (54) \) hold.

Although existence and local stability is obtained under an intuitive sufficient condition, global stability is harder to prove. A sufficient condition is that \( \omega_3 \) be small enough.

Proposition 3.6. Existence and Uniqueness of Equilibrium Path

- If \( \lim_{\Sigma_h \to \infty} \alpha_h(\Sigma_h) < 1 \), there exists at least one locally stable steady-state.
- For \( \omega_3 \) small enough, there exists a unique globally stable steady-state and a unique equilibrium path.

with \( \lim_{\Sigma_h \to \infty} \alpha_h(\Sigma_h) = \alpha_1 + \alpha_1 \alpha_2(\omega_2 + \tau_m(1 - \tau_u)\omega_1) + \alpha_2[\omega_1(1 - \tau_u)(1 - \tau_n) - \omega_3](1 - \tau_y)\lambda \)

A high \( \omega_3 \) might lead to multiple equilibria by making inequality \( \Sigma_h \) potentially grow too fast in some parts of the state-space, i.e. by making the derivative of the right-hand-side of \( (28) \) higher than 1, thus failing to meet the crucial defining feature of a contraction mapping. This stems from the fact that \( \nu \) is increasing in \( \Sigma_h \), hence that \( \varepsilon_l \) for \( l = 1, 2, 3 \), \( \alpha_h \) and \( X_2 \) are increasing in \( \Sigma_h \).

4 Rationalizing Recent Trends in Higher Education

This section derives what I consider to be the most important analytical result of the paper. The increase in the market returns to education \( \lambda \) is able to generate (a) the increase in
income inequality [Piketty and Saez, 2003]; (b) the increase in the dispersion of expenditures per students across colleges [Capelle, 2019c]; (c) the stagnation of the share of students from the lowest income quintile in top colleges despite the increase in financial aid [Bailey and Dynarski, 2011, Chetty et al., 2017]; (d) the increase in real terms of tuition fees before and after financial aid (figure A4 in appendix); and (e) the slight increase in the intergenerational elasticity of income mobility [Black and Devereux, 2011, Davis and Mazumder, 2017, Urrutia and Tavares, 2016]. It is natural to focus on the increase in the returns to education as it is widely recognized to be one of the main sources of the increase in inequality [Katz and Murphy, 1992, Autor et al., 2008].\footnote{I do not take a stand on the exact source of increase in the returns to human capital. It may come from a combination of a skill-biased technological change [Acemoglu, 2002], an improvement in the assortative matching of workers and firms [Kremer and Maskin, 1996, Song et al., 2018], an increase in assortative mating and in the number of single households [Greenwood et al., 2014, Heathcote et al., 2010] and an increase in the substitutability across skills due to international trade [Grossman and Maggi, 2000] or due to better communication technology [Garicano and Rossi-Hansberg, 2006].}

I also show that a decline in the progressivity of government subsidies to colleges—an observed feature of the data—is able to rationalize the same facts. However, the quantitative analysis in section 7 suggests that this change would fail to match the large changes in income inequality observed in the data. Additionally, from a theoretical perspective, the first comparative static is more complex and less intuitive and therefore more interesting.

4.1 An Increase in the Returns to Human Capital

The following proposition formally states the key comparative static result.

**Proposition 4.1.** Assume the economy starts from a steady-state at \( t = 0 \). Consider a weakly increasing sequence \( \{\lambda_t\}^\infty_0 \). If \( \omega_1(1-\tau_n)(1-\tau_u) > \omega_3 \), along the equilibrium path,

a) The Gini coefficient of human capital and income increase.

b) The Gini coefficient of colleges’ (log) expenditures per student (and quality) increase.

c) The ratio of variance of (log) income within a college over variance of (log) income in economy decreases.

d) The intergenerational elasticity increases.

e) The average expenditure for college as a share of income increases.

The formal proof of this proposition is contained in appendix A.8. Here I present intuition for the stated effects.

Intuitively, when the returns to human capital, \( \lambda \), increase, the dispersion of households’ income rises for a given distribution of human capital (fact (a)). Given that households all
spend the same share of their income for the higher education of their child, it implies an increase in the dispersion of desires to pay for college. Following this change on the demand side of the higher education market, colleges react: top colleges take advantage of the rising willingness to pay of their pool of students by increasing their fees and their spending relative to colleges at the bottom. Inequality of revenues and spending across colleges rises (fact (b)).

Poor but high ability students get priced out of top colleges following the relative rise of their tuition fees. More generally this rise in the dispersion of tuition for colleges implies that parental income matters, from now on, even more to access a higher quality college than it used to, relative to ability. It corresponds to an increase in the elasticity of college quality to income, \( \varepsilon_I \), what I described earlier as a strengthening of the income-sorting channel.

Consequently, top colleges become less diverse in terms of economic background because poor students are priced out and students from rich families can buy their way to the top. More generally, colleges become more segregated and homogeneous in terms of parental income (fact (c)). Another implication is that intergenerational mobility decreases, as parental income become increasingly determinant for the opportunities of children (fact (d)). This is a direct manifestation of the Great Gatsby curve [Corak, 2013, Krueger, 2012], whereby an increase in income inequality leads to a strengthening of the transmission of economic status, here through access to better quality higher education, which feeds back into higher inequality...

Overtime, the initial increase in inequality gets amplified through the higher education system. Students from richer background get relatively higher quality higher education, which increases the dispersion of human capital and therefore of income once their generation becomes adult. This circle continues at the next generation as this increased dispersion of human capital translates into a higher dispersion of children abilities which gets amplified by the increasingly unequal ability and desires of parents to pay for college. And so on and so forth.\(^{19}\)

The amplification through colleges happens through two channels: the reallocation of resources and the reallocation of students. Financial resources and expenditures become increasingly concentrated at the top of the college distribution. In contrast, able students become slightly less concentrated at the top of the college ladder, partially mitigating the amplification.

Why do colleges accommodate the increased dispersion in desires to pay for colleges? They are led to do so by their desire to maximize the quality they provide, and so despite their social objective. Even if an individual college at the top of the distribution didn’t raise

---

\(^{19}\)The quantitative section provides a structural decomposition of the increase in the Gini coefficient and in the IGE into the direct impact and the amplification through the higher education system.
its tuition fees relative to, say, the median college, another college would fill up this gap, offering higher quality for higher tuition fees. This mechanism is akin to the revenue theory of cost by Bowen [1980], but now applied to a hierarchy of colleges.\textsuperscript{20}

Finally, the reason why the average tuition fees and the share of total income devoted to higher education increase has to do with the rising incentives of households to accumulate human capital, in an economy in which its returns have increased (fact (e)). It is therefore the same demand-driven mechanism that drives both the average increase in tuition and the rise in inequality across colleges.

### 4.2 A Decrease in Public Transfers Progressivity across Colleges

In a companion paper Capelle [2019a], I have documented the large decline in the average rate of government subsidy to colleges, $a_u$, as well as in its progressivity, $\tau_u$. In this section, I explain intuitively why a weakly decreasing sequence $\{\tau_{u,t}\}_{0}^{\infty}$ has the same qualitative effects as an increasing sequence $\{\lambda_t\}_{0}^{\infty}$ as states in proposition 4.1.\textsuperscript{21}

For a given distribution of tuition fees across colleges, a decrease in the progressivity of public subsidies, $\tau_u$, leads to an increase in the dispersion of financial resources and therefore quality across colleges (fact (b)). Mechanically, because of this decline in the progressivity of public subsidies, the college quality ladder becomes steeper and the sensitivity of the quality of higher education to parental income increases. Moreover, the decline in $\tau_u$ implies that, from the point of view of colleges, the marginal productivity of tuition fees in terms of quality has increased relative to students ability. This gives incentives to colleges to target students with a higher desire to pay, at the expenses of high ability students. Overall, this implies a strengthening of the income-sorting channel.\textsuperscript{22}

Similarly to the case with the increase in $\lambda$, top colleges become less diverse in terms of family background and colleges become more segregated by parental income (fact (c)). Intergenerational mobility decreases, as parental income become increasingly determinant for the opportunities of children (fact (d)). The initial change in the allocation of college quality to students gets amplified over generations: because they experienced a more unequal

\textsuperscript{20}Bowen summarizes his theory page 19:

1) The dominant goals of institutions are educational excellence, prestige, and influence. 2) In quest of excellence, prestige, and influence, there is virtually no limit to the amount of money an institution could spend for seemingly fruitful educational ends. 3) Each institution raises all the money it can. 4) Each institution spends all it raises. 5) The cumulative effect of the preceding four laws is toward ever increasing expenditure.

\textsuperscript{21}The formal proposition and proof are given in appendix A.8.

\textsuperscript{22}Formally, it is easy to see that it also strengthens the ability-sorting channel—by increasing the impact of merit-based subsidy.
distribution of higher education, the next generation of households is more unequal in terms of human capital and therefore income. This translates into higher inequality of abilities of their children, and so on and so forth... The economy gradually shifts rightwards (higher inequality, lower mobility) along the Great Gatsby curve.

Finally, the decrease in $\tau_u$ incentives households to invest more in higher education, since the elasticity of quality to tuition is now higher, which leads to an increase in the average spending rate $s$ and in average tuition.

5 Normative Analysis: Sorting, Efficiency and Welfare

One important dimension of efficiency in this economy is how students and educational resources are allocated across colleges: a sub-optimal allocation of students and resources across colleges today entails lower output for the next generation. Is there misallocation in the competitive equilibrium? If so, what are the nature and the roots of the misallocation? Is it possible for a government deciding upon taxes to restore the efficient sorting of students and resources?

This section shows that in any social welfare optimum, there is perfect positive assortative matching of students by ability across colleges. Sorting should be independent of parental income, holding student ability fixed. The competitive equilibrium with incomplete financial markets displays too many rich and not so high ability children in good colleges and too many poor but high ability children in low quality colleges. As peer-effects become relatively more important than educational services in the production function of quality by colleges, the sorting of students improves because universities have a stronger incentive to attract high ability children. Perfect positive assortative matching is achieved in the extreme case where only peer-effects matter.

Financial market incompleteness is at the root of this misallocation. A competitive economy with complete financial markets generated Pareto optimal allocations: high ability but poor children lack access to liquidity to go to good colleges and rich but not very able children over-accumulate human capital because they can’t get direct financial transfers from their parents. Solving a second-best Ramsey problem, the government could restore the static optimal sorting with a mix of progressive need-based and merit-based tuition subsidies. But in a dynamic setting, the government trades off the optimal allocation of students across colleges, the insurance of individuals against birth and labor market shocks and the (dis-)incentives created by distortive taxes. In our framework, it turns out that replicating the first-best sorting is never optimal.

\footnote{In particular, the presence of a within-college externality (the peer-effect) doesn’t introduce any inefficiency.}
5.1 Social Welfare Optima

I consider a social planner who maximizes a linear social welfare function:

\[ \ln \mathcal{U}^P(\psi(h, \xi_y)) = \max_{(q,c,\ell)} \int_{i} \omega(i) \left[ \ln c(i) - \ell^0(i) \right] di + \beta \ln \mathcal{U}^P(\psi'(h, \xi_y)) \]

where \( \psi(h, \xi_y) \) denotes the joint distribution over human capital and birth shock and \( \omega(i) \) denotes the Pareto weights for dynasty \( i \). It is implicitly assumed that the social planner cares only about household’s utility and not college’ payoffs.\(^{24}\)

The planner chooses for each type \( (h, \xi_y) \) how much to save, work and the quality of college for the child as well as how much to spend in each college, subject to the technological constraints for producing output, higher education quality, as well as the resource constraint and the individual law of motion for human capital.

This seemingly complicated dynamic problem can be decomposed into a static sorting problem of individuals and resources across colleges and a dynamic problem of choosing how many aggregate resources to save and how much to work. In the optimal allocation, there is perfect insurance of consumption across households and there is perfect positive assortative matching of students. The allocation of educational resources is harder to characterize in general, but one can obtain closed-forms expressions in two cases. This is described in the following proposition.

**Proposition 5.1.** Assume \( \bar{\lambda} = \lambda, A_I = A, \eta = +\infty \). Any social pareto optimum features:

1. Consumption: there is perfect insurance of consumption across individuals within a generation. If \( C_t \) denotes aggregate consumption, then \( \forall (i, t) \quad c_t(i) = \omega(i)C_t \)

2. Allocation of students: there is perfect positive assortative matching of students across colleges. The sorting rule is completely independent of parental income and depends only on the student’s ability. Each college is populated by only one type of student.

3. Allocation of resources

\(^{24}\)The view is that colleges are purely instrumental institutions in the process of human capital accumulation. In the same way that firms in competitive markets can be omitted in traditional social planner’s problems, competitive non-profit colleges can be omitted in the present analysis.
In the particular case where \( \frac{1 + \alpha_2 \omega_2}{1 - \frac{\lambda m}{\eta - \mu} \alpha_2 \omega_1} = \frac{1}{\alpha_1} \),

\[
I(h_s) = I_0 h_s^{\epsilon_{ih_s}}
\]

with \( I_0 = \frac{s \alpha_1 \epsilon_{ih_s} \left( m - \frac{\sigma^2}{2} \right) + \left( \alpha_1 \epsilon_{ih_s} \right)^2 \left( \Sigma^2_h + \sigma^2 \right)}{e^{\lambda m + \frac{\lambda^2 \Sigma^2_h}{2} \epsilon_{ih_s}}}
\)

and \( \epsilon_{ih_s} = \frac{\eta \lambda}{\eta - \mu} \left( \frac{1 + \alpha_2 \omega_2}{1 - \frac{\lambda m}{\eta - \mu} \alpha_2 \omega_1} \right)
\)

where \( m \) and \( \Sigma^2_h \) denote respectively the mean and the variance of the distribution of (log) human capital and \( s \) is the aggregate rate of investment in higher education:

\[
s = \frac{\beta \omega_1 \alpha_2 \gamma_{1,t+1}}{1 - \beta + \beta \omega_1 \alpha_2 \gamma_{1,t+1}}
\]

with \( \gamma_{1,t} = (1 - \beta) \sum_{k=0}^{\infty} \beta^k \lambda_t \prod_{m=0}^{k-1} \left( \alpha_1 + \omega_1 \alpha_2 \lambda_{t+m} + \alpha_1 \alpha_2 \omega_2 \right)\)

Or if there are only two periods, the allocation is characterized by:

\[
I(i) = I_0 h_s(i)^{\epsilon_{ih_s}}
\]

\[
I_0 = \beta \frac{C}{\lambda} \frac{\alpha_1 \omega_1 E \left( \xi_y^\lambda \right) \lambda}{\gamma_{1,t+1} \prod_{m=0}^{k-1} \left( \alpha_1 + \omega_1 \alpha_2 \lambda_{t+m} + \alpha_1 \alpha_2 \omega_2 \right)}
\]

where \( \epsilon_{ih_s} \) is the same as in the previous case and \( Y' \) is output in the second period.

**Proof.** See appendix B.1.

\[\square\]

### 5.2 The Role of Financial Constraints in Misallocation

In the first best, the elasticity of college quality to parental income is zero: there is therefore perfect sorting of students across colleges by ability and there is no dispersion of abilities within any college. This is in sharp contrast with the decentralized economy.

What are the roots of this different allocation rule? The first one, the main one, is the incompleteness of financial markets, which includes the impossibility to transfer financial wealth across generations and the uninsurability of birth and labor market shocks. The second one is the social objective of universities. But neither the quality-maximizing behavior of colleges nor the externalities due to the peer-effect introduce additional inefficiencies.

**Proposition 5.2.** Assume \( \bar{\lambda} = \lambda, A_I = A, \eta = +\infty \). The decentralized equilibrium with complete financial markets is Pareto-efficient if and only \( \omega_3 = 0 \).
The reason why the externality embedded in the peer-effect doesn’t cause additional inefficiencies is because it is internalized by the universities and therefore correctly priced in the equilibrium tuition schedule. This proposition implies that there are only two sources of inefficiencies in this framework: the social objective of universities and the incompleteness of financial markets. Regarding the latter, it means that high ability and poor children lack access to liquidity to go to good colleges and parents of rich but not so high ability children over-accumulate human capital because investment in human capital acts as a substitute for precluded financial bequests. This is the source of both the mismatch of students types across colleges but also the misallocation of educational resources. The equilibrium allocation will feature rich but not so high ability children going to the same colleges as higher quality but poorer children. High quality colleges will feature too many rich but no so high ability children while low quality colleges will feature too many high ability but poor children.

5.3 Peer-effect and Mismatch

How does the imperfect positive matching of students depend on the strength of peer-effects, $\omega_2$? How does the incompleteness of financial markets interact with the technology to produce higher education to shape the imperfect sorting? From the equilibrium distribution of abilities within colleges, we know that the degree of mismatch of students across colleges depends on the relative strength of the income-sorting variable $\varepsilon_I$ and the ability-sorting variable $\varepsilon_A$ which depend positively respectively on $\omega_1$ and $\varepsilon_2$.

**Proposition 5.3.** Assume $\bar{\lambda} = \lambda, A_I = A, \eta = +\infty$. In the competitive equilibrium with financial constraint,

1. The variance of abilities within colleges—a measure of mismatch—is decreasing in $\omega_2$ if and only if $\omega_I > 0$.

2. In the extreme case where colleges are pure clubs with no social objective $\omega_2 > 0, \omega_1 = \omega_3 = 0$,
   - the allocation of students and (absence of) resources across colleges is the same as in the first best
   - the incompleteness of financial markets causes inefficiencies only through the lack of insurance of consumption.

**Proof.** See appendix B.4.
As colleges become a pure club (which corresponds to the case \( \omega_2 = 1 \),) educational services and material resources lose their usefulness and matching students of similar abilities is the only thing that matters from an efficiency point of view. Colleges have no incentive anymore to look for rich children and look instead only for high ability students independently of their income which in turn makes the equilibrium tuition schedule infinitely elastic to abilities. An infinitely elastic tuition schedule in the dimension of abilities means that being rich can’t buy a household a good college, only ability can. The resulting allocation is therefore a perfect positive assortative matching of students across colleges.

5.4 Second-best Allocation: Optimal Financial Aid

The government faces a Ramsey problem. It has control over three parameters of financial aid policies: the average aid \( a_h \), the progressivity with respect to parental income \( \tau_n \) and the progressivity with respect to student ability \( \tau_m \). It takes the income tax schedule as given \( (a_y, \tau_y) \) and I ignore, for conciseness, subsidies to colleges \( (a_u, \tau_u) \). It maximizes a similar objective as the social planner:

\[
\ln \mathcal{U}^g(m, \Sigma_h) = \max_{\tau_m, \tau_n, a_h} (1 - \beta) \left[ \ln \left( \int \left( \frac{(1 - s)y}{1 + a_c} \right)^{\frac{\sigma - 1}{\sigma}} \right)^{\frac{\sigma}{\sigma - 1}} - \ell' \right] + \beta \ln \mathcal{U}^g(m', \Sigma_h') \tag{29}
\]

where \( \sigma \) indexes the degree to which the government cares about equity.\(^{25}\) The problem of the government consists in maximizing (29) subject to the optimal saving rate for higher education (22), the optimal labor supply decision (23), the aggregate budget constraint of the government (17) and the law of motion for \( m \) and \( \Sigma^2 \) given respectively by (27) and (28). In appendix B.5, I show that the optimal average financial aid (and the implied optimal consumption tax) is given by

\[
(1 + a_u^*)(1 + a_h^*) = \frac{a_y}{s(1 - a_y)} + \left( a_c^* \frac{(1 - s)}{(1 + a_c^*)s} + 1 \right)
\]

with \( a_c^* = \alpha_2 \omega_1 (1 - s)(1 - a_y) \left( \frac{1}{1 - \rho} - s - a_y \right) \).

It is naturally decreasing in \( s \) the private rate of spending in higher education and increasing in \( \omega_1 \), the elasticity of quality to educational resources.

It is much harder to characterize analytically the optimal progressivity \( \tau_n, \tau_m \). One can show that the government could replicate the perfect positive assortative matching of students

\(^{25}\)This welfare function is the same as the social planner’s when \( \sigma = 1 \) and when the Pareto weights are all equal \( \varepsilon_I = 1 \).
with a sufficiently progressive financial aid policy

\[ \tau^*_n = 1 - \frac{\omega_3}{\omega_1}. \]

In general there is no \( \tau^*_m \) that restores the first best allocation of resources across colleges. In the specific cases where \( \frac{1 + \alpha_2 \omega_2}{1 - \frac{\alpha_1}{\alpha_2} \omega_1} = \frac{1}{\alpha_1} \), it could be achieved by setting

\[ \tau^*_m = \varepsilon h_s. \]

Even in this very special case, the government would not be willing to set this \( \tau_n \) and \( \tau_m \) and implement the sorting of the first best, for two reasons: 1-they have distortive implications for labor supply/leisure if \( \eta < +\infty \), 2-the lack of insurance of consumption might urge the government to use the higher education system as an insurance mechanism, by mixing types—sending unlucky children, those from low human capital background, and those with low birth shock, to relatively better colleges—and decreasing the steepness of the allocation of educational resources along the quality ladder. The latter motive is reinforced by the equity concerns of the government.

Figure 1 illustrates the policy trade-offs faced by the government. The dotted line captures a measure of welfare that takes into account only the aggregate stock of human capital, the dashed line takes into account aggregate output net of leisure cost and the solid line is a measure of welfare that also takes into account the cost stemming from the lack of insurance as well as the equity motive. Maximizing the stock of human capital requires an interior \( \tau_n \) and \( \tau_m \) to correct for the misallocation of resources and students in the laissez-faire equilibrium. Taking into account labor supply and leisure implies a further increase in \( \tau_n \)—because individuals work too much in the laissez-faire—and taking into account the need for insurance and the motive for equity further requires an increase in \( \tau_n \) to send children from poor background to better colleges and a decrease in \( \tau_m \) to avoid much inequality in college experiences.

The third panel shows how welfare varies when both \( \tau_n \) and \( \tau_m \) are allowed to move. These two kinds of subsidies display complementarities: the increase of one ask for an increase in the other, because \( \tau_m \) increase inequality and decrease insurance, while \( \tau_n \) correct for it.
Taking the Model to the Data

In the previous sections I developed a tractable model of human capital accumulation with a ladder of colleges which allows for a sharp analytical characterization of the equilibrium sorting of students across colleges, income inequality, intergenerational mobility and aggregate production. The next two sections are devoted to the quantitative assessment of the role of higher education in shaping inequality and intergenerational mobility. This section has two goals. First I relax some of the assumptions I adopted for tractability in the previous sections and extend the model to a richer quantitative environment. Second I explain how I estimate this richer model, which I then use in section 7 to assess the quantitative relevance of the higher education system.
6.1 Quantitative Extension

I extend the model in two dimensions. First the restrictions on intergenerational financial transfers are partially relaxed: negative transfers up to a limit are allowed (student debt not offset by parental transfers) as well as positive transfers (bequests):

\[
\ln \mathcal{U}(h, h_s, a) = \max_{c, \ell, q, a'} \{ (1 - \beta) [\ln c - \ell^0] + \beta E \ln (\mathcal{U}(h', h'_s, a')) \} \tag{30}
\]

\[y + e^{rH} a = c(1 + a_c) + e(q, y, h_s) + a' \tag{31}\]

\[a' \geq a \tag{32}\]

where \(e^{rH}, r, H\) denote respectively the "generational" rate of return, the annual rate of interest and generation length and \(a\) is the exogenous borrowing limit.

Households therefore face a portfolio problem: they have to decide upon the optimal combination of bequest and higher education for their offspring. High ability children from a poor background will take up loans and rich families with low ability children will choose to transmit financial wealth instead of buying a very high quality college for their kid. Overall, allowing for financial transfers should weaken the link between parental income and the child’s position on the college ladder.

Secondly, there is an outside option delivering \(q\) for free:

\[e(q, y, h_s) = 0 \quad \forall (h_s, y) \tag{33}\]

Some individuals will find it optimal not to go to college and take up the free outside option. This gives rise to a meaningful enrollment decision that was absent from the previous framework where all individuals got at least some arbitrarily low quality of higher education. A direct implication of equation (33) is that, if \(q > 0\), in equilibrium no individual ever chooses \(q < \bar{q}\) and there is a Dirac peak at \(\bar{q}\). It is natural to define the enrollment rate as the share of individuals with \(q > \bar{q}\).

The set of technological constraints faced by the household is otherwise similar to the original problem described in section 2. Formally, the problem of the household consists in maximizing (30) subject to (13), (14), (2)-(6) and the new constraints (31)-(33). The rest of the model remains the same.\(^{26}\) The original problem is the special case when

\[a' = q = 0\]

\(^{26}\)I explain how the problem of the colleges is kept tractable in this more complicated framework in appendix C.1.
For future reference, I call M1 the model studied in sections 2 and 3 and M2 the augmented model described in this section. In the remainder of the paper, and unless otherwise stated, the expression "income Gini coefficient" refers to the "labor earnings Gini coefficient", the key object of interest of our analysis.

6.2 Data

The core dataset is the restricted-use version of the NLSY-1997, a representative panel of individuals who were 12 to 17 years-old in 1997, whom I follow every year up to now. It features data on parental income, abilities measured by a common comprehensive test-score, the Armed Services Vocational Aptitude Battery or ASVAB, a detailed description of their journey through the higher education system—each college they attended, the time spent and the degree obtained—and their labor earnings.

To estimate the parameters related to financial aid, I use the restricted-use NCES-NPSAS in 2000, which is the closest survey to the average year when individuals in the NLSY go to college. It is a representative survey of students that features detailed information about parental income, out-of-pocket college costs and financial aid disaggregated by source—federal government, state, private and institutional.

The publicly available NCES-IPEDS annual surveys provide college-level information on expenditures, revenues, enrollment and the distribution of test scores within each college. I use the 2000 to 2004 surveys. Finally I complement these data with statistics on enrollments from the NCES and measures of aggregate spending for higher education from the OECD.

6.3 External Calibration

The full list of the nineteen parameters that need to be calibrated is given in the first column of table 1. Seven of them I set without solving the model while the remaining twelve are calibrated solving it. I provide here a quick overview of the procedure for the former. Appendix C.2 provides more details.

The income tax schedule parameters $a_y, \tau_y$ are informed by the average income tax rate and the slope of the income tax schedule estimated by Heathcote et al. [2017a]. In a companion paper, I estimate the college subsidy parameters $a_u, \tau_u$, Capelle [2019c]. They are respectively the average subsidy received by college per student and the slope of the transfers to colleges schedule. I use the average financial aid received by students to calibrate $a_h$.

I use estimates of the Frisch elasticity of labor supply from the literature to calibrate $\eta$. The elasticity of substitution in the educational services sector is set to $\bar{\lambda} = \lambda$ so that the
price of educational services is given by \( p_t = \frac{A_t}{M_t} \). The generation length is set to \( H = 30 \) years. The lower limit, \( \underline{a} \) is set to match the official borrowing limit for student loan.

### 6.4 Internal Calibration: General Strategy

The algorithm used to estimate the parameters is akin to a Simulated Method of Moments. If \( \theta \) denotes the vector of the twelve parameters to be estimated, \( M(\theta) \) the vector of model-generated moments, \( \hat{M}(\theta) \) one such realization and \( \hat{m} \) the vector of empirical moments, one seeks to find \( \hat{\theta} \) such that

\[
\hat{\theta} = \arg \min_{\theta} \left[ \hat{m} - \hat{M}(\theta) \right]^T W \left[ \hat{m} - \hat{M}(\theta) \right]
\]  

(34)

with \( E \left[ \hat{M}(\theta) \right] = M(\theta) \) and \( W \) a weighting matrix. I depart from this standard expression in two ways. Firstly the empirical moments themselves \( \hat{m} \) are function of a subset of parameters, \( \theta_1 = \{ \alpha_1, \omega_2 \} \). Many of the targeted moments are coefficients in a regression involving one of two variables that need to be constructed using \( \theta_1 \): students ability \( h_s \) and college quality \( q \). Secondly, as is explained in details in appendix C.7, the procedure to generate the student ability variable, based on test scores, isn’t a deterministic function of the observables and the parameters but involves some randomness. Our estimation procedure actually takes the following form:

\[
\hat{\theta} = \arg \min_{\theta} \left[ \hat{m}(\theta_1, \varepsilon) - \hat{M}(\theta) \right]^T W \left[ \hat{m}(\theta_1, \varepsilon) - \hat{M}(\theta) \right]
\]  

(35)

where \( \varepsilon \) denotes the noise introduced in the process of constructing the student ability variable.

Despite the large dimension of the parameter space, the algorithm is quick and the global minimum to the loss function can be easily and with certainty found. I proceed in two steps. First I estimate M1, the version of the model without outside option and financial asset. The closed-form solutions to the model enables me to compute the exact value for the moments—\( \hat{M}(\theta) = M(\theta) \)—and to run the estimation on a fine and large grid. This in turn allows me to check numerically that the parameters are well identified, in the sense that the loss function is steep at the global minimum, which I do in C.9. The estimates are reported in the third column of table 1. I then estimate the augmented version M2, the version of the

---

\( ^{27} \)It would certainly have been possible to find a transformation of the empirical moments that would have been independent of the parameters \( \theta_1 \), such that (35) could have been rewritten according to (34). However it was difficult given administrative restrictions associated with using multiple restricted-use data located on different servers to extract these moments. Moreover, the present estimation strategy also has the advantage of targeting moments that have a closed-form expression and interpretation in M1.
model with an outside option and a financial asset. In this second step, I simulate the model over a smaller grid chosen close to the estimates of the first step. The key assumption here is that the parameters estimated with M1 are not too far from the true parameters in M2. It turns out that almost all estimates but one, $\omega_2$, are identical to their counterpart in M1. The results are reported in the fourth column of table 1.

6.5 Internal Calibration: Moments for Identification

I next make a heuristic identification argument that justifies the choice of moments used in the estimation. Although no parameter can be identified out of a single moment, I will stress in this section which moment is important for each parameter. Thanks to the closed-form expressions of these moments in terms of structural parameters in M1, it is possible to formalize this argument, which I do in appendix C.4.\textsuperscript{28}

Assume for a moment that one perfectly observes child ability and college quality $\{h_{s,i}, q_i\}$.

I first estimate the financial aid schedule (15), and use the elasticity of government financial aid to parental income and to students ability to inform the slopes of the financial aid schedule, which pins down $\tau_n$ and $\tau_m$. I estimate the tuition schedule by running a regression of before-government-aid tuition fees on a college fixed-effect, ability and parental income. I use the elasticity with respect to parental income to inform the social objective parameter $\omega_3$. I then estimate the sorting rule by running a regression of college quality on students ability and parental income. The elasticity of college quality to students ability has a first-order effect on the peer-effect parameter $\omega_2$.

I then estimate the human capital accumulation function (4), and the market earnings function (13), to recover $\alpha_1, \alpha_2, \alpha_3$ and the returns to human capital $\lambda$. I use the intergenerational elasticity of income to inform $\alpha_1$. I run a regression of (log) child earnings on (log) abilities, (log) college quality and (log) parental income. The elasticity of a child’s income to their ability identifies $\lambda$. Conditional on $\lambda$, the elasticity of child’s income to college quality (resp. parental income) identifies $\alpha_2$ (resp. $\alpha_3$). Similarly, the elasticity of child’s income to parental income identifies $\alpha_3$.

Given how important $\alpha_2$ is in the propagation of shocks from higher education onto the macroeconomic system and the uncertainty surrounding its value, I will present results for a range of plausible $\alpha_2$. For every $\alpha_2$ in this range, I re-estimate the model to minimize the

\textsuperscript{28}Another advantage of the closed-form expressions in M1 is the ability to investigate the invertibility of the model, given a set of targeted moments. It is possible to show that if abilities are directly observable and there is no social objective $\omega_3 = 0$, then the parameters are exactly identified. Although child’s ability $h_s$ are non-observable, it helps build confidence in the identification of the parameters in my procedure. The proposition and more details are provided in appendix C.10.
distance with the targeted moments, and forcing the elasticity of the child’s income to their college quality to be compatible with the assumed $\alpha_2$. The range of $\alpha_2$ considered is $[0, .35]$.

However, child ability and college quality are not observable. I first explain how I construct child’s ability. All children in the NLSY take the same test in high school. I assume that the resulting test scores are ranked in the same order as ability, $h_s$. Conditional on a $(\alpha_1, \lambda)$ it is possible to show that the correlation between the rank of the test scores and parental income $\text{corr}(\text{rank}(h_s), y)$ identifies the variance of the birth shock, $\sigma_b^2$. I then generate model-consistent abilities that have the following properties (i) they preserve the ranking of test scores, (ii) they are compatible with the distributional assumption for the birth shock (5) where $\sigma_b^2$ has just been estimated and (iii) they are compatible with the functional form for the transmission process (2). The construction of college quality is more direct. I use the information about which college each child has attended, average test scores and educational spending in each college and the assumed production function for quality (9) to construct the model-consistent variable $q_i$.

The Gini coefficient of income is used to inform the variance of labor market shocks, $\sigma_y^2$. The intergenerational rate of preference, $\beta$, is strongly related to the share of private spending in higher education in GDP. The outside option to college, $q$, is directly related to the enrollment rate, the lower $q$, the stronger the incentives to go to college. Recall that my model takes $r$ as exogenous. Changing $r$ has an impact on the incentives to accumulate the financial asset and in steady-state on the mass of households close to the borrowing constraint. $r$ has consequently a first order effect on the elasticity of college quality to parental income.\textsuperscript{29}

The following table summarizes the targeted moments and the resulting parameters in M1 and M2.

\textsuperscript{29}Notice that this elasticity is otherwise too high in the M1 calibration.
### Table 1: Parameters and Moments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Target/Source</th>
<th>Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta$</td>
<td>(Inv.) elast. labor</td>
<td>2</td>
<td>Chetty et al. [2011], Own Comput.</td>
<td>Data</td>
</tr>
<tr>
<td>$\tau_y$</td>
<td>Income Tax Slope</td>
<td>.23</td>
<td>Heathcote et al. [2017b], Own Comput.</td>
<td>M1</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Return to human capital</td>
<td>.67</td>
<td>Own Comput.</td>
<td>M2</td>
</tr>
<tr>
<td>$a_u$</td>
<td>Av. Transfer to College</td>
<td>.4</td>
<td>Av. Transfer to College</td>
<td>id.</td>
</tr>
<tr>
<td>$a_y$</td>
<td>Av. Income Tax Rate</td>
<td>.2</td>
<td>Av. Income Tax Rate</td>
<td>id.</td>
</tr>
<tr>
<td>$a_h$</td>
<td>Av. Financial Aid</td>
<td>.2</td>
<td>Av. Financial Aid</td>
<td>id.</td>
</tr>
<tr>
<td>$\tau_u$</td>
<td>Elas. Transfers to Coll.</td>
<td>.35</td>
<td>Elas. Transfers to Coll. [Capelle, 2019c]</td>
<td>M1</td>
</tr>
<tr>
<td>$\bar{a}$</td>
<td>Borrowing Limit</td>
<td>(0)</td>
<td>Bequest Limit</td>
<td>M2</td>
</tr>
<tr>
<td>$\bar{\alpha}$</td>
<td>Bequest Limit</td>
<td>+\infty</td>
<td></td>
<td>id.</td>
</tr>
<tr>
<td>$\tau_n$</td>
<td>Elas. Gov. Fin. Aid to $y$</td>
<td>.195</td>
<td>Elas. Gov. Fin. Aid to $y_m$</td>
<td>M1</td>
</tr>
<tr>
<td>$\tau_{m}$</td>
<td>Elas. Gov. Financial Aid to $h_s$</td>
<td>.07</td>
<td>Elas. Gov. Financial Aid to $h_s$</td>
<td>M2</td>
</tr>
<tr>
<td>$\omega_3$</td>
<td>Social Obj. Param. of Coll.</td>
<td>0</td>
<td>Elas. Tuition to $y$</td>
<td>id.</td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>Elas. $q$ to Average Ability</td>
<td>.84</td>
<td>Elas. $q$ to $h_s$ in sorting rule</td>
<td>M1</td>
</tr>
<tr>
<td>$\omega_1$</td>
<td>Elas. $q$ to $I$</td>
<td>1</td>
<td>Normalization</td>
<td>id.</td>
</tr>
<tr>
<td>$\sigma_b^2$</td>
<td>Var. birth shock</td>
<td>6.6</td>
<td>$\rho(y_{m,i}, \text{rank}(h_{s,i}))$</td>
<td>M1</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>Elas. $h'$ to $h_s$</td>
<td>.21</td>
<td>InterGen. Elas. [Mazumder, 2015]</td>
<td>M2</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>Elas. $h'$ to $q$</td>
<td>.2</td>
<td>Elas. $y'_m$ to $q$</td>
<td>id.</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>Elas. $h'$ to $h$</td>
<td>.2</td>
<td>Elas. $y'_m$ to $y_m$</td>
<td>id.</td>
</tr>
<tr>
<td>$\sigma_y^2$</td>
<td>Var. Lab. Mkt. shock</td>
<td>.74</td>
<td>Income Gini Coef. [Kopczuk et al., 2010]</td>
<td>M1</td>
</tr>
<tr>
<td>$q$</td>
<td>Outside Option</td>
<td>0.0278</td>
<td>Enrollment Rate (NCES)</td>
<td>id.</td>
</tr>
<tr>
<td>$\bar{r}$</td>
<td>Interest Rate</td>
<td>-</td>
<td>Elas. $q$ to $y$ in sorting rule</td>
<td>M1</td>
</tr>
</tbody>
</table>

1 Based on Barro et al. [1995], Krueger [1999], Autor et al. [2008], Borjas et al. [2012]. See appendix C for more details.
7 The Role of Higher Education: Quantitative Results

In this section I use the calibrated model from the previous section to assess the quantitative relevance of the higher education system in shaping inequality, intergenerational mobility and the efficiency of the accumulation of human capital. I first present a set of policy experiments that shed light on different aspects of the system and/or are of a specific political or historical interest. Second I assess the quantitative effects of a rise in the return to human capital and decompose the rise in inequality into a direct effect and the endogenous amplification through the higher education sector.

As stated in the previous section, given the uncertainty surrounding the value of the elasticity of income to college quality, $\alpha_2$, the results presented below are always upper bounds. The lower bound, corresponding to $\alpha_2 = 0$, will be 0, unless otherwise stated.

7.1 Policy Experiments in Higher Education

In table 2, I gather the results of the six policy experiments discussed below. I provide the percentage change from the status-quo steady-state to counterfactual steady-state of the Gini coefficient of labor earnings, expenditures per students, the intergenerational elasticity, GDP and a measure of welfare. The social welfare function is given by (29) with elasticity of substitution across households, $\sigma$, in the range $[.2, 1]$.\(^{30}\)

How would the equilibrium outcome change if everyone got the exact same higher education? The first policy experiment consists in randomly allocating students across colleges. This leads to an equalization of spending per student and average student ability across all colleges. It is therefore a natural way to quantify the total effect that higher education has on income inequality, intergenerational persistence and GDP. Formally, I provide everyone with the same college quality, $\bar{q}$, where the latter is compatible with the production function of quality (9), the average children ability in society and average government transfers per student.\(^{31}\) I find that doing so would reduce the income Gini by 8.5% and the IGE by 24.3%.

\(^{30}\)Recall that the cases $\sigma \to 0, \sigma = 1$ correspond to a Rawlsian and s utilitarian social welfare functions, respectively and the case $\sigma \to +\infty$ to a social welfare function that is a monotonic transformation of GDP. In the context of my model, with missing insurance markets for birth and labor market shocks, a concern for equity also captures a concern for intergenerational insurance.

\(^{31}\)The random allocation of students not only equalizes college experiences among college-goers but implies that everyone goes to college. It thus neutralizes both the extensive (going or not) and the intensive (quality) margin. Notice also that households optimally stop spending for higher education and all the resources spent in the higher education system have to be financed through taxes and transfers to colleges. One therefore needs to take a stand on the level of government spending. I choose to keep the aggregate share of GDP going to higher education constant to its status-quo level. This choice does not influence inequality or mobility, but it does have a first order effect on the aggregate level of production. This assumption allows to focus on the effect of misallocation on aggregate production.
(see line no. 1 in table 2). To get a sense of the magnitude, a reduction by 8.5% of the income Gini corresponds to a reduction of 4 p.p., which is half of the total increase since 1980. It is therefore a sizable effect. GDP however drops by 7% because of the increase in the misallocation of students and resources across colleges.

The second policy experiment consists in neutralizing the effect that parental income has on the sorting of students across colleges conditional on child ability. Different policies can implement this “income-neutral sorting” within the model. For example, one can design a policy that redistributes and equalizes resources across colleges, or a very progressive need-based financial aid schedule. Formally, any combination of \((\tau_n, \tau_u)\) such that \(\omega_3 = \omega_1(1 - \tau_n)(1 - \tau_u)\) would generate such an income-neutral sorting. Recall that in M1, this perfectly shuts down the income-sorting channel. More generally, it ensures that all households have to pay the same share of their income to get into a given college, keeping the ability of the child constant. Importantly, a general implication of such a policy is an equalization of resources and spending across all colleges. This experiment therefore isolates the specific contribution of the peer-effects.\(^{32}\) I find that in the counterfactual allocation, the income Gini is reduced by 3% (or 1.4 p.p.) and the IGE by 20% (line no 2). On the one hand the mismatch of student abilities is reduced with the elimination of the income-sorting channel, but the equilibrium equalization of spending across colleges leads to a less efficient accumulation of human capital. Overall, GDP falls by 1.8%.\(^{33}\)

The third experiment requires colleges to adopt need-blind admissions policies. Like in the previous experiment, it implies \(\omega_3 = \omega_1(1 - \tau_n)(1 - \tau_u)\) and that all households pay the same share of their income to get into a given college conditional on ability, which, in M1, perfectly shuts down the income-sorting channel. By contrast, it doesn’t lead to an equalization of resources and spending across colleges. The segregation of colleges by abilities and the positive correlation between abilities and parental income leads to a positive sorting of resources across colleges. Colleges with higher ability students have more resources and

\(^{32}\)The distribution of student abilities across colleges changes in the counterfactual. In particular, colleges become more homogeneous in terms of ability. One could therefore argue that this counterfactual isn’t capturing only the effect of neutralizing the role of parental income in the sorting across colleges. For example, Chetty et al. [2019] do a counterfactual exercise along these lines: they reallocate students across colleges so that the distribution of abilities remain unchanged but so that conditional on ability, the allocation becomes independent of parental income. The problem of such a counterfactual is that it is not compatible with a general equilibrium allocation of students and resources. Reallocating students entails a reallocation of financial resources, and therefore of value-added. The advantage of the counterfactual I propose is therefore its implementability and compatibility with a decentralized equilibrium.

\(^{33}\)Like in the previous counterfactual, one needs to take a stand on how the average government subsidies, \(a_u, a_h\), react, in particular following the fall of household spending rates for higher education implied by the increase in \(\tau_u, \tau_n\). Here again, I assume that government policies exactly offset the decrease in average private spending, so that the aggregate spending rate in higher education remains constant in the two counterfactuals.
spend more. This policy has, maybe surprisingly, a positive effect on income inequality (+3%) because it improves the matching of students, like the previous policy, but doesn’t reduce the dispersion of expenditures across colleges, which actually increases by 13%. The perfect positive assortative matching of students and the positive sorting of resources lead to a large increase in GDP (+22%). The IGE falls by less than in the previous experiment for the same reason: educational spending remains increasing with college equality. As a result, the welfare gains of this policy are very large, around 8%, irrespective of the strength of the concern for equity, σ.

The fourth policy experiment consists in eliminating all current government interventions in higher education. It is aimed at quantifying the extent to which the current government interventions affect inequality and mobility. Formally, a *laissez-faire* equilibrium corresponds to the case $\tau_u = \tau_n = \tau_m = a_u = a_h = 0$. I find that doing so leads to an increase in the income Gini by 2%, an increase in the IGE by 12% and an increase in the Gini coefficient of college expenditures by 70% (line no. 4a). Across all measures, two-third of these changes are due to the transfers to colleges (line no. 4b) and one-third to need-based financial aid (line no. 4c). Merit-based aid plays virtually no role (line no. 4d). While most of the policy debates have been, in recent decades, centered around the issue of federal financial aid and income tax credit and as transfers to colleges have been significantly cut [Mc Guinness, 2011], these results highlights the importance, in the current system, of transfers to colleges.

The fifth policy experiment evaluates a (conservative) version of the recent proposal by democratic candidates to make college free for all. Although not fully specified as of now, the plan envisions (i) setting tuition to a minimum fee at public institutions (ii) offsetting the implied revenue losses with federal and state subsidies to colleges. The proposal states that on aggregate the loss of tuition will be offset by subsidies to colleges, but it doesn’t specify how much redistribution of resources across colleges shall occur. The most progressive option, where resources are fully equalized across colleges, corresponds to the second counterfactual (line no. 2) derived earlier. In contrast, I present here the most conservative option, where government transfers exactly offset the loss of tuition revenues at the college level. Formally, I assume (i) $\tau_n = 1$, (ii) $a_h$ and $a_u$ are such that the share of GDP going to higher education remains the same as in the benchmark allocation—as in counterfactual 1 and 2—and (iii) $\tau_m$ is such that the Gini of expenditures per students remain constant. The equilibrium allocation therefore features perfect stratification of colleges by student ability and unchanged sorting of expenditures per students across colleges. Such a policy would lead to an increase in the income Gini by 2.4% and an decrease in the IGE by 7.4% (line no. 5). GDP would increase by 2.8% thanks to the improvement in the matching of students.

---

34See here for a detailed version of the Act, and here for a summary.
The sixth policy experiment sets the level and progressivity of government transfers to colleges to what they were in 1980. It aims at quantifying the impact on inequality and mobility of the sharp decline in both the average and progressivity of government transfers to colleges over the past forty years, documented in Capelle [2019b]. I have shown in section 4 that qualitatively such policy changes have very likely contributed to the trends (a)-(e). Quantitatively, I find that setting the parameters of the subsidies to colleges schedule to what they used to be in 1980 implies a decrease in the income Gini by .6% and a decrease in the IGE by 3.4% (line no. 5). The decline in public transfers to colleges can quantitatively account for a very small share (2.7%) of the total increase in the income Gini but a large share of the total increase in the Gini of expenditures per student (90%). Although qualitatively consistent with fact (a)-(e), the latter finding makes the decline in public transfers a less compelling explanation than the increase in the returns to education, to which I now turn.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Random Admission</td>
<td>-8.5</td>
<td>-100</td>
<td>-24.3</td>
<td>-7.1</td>
<td>[-9.6, -0.8]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Income-neutral Admission</td>
<td>-2.5</td>
<td>-100</td>
<td>-14.7</td>
<td>-1.8</td>
<td>[4.1, 0.5]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Need-blind Admission</td>
<td>3.0</td>
<td>13</td>
<td>-6.4</td>
<td>22.0</td>
<td>[8.3, 7.9]</td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>Laissez-faire</td>
<td>2.0</td>
<td>70.3</td>
<td>12</td>
<td>2.9</td>
<td>[-6.1, -1.9]</td>
<td></td>
</tr>
<tr>
<td>4b</td>
<td>No Transfer to College</td>
<td>1.5</td>
<td>48.2</td>
<td>7.9</td>
<td>.6</td>
<td>[-2.5, 0.2]</td>
<td></td>
</tr>
<tr>
<td>4c</td>
<td>No Need-based Aid</td>
<td>.5</td>
<td>18.2</td>
<td>3.0</td>
<td>2.4</td>
<td>[-2.7, -1.2]</td>
<td></td>
</tr>
<tr>
<td>4d</td>
<td>No Merit-based Aid</td>
<td>-.02</td>
<td>-3.1</td>
<td>-0.03</td>
<td>.2</td>
<td>[0.0, 2.1]</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>College for All (conservative)</td>
<td>2.4</td>
<td>0</td>
<td>-7.4</td>
<td>2.8</td>
<td>[-2.2, 0.8]</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Transfers to College, 1980</td>
<td>-.6</td>
<td>-21.6</td>
<td>-3.4</td>
<td>-3.0</td>
<td>[-1.0, -1.1]</td>
<td></td>
</tr>
</tbody>
</table>

Status-quo Levels: .45 .22 .4

### 7.2 Increase in the Returns to Education and Propagation through Higher Education

I now quantify the extent to which a reasonably parametrized increase in the returns to education, $\lambda$, can explain the stylized facts (a)-(e) presented in the introduction. Let’s denote $\lambda_{1980}$ and $\lambda_{2010}$ the value of the returns to education in the original steady-state (1980) and in the final steady-state (2010). The value in 2010, $\lambda_{2010}$, is the one estimated in the previous section. In order to calibrate the value in the old steady-state, $\lambda_{1980}$, I target the change in the
college premium across the two periods, keeping all other parameters constant.\footnote{\textsuperscript{35}As I discuss in appendix XX, the model overestimates the level of the college premium because of the fat tail of income in the data. It is therefore natural to target the relative increase in the college premium rather than its level in 1980. I use the value provided by Autor et al. [2008] for the college premium: $\Delta \log(\frac{w_{\text{college}}}{w_{\text{HS}}}) = .65 - .45 = .2$, see figure 2 in their paper.} Comparing the two steady-states corresponding to the two values, $\lambda_{1980} < \lambda_{2010}$. I find that the model generates an increase in the income Gini coefficient by 13 p.p., which corresponds to 130\% of the empirical change, an increase in the expenditure per student Gini by 5 p.p. corresponding to 100\% of the empirical change and an increase in the IGE by 6\%.

I then isolate the contribution of the higher education system to the increase in income inequality. I compute the Gini coefficient of a counterfactual distribution of labor incomes where (i) the underlying distribution of human capital is the one in the 1980 steady-state, (ii) the returns to human capital parameter, $\lambda$, is set at its 2010 level, $\lambda_{2010}$, and (iii) the labor supply policy function is the one in the 2010-steady-state. This gives the level of inequality if there were no propagation through higher education. I provide more details in appendix D. I find that the higher education sector accounts for 6\% of the seven percentage points increase generated by the model, which corresponds to a little bit more than half a percentage point increase in the Gini coefficient of income.\footnote{\textsuperscript{36}In M1, the version of the model without intergenerational transfers and without outside option to colleges, the endogenous amplification through the higher education is even lower, 3\% of the total increase in inequality. The additional amplification in M2 comes from the extensive margin: the fact that more students get enrolled into college when the returns to human capital increases reinforces the feedback mechanism of higher education on inequality. That the extensive margin plays an amplifying role is not a qualitative feature of the model but depends on the initial enrollment rate.}

I now decompose the propagation through the higher education system into two channels: the reallocation of resources and the reallocation of students quality across clubs. Formally, I compute a counterfactual steady-state in which (i) the tuition schedule and the expenditures per student by college rank is fixed at what they were in the initial steady-state, in 1980, (ii) the returns to education parameter $\lambda$ is set at its 2010 value. Notice that this counterfactual is not compatible with a decentralized equilibrium, because colleges are no longer on their F.O.C. and their budget constraint don’t hold, so that some have deficits and some have surpluses. See appendix D.2 for more details. I find that more than 100\% of the total effect stems from the increased dispersion of financial resources. The reallocation of students across the quality ladder of colleges dampens very slightly the effect.\footnote{\textsuperscript{37}This dampening effect is stronger in the short-run. In the long-run, low ability kids from rich family 'catch up'.}

Intuitively, in the counterfactual where expenditures have been fixed, top colleges display surpluses, because the willingness to pay of their equilibrium pool of students has increased with the rise of inequality pushing tuition fees up, while bottom colleges have deficits. The

\[ \Delta \log(\frac{w_{\text{college}}}{w_{\text{HS}}}) = .65 - .45 = .2, \text{ see figure 2 in their paper.} \]
dispersion of average student ability however declines as rich but not so smart children manage to buy their way to top colleges while smart but poor children are being priced-out. From the counterfactual allocation to the final steady-state in the year 2010, expenditures per students adjust according to revenues and tuition fees, thus keeping up with the willingness to spend of their respective pool of families, which increases revenues and thus quality at the top and decreases it at the bottom. In the counterfactual, the Gini coefficient of income barely moves and if anything slightly declines. This allows us to conclude that it is really the increase in the dispersion of revenues and therefore expenditures per students across colleges that is the root of the increase in the amplification by colleges.\textsuperscript{38} One natural test of this mechanism is to look at the evolution of the dispersion of tuition fees. Davies and Zarifa [2012] indeed find that the Gini coefficient for tuition fees has increased over the period 1971-2006.

8 Conclusion

This paper studies the extent to which the higher education system shapes economic inequality and intergenerational mobility in the U.S., how it has propagated and amplified macroeconomic shocks such as an increase in the returns to human capital and how government policies may affect these responses. An increase in the market returns to human capital increase inequalities directly but also indirectly over time through a more unequal sorting of resources across colleges and a (more unequal) accumulation of human capital.

The tractability of the model studied at the beginning of the paper has a number of advantages. The closed-form solutions allow to derive clear insights about the mechanisms through which colleges shape inequalities and social mobility in a model with a very rich heterogeneity on both the household and the college side. Another major advantage is the ability to estimate the model in a very transparent and efficient fashion. A third is to allow for a detailed welfare analysis.

However it goes with strong and simplifying assumptions. I partially relax two of them in the quantitative section: I allow for intergenerational financial transfers, in particular student debt, and I allow for an outside option to college, so that some students optimally choose not to enroll. Four assumptions would need further investigations: first, the allocation of students across colleges in the model works through a system of clearing markets, while the real world displays a mix of price mechanism and quantity restrictions. Second, the recent literature on human capital has shown the importance of strong complementarities between parental income and school inputs in the production function of the child’s human

\textsuperscript{38}A graphical representation of the counterfactual allocations is shown in figure A1 in appendix D.2.
capital, while I maintain the Cobb-Douglas functional form. Exploring the implications of these more realistic assumptions are interesting avenues for future research. Third, higher education in the U.S. is more than a tool for accumulating human capital. It is a prize, whose value goes far beyond labor market returns. Some papers have shown that there are non-pecuniary returns to education [Oreopoulos and Salvanes, 2011, Hout, 2012] and that there is a consumption value to going to college [Gong et al., 2019]. Taking into account these dimensions has potentially far reaching implications for welfare and policy analysis. Finally, as highlighted in the quantitative section, the model presented here doesn’t generate enough dispersion of college resources compared to the data. In a forthcoming paper [Capelle, 2019b], I show that taking into account donations and endowments help in matching the data and has the potential to generate additional amplification.
References


A Analytical Model - Details

I solve the model using a guess and verify. I guess that the tuition function in a generalized equilibrium with government intervention is given by:

\[ e_u(q, h_s, y) = \left( \frac{pI}{(1 + a_u)T_u} q^{\frac{1}{\epsilon_1}} h_s^{\frac{\epsilon_2}{\epsilon_1}} \left( \frac{y}{\kappa_2} \right)^{\frac{\epsilon_3}{\epsilon_1}} \right)^{\frac{1}{1 - \tau_u}} \]  

(36)

Households face the following after government financial aid tuition schedule:

\[ e(q, h_s, y) = h_s^{-\tau_n} y^{\tau_n} \frac{T_e}{(1 + a_h)} \left( \frac{pI}{(1 + a_u)T_u} q^{\frac{1}{\epsilon_1}} h_s^{\frac{\epsilon_2}{\epsilon_1}} \left( \frac{y}{\kappa_2} \right)^{\frac{\epsilon_3}{\epsilon_1}} \right)^{\frac{1}{1 - \tau_u}} \]  

(37)

which implies that they choose the following quality for their child:

\[ q^i = \left( \frac{s_t(y_{i,t})^{1 - \tau_n} (h_{s,t}^{\tau_n} (1 + a_{h,t}))}{T_{e,t}} \right)^{\frac{\epsilon_1}{1 - \tau_u}} \left( \frac{(1 + a_{u,t})T_{u,t}}{pI} \right)^{\frac{\epsilon_1}{1 - \tau_u}} (h_{s,t}^{\epsilon_2})^{\frac{\epsilon_3}{1 - \tau_u}} \]  

A.1 Solution to the Household Problem

Using the guess (37), the problem of the Households rewrites:

\[ \ln \mathcal{W} = \max_{s,\ell} (1 - \beta) [\ln(1 - s)/(1 + a_c) + \ln(1 - a_y) + \ln T_y] + (1 - \tau_y)(\lambda \ln h + \mu \ln \ell + \ln A) - \ell^\theta] + \beta E \ln \left( \mathcal{W}' \right) \]  

(38)

\[ \ln h' = \ln \xi_y + \ln \kappa + \alpha_1 (1 + \alpha_2 (\epsilon_2 + \tau_n (1 - \tau_u) \epsilon_1)) \ln \xi_b + \alpha_h \ln h 
+ \alpha_2 \epsilon_1 ((1 - \tau_u) (\ln s - \ln T_e + \ln(1 + a_h)) + \ln(1 + a_u) T_u - \ln pI) 
+ \alpha_2 \epsilon_1 (1 - \tau_u) (1 - \tau_n,t - \epsilon_3) (1 - \tau_y) \mu \ln \ell 
+ \alpha_2 \epsilon_1 (1 - \tau_u) (1 - \tau_n) - \epsilon_3) (1 - \tau_y,t) \ln A + \ln(1 - a_y) + \ln T_y) 
+ \alpha_2 \epsilon_3 \ln \kappa_2 \]  

(39)

with \( \alpha_h = \alpha_1 + \alpha_3 + \alpha_1 \alpha_2 \epsilon_2 + \tau_n (1 - \tau_u) \epsilon_1 + \alpha_2 \epsilon_1 (1 - \tau_u) (1 - \tau_n) - \epsilon_3) (1 - \tau_y) \lambda \) and \( s \) the saving rate, i.e. the amount of spending for college over income. I then guess that

\[ \ln \mathcal{W}_t = U_t \ln h_t + Z_t \ln \xi_{b,t} + B_t \]  

(40)

Replacing (40) into (38), one gets:

\[ U_t \ln h_t + Z_t \ln \xi_{b,t} + B_t = \max_{s,\ell} (1 - \beta) [\ln(1 - s)/(1 + a_{c,t}) + \ln(1 - a_{y,t}) + \ln T_{y,t}] + (1 - \tau_{y,t})(\lambda \ln h + \mu \ln \ell + \ln A) - \ell^\theta] 
+ \beta E_{t} ((U_{t+1} \ln h_{t+1} + Z_{t+1} \ln \xi_{b,t+1} + B_{t+1})) \]
Then using (39) to substitute for \( \ln h_{t+1} \) and using (5) and (6):

\[
U_t \ln h_t + Z_t \ln \xi_{b,t} + B_t = \max_{s,\ell}(1 - \beta)[\ln(1 - s)/(1 + a_{c,t}) + \ln(1 - a_{y,t}) + \ln T_{y,t} + (1 - \tau_{y,t})(\lambda \ln h + \mu \ln \ell + \ln A) - \ell^\eta] \\
+ \beta \left[ U_{t+1} \left( \mu_y + \ln \kappa + \alpha_1(1 + \alpha_2(\varepsilon_2 + \tau_{m,t}(1 - \tau_u)) \right) \ln \xi_{b,t} + \alpha_{h,t} \ln h_t \\
+ \alpha_2 \varepsilon_1 ((1 - \tau_u) (\ln s - \ln T_e + \ln(1 + a_{h,t})) + \ln(1 + a_{u,t})T_u - \ln p_{I,t}) \\
+ \alpha_2 (\varepsilon_1 (1 - \tau_u)(1 - \tau_{n,t}) - \varepsilon_3) (1 - \tau_{y,t}) \mu \ln \ell \\
+ \alpha_2 (\varepsilon_1 (1 - \tau_u)(1 - \tau_{n,t}) - \varepsilon_3) ((1 - \tau_{y,t}) \ln A + \ln(1 - a_{y,t}) + \ln T_{y,t}) \\
+ \alpha_2 \varepsilon_3 \ln \kappa_{2,t} \right] + Z_{t+1} \mu_b + B_{t+1}
\]

Gathering all the terms in \( \ln h_t \) one gets that \( U_t \) has to verify the following forward first order linear recursive equation:

\[
U_t = (1 - \beta)(1 - \tau_{y,t}) \lambda_t + \beta U_{t+1} \alpha_{h,t} \\
\Rightarrow U_t = (1 - \beta) \sum_{k=0}^{\infty} \beta^k (1 - \tau_{y,t+k}) \lambda_{t+k} \prod_{m=0}^{k-1} \alpha_{t+m}^h
\]

Gathering all the terms in \( \ln \xi_{b,t} \) and using the recursive equation for \( U_t \), one gets that \( Z_t \) has to verify the following equation:

\[
Z_t = \beta U_{t+1} \alpha_1(1 + \alpha_2(\varepsilon_2 + \varepsilon_1 \tau_{m,t})) \\
\Rightarrow Z_t = \left( U_t - (1 - \beta)(1 - \tau_{y,t}) \lambda \right) \frac{\alpha_1(1 + \alpha_2(\varepsilon_2 + \varepsilon_1 (1 - \tau_u) \tau_{m,t}))}{\alpha_t^h}
\]

Finally gathering the independent terms:

\[
B_t = \max_{s,\ell}(1 - \beta)[\ln(1 - s)/(1 + a_{c,t}) + \ln(1 - a_{y,t}) + \ln T_{y,t} + (1 - \tau_{y,t})(\mu \ln \ell + \ln A) - \ell^\eta] \\
+ \beta \left[ U_{t+1} \left( \mu_b + \ln \kappa + \alpha_2 \varepsilon_1 ((1 - \tau_u) (\ln s - \ln T_e + \ln(1 + a_{h,t})) + \ln(1 + a_{u,t})T_u - \ln p_{I,t}) \\
+ \alpha_2 \varepsilon_2 (1 - \tau_u)(1 - \tau_{n,t}) - \varepsilon_3) ((1 - \tau_{y,t}) \ln A + \mu \ln \ell) + \ln(1 - a_{y,t}) + \ln T_{y,t}) \\
+ \alpha_2 \varepsilon_3 \ln \kappa_{2,t} \right] + Z_{t+1} \mu_b + B_{t+1}
\]
The F.O.C for \( s \) is:

\[
0 = -\frac{1 - \beta}{1 - s} + \beta U_{t+1} \frac{\alpha_2 \varepsilon_1 (1 - \tau_u)}{s}
\]

\( \iff \)

\[
s_t = \frac{\beta \alpha_2 \varepsilon_1 (1 - \tau_u) U_{t+1}}{1 - \beta + \beta \alpha_2 \varepsilon_1 (1 - \tau_u) U_{t+1}}
\]

And the F.O.C w.r.t. \( \ell \) is:

\[
0 = (1 - \beta) \left( \frac{1 - \tau_{y,t}}{\ell} - \eta \ell^{\alpha - 1} \right) + \beta U_{t+1} \alpha_2 (\varepsilon_1 (1 - \tau_u) (1 - \tau_{n,t}) - \varepsilon_3) \frac{(1 - \tau_{y,t}) \mu}{\ell}
\]

\( \iff \)

\[
\ell = \left[ (1 - \tau_{y,t}) \frac{\mu}{\eta} \left( 1 + \frac{\beta}{1 - \beta} \alpha_2 (\varepsilon_1 (1 - \tau_u) (1 - \tau_{n,t}) - \varepsilon_3) U_{t+1} \right) \right]^{\frac{1}{\eta}}
\]

Although \( \kappa_{2,t} \) depends on the aggregate saving rate and labor in the economy, individual households don’t internalize it.

### A.2 University problem

I first provide a generalized definition of \( \sigma_u \) to the framework with government policies ((7) abstracted from them).

\[
\sigma_u^2 = E \left( \left( \ln \left( \theta^{\frac{\omega_2}{\omega_1(1 - \tau_u)}} D^{-\frac{\omega_3}{\omega_1(1 - \tau_u)}} \right) - \ln h_s^{\frac{\omega_2}{\omega_1(1 - \tau_u)}} y^{-\frac{\omega_3}{\omega_1(1 - \tau_u)}} \right)^2 \right)
\]

Using this definition and our guess for tuitions (36), one gets a new expression for \( \sigma_u^2 \):

\[
\sigma_u^2 = \int \phi(h_s, \ln y) \left( \ln \left( \frac{p_I \bar{I}}{(1 + a_u) T_u} \right)^{\frac{1}{1 - \tau_u}} \frac{\theta}{h_s}^{\frac{\omega_2}{\omega_1(1 - \tau_u)}} D^{-\frac{\omega_3}{\omega_1(1 - \tau_u)}} y^{-\frac{\omega_3}{\omega_1(1 - \tau_u)}} \right)^2 d \ln h_s d \ln y
\]

\[
= \int \phi(h_s, \ln y) \left( \ln e_u(h_s, y) - \ln \left( \frac{p_I \bar{I}}{(1 + a_u) T_u} \right)^{\frac{1}{1 - \tau_u}} \right)^2 d \ln h_s d \ln y
\]

where I define \( \ln \bar{I} = \ln I - (1 - \tau_u) \frac{\sigma_u^2}{2} \). The next step shows that \( \sigma_u^2 \) is the within-university variance of log tuition, or in other words, that

\[
E \ln e_u(h_s, y) = \ln \left( \frac{p_I \bar{I}}{(1 + a_u) T_u} \right)^{\frac{1}{1 - \tau_u}}
\]

I first guess that tuitions are log-normally distributed within the university. Denoting \( \mu_{eq}, \sigma_{eq} \) the mean and standard deviation of log tuition within the university of quality \( q \), the budget constraint
of the university - given by (16) - rewrites:

\[ p_I I = T_u (1 + a_u) (E_{h_s,y} [e_u(q,h_s,y)])^{1 - \tau_u} \]

\[ \iff p_I I = T_u (1 + a_u) e^{(1 - \tau_u) \mu_{eq} + (1 - \tau_u) \frac{\sigma^2_{eq}}{2}} \]

\[ \iff \ln \left( \frac{p_I}{1 + a_u T_u} \right) + \ln I - (1 - \tau_u) \frac{\sigma^2_{eq}}{2} = (1 - \tau_u) \mu_{eq} \]

\[ \iff \frac{1}{1 - \tau_u} \ln \left( \frac{p_I}{1 + a_u T_u} \right) I + \frac{\sigma^2_u}{2} - \frac{\sigma^2_{eq}}{2} = \mu_{eq} = E \ln e_u(h_s, y) \]

where the last line uses the definition of \( \ln \tilde{I} \). Substituting this last line into the expression of \( \sigma^2_u \) above gives:

\[ \sigma^2_u = \int \phi(h_s, y) \left( \ln e_u(h_s, y) - E \ln e_u(h_s, y) + \frac{\sigma^2_{eq} - \sigma^2_u}{2} \right)^2 d h_s d y \]

\[ \iff \sigma^2_u = \sigma^2_{eq} + \left( \frac{\sigma^2_{eq} - \sigma^2_u}{2} \right)^2 + 0 \]

\[ \Rightarrow \sigma^2_u = \sigma^2_{eq} \text{ or } \sigma^2_u = \sigma^2_{eq} + 4 \]

\( \sigma_u = \sigma_{eq} \) is a solution to the quadratic equation. I select this solution and ignore the other one.

From now on

\[ E \ln e_u(h_s, y) = \ln \left( \frac{p_I \tilde{I}}{(1 + a_u)T_u} \right)^{1 - \tau_u} \text{ and } \sigma^2_u = \sigma^2_{eq} \]

are respectively the mean and standard deviation of within-university log tuitions.

Therefore I can now rewrite the problem of the university substituting the primitive expression for the budget constraint/production of effective educational services with its transformed expression

\[ \ln I - (1 - \tau_u) \frac{\sigma^2_u}{2} = \ln \tilde{I} \]

\[
\max_{\tilde{I}, \theta, D, \omega} \tilde{I}^\omega_1 \theta^\omega_2 D^{-\omega_3} \quad (42)
\]

\[ \ln \tilde{I} \int_0^1 r(h_s, y) d h_s d y = \int r(h_s, y) \left( (1 - \tau_u) \ln(e_u)^i + \ln(1 + a_u)T_u/p_I \right) d h_s d y \]

\[ \ln \theta \int_0^1 r(h_s, y) d h_s d y = \int_0^1 r(h_s, y) \ln h_s d h_s d y \]

\[ \ln D \int_0^1 r(h_s, y) d h_s d y = \int_0^1 r(h_s, y) \ln y d h_s d y \]
The F.O.Cs write:

\[
\begin{align*}
\frac{\omega_1}{t} + \frac{\lambda_1}{t} &= 0 \\
\frac{\omega_2}{\theta} + \frac{\lambda_2}{\theta} &= 0 \\
-\frac{\omega_3}{D} + \frac{\lambda_3}{D} &= 0
\end{align*}
\]

and for all \((h_s, y)\)

\[
r(ln h_s, ln \hat{y}) = \begin{cases} 
0 & \text{if } \lambda_1 \left( \ln \left( \frac{p_1 I}{(1+a_u)T_u} \right) - (1 - \tau_u) \ln e_u \right) + \lambda_2 \ln \left( \frac{\theta}{h_s} \right) + \lambda_3 \ln \left( \frac{D}{y} \right) < 0 \\
r \in \mathbb{R} & \text{if } 0 \\
+\infty & \text{if } > 0
\end{cases}
\]

After solving out for the Lagrange multipliers, this last system rewrites:

\[
r = \begin{cases} 
0 & \text{if } \left( \frac{p_1 I}{(1+a_u)T_u} q^\frac{\omega_2}{\omega_1} h_s - \frac{\omega_2}{\omega_1} y \right) \frac{1}{1-\tau_u} < e_u(q, h_s, y) \\
r \in \mathbb{R} & \text{if } = \\
+\infty & \text{if } > 
\end{cases}
\]

I guess that in equilibrium, \(D = \kappa_2 q^\nu\). Therefore whenever a college admits a certain student type, the tuition formula rewrites:

\[
e_u(q, h_s, y) = \left( \frac{p_1 I}{(1+a_u)T_u} q^\frac{1-\nu \varepsilon_3}{\omega_1} h_s - \frac{\omega_2}{\omega_1} y \right) \frac{1}{(1-\tau_u)} \varepsilon_1
\]

\[
\Leftrightarrow e_u(q, h_s, y) = \left( \frac{p_1 I}{(1+a_u)T_u} q^\frac{1}{\omega_1} h_s - \frac{\varepsilon_2}{\omega_1} y \right) \frac{1}{(1-\tau_u)} \varepsilon_1
\]

with \(\varepsilon_1 = \frac{\omega_1}{1-\nu}, \varepsilon_2 = \frac{\varepsilon_2}{1-\nu}, \varepsilon_3 = \frac{\varepsilon_3}{1-\nu}\)

I can solve for \(\nu\) and \(\kappa_{2,t}\) using the equilibrium outcome given by the mean income in proposition A.2. I do this later in appendix A.4.5.

This confirms our guess for tuitions (36). Given this guess for tuition, a university is always at the interior solution, therefore always indifferent between all types.
A.3 Government Budget, Educational Sector and Market Clearing

A.3.1 Government Budget Constraints

Lemma 2. Along the equilibrium path, the government budget constraints (17), (18), (19) and (20) rewrite

\[
\frac{a_{c,t}(1-s_t)}{(1+a_{c,t})} = s_t(1 + a_{u,t})(1 + a_{h,t}) - \frac{a_{y,t}}{1-a_{y,t}} - s_t \tag{43}
\]

\[
\ln T_y = \tau_y \ln \alpha + \tau_y \mu \ln \ell + \tau_y \lambda m_h + \frac{\lambda^2}{2} (2 - \tau_y) \tau_y \Sigma_{h,t} \tag{44}
\]

\[
\ln T_u = (-\tau_n \lambda + \alpha_1 \tau_m) m_h + \frac{\alpha_1 \tau_m}{2} (\alpha_1 \tau_m - 1) \sigma_b^2 - \tau_n (\ln \alpha^{\mu}(1 - a_y))
\]

\[
+ \left[ \lambda^2 (1 - \tau_y)^2 (\tau_n - 2) \tau_n + 2 \lambda (1 - \tau_n)(1 - \tau_y) \tau_m \alpha_1 + (\alpha_1 \tau_m)^2 - \tau_n \lambda^2 (2 - \tau_y) \tau_y \right] \frac{\Sigma_{h,t}^2}{2} \tag{45}
\]

\[
\ln T_a = \tau_u \left( \ln \alpha^{\mu} s(1 + a_h)(1 - a_y) + \lambda m_h + \frac{\lambda^2 \Sigma_{h,t}^2 }{2} \right) + \frac{\tau_u}{1-\tau_u} \frac{\sigma_b^2}{2} \tag{46}
\]

Proof. 1. Solving for the aggregate state budget constraint.

The revenues of the state come from average taxes on market income \(a_y y_m\), average taxes rate on consumption \(a_c\) and the collection of tuitions \(e^i\) before all subsidies. The expenses consist in tuitions redistributed to colleges after subsidies \(e^i(1 + a_{u,t})(1 + a_h)\).

\[
\int_0^1 a_y y_m^i + a_c e^i + e^i di = \int_0^1 e^i(1 + a_u)(1 + a_h) di
\]

\[
\int_0^1 a_y y_m^i + a_c \left( \frac{1-s}{1+a_c} \right) y_I + sy_I di = \int_0^1 sy_I(1 + a_u)(1 + a_h) di
\]

\[
a_y \int_0^1 y_m^i di + \left( a_c \left( \frac{1-s}{1+a_c} \right) + s \right) \int_0^1 y_I di = s(1 + a_u)(1 + a_h) \int_0^1 y_I di
\]

\[
= \frac{a_y}{1-a_y} + \left( a_c \left( \frac{1-s}{1+a_c} \right) + s \right) = s(1 + a_u)(1 + a_h)
\]

2. Solving for \(T_y\)

Using (18), and the expression for market income \(y_m\), (14), one gets:

\[
\int_0^1 y_m^{1-\tau_y} T_y di = \int_0^1 y_m di
\]

\[
\int_0^1 (\alpha^{\mu} h^\lambda)^{1-\tau_y} T_y di = \int_0^1 \alpha^{\mu} h^\lambda di
\]
Using the guess that $\ln h$ is normally distributed along the equilibrium path, we get:

$$T_y e^{\lambda (1 - \tau_y) m_h + \left(\frac{\lambda (1 - \tau_y)\Sigma^2_b}{2}\right)} = A^\tau_y e^{\lambda m_h + \frac{\lambda^2 \Sigma^2_b}{2}}$$

$$\iff \ln T_y + \lambda (1 - \tau_y) m_{h,t} + \frac{\lambda^2 (1 - \tau_y^2) \Sigma^2_h}{2} = \tau_y \ln A + \tau_y \mu \ln \ell + \lambda m_h + \frac{\lambda^2 \Sigma^2_h}{2}$$

$$\iff \ln T_y = \tau_y \ln A + \tau_y \mu \ln \ell + \tau_y \lambda m_h + \frac{\lambda^2 (2 - \tau_y) \tau_y \Sigma^2_{h,t}}{2}$$

3. Solving for $T_e$

Similarly, using (19), one gets:

$$T_e \int_0^1 (1 - a_y)(A\ell h^\lambda)^{1 - \tau_y} y_T di = \int_0^1 \left((1 - a_y)(A\ell h^\lambda)^{1 - \tau_y} y_T\right)^{1 - \tau_n} h_s^{\tau_n} di$$

$$T_e (1 - a_y)^{\tau_n} (A\ell)^{\tau_n (1 - \tau_y)} (T_y)^{\tau_n} e^{\lambda (1 - \tau_y) m_h + \left(\frac{\lambda (1 - \tau_y)\Sigma^2_b}{2}\right)}$$

$$= e^{\lambda (1 - \tau_y) + \tau_n(1 - \tau_n) + \tau_n \alpha_1 m_h - \alpha_1 \tau_n \sigma_b^2 + \left(\frac{\lambda (1 - \tau_y)}{2}\right)^2 + \left(\frac{\alpha_1 \tau_n}{2}\right)^2}$$

Taking logs gives

$$\ln T_e = (-\tau_n \lambda + \alpha_1 \tau_m) m_h + \frac{\alpha_1 \tau_m}{2} (\alpha_1 \tau_m - 1) \sigma_b^2 - \tau_n (\ln A\ell \mu (1 - a_y))$$

$$+ \left[\left(\tau_n (1 - \tau_y) + \tau_n \alpha_1\right)^2 + \tau_n (\lambda (1 - \tau_y))^2 - \tau_n \lambda^2 - (\lambda (1 - \tau_y))^2\right] \Sigma^2_h$$

$$\iff \ln T_e = (-\tau_n \lambda + \alpha_1 \tau_m) m_h + \frac{\alpha_1 \tau_m}{2} (\alpha_1 \tau_m - 1) \sigma_b^2 - \tau_n (\ln A\ell \mu (1 - a_y))$$

$$+ \left[\lambda^2 (1 - \tau_y)^2 (\tau_n - 2) \tau_n + 2 \lambda (1 - \tau_n) (1 - \tau_y) \tau_n \alpha_1 + (\alpha_1 \tau_m)^2 - \tau_n \lambda^2 (2 - \tau_y) \tau_y\right] \Sigma^2_h$$

4. Solving for $T_u$
Substituting (16) into (20), one gets

\[ \int E_{h,y} [e_u(q, h_s, y)] f_q dq = \int T_u (E_{h,y} [e_u(q, h_s, y)])^{1-\tau_u} f_q dq \]

\[ \int \left( \frac{pI_q}{(1 + a_u)T_u} \right)^{1-\tau_u} f_q dq = \int \frac{pI_q}{(1 + a_u)T_u} f_q dq \]

\[ \left( \frac{pI}{(1 + a_u)T_u} \right)^{1-\tau_u} \int T_u^{1-\tau_u} f_q dq = \int T_u^{1-\tau_u} f_q dq \]

\[ \left( \frac{pI}{(1 + a_u)T_u} \right)^{1-\tau_u} \int I_q^{1-\tau_u} f_q dq = \int I_q f_q dq \]

\[ \left( \frac{pI}{(1 + a_u)T_u} \right)^{1-\tau_u} \int I_i^{1-\tau_u} f_i dq = \int I_i f_i dq \]

where \( i \) indexes households. I then guess that \( I_i \) is log-normally distributed with mean \( \mu_I \) and variance \( \sigma_I^2 \). I give expression for these variables in appendix A.6. Hence one gets:

\[ \left( \frac{pI}{1 + a_u} \right)^{\frac{\tau_u}{\tau_a}} e^{\mu_I + \frac{\sigma_I^2}{2(1-\tau_u)}} = (T_u)^{\frac{1}{1-\tau_u}} e^{\mu_I + \frac{\sigma_I^2}{2}} \]

\[ \left( \frac{pI}{1 + a_u} \right)^{\frac{\tau_u}{\tau_a}} e^{\mu_I + \frac{\sigma_I^2}{2} (\frac{\tau_u(2-\tau_u)}{1-\tau_u})} = (T_u)^{\frac{1}{1-\tau_u}} \]

\[ \Rightarrow \ln T_u = \tau_u \ln \left( \frac{pI}{1 + a_u} \right) + \mu_I + \frac{\sigma_I^2}{2} \left( \frac{\tau_u(2-\tau_u)}{1-\tau_u} \right) \]

From appendix A.6 \( I_q \) is log-normal, hence

\[ \ln E(I) = \mu_I + \frac{\sigma_I^2}{2} \]

From appendix A.3.4,

\[ \ln E(I) = \ln \frac{A\ell \mu s(1 + a_h)(1 + a_u)(1 - a_y)}{pI} + \lambda m_h + \lambda^2 \frac{\Sigma_h^2}{2} \]

Hence combining both expression together,

\[ \mu_I = \ln \frac{A\ell \mu s(1 + a_h)(1 + a_u)(1 - a_y)}{pI} + \lambda m_h + \lambda^2 \frac{\Sigma_h^2}{2} - \frac{\sigma_I^2}{2} \]

Substituting back into the expression for \( T_u \) gives

\[ \ln T_u = \tau_u \ln \left( \frac{pI}{1 + a_u} \right) + \tau_u \ln \left( \frac{A\ell \mu s(1 + a_h)(1 + a_u)(1 - a_y)}{pI} + \lambda m_h + \lambda^2 \frac{\Sigma_h^2}{2} \right) + \frac{\tau_u}{1 - \tau_u} \frac{\sigma_I^2}{2} \]

\[ = \tau_u \left( \ln A\ell \mu s(1 + a_h)(1 - a_y) + \lambda m_h + \lambda^2 \frac{\Sigma_h^2}{2} \right) + \frac{\tau_u}{1 - \tau_u} \frac{\sigma_I^2}{2} \]
From appendix A.6, I have an expression for $\sigma_I^2$

$$
\sigma_I^2 = \left( \frac{\alpha_1(1-\tau_u)}{(1-\tau_m)(1-\nu\omega_3)}(\tau_m + \omega_2(1-\nu\omega_3)\mu) \right)^2 \left( \sigma_b^2 + \left( \frac{\omega_I}{\omega_A} + 1 \right)^2 \Sigma_h^2 \right)
$$

\[\square\]

A.3.2 Educational Sector: Generalization

In this section I present an extension of the very simple educational sector presented in the main text. I start by assuming that the production function for educational services is given by

$$
y_I = A_I h^{\bar{\lambda} \ell}^{\mu}
$$

with $\bar{\lambda} \geq \lambda$

The assumption that $\bar{\lambda} \geq \lambda$ says that households with a lot of human capital have a comparative advantage in working for the educational services sector—a feature of the data.

Colleges buy services at price $p_I$ from the educational sector. The latter has two peculiarities. First of all it is made of one non-profit agency that aggregates heterogeneous inputs to produce an homogeneous educational service and whose objective is to minimize costs subject to a non-negative profit condition. Secondly it has full power in the setting of wages. I discuss these two assumptions below. The problem of the educational service agency willing to produce $\bar{I}$ is:

$$
\min_{d(h)} p_I \text{ s.t. } \int A_I h^{\bar{\lambda} \ell}^{\mu} d(h) dh \geq \bar{I} \tag{49}
$$

$$
p_I \int A_I h^{\bar{\lambda} \ell}^{\mu} d(h) dh \geq \int h^{\lambda \ell}^{\mu} d(h) dh \tag{50}
$$

where the right hand side of the last line embeds the assumption of bilateral bargaining with full power to the agency and $d(h)$ is the indicator function equal to 1 if a household with human capital $h$ is employed in the sector.

**Discussions of assumptions.** The assumption that wages are set through a bilateral negotiation where the agency has full bargaining power implies that wages are determined by the marginal product in the final good sector. These assumptions are not particularly unrealistic and buy us some tractability. The assumption of minimizing the price of educational services will imply that the price of educational services will not reflect the marginal cost of producing these services but rather their average cost. The joint assumption of price-minimizing monopoly allows us to get rid of profit in this sector, that would otherwise need to be given back to either the household sector or
the government.

The main reason why I model the educational sector separately from the final good sector is to allow for the supply-side Baumol-Bowen disease effect, widely discussed in the literature on the rise of college tuition. All the simplifying assumptions are to ensure that the model remains tractable with this additional block.

### A.3.3 Educational Sector: Hiring Rule

In the generalized version of the model, the equilibrium definition also includes a hiring rule \( d_t(h) \). The hiring rule of the agency problem (48) takes the form of the threshold rule where the most educated individuals work in the educational sector. The solution of the agency problem managing the educational sector implies that the agency always hires first the most educated individuals. This stems from the assumption that \( \tilde{\lambda}_t > \lambda_t \), which says that individuals with high human capital have a comparative advantage working for the agency.

**Proposition A.1.** Provided \( a_{u,t}, a_{h,t}, \beta \) are not too high, there exists a unique \( h_t \in \mathbb{R} \) such that

\[
d_t(h) = 1 \iff h \geq h_t
\]  

(51)

Let’s look at the difference between the product and cost of hiring human capital \( h \).

\[
p_I A_I h^{\bar{\lambda}} - Ah^{\lambda}
\]  

(52)

**Lemma 3.** (52) is increasing in \( h \) if

\[
\frac{E(h_{\bar{\lambda}} | h_{\bar{\lambda}} \geq 1)}{E(h_{h_{\lambda}} | h_{\lambda} \geq 1)} = \frac{m}{m} \left( \frac{m_{h_{\lambda}} - \log(h)}{\Sigma_h} - \bar{\lambda} \Sigma_h \right) \leq \frac{\bar{\lambda}}{\lambda} 
\]

where \( h_{\bar{\lambda}} = \left( \frac{h}{h} \right)^{\bar{\lambda}} \)

\[
h_{\lambda} = \left( \frac{h}{h} \right)^{\lambda}
\]

and \( m(.) \) denotes the Mills ratio.

**Proof.** We want to show that this is increasing in \( h \) so that the agency—willing to minimize the average cost—wants to hire first the individuals with highest human capital. The condition is:

\[
p_I A_I \bar{\lambda} h^{\bar{\lambda} - 1} > A \lambda h^{\lambda - 1}
\]

\[
\iff \bar{\lambda} h^{\bar{\lambda} - \lambda} > \frac{A}{A_{IPI}}
\]

65
It is sufficient to show that this is true for $h = \bar{h}$ since $\bar{h} > \lambda$, i.e.

$$
\frac{\bar{\lambda}}{\bar{\lambda}} \frac{\lambda - \lambda}{e^{\lambda h}} > \frac{A}{\mu p_I}
$$

$$
\frac{\bar{\lambda}}{\lambda} > \frac{\int_{\bar{h}}^{\lambda} \frac{A}{\mu p_I}}{\int_{\bar{h}}^{\lambda} \frac{A}{\mu p_I}} = \frac{E(h_{\bar{\lambda}} | h_{\lambda} > 1)}{E(h_{\lambda} | h_{\lambda} > 1)}
$$

where the last line uses the expression for the price $p_I$ given by (??). Let’s now show that the RHS of the last line is equal to the ratio of the Mills ratio. By the well-known formula for conditional expectation of log-normal distributions, one gets

$$
E(h_{\bar{\lambda}} | h_{\lambda} > 1) = e^{\lambda (m_h - \log h) + \frac{\lambda^2}{2} \Sigma_h} \Phi \left( \frac{\lambda (m_h - \log h) + \lambda \Sigma_h}{\lambda \Sigma_h} \right)
$$

$$
\Rightarrow \frac{E(h_{\bar{\lambda}} | h_{\lambda} > 1)}{E(h_{\lambda} | h_{\lambda} > 1)} = e^{\lambda - \lambda} (m_h - \log h) + \frac{\lambda^2}{2} \Sigma_h \Phi \left( \frac{m_h - \log h}{\Sigma_h} + \bar{\lambda} \Sigma_h \right)
$$

$$
= \frac{e^{\lambda - \lambda} (m_h - \log h) + \frac{\lambda^2}{2} \Sigma_h \Phi \left( \frac{m_h - \log h}{\Sigma_h} + \bar{\lambda} \Sigma_h \right)}{1 - \Phi \left( \frac{\log h - m_h}{\Sigma_h} - \lambda \Sigma_h \right)}
$$

$$
= \frac{e^{\lambda - \lambda} (m_h - \log h) + \frac{\lambda^2}{2} \Sigma_h \Phi \left( \frac{m_h - \log h}{\Sigma_h} - \bar{\lambda} \Sigma_h \right)}{m \frac{1}{\Sigma_h - \lambda \Sigma_h}}
$$

$$
= \frac{m \left( \frac{\log h - m_h}{\Sigma_h} - \bar{\lambda} \Sigma_h \right)}{m \left( \frac{\log h - m_h}{\Sigma_h} - \lambda \Sigma_h \right)}
$$

\[ \square \]

Lemma 4. For $h$ high enough,

$$
\frac{m \left( \frac{\log h - m_h}{\Sigma_h} - \bar{\lambda} \Sigma_h \right)}{m \left( \frac{\log h - m_h}{\Sigma_h} - \lambda \Sigma_h \right)} \leq \frac{\bar{\lambda}}{\lambda}
$$

Proof. First denote $m \left( \frac{\log h - m_h}{\Sigma_h} - \bar{\lambda} \Sigma_h \right) = \frac{m}{\lambda}$ and $m \left( \frac{\log h - m_h}{\Sigma_h} - \lambda \Sigma_h \right) = m \lambda$. It is well known that the Mills ratio associated with a standard normal is decreasing and convex. This implies that

$$
\frac{m}{m \lambda} > 1 \iff \frac{\bar{\lambda}}{\lambda}
$$

and

$$
\frac{\partial m}{\partial \log h} = \frac{m \lambda \left( \frac{m \lambda - m \lambda m \lambda}{m \lambda} \right)}{m \lambda} < 0
$$

66
since by convexity and by negativity

\[ m'_\lambda < m'_\bar{\lambda} \]

and

\[ -m'_\lambda < -m'_\lambda \frac{m_\bar{\lambda}}{m_\lambda} \]

\[ \Rightarrow m'_\bar{\lambda} - m_\lambda m'_\lambda < 0 \]

I now show that the limit of the ratio of Mills ratio tends to $+\infty$ when $h$ tends to 0. Denoting $X_\lambda(h) = \frac{\log h - m_h}{\Sigma_h} - \lambda \Sigma_h$ and similarly for $\bar{\lambda}$, one has

\[
\lim_{\log h \to -\infty} \frac{m(X_\lambda(h))}{m(X_\bar{\lambda}(h))} = \lim_{\log h \to +\infty} \frac{m(\bar{X}_\lambda(h))}{m(-X_\lambda(h))} \\
= \lim_{\log h \to +\infty} \frac{1 - \Phi(X_\lambda(h))}{\Phi(X_\bar{\lambda}(h))} m(X_\lambda(h)) \times 1 \\
= \lim_{\log h \to +\infty} \frac{-\phi(X_\lambda(h))}{-\phi(X_\bar{\lambda}(h))} \\
= \lim_{\log h \to +\infty} e^{\Sigma_h (\bar{\lambda} - \lambda) \left( \frac{\log h - m_h}{\Sigma_h} - (\lambda + \bar{\lambda}) \Sigma_h \right)} \\
= +\infty
\]

where the second line uses the symmetry of the standard normal cdf and pdf, the third uses l'Hôpital’s rule and the fourth that $\bar{\lambda} > \lambda$.

Hence the ratio of Mills ratios is decreasing, continuous, goes to $+\infty$ when $h$ approaches 0 and to 1 when it goes to $+\infty$. Therefore there exists a unique threshold at which it crosses $\frac{\bar{\lambda}}{\lambda}$ and is below it for $h$ higher. This finishes the proof.

Given the market clearing condition (53) pinning down $\log h$, it is clear that $a_u, a_h, s$ (and hence $\beta$) low enough ensures that $\log h$ will be high enough in equilibrium. I have checked that in our calibration this property holds.

### A.3.4 Market Clearing

It remains to be checked that the final good and the educational services market clear. By Walras’ law, if the latter clears, since college admission markets already clear, the final good market should clear. The market clearing condition in the educational sector requires that the demand for educational services coming from colleges—the left-hand-side of equation (53)—be equal to the supply of human capital supplied by the agency. It implicitly pins down the threshold $h_t$. The price of educational services is then given by the no-profit condition (50) of the agency. The following proposition summarizes these results.

**Lemma 5.** 1. The human capital threshold to work in the educational sector $h_t$ is implicitly given
Taking the logarithm of (25):

\[(1 + a_{u,t})(1 + a_{h,t})s_t e^{\lambda m_{h,t} + \frac{\lambda^2}{2} \Sigma_{h,t}^2 (1 - a_{y,t})} = H_t = e^{\lambda m_{h,t} + \frac{\lambda^2}{2} \Sigma_{h,t}^2 (1 - a_{y,t})} \]

where \(\Phi(.)\) denotes the c.d.f. of a standard normal.

2. Defining \(H_t^I = \int_{h_t}^{\infty} h^\lambda dh\), the equilibrium price of educational services is given by

\[
p_{I,t} = \frac{A \int_{h_t}^{1} h^\lambda dh}{A_I \int_{h_t}^{1} h^\lambda dh} = \frac{A (\lambda - \bar{\lambda}) m_{h,t} + (\lambda^2 - \bar{\lambda}^2) \frac{\Sigma_{h,t}^2}{2}}{\lambda - \bar{\lambda}} \Phi \left[ \frac{m_{h,t} - \lambda \Sigma_{h,t}}{\Sigma_{h,t}} + \lambda \Sigma_{h,t} \right] \\
(53)
\]

Proof. Demand for educational services is given by \(\frac{(1+a_u)(1+a_h)}{p_I} s \int_0^1 y_I d\bar{y}\) and supply is given by \(A_I \ell \int_2^{\infty} (h^\lambda) \bar{h} dh\) hence market clearing writes:

\[
(1 + a_u)(1 + a_h)s \int_0^1 y_I d\bar{y} = A_I \ell \mu \int_2^{\infty} (h^\lambda) \bar{h} dh
\]

To derive the expression for the equilibrium price (54), I first use the no-profit condition (50) and the threshold condition (51). I then make use of the well-known formula for conditional expectation of log-normal distribution to express in closed-form.

\[
A \int_{h_t}^{1} h^\lambda dh \\
A \int_{h_t}^{1} h^\lambda dh
\]

A.4 Within university distribution

A.4.1 Parent’s education and income

Taking the logarithm of (25):

\[
\ln q = (\varepsilon_I + \varepsilon_A) \ln h + \varepsilon_A \ln \xi_b + x
\]

with

\[
x = \varepsilon_1 \left( (1 - \tau_u) \ln \left( \frac{(1 + a_h)}{T_e} \right) + \ln \left( \frac{(1 + a_u)}{p_I} T_u \right) \right) + \varepsilon_1 (1 - \tau_u) (1 - \tau_n) - \varepsilon_3 \left[ (1 - \tau_y)(\ln A + \mu \ln \ell) + \ln T_y + \ln (1 - a_y) \right]
\]

with \(\varepsilon_I = (\varepsilon_1 (1 - \tau_u) (1 - \tau_n) - \varepsilon_3 (1 - \tau_y) \lambda + \varepsilon_A = \alpha_1 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)\). The subscripts 'I' ('A') stands for Income-sorting (Ability-sorting) channel as labeled in the main text. All pairs
(h, ξ_b) that verify this condition will go to a university with quality q. The distribution of parents human capital within a given university of quality q can therefore be computed explicitly. The mass of individuals with ln h and going to ln q is given by:

\[
\begin{align*}
& f\left(\frac{1}{\varepsilon_A} (\ln q - x - (\varepsilon_I + \varepsilon_A) \ln h) \cap \ln h\right) \\
& = f_{\xi_b} \left(\frac{1}{\varepsilon_A} (\ln q - x - (\varepsilon_I + \varepsilon_A) \ln h)\right) f_h (\ln h) \\
& = \phi \left(\frac{1}{\varepsilon_A} (\ln q - x - (\varepsilon_I + \varepsilon_A) \ln h)\right), \mu_b, \sigma^2_b \phi \left(\ln h, m_h, \Sigma^2_h\right) \\
& = \phi \left(\ln h, \frac{1}{\varepsilon_I + \varepsilon_A} (\ln q - x - \varepsilon_A \mu_b)\right), \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^2 \frac{\sigma^2_b}{\sigma^2_I} \phi \left(\ln h, m_h, \Sigma^2_h\right) \\
& = \phi \left(\ln h, \mu^q_1, \sigma^2_I \right) \phi \left(\ln h, m_h, \Sigma^2_h\right) \\
& = \phi \left(\mu^q_1, m_h, \sigma^2_I + \Sigma^2_h\right) \phi \left(\ln h, \mu^q_2, \sigma^2_2\right)
\end{align*}
\]

where the RHS is the mass of individuals going to quality q and the LHS is the density of people whose parents have human capital h conditional on college q.

\[
\begin{align*}
\mu^q_2 &= \frac{\Sigma^{-2}_h m_h + \sigma^{-2}_I \mu^q_1}{\sigma^{-2}_I + \Sigma^{-2}_h} \\
&= \frac{\Sigma^{-2}_h m_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^{-2} \sigma^{-2}_b \frac{(\ln q - x - \varepsilon_A \mu_b)}{\varepsilon_I + \varepsilon_A}}{\Sigma^{-2}_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^{-2} \sigma^{-2}_b} \\
&= \mu_{2,1} m_h + \mu_{2,2} (\ln q - x - \varepsilon_A \mu_b)
\end{align*}
\]

with

\[
\begin{align*}
\mu_{2,1} &= \frac{\Sigma^{-2}_h}{\Sigma^{-2}_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^{-2} \sigma^{-2}_b} \\
\mu_{2,2} &= \left[\Sigma^{-2}_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^{-2} \sigma^{-2}_b\right] (\varepsilon_I + \varepsilon_A)
\end{align*}
\]

and where

\[
\sigma^2_2 = \frac{\sigma^2_I \Sigma^2_h}{\sigma^2_I + \Sigma^2_h} = \left[\Sigma^2_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^2 \sigma^2_b\right] \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^2 \sigma^2_b + \frac{\Sigma^2_h \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^2 \sigma^2_b}{\Sigma^2_h + \left(\frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A}\right)^2 \sigma^2_b}
\]

where the second line stems from independence of h and ξ_b. The mass of individuals studying in universities of quality q is given by \(\phi \left(\mu^q_1, m_h, \sigma^2_I + \Sigma^2_h\right)\) and the density conditional on being in the
university of quality $q$ of $\ln h$ is therefore given by:

$$\phi \left( \ln h, \mu_2, \sigma_2^2 \right)$$

It depends on $q$ because $\mu_2$ through $\mu_1$ is a function of $q$. More precisely $\mu_2$ is increasing in $q$, hence as one go through the quality ladder of colleges, students comes from richer and richer/more and more educated background.

From the distribution of parents’ human capital within a university one can get the distribution of parents’ after tax income within the university.

$$\ln y = \ln(1 - a_y) + (1 - \tau_y) \left[ \ln A + \lambda \ln h + \mu \ln \ell \right] + \ln T_y$$

$$\Rightarrow \ln y \sim \mathcal{N} \left( \ln(1 - a_y) + (1 - \tau_y) \left[ \ln A + \lambda \mu_2 + \mu \ln \ell \right] + \ln T_y, (1 - \tau_y)^2 \lambda^2 \sigma^2_2 \right)$$

### A.4.2 Distribution of college quality

Since $\phi \left( \mu_1^q, m_h, \sigma_1^2 + \Sigma^2_h \right)$—with $\mu_1 = \frac{1}{\varepsilon_I + \varepsilon_A} (\ln q - x - \varepsilon_A \mu_b)$ and $\sigma_1^2 = \left( \frac{\varepsilon_A}{\varepsilon_A + \varepsilon_I} \right)^2 \sigma_b^2$—is the mass of students in college of quality $q$, the distribution of quality is given by:

$$\ln q \sim \mathcal{N} \left( (\varepsilon_I + \varepsilon_A) m_h + x + \varepsilon_A \mu_b, \varepsilon_A^2 \sigma_b^2 + (\varepsilon_I + \varepsilon_A)^2 \Sigma^2_h \right)$$

### A.4.3 Students’ abilities

From the definition of abilities and the condition used above:

$$\ln h_s = \alpha_1 \ln h + \alpha_1 \ln \xi_b$$

and

$$\ln q = (\varepsilon_I + \varepsilon_A) \ln h + \varepsilon_A \ln \xi_b + x$$

$$\Rightarrow \ln q = \varepsilon_I \ln h + \frac{\varepsilon_A}{\alpha_1} \ln h_s + x$$

Fixing $\ln h_s$ to some value, the students with ability $\ln h_s$ are the ones whose parents have human capital

$$\ln h = \frac{1}{\varepsilon_I} \left( \ln q - \frac{\varepsilon_A}{\alpha_1} \ln h_s - x \right)$$

Therefore the mass of students with this ability stems from the distribution of parents’ human capital within the university:

$$\ln h_s \sim \mathcal{N} \left( \frac{\alpha_1}{\varepsilon_A} (\ln q - \varepsilon_I \ln h - x), \left( \frac{\alpha_1 \varepsilon_I}{\varepsilon_A} \right)^2 \sigma_2^2 \right)$$
A.4.4 Summary of quality distribution and within college distributions

Proposition A.2. 1. The distribution of college quality is given by

\[ \ln q \sim \mathcal{N}\left( (\varepsilon_I + \varepsilon_A)m_h + x + \varepsilon_A \mu_b, \varepsilon_A^2 \sigma_b^2 + (\varepsilon_I + \varepsilon_A)^2 \Sigma^2 \right) \]

and \( x \) has been defined earlier.

2. Within a college of quality \( q \), the distribution of parents’ (log) human capital is given by:

\[
f(\ln h|q) \sim \mathcal{N}\left( \sum^{-2}_h m_h + \left( \frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A} \right)^{-2} \sigma_b^{-2} \frac{\ln q - x - \varepsilon_A \mu_b}{\varepsilon_I + \varepsilon_A}, \frac{\Sigma^2_h}{\sum^{-2}_h + \left( \frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A} \right)^{-2} \sigma_b^{-2}} \right)
\]

\[
\sim \mathcal{N}\left( \mu_2^q, \sigma_2^q \right)
\]

3. And the distribution of students’ (log) abilities is given by

\[
f(\ln h_s|q) \sim \mathcal{N}\left( \frac{\alpha_1}{\varepsilon_A} \ln q - \varepsilon_I \mu_2^q - x, \left( \frac{\alpha_1 \varepsilon_I}{\varepsilon_A} \right)^2 \sigma_2^q \right)
\]

A.4.5 Solving for \( \kappa_2 \) and \( \nu \)

The initial guess was that \( D = \kappa_2 q^\nu \). Let’s solve for \( \kappa_2 \) and \( \nu \) since \( \ln D \) is the mean log (after tax) income within a university.

\[
\ln D = \ln(1 - a_y) (A \ell^1 - \tau_y) T_y + (1 - \tau_y) \lambda \ln \mu_2^q
\]

\[
= \ln(1 - a_y) (A \ell^1 - \tau_y) T_y + (1 - \tau_y) \lambda (\mu_{2,1} m_h + \mu_{2,2} (\ln q - x - \varepsilon_A \mu_b))
\]

Identifying coefficients with the guess \( \ln D = \ln \kappa_2 + \nu \ln q \), one gets:

\[
\nu = (1 - \tau_y) \lambda \mu_{2,2}
\]

\[
\Leftrightarrow \nu = (1 - \tau_y) \lambda \sum^{-2}_h + \left( \frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A} \right)^{-2} \sigma_b^{-2} \frac{\varepsilon_I + \varepsilon_A}{\varepsilon_I + \varepsilon_A}
\]

\[
\Leftrightarrow \nu = (1 - \nu \omega_3) (1 - \tau_y) \lambda \left[ \sum^{-2}_h + \left( \frac{\omega_A}{\omega_I + \omega_A} \right)^{-2} \sigma_b^{-2} \right] (\omega_I + \omega_A)
\]

\[
\Leftrightarrow \nu = \frac{(1 - \tau_y) \lambda \left( \frac{\omega_A}{\omega_I + \omega_A} \right)^{-2} \sigma_b^{-2}}{\sum^{-2}_h + \left( \frac{\omega_A}{\omega_I + \omega_A} \right)^{-2} \sigma_b^{-2} \left[ \omega_I + \omega_A \right] + \omega_3 (1 - \tau_y) \lambda \left( \frac{\omega_A}{\omega_I + \omega_A} \right)^{-2} \sigma_b^{-2}}
\]
Dividing the numerator and denominator by \((1 - \tau_y)\lambda \) and then by \(\left(\frac{\omega_A}{\omega_f + \omega_A}\right)^{-2} \sigma_b^{-2} \) gives

\[
\nu = \frac{\left(\frac{\omega_A}{\omega_f + \omega_A}\right)^{-2} \sigma_b^{-2}}{\Sigma_h^{-2} + \left(\frac{\omega_A}{\omega_f + \omega_A}\right)^{-2} \sigma_b^{-2} \left[\left(\omega_1(1 - \tau_u)(1 - \tau_n) - \omega_3\right) + \omega_3 \left(\frac{\omega_A}{(1 - \tau_y)\lambda}\right) + \omega_3\right]}
\]

\[
= \frac{\Sigma_h^{-2} \left(\frac{\omega_A}{\omega_f + \omega_A}\right)^{2} \sigma_b^{2} + 1}{\Sigma_h^{-2} \left(\frac{\omega_A}{\omega_f + \omega_A}\right)^{2} \sigma_b^{2} + 1 \left[\left(\omega_1(1 - \tau_u)(1 - \tau_n) - \omega_3\right) + \omega_3 \left(\frac{\omega_A}{(1 - \tau_y)\lambda}\right) + \omega_3\right]}
\]

\(\nu\) is therefore only a function of \(\Sigma_h^2\). Identifying the coefficient independent of \(\ln q\), one gets:

\[
\kappa_2 = (1 - a_y) (A\ell^\mu)^{1 - \tau_y} T_y e^{(1 - \tau_y)\lambda}(\mu_{2,1}m_h - \mu_{2,2}(x + \varepsilon_A\mu_b))
\]

\(x_t\) is a linear function of \(\ln \kappa_2\), hence one gets:

\[
\kappa_2 = (1 - a_y) (A\ell^\mu)^{1 - \tau_y} T_y e^{(1 - \tau_y)\lambda}(\mu_{2,1}m_h - \mu_{2,2}(x + \varepsilon_A\mu_b))
\]

\[
\leftrightarrow \kappa_2 = (1 - a_y) (A\ell^\mu)^{1 - \tau_y} T_y e^{(1 - \tau_y)\lambda}(\mu_{2,1}m_h - \mu_{2,2}(\tilde{x} + \varepsilon_A\mu_b))\kappa_2^{-(1 - \tau_y)\lambda}\mu_{2,2}\varepsilon_3
\]

\[
\leftrightarrow \kappa_2 = \left((1 - a_y) (A\ell^\mu)^{1 - \tau_y} T_y e^{(1 - \tau_y)\lambda}(\mu_{2,1}m_h - \mu_{2,2}(\tilde{x} + \varepsilon_A\mu_b))\right)^{\frac{1}{1 - (1 - \tau_y)\lambda}\mu_{2,2}\varepsilon_3}
\]

\[
\leftrightarrow \kappa_2 = \left((1 - a_y) (A\ell^\mu)^{1 - \tau_y} T_y e^{(1 - \tau_y)\lambda}(\mu_{2,1}m_h - \mu_{2,2}(\tilde{x} + \varepsilon_A\mu_b))\right)^{1 - \nu\omega_3}
\]

with \(\tilde{x} = x - \varepsilon_3 \ln \kappa_2\) independent of \(\kappa_2\)

### A.5 Law of motion

Replacing \(\kappa_2\) (second line), \(\ln T_y\) (third line), \(\ln T_o\) (fourth line) and \(\ln T_u\) (fifth line) obtained above in the law of motion for human capital one gets:
\[
\ln h' = \ln \xi_y + \ln \kappa + \alpha_1 (1 + \alpha_2 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)) \ln \xi_b + \alpha_h \ln h \\
+ \alpha_2 \varepsilon_1 \left( (1 - \tau_u) (\ln s + a_h) - \ln T_e + \ln \left( \frac{1 + a_u T_u}{p_f} \right) \right) \\
+ \alpha_2 \varepsilon_1 (1 - \tau_u) (1 - \tau_n - \varepsilon_3) \left( (1 - \tau_y) (\ln \Lambda T^\mu) + \ln (1 - a_y) T_y \right) + \alpha_2 \varepsilon_3 \ln \kappa_2
\]

\[
\ln h' = \ln \xi_y + \ln \kappa + \alpha_1 (1 + \alpha_2 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)) \ln \xi_b + \alpha_h \ln h \\
+ \alpha_2 (1 + \nu \omega_3) \varepsilon_1 \left( (1 - \tau_u) (\ln s + \ln T_e + \ln (1 + a_h)) + \ln \left( \frac{1 + a_u T_u}{p_f} \right) \right) \\
+ \alpha_2 (\omega_3 + (1 - \nu \omega_3) \varepsilon_1 (1 - \tau_u) (1 - \tau_n - \varepsilon_3)) \left( (1 - \tau_y, t) (\ln \Lambda T^\mu) + \ln (1 - a_y) + \ln T_y \right) \\
+ \alpha_2 \omega_3 \nu \varepsilon A \sigma_b^2 + \alpha_2 \omega_3 (1 - \tau_y) \lambda \mu_{2,1} m_h
\]

\[
\ln h' = \ln \xi_y + \ln \kappa + \alpha_1 (1 + \alpha_2 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)) \ln \xi_b + \alpha_h \ln h \\
+ \alpha_2 \omega_1 \left( (1 - \tau_u) (\ln s + \ln T_e + \ln (1 + a_h)) + \ln \left( \frac{1 + a_u T_u}{p_f} \right) \right) \\
+ \alpha_2 (\omega_3 + (\omega_1 (1 - \tau_u) (1 - \tau_n) - \omega_3)) \left( \ln \Lambda T^\mu + \ln (1 - a_y) + \tau_y \lambda m_h + \frac{\chi^2}{2} (2 - \tau_y) \tau_y \Sigma_h^2 \right) \\
+ \alpha_2 \omega_3 \nu \varepsilon A \sigma_b^2 + \alpha_2 \omega_3 (1 - \tau_y) \lambda \mu_{2,1} m_h
\]
\[
\ln h' = \ln \xi_y + \ln \kappa + \alpha_1(1 + \alpha_2(\varepsilon_2 + \tau_m(1 - \tau_u)\varepsilon_1)) \ln \xi_h + \alpha_h \ln h \\
+ \alpha_2 \omega_1 \left( (1 - \tau_u) \left( \ln s + \ln(1 + a_h) \right) + \ln \frac{(1 + a_u)I_u}{P_I} \right) \\
+ \alpha_2 \omega_1(1 - \tau_u)(1 - \tau_n) \left( \ln A_\ell \mu + \ln(1 - a_y) + \tau_y \lambda m_h + \frac{\lambda^2}{2}(1 - \tau_y) \tau_y \Sigma_h^2 \right) \\
+ \alpha_2 \omega_3 \nu A \sigma_b^2 + \alpha_2 \omega_3(1 - \tau_y) \lambda \mu_{2,1} m_h \\
+ \alpha_2 \omega_1(1 - \tau_u)(1 - \tau_n) \left[ (\tau_n \lambda - \alpha_1 \tau_m) m_h + \frac{\alpha_1 \tau_m}{2} (1 - \alpha_1 \tau_m) \sigma_b^2 \right]
\]

\[
\ln h' = \ln \xi_y + \ln \kappa + \alpha_1(1 + \alpha_2(\varepsilon_2 + \tau_m(1 - \tau_u)\varepsilon_1)) \ln \xi_h + \alpha_h \ln h \\
+ \alpha_2 \omega_1 \left( \ln s + \ln(1 + a_h) + \ln \frac{(1 + a_u)I_u}{P_I} + \tau_u \left( \ln A_\ell \mu(1 - a_y) + \lambda m_h + \frac{\lambda^2}{2} (1 - \tau_y) \tau_y \Sigma_h^2 \right) + \frac{\tau_u}{1 - \tau_u} \sigma_f^2 \right) \\
+ \alpha_2 \omega_1(1 - \tau_u)(1 - \tau_n) \left( \ln A_\ell \mu + \ln(1 - a_y) + \tau_y \lambda m_h + \frac{\lambda^2}{2}(1 - \tau_y) \tau_y \Sigma_h^2 \right) \\
+ \alpha_2 \omega_3 \nu A \sigma_b^2 + \alpha_2 \omega_3(1 - \tau_y) \lambda \mu_{2,1} m_h \\
+ \alpha_2 \omega_1(1 - \tau_u)(1 - \tau_n) \left[ (\tau_n \lambda - \alpha_1 \tau_m) m_h + \frac{\alpha_1 \tau_m}{2} (1 - \alpha_1 \tau_m) \sigma_b^2 \right]
\]

\[
- \left[ \lambda^2 (1 - \tau_y)^2 (\tau_n - 2) \tau_n + 2 \lambda (1 - \tau_n)(1 - \tau_y) \tau_m \alpha_1 + (\alpha_1 \tau_m)^2 - \tau_n \lambda^2 (2 - \tau_y) \tau_y \right] \frac{\Sigma_h^2}{2}
\]
This implies the following law of motion for $m_h$:

\[
m_h' = \rho m_h - \frac{\sigma_y^2}{2} + \ln \kappa - \alpha_1 (1 + \alpha_2 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)) + \frac{\tau_u - \sigma_i^2}{2} + \alpha_2 \omega_1 \left( \ln s (1 + a_u) (1 + a_h) - \ln p_I + \tau_u \left( \ln A \ell^\mu (1 - a_y) + \frac{\lambda^2 \Sigma_h^2}{2} \right) + \frac{\tau_u - \sigma_i^2}{2} \right) + \alpha_2 \omega_1 (1 - \tau_u) (1 - \tau_u) \left( \ln A \ell^\mu (1 - a_y) + \frac{\lambda^2}{2} (2 - \tau_y) \tau_y \Sigma_h^2 \right) + \alpha_2 \omega_2 (1 - \tau_u) \left[ \frac{\alpha_1 \tau_m}{2} (1 - \alpha_1 \tau_m) \sigma_b^2 + \frac{\tau_u - \sigma_i^2}{2} \ln A \ell^\mu (1 - a_y) \right] - \left[ \lambda^2 (1 - \tau_y)^2 (\tau_n - 2) \tau_n + 2 \lambda (1 - \tau_n) (1 - \tau_y) \tau_m \alpha_1 + (\alpha_1 \tau_m)^2 - \tau_n \lambda^2 (2 - \tau_y) \tau_y \frac{\Sigma_h^2}{2} \right]
\]

\[
m_h' = \rho m_h - \frac{\sigma_y^2}{2} + \ln \kappa + \left[ \alpha_2 \omega_1 \alpha_1 \tau_m (1 - \tau_u) (1 - \alpha_1 \tau_m) + \alpha_2 \omega_2 (1 + \alpha_2 (\varepsilon_2 + \tau_m (1 - \tau_u) \varepsilon_1)) \right] + \frac{\tau_u - \sigma_i^2}{2} + \alpha_2 \omega_1 \left( \ln s (1 + a_u) (1 + a_h) - \ln p_I + \tau_u \left( \ln A \ell^\mu (1 - a_y) + \frac{\lambda^2 \Sigma_h^2}{2} \right) + \frac{\tau_u - \sigma_i^2}{2} \right) + \alpha_2 \omega_1 (1 - \tau_u) (1 - \tau_u) \left( \ln A \ell^\mu (1 - a_y) + \frac{\lambda^2}{2} (2 - \tau_y) \tau_y \Sigma_h^2 \right) + \alpha_2 \omega_1 (1 - \tau_u) \left[ \tau_n (\ln A \ell^\mu (1 - a_y)) \right] - \left[ \lambda^2 (1 - \tau_y)^2 (\tau_n - 2) \tau_n + 2 \lambda (1 - \tau_n) (1 - \tau_y) \tau_m \alpha_1 + (\alpha_1 \tau_m)^2 - \tau_n \lambda^2 (2 - \tau_y) \tau_y \frac{\Sigma_h^2}{2} \right]
\]

where the first line factorizes out all the terms in $m_h$ and the second all the terms in $\sigma_y^2$. The next steps consist in simplifying the coefficient in front of $\sigma_b^2$, of factorizing out all the terms in $\Sigma_h^2$ and
in using the expression in (A.6) for $\sigma^2_f$.

\[
m'_h = \rho m_h - \frac{\sigma^2_y}{2} + \ln \kappa - \alpha_1 \left( \alpha_2 \left( \omega_2 + \omega_1 (1 - \tau_u) \right) (\tau_m^2 \alpha_1) + 1 \right) \frac{\sigma^2_h}{2} \\
+ \alpha_2 \omega_1 \left( \ln s(1 + a_u)(1 + a_h) - \ln p_I + \tau_u \left( \ln A\ell^\mu(1 - a_y) + \lambda^2 \frac{\Sigma^2_h}{2} \right) + \frac{\tau_u}{1 - \tau_u} \frac{\sigma^2_I}{2} \right) \\
+ \alpha_2 \omega_1 \ln A\ell^\mu(1 - a_y) \\
+ \alpha_2 \omega_1 (1 - \tau_u) \left[ \lambda^2 - \left( \lambda(1 - \tau_y) (1 - \tau_n) + (\alpha_1 \tau_m) \right) \right] \frac{\Sigma^2_h}{2}
\]

\[
m'_h = \rho m_h - \frac{\sigma^2_y}{2} + \ln \kappa \\
+ \left[ \frac{\tau_u}{1 - \tau_u} \left( \frac{\alpha_1(1 - \tau_u)}{(1 - \tau_n) \nu} \right) (\tau_m + \omega_2 (1 - \tau_n) \nu) \right] - \alpha_1 \left( \alpha_2 \left( \omega_2 + \omega_1 (1 - \tau_u) \right) (\tau_m^2 \alpha_1) + 1 \right) \frac{\sigma^2_h}{2} \\
+ \alpha_2 \omega_1 \left( \ln A\ell^\mu(1 - a_y) s(1 + a_u)(1 + a_h) - \ln p_I \right) \\
+ \alpha_2 \omega_1 \left[ \lambda^2 + \frac{\tau_u}{1 - \tau_u} \left( \frac{\alpha_1(1 - \tau_u)}{(1 - \tau_n) \nu} \right) (\tau_m + \omega_2 (1 - \tau_n) \nu) \right] \left( \frac{\omega_I}{\omega_A} + 1 \right)^2 \\
\quad - \left( \frac{\tau_u}{1 - \tau_u} \left( \lambda(1 - \tau_y) (1 - \tau_n) + (\alpha_1 \tau_m) \right) \right) \frac{\Sigma^2_h}{2}
\]

with

\[
\rho = \alpha_h + \alpha_2 \omega_1 \tau_u \lambda + \alpha_2 \omega_1 (1 - \tau_u) (1 - \tau_n) \tau_y \lambda + \alpha_2 \omega_1 (1 - \tau_u) (\tau_n \lambda - \alpha_1 \tau_m) + \alpha_2 \omega_3 (1 - \tau_y) \lambda \mu_{2,1}
\]

\[
= \alpha_h + \alpha_2 \left[ \omega_1 \left( (1 - \tau_u) \left( (1 - \tau_n) \tau_y \lambda + \tau_n \lambda - \alpha_1 \tau_m \right) + \tau_u \lambda \right) + \omega_3 (1 - \tau_y) \lambda (1 - \mu_{2,2}(\varepsilon_I + \varepsilon_A)) \right] \\
= \alpha_1 + \alpha_3 + \alpha_2 \left[ \omega_1 \left( (1 - \tau_u) \left( (1 - \tau_n) \tau_y \lambda + \tau_n \lambda - \alpha_1 \tau_m \right) + \tau_u \lambda \right) + \omega_3 (1 - \tau_y) \lambda (1 - \nu \omega_3) \right] \\
= \alpha_1 + \alpha_3 + \alpha_2 \left[ \omega_1 \left( (1 - \tau_u) \left( (1 - \tau_n) \tau_y \lambda + \tau_n \lambda - \alpha_1 \tau_m \right) + \tau_u \lambda \right) + \omega_3 (1 - \tau_y) \lambda (1 - \nu \omega_3) \right]
\]

\[
\rho = \alpha_1 + \alpha_3 + \alpha_1 \alpha_2 \omega_2 + \alpha_2 \omega_1 \lambda
\]

where the second line stems from $\mu_{2,1} = 1 - \mu_{2,2}(\varepsilon_I + \varepsilon_A)$. Finally taking the variance immediately gives the expression for $\Sigma^2_h$:

\[
\Sigma^2_h = \sigma^2_y + (\alpha_1(1 + \alpha_2(\varepsilon_2 + \tau_m(1 - \tau_u)\varepsilon_1))) \sigma^2_h + (\alpha_h)^2 \Sigma^2_h
\]

76
A.6 From the distribution of $\ln q$ to the distribution of $\ln I$.

Using the definition of $q$ and the expression for $\theta$ obtained earlier,

$$\ln q = \omega_1 \ln \tilde{I} + \omega_2 \ln \theta$$

$$\ln q = \omega_1 \ln \tilde{I} + \omega_2 \left( \frac{\alpha_1}{\varepsilon_A} (\ln q - \varepsilon_I \mu_2^q - x) \right)$$

$$\ln \tilde{I} = \frac{1}{\omega_1} \left( \ln q \left( 1 - \alpha_1 \frac{\omega_2}{\varepsilon_A} + \alpha_1 \frac{\omega_2}{\varepsilon_A} (\varepsilon_I \mu_2 + x) \right) \right).$$

Given the expression for $\mu_2^q$:

$$\mu_2^q = \mu_{2,1} m_h + \mu_{2,2} (\ln q - \varepsilon_A \mu_b)$$

one obtains

$$\ln \tilde{I} = \frac{1}{\omega_1} \left( \ln q \left( 1 - \alpha_1 \frac{\omega_2}{\varepsilon_A} + \alpha_1 \frac{\omega_2}{\varepsilon_A} \varepsilon_I \mu_2 + (1 - \varepsilon_I \mu_2) x - \varepsilon_I \mu_2 \varepsilon_A \mu_b \right) \right)$$

Hence from the distribution $\ln q$ I can recover the distribution of $\ln \tilde{I}$:

$$\ln q \sim \mathcal{N} \left( (\varepsilon_I + \varepsilon_A) m_h + x + \varepsilon_A \mu_b, \varepsilon_A^2 \sigma_b^2 + (\varepsilon_I + \varepsilon_A)^2 \Sigma_h^2 \right)$$

$$\sim \mathcal{N} \left( \mu_{\tilde{I}}, \sigma_{\tilde{I}}^2 \right)$$

$$\Rightarrow \ln \tilde{I} \sim \mathcal{N} \left( \mu_{\tilde{I}}, \sigma_{\tilde{I}}^2 \right)$$

with

$$\mu_{\tilde{I}} = \frac{1}{\omega_1} \left( \mu_q \left( 1 - \alpha_1 \frac{\omega_2}{\varepsilon_A} + \alpha_1 \frac{\omega_2}{\varepsilon_A} \varepsilon_I \mu_2 + (1 - \varepsilon_I \mu_2) x - \varepsilon_I \mu_2 \varepsilon_A \mu_b \right) \right)$$

and

$$\sigma_{\tilde{I}}^2 = \left( \frac{\alpha_1 (1 - \tau_m)}{(1 - \nu \omega_3)} (\tau_m + \omega_2 (1 - \tau_n) \nu) \right)^2 \left( \frac{\omega_I}{\omega_A + 1} \right)^2 \Sigma_h^2$$

The last line stems from

$$\frac{1}{\omega_1} \left( 1 - \alpha_1 \frac{\omega_2}{\varepsilon_A} + \alpha_1 \frac{\omega_2}{\varepsilon_A} \varepsilon_I \mu_2 + (1 - \varepsilon_I \mu_2) x - \varepsilon_I \mu_2 \varepsilon_A \mu_b \right)$$

$$= \frac{1}{\varepsilon_A \omega_1} \left( \varepsilon_A - \alpha_1 \omega_2 + \alpha_1 \omega_2 \varepsilon_I \mu_2 + (1 - \varepsilon_I \mu_2) x - \varepsilon_I \mu_2 \varepsilon_A \mu_b \right)$$

$$= \frac{1}{\varepsilon_A \omega_1} \left( \varepsilon_I (1 - \tau_m) \tau_m + \varepsilon_2 v \omega_3 + \omega_2 \varepsilon_I \mu_2 \right)$$

$$= \frac{1}{\varepsilon_A \omega_1} \left( \varepsilon_I (1 - \tau_u) \tau_m + \omega_2 \varepsilon_1 (1 - \tau_u) (1 - \tau_n) \nu \right)$$

$$= \frac{\alpha_1 (1 - \tau_u)}{\varepsilon_A (1 - \nu \omega_3)} (\tau_m + \omega_2 (1 - \tau_n) \nu)$$
where I used $\nu = (1 - \tau_y)\lambda\mu_{2,2}$ and $\varepsilon_I = \frac{\omega_I}{1 - \nu\omega_3}$.

$$
\sigma^2_I = \left( \frac{\alpha_1(1 - \tau_u)}{\varepsilon_A(1 - \nu\omega_3)} \right) \left( \tau_m + \omega_2(1 - \tau_n)\nu \right)^2 \sigma^2_q
$$

$$
= \left( \frac{\alpha_1(1 - \tau_u)}{\varepsilon_A(1 - \nu\omega_3)} \right) \left( \tau_m + \omega_2(1 - \tau_n)\nu \right)^2 \left( \varepsilon_A\sigma^2_h + (\varepsilon_I + \varepsilon_A)^2\Sigma^2_h \right)
$$

$$
= \left( \frac{\alpha_1(1 - \tau_u)}{(1 - \nu\omega_3)} \right) \left( \tau_m + \omega_2(1 - \tau_n)\nu \right)^2 \left( \sigma^2_h + \left( \frac{\omega_I}{\omega_A} + 1 \right)^2\Sigma^2_h \right)
$$

Since $\ln \tilde{I} = \ln I - (1 - \tau_u)\frac{\sigma^2_u}{2}$ and $\sigma^2_u$ is common across all colleges, I have

$$
\ln I \sim \mathcal{N} \left( \mu_I, \sigma^2_I \right)
$$

with $\mu_I = \mu_I + (1 - \tau_u)\frac{\sigma^2_u}{2}$

and $\sigma^2_I = \sigma^2_I$

**Expression for $\sigma^2_u$**

Given that all households save a fraction $s$ of their disposable income (first line) and the selection equation into college (second line), one gets

$$
\ln e_u = \ln \left( \frac{1 + a_h}{T_e} \right) s + \tau_m \ln h + (1 - \tau_n) \ln y
$$

$$
= \ln \left( \frac{1 + a_h}{T_e} \right) s + \tau_m \frac{\alpha_1}{\varepsilon_A} \ln q - \varepsilon_I \ln h - x + (1 - \tau_n)(1 - \tau_y)\lambda\ln h
$$

$$
+ (1 - \tau_n)\ln \left( 1 - a_y \right) A^{1 - \tau_y} \lambda^{1 - \tau_y}\mu
$$

$$
= \ln \left( \frac{1 + a_h}{T_e} \right) s + \tau_m \frac{\alpha_1}{\varepsilon_A} \ln q - x + (1 - \tau_n)(1 - \tau_y)\lambda - \tau_m \frac{\alpha_1}{\varepsilon_A} \varepsilon_I \ln h
$$

$$
+ (1 - \tau_n)\ln \left( 1 - a_y \right) A^{1 - \tau_y} \lambda^{1 - \tau_y}\mu
$$

Hence the within-university variance of tuitions is given by:

$$
\sigma^2_u = \left( (1 - \tau_n)(1 - \tau_y)\lambda - \tau_m \frac{\alpha_1}{\varepsilon_A} \varepsilon_I \right)^2 \sigma^2_2
$$

$$
= \left( (1 - \tau_y)\lambda \frac{(1 - \tau_n)\omega_2 + \tau_m\omega_3}{\omega_2 + \omega_1(1 - \tau_n)\tau_m} \right)^2 \sigma^2_2
$$

which is indeed constant across universities wince $\sigma^2_2$ is an aggregate constant.

**A.7 Existence and Uniqueness of Equilibrium Path**

The set of equations defining an equilibrium path (1) is block-recursive. In particular, the law of motion of $\Sigma_h$, is independent and the path of all other variables are uniquely pinned-down by the
path of $\Sigma_h$. It is therefore necessary and sufficient to focus on the existence and uniqueness of the path of $\Sigma_h$. Let’s define new notations to simplify the algebra:

$$\Sigma_h^2 = f(\Sigma_h^2)$$

$$= (\alpha_h (\Sigma_h^2))^2 \Sigma_h^2 + X_2 (\Sigma_h^2)$$

$$= \left[ \alpha_1^2 + \left( \frac{A}{1-\nu \omega_3} \right)^2 + \frac{2\alpha_1 A}{1-\nu \omega_3} \right] \Sigma_h^2 + \sigma_y^2 + \left[ \alpha_1^2 + \frac{B^2}{(1-\nu \omega_3)^2} + \frac{2B\alpha_1}{1-\nu \omega_3} \right] \sigma_b^2$$

with $A = \alpha_1 \alpha_2 (\omega_2 + \tau_m (1-\tau_u) \omega_1) + \alpha_2 (\omega_1 (1-\tau_u) (1-\tau_n) - \omega_3) (1-\tau_y) \lambda$

$$B = \alpha_1 \alpha_2 (\omega_2 + \tau_m (1-\tau_m) \omega_1)$$

$$\nu = \frac{CE \Sigma_h^2}{(E + \omega_3) C}$$

$$C = \left( \frac{\omega A}{\omega I + \omega A} \right)^{-2} \sigma_b^{-2}$$

$$E = (\omega_1 (1-\tau_u) (1-\tau_n) - \omega_3) + \frac{\omega A}{(1-\tau_y) \lambda}$$

$f(.)$ is differentiable for $\Sigma_h^2 \in (0, \infty)$ and

$$\lim_{\Sigma_h^2 \to 0} f(\Sigma_h^2) = \sigma_y^2 + \left[ \alpha_1^2 + B^2 + 2B\alpha_1 \right] \sigma_b^2 > 0$$

The derivative $f'(.)$ is:

$$\frac{\partial f}{\partial \Sigma_h^2} = \left[ \alpha_1^2 + \left( \frac{A}{1-\nu \omega_3} \right)^2 + \frac{2\alpha_1 A}{1-\nu \omega_3} \right]$$

$$+ \left[ \left( \frac{A}{1-\nu \omega_3} \right)^2 + \frac{\alpha_1 A}{1-\nu \omega_3} \right] \Sigma_h^2 + \left[ \frac{B^2}{(1-\nu \omega_3)^2} + \frac{B\alpha_1}{1-\nu \omega_3} \right] \sigma_b^2 \frac{2\omega_3}{1-\nu \omega_3} \frac{\partial \nu}{\partial \Sigma_h^2}$$

with $\frac{\partial \nu}{\partial \Sigma_h^2} = \frac{CE}{(E + \Sigma_h^2 (E + \omega_3) C)^2}$

Hence

$$\lim_{\Sigma_h^2 \to \infty} \frac{\partial f}{\partial \Sigma_h^2} = \left[ \alpha_1^2 + \left( \frac{A}{1-\nu \omega_3} \right)^2 + \frac{2\alpha_1 A}{1-\nu \omega_3} \right]$$

$$= [\alpha_1 + \alpha_1 \alpha_2 (\omega_2 + \tau_m (1-\tau_m) \omega_1) + \alpha_2 (\omega_1 (1-\tau_u) (1-\tau_n) - \omega_3) (1-\tau_y) \lambda]^2$$

Therefore if $[\alpha_1 + \alpha_1 \alpha_2 (\omega_2 + \tau_m (1-\tau_m) \omega_1) + \alpha_2 (\omega_1 (1-\tau_u) (1-\tau_n) - \omega_3) (1-\tau_y) \lambda]^2 < 1$, the equation $\Sigma_h^2 = f(\Sigma_h)$ has at least one solution since $f$ is continuous and $\lim f(0) > 0$. Moreover, it has to be that an odd number of these solutions are characterized by $f'(\Sigma_h) < 1$, which guarantees local stability of the equilibrium path around these solutions.
Let’s now show that the equilibrium path is unique for $\omega_3$ small enough. A first order approximation of $f$ around $\omega_3 = 0$ writes:

\[
\begin{align*}
\frac{\partial f(\Sigma_h^2)}{\partial \Sigma_h^2} &\approx_{\omega_3 = 0} \left[ \alpha_1^2 + A^2 + 2\alpha_1 A \right] \Sigma_h^2 + \sigma_y^2 + \left[ \alpha_1^2 + B^2 + 2B\alpha_1 \right] \sigma_b^2 \\
&\quad + \left[ \left[ A^2 + \alpha_1 A \right] \Sigma_h^2 + \left[ B^2 + \alpha_1 B \right] \sigma_b^2 \right] 2\omega_3
\end{align*}
\]

with $\nu = \frac{C}{E \Sigma_h^{-2} + EC}$

which gives:

\[
\begin{align*}
\frac{\partial f(\Sigma_h^2)}{\partial \Sigma_h^2} &\approx_{\omega_3 = 0} \left[ \alpha_1^2 + A^2 + 2\alpha_1 A \right] \\
&\quad + \left[ \left[ A^2 + \alpha_1 A \right] \Sigma_h^2 + \left[ B^2 + \alpha_1 B \right] \sigma_b^2 \right] 2\omega_3 \frac{CE}{(E + ECS_h^2)^2} \\
&\quad + \left[ \left[ A^2 + \alpha_1 A \right] \Sigma_h^2 + \left[ B^2 + \alpha_1 B \right] \sigma_b^2 \right] \frac{E}{E + ECS_h^2} + \left[ A^2 + \alpha_1 A \right] \frac{C}{E + ECS_h^2} 2\omega_3
\end{align*}
\]

Since I have assumed that $\left[ \alpha_1^2 + A^2 + 2\alpha_1 A \right] < 1$, and $F(\Sigma_h^2)$ is bounded for $\Sigma_h^2 \in (0, \infty)$, there exists an $\omega_3$ small enough such that for all $\Sigma_h^2$, $\frac{\partial f(\Sigma_h^2)}{\partial \Sigma_h^2} < 1$. This is sufficient for the existence and uniqueness of a globally stable steady-state.

**A.8 Rise in returns to human capital**

The steady-state IGE is:

\[
\alpha_h = \alpha_1 + \alpha_3 + \alpha_1 \alpha_2 (\varepsilon_2 + \tau_m \varepsilon_1) + \alpha_2 (\varepsilon_1 (1 - \tau_n) - \varepsilon_3) (1 - \tau_y) \lambda
\]

So a small increase in $\lambda$ has the following long-run effect on the IGE:

\[
\frac{\partial \alpha_h}{\partial \lambda} = \frac{\partial \nu}{\partial \lambda} \left[ \alpha_1 \alpha_2 \left( \frac{\partial \varepsilon_2}{\partial \nu} + \tau_m \frac{\partial \varepsilon_1}{\partial \nu} \right) + \alpha_2 \left( \frac{\partial \varepsilon_1}{\partial \nu} (1 - \tau_n) - \frac{\partial \varepsilon_3}{\partial \nu} \right) (1 - \tau_y) \lambda \right] \\
\quad + \alpha_2 (\varepsilon_1 (1 - \tau_n) - \varepsilon_3) (1 - \tau_y)
\]

\[
\begin{align*}
= &\left[ \frac{\partial \nu}{\partial \lambda} \right] \left[ \frac{\partial \Sigma_h^2}{\partial \lambda} \right] \left[ \alpha_1 \alpha_2 \left( \frac{\partial \varepsilon_2}{\partial \nu} + \tau_m \frac{\partial \varepsilon_1}{\partial \nu} \right) + \alpha_2 \left( \frac{\partial \varepsilon_1}{\partial \nu} (1 - \tau_n) - \frac{\partial \varepsilon_3}{\partial \nu} \right) (1 - \tau_y) \lambda \right] \\
&\quad + \alpha_2 (\varepsilon_1 (1 - \tau_n) - \varepsilon_3) (1 - \tau_y)
\end{align*}
\]
where \( \frac{\partial \nu}{\partial \lambda} \) denotes the total derivative of \( \nu \) w.r.t. \( \lambda \). I can then compute the different derivatives:

\[
\begin{align*}
\frac{\partial \varepsilon_l}{\partial \nu} &= \varepsilon_l \varepsilon_3 > 0 \\
\frac{\partial \nu}{\partial \Sigma_h^2} &= \frac{CE}{(E + (E + \omega_3)C \Sigma_h^2)^2} > 0 \\
\text{with } C \text{ and } E \text{ have been defined in the proof of existence and uniqueness.}
\end{align*}
\]

\[
\frac{\partial \nu}{\partial \lambda} = 2C \left( \frac{\omega_A}{\omega_A + \omega_l} \right) \frac{1}{\omega_l} \left[ E \Sigma_h^{-2} + \omega_3 C \right] + C \left( \frac{\omega_A}{1 - \tau_y} \right)^2 > 0
\]

\[
\frac{\partial X_2}{\partial \lambda} = \sigma_1^2 \alpha_1 (1 + \alpha_2 (\varepsilon_2 + \tau_m \varepsilon_1)) \alpha_1 \alpha_2 \varepsilon_3 (\varepsilon_2 + \tau_m \varepsilon_1) \varepsilon_1 \frac{\partial \nu}{\partial \lambda} > 0
\]

\[
\frac{\partial \Sigma_h^2}{\partial \lambda} = -\frac{\partial X_2}{\partial \lambda} + \Sigma_h^2 \frac{\partial \alpha_h}{\partial \Sigma_h^2} \alpha_h - \frac{\partial X_2}{\partial \Sigma_h^2} > 0
\]

where \( \frac{\partial \alpha_h}{\partial \lambda} \) has to be understood as the partial derivative of \( \alpha_h \) w.r.t. \( \lambda \) keeping \( \Sigma_h^2 \) constant. The last line stems from the fact that the steady-state is locally stable - which requires that

\[
1 - (\alpha_h)^2 - \Sigma_h^2 \frac{\partial \alpha_h}{\partial \Sigma_h^2} \alpha_h - \frac{\partial X_2}{\partial \Sigma_h^2} = \frac{\partial (\Sigma_h^2)^2}{\partial \Sigma_h^2} > 0.
\]

Hence, putting everything together yields:

\[
\frac{\partial \alpha_h}{\partial \lambda} = \left[ \frac{\partial \nu}{\partial \lambda} + \frac{\partial \nu}{\partial \Sigma_h^2} \frac{\partial \Sigma_h^2}{\partial \lambda} \right] \varepsilon_3 [\alpha_1 \alpha_2 (\varepsilon_2 + \tau_m \varepsilon_1) + \alpha_2 (\varepsilon_1 (1 - \tau_n) - \varepsilon_3) (1 - \tau_y) \lambda]
\]

\[
+ \alpha_2 (\varepsilon_1 (1 - \tau_n) - \varepsilon_3) (1 - \tau_y) > 0
\]

This proves not only that the steady-state IGE is increasing in \( \lambda \) but that the variance of human capital in the economy is as well. Given that the variance of market income is given by \( \lambda^2 \Sigma_h^2 \) it is immediate that it increases too. Turning to the private spending on higher education, given by \( s \), it is immediate to see from the expressions (22) and (24) that it is increasing in the future path of \( \lambda \).

Let’s now turn to the ratio of within college variance of \( (\log) \) parental income over economy-wide
variance of (log) income:

\[
\frac{\nabla \text{Var}(\log y \mid q)}{\nabla \text{Var}(\log y)} = 1 \lambda^2 \sum_h \frac{\varepsilon_A^2}{(\varepsilon_I + \varepsilon_A)^2} \sigma_b^2 \left( \varepsilon_A^2 + \varepsilon_I^2 + 2 \varepsilon_A \varepsilon_I \right) \sigma_h^2 \left( \varepsilon_A^2 + \varepsilon_I^2 + 2 \varepsilon_A \varepsilon_I \right) \sigma^2_h \nu_h + \left( \varepsilon_A^2 + \varepsilon_I^2 + 2 \varepsilon_A \varepsilon_I \right) \sigma^2_h \nu_h + \left( \varepsilon_A^2 + \varepsilon_I^2 + 2 \varepsilon_A \varepsilon_I \right) \sigma^2_h \nu_h
\]

\[
\Rightarrow \frac{\partial \text{Var}(\log y \mid q)}{\partial \lambda} = \sigma_b^2 \frac{\partial B}{\partial \lambda} \left[ \sum_h^2 + \left( \frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A} \right)^2 \sigma_b^2 \right] - B \sigma_b^2 \left[ \frac{\partial \sigma_h^2}{\partial \lambda} + \sigma_b^2 \frac{\partial B}{\partial \lambda} \right]
\]

\[
= \sigma_b^2 \frac{\partial B}{\partial \lambda} \sum_h^2 - B \sigma_b^2 \left( \frac{\partial \sigma_h^2}{\partial \lambda} + \sigma_b^2 \frac{\partial B}{\partial \lambda} \right) < 0
\]

with \( B = \left( \frac{\varepsilon_A}{\varepsilon_I + \varepsilon_A} \right)^2 \) and since \( \frac{\partial \sigma_h^2}{\partial \lambda} > 0 \) and \( \frac{\partial B}{\partial \lambda} < 0 \).

Finally the variance of (log) college quality is given by \( \varepsilon_A^2 \sigma_b^2 + (\varepsilon_I + \varepsilon_A)^2 \Sigma_h^2 \). It is immediate that it increases with \( \lambda \) since \( \varepsilon_I, \varepsilon_A, \Sigma_h^2 \) increase with \( \lambda \).

**Monotonic transition path**  We now show that the transition is monotonic in the sense that the intergenerational elasticity, the variance of log income, the saving rate and the dispersion of quality are increasing over the transition path and the ratio of variances of log income within and in the overall economy is decreasing. From the law of motion of \( \Sigma_h^2 \), in the first period the initial increase in \( \lambda \) raises \( \alpha_h \) and triggers the initial increase in the dispersion of human capital. Since \( X_2(\Sigma_h) \) and \( \alpha_h(\Sigma_h) \) are both increasing in \( \Sigma_h \) it further increases \( \Sigma_h^2 \) at the following period and so on... This establishes that \( \Sigma_h^2 \) is strictly increasing over the transition path. This also establishes the monotonic increase in \( \alpha_h \) and all \( \omega \)'s.

Turning to the private spending on higher education, given by \( s \), it is immediate to see from its expression (??) that it is increasing in the future path of \( \lambda, \alpha_h \) and \( \varepsilon_1 \). Since these three variables are increasing over the transition path, \( s \) also increases.

The variance of log college quality is also increasing because \( \varepsilon_I, \varepsilon_A, \Sigma_h^2 \) are increasing over the transition path.

The ratio of within college variance of (log) parental income over economy-wide variance of (log) income will decrease monotonically over the transition path because of the initial increase in \( \lambda \), this is the first term in the derivative \( \sigma_b^2 \frac{\partial B}{\partial \lambda} \Sigma_h^2 \), and then decreases further as \( \Sigma_h \) increases, this is the second term \( B \sigma_b^2 \frac{\partial \Sigma_h^2}{\partial \lambda} \).
B Appendix for Efficiency and Welfare Analysis

B.1 Social Optimum

In this section, I derive the optimal allocation of a social planner seeking to maximize a social welfare function and who has access to a full set of instruments to implement it. The difficult part of the derivation lies in the static allocation of students and total amount of educational resources across colleges, because of the peer-effect, which introduces an externality. The consumption and labor decision is otherwise standard.

I decompose the analysis of this static allocation in two steps. First, I characterize, given an allocation of students across colleges, the optimal allocation of real resources across these clubs. Second, I show that the optimal allocation of students should feature perfect assortative matching, under some restrictions on the Pareto weights.

For simplicity in this section I will assume that the technology to produce the consumption good and the educational services is the same, i.e. \( \lambda = \gamma \) and \( A_I = A \).

B.1.1 Optimal Consumption, Labor and Investment given Sorting of Students

A student type is a pair \((h_s, y)\) where \(h_s\) and \(y\) are respectively ability and parental income. The optimal sorting of student types across colleges will follow a perfect positive assortative matching: colleges will be completely homogeneous ability-wise (also parental-income-wise). Given this allocation of students, I solve for the optimal spending.

Assuming that each college is homogeneous ability-wise (and parental income-wise) implies that the quality production function at college \( j \) can be rewritten: \( q(i) = I(i)^{\omega_1} h_s(i)^{\omega_2} \). Starting with a distribution of birth shock for children, labor market shocks and post-college human capital for parents \( \{\xi_y(i), \xi_b(i), h_e(i)\}_i \) and a set of generation-invariant Pareto weights \( \omega(i) \), the problem of the social planner is:

\[
\begin{align*}
\ln U^{SP} \left( \{\xi_y(i), \xi_b(i), h_e(i)\}_i \right) &= \max_{c(i), \ell(i), I(i)} (1 - \beta) \int \omega(i) \left( \ln c(i) - \ell(i) \right) di \\
&\quad + \beta E \ln U^{SP} \left( \{\xi_y'(i), \xi_b'(i), h_e'(i)\}_i \right) \\
A \int h^\lambda(i) \ell^\mu(i) di &= \int c(i) + \int I(i) di \quad (\lambda_1) \\
\forall i \quad h_e'(i) &= \kappa(h(i)\xi_b(i))^\alpha q(i)^{\alpha_2} \quad (\lambda_2(i)) \\
\forall i \quad h(i) &= \xi_y h_e(i) \\
\forall i \quad q(i) &= (h(i)\xi_b(i))^{\alpha_1}\omega_2(I(i))^{\omega_1}.
\end{align*}
\]

The social planner has to choose the optimal allocation of a given amount of resources

\[
\int I(i) di
\]
across all colleges, the optimal of consumption and labor assuming that all colleges are homogeneous in terms of ability so that the peer effect is equal to the common ability of the students and there is no loss of quality from within-college heterogeneity. Since the composition of the college doesn’t matter anymore, one can index the quality by the household index $i$.

The F.O.C. and the Envelope theorem associated with this problem write

$$c(i) \quad \omega(i) \frac{1 - \beta}{c(i)} = \lambda_1$$

$$\ell(i) \quad A h(i)^{\lambda} \lambda_1 \mu \ell(i)^{\mu - 1} = (1 - \beta) \omega(i) \eta \ell(i)^{\eta - 1}$$

$$I(i) \quad \lambda_1 = \lambda_2(i) \frac{h_e'(i)}{I(i)}$$

$$h_e'(i) \quad \beta E \frac{U h_e'}{U} = \lambda_2(i)$$

$$h_e(i) \quad \frac{U h_e}{U} = \lambda_1 A \xi_g(i)^{\lambda} h_e(i)^{\lambda - 1} \lambda \ell(i)^{\mu} + \lambda_2(i) \alpha_1 (1 + \alpha_2 \omega_2) \frac{h_e'}{h_e}$$

**Lemma 6.** Denote $C = \int c(i)di$ aggregate consumption for the current generation. In any social planner’s allocation, the optimal allocation of consumption and labor follows:

$$c(i) = \omega(i) C$$

$$\ell(i) = \left( \frac{\mu A h(i)^{\lambda}}{\eta \omega(i) C} \right)^{\frac{1}{\eta - \mu}}$$

**Proof.** Using the first F.O.C., it is easy to see that the planner chooses consumption for individual $i$ in the current generation as a time-invariant share of aggregate consumption

$$c(i) = \frac{\omega(i)}{\omega(1)} c(1)$$

$$\Rightarrow C = \frac{1}{\omega(1)} c(1)$$

$$c(i) = \omega(i) C$$

where $\omega(1)$ is the weight associated with an arbitrary dynasty with index '1'. Notice that this from the first F.O.C. implies $\lambda_1 = \frac{1 - \beta}{\xi}$ and $\frac{(1 - \beta) \omega(i)}{\lambda_1} = \omega(i) C$. From this and the the second F.O.C., one gets that the labor supply is an increasing function of human capital

$$\ell(i) = \left( \frac{\mu A h(i)^{\lambda}}{\eta \omega(i) C} \right)^{\frac{1}{\eta - \mu}}$$

**Lemma 7.** In the social planner’s allocation, if all colleges are perfectly homogeneous ability and parental-income-wise, $I(i)$ is solely a function of $h_s(i)$, the dynasty-specific Pareto weight $\omega(i)$ and
the path of aggregates.

Proof. Combining the conditions for $I$, $h'_e$ and the E.T. for $h_e$ one gets

$$I(i) = \beta \frac{C}{C'} \left[ \chi_1 \chi_2(i) h_s(i) \left( \frac{\mu}{\eta} \frac{1 + \alpha_2 \omega_2}{\eta^{\alpha_2}} I(i) \right) \lambda + \alpha_1 (1 + \alpha_2 \omega_2) E (I'(i)|h'_e(i)) \right]$$

(55)

where the expectation on the RHS of (55) is conditional on $h'_e(i) = \kappa h_s(i)$.

Notice first that there is no closed-form solution for this equation in general. Assume it admits a solution for any $\omega(i)$, $h_s(i)$. (55) has to be solved forward. Because the conditional expectation of $I'(i)$ on the RHS of (55) depends only on $h_s(i)$ and $I(i)$, the entire RHS of equation (55) depends only on $h_s(i)$ and $I(i)$ and the path of aggregate consumption and technology as well as the dynastic-specific Pareto weight $\omega(i)$. Hence the lemma.

\[ \square \]

B.1.2 Sorting of Students

In this section, I show that it is optimal for the planner to perfectly sort students according to their abilities and social origin and therefore build perfectly homogeneous colleges. I proceed in three steps:

1. I show first that maximizing future output is the right objective when choosing the optimal sorting of students within colleges

2. I show that any optimal allocation should feature stratified colleges by abilities - i.e. that in any college there would be only two consecutive types.\(^{39}\)

3. I then show that when looking at the optimal allocation with only two types necessarily features perfect segregation across two homogeneous colleges. This part of the proof is actually the most intuitive and most insightful. It resembles many of the proofs in optimal matching that using the property of supermodality.

1. Define aggregate output:

$$Y = \int Ah^\lambda(i)\ell(i)^\mu di$$

Define $s = 1 - C'$ the share of output devoted to human capital accumulation. Given lemma

\(^{39}\)The proof here is done for a discrete distribution of abilities. I believe the proof goes through in a setting with a continuum of types. I leave it for future research.
(6), the optimal labor can be rewritten
\[ \ell(i) = \left( \frac{\mu A h(i)^\lambda}{\eta \omega(i) C} \right)^\frac{1}{\eta - \mu} \]

\[ \Rightarrow \ell(i)^\eta = \left( \frac{A h(i)^\lambda}{\eta \omega(i) C} \right) \ell(i)^\mu \]

and the flow of social welfare for the current generation is
\[
\int \omega(i) \left[ \ln(c(i)) - \ell(i)^\eta \right] \, di = \int \omega(i) \left[ \ln(C) + \ln(\omega(i)) - \left( \frac{A h(i)^\lambda}{\eta \omega(i) C} \right) \ell(i)^\mu \right] \, di \\
= \ln(C) - \frac{Y}{\eta C} + \int \omega(i) \ln(\omega(i)) \, di \\
= \ln((1 - s)Y) - \frac{1}{\eta(1 - s)} + \int \omega(i) \ln(\omega(i)) \, di \\
= \ln(1 - s) + \ln(Y) - \frac{1}{\eta(1 - s)} + \int \omega(i) \ln(\omega(i)) \, di
\]

The sorting of students across colleges doesn’t change the current flow of social welfare, but changes output at the following period:
\[
\int \omega(i) \left[ \ln(c(i)) - \ell(i')^{\eta} \right] \, di = \ln(1 - s') + \ln(Y') - \frac{1}{\eta(1 - s')} + \int \omega(i) \ln(\omega(i)) \, di
\]

The best the social planner can do when allocating individuals across colleges is to maximize future output \( Y' \).

2. Let’s now choose the sorting of students that maximizes future output.

\[
Y' = \int A'h^{\tilde{\lambda}}(i) \ell'^\mu(i) \, di \\
= \int A'h^{\tilde{\lambda}}(i) \left( \frac{\mu A' h(i)^{\tilde{\lambda}}}{\eta \omega(i) C'} \right)^{\frac{\mu}{\eta - \mu}} \, di \\
= \left( \frac{\mu}{\eta C'} \right)^{\frac{\mu}{\eta - \mu}} \cdot \frac{A'}{\eta^{\frac{\mu}{\eta - \mu}}} \int \frac{h^{\tilde{\lambda}}(i)}{\omega(i)^{\frac{\mu}{\eta - \mu}}} \, di
\]

with \( \tilde{\lambda} = \frac{\lambda \eta}{\eta - \mu} \).

**Assumption** At this point, one needs to make one of two assumptions.

(a) Either one assumes \( \omega(i) = 1 \) for all \( i \)

(b) Or one assumes \( \eta \to +\infty \), i.e. that labor is inelastic.

Both assumption makes optimal labor to be only a function of aggregates and the individual’s human capital. Otherwise, labor depends also on the weight put by the social planner on the
dynasty and this would incentivize the social planner to make a dynasty with lower weight to get to a better college education because it will be working more later on. For the sake of expositional clarity, I prefer to shut down this mechanism. For the rest of this section, I assume \( \eta \to +\infty \). This implies \( \ell = 1 \) for all individuals and

\[
Y' = A' \int h^\lambda(i)di
\]

Armed with this assumption, dropping \( A' \) and given that the labor market shock is i.i.d. across households, the Lagrangian associated with the maximization problem is:

\[
\max_{\mu_i} \int f_j^{\alpha_1 \omega_1} \theta_j^{\alpha_2 \omega_2} h_s(l)^{\alpha_1 \lambda} \mu_i^j - \int_{\ell} \lambda_j \left[ \ln(h_s(i)) - \int_{\ell} \frac{\mu_j^j}{\mu_i^j} \ln(h_s(i)) \right] dj + \int_j \gamma_i \left[ \bar{\mu}_i - \int_{\ell} \mu_i^j dj \right] di + \int_j \int_{\ell} \Sigma_j^j \mu_i^j dj di
\]

where \( \mu_i^j \) denotes the mass of students of type \( i \) in college \( j \), \( \bar{\mu}_i \) is the total mass of students of type \( i \) and \( \lambda_j, \gamma_i, \Sigma_j^j \) are the lagrange multipliers associated respectively to the definition of the peer-effect, the resource constraint of students of type \( i \) and the non-negativity constraint of \( \mu \)'s.

The F.O.Cs w.r.t. \( \mu_i^j \) and \( \theta_j \) write:

\[
\int f_j^{\alpha_1 \omega_1} \theta_j^{\alpha_2 \omega_2} h_s(l)^{\alpha_1 \lambda} + \lambda_j \left[ \ln(h_s(i)) \left[ \frac{f_i^{\mu_i^j} di - \mu_i^j}{\left( \int_{\ell} \mu_i^j di \right)^2} \right] - \int_{\ell} \frac{\mu_j^j}{\left( \int_{\ell} \mu_i^j di \right)^2} \ln(h_s(i)) \right]
\]

\[
\gamma_l - \Sigma_i^l
\]

\[
\lambda_j = \alpha_2 \lambda_2 I_j^{\alpha_1 \omega_1} \theta_j^{\alpha_2 \omega_2} \int_{\ell} \mu_i^j h_s(i)^{\alpha_1 \lambda} di
\]

Combining the first F.O.C for two different agents \( l, l' \) and two different colleges \( j, j' \) gives:

\[
\left[ h_s(l)^{\alpha_1 \lambda} - h_s(l')^{\alpha_1 \lambda} \right] \left[ f_j^{\alpha_1 \omega_1} \theta_j^{\alpha_2 \omega_2} - f_{j'}^{\alpha_1 \omega_1} \theta_{j'}^{\alpha_2 \omega_2} \right] + \ln \left( \frac{h_s(l)}{h_s(l')} \right) \left[ \frac{\lambda_j}{\mu_j} - \frac{\lambda_{j'}}{\mu_{j'}} \right]
\]

\[
\Sigma_j^l + \Sigma_{l'}^j - \Sigma_{j'}^l - \Sigma_i^l
\]  

(57)

where \( \mu_j \) denotes the total mass of students at college \( j \).

Given the F.O.C. for \( \lambda_j \), the last parenthesis rewrites:

\[\text{Setting all weights to 1 would amount to selecting one particular point on the Pareto frontier. This is an undesirable assumption since the goal of this section is to derive general characteristics of the allocations of students and resources at any point on the Pareto frontier.}\]
\[
\left[ \frac{\lambda_j}{\mu_j} - \frac{\lambda'_{j'}}{\mu'_{j'}} \right] = \alpha_2 \lambda \omega_2 \left[ I^{\alpha_2 \lambda \omega_1} \theta_j^{\alpha_2 \lambda \omega_2} \int s_i^j h_s(i)^{\alpha_1 \lambda} di - I^{\alpha_2 \lambda \omega_1} \theta_j^{\alpha_2 \lambda \omega_2} \int s_i^j' h_s(i)^{\alpha_1 \lambda} di \right]
\]

where \( s_i^j \) is the share of students of type \( i \) out of the total student population of college \( j \).

Assume the \( I^{\alpha_2 \lambda \omega_1} \theta_j^{\alpha_2 \lambda \omega_2} \int s_i^j h_s(i)^{\alpha_1 \lambda} di \) are ordered in the same way as \( I^{\alpha_2 \lambda \omega_1} \theta_j^{\alpha_2 \lambda \omega_2} \int s_i^j' h_s(i)^{\alpha_1 \lambda} di \). Then from (57) it is impossible that all \( \Sigma_i^j \) be all zero at the same time. Without loss of generality assume \( j \) provides a higher quality than \( j' \) and that \( l \) is of a higher type than \( l' \), then it has to be that \( \Sigma_i^j + \Sigma_i^{j'} > 0 \), i.e. that either there is no low type type \( l' \) at high quality college \( j \) or no high quality type \( l \) at low quality college \( j' \).

Moreover it is immediate to see that if two colleges have in common two different student types, then they must be of the same quality. Together with the previous property implies that any optimal allocation has to be totally stratified, i.e. there is only two consecutive student types in any college.

3. I now show that looking at the subsample of two consecutive types that potentially go to the same college, it is always better to separate them in two different homogeneous colleges than to keep them together. Consider a now simple framework with only two types and two colleges - I know they can’t be more with two types according to the previous statements. Denote \( \bar{\mu}_1, \bar{\mu}_2 \) the total mass of type 1 (2) in that subsample. Keeping the same notations as before, 'local' output is:

\[
\tilde{y} = \mu_1^1 h_{s,1} I_1^{\alpha_1 \lambda \omega_1} \theta_1^{\alpha_1 \lambda \omega_2} + \mu_2^1 h_{s,2} I_2^{\alpha_1 \lambda \omega_1} \theta_2^{\alpha_1 \lambda \omega_2} + (\bar{\mu}_1 - \mu_1^1) h_{s,1} I_2^{\alpha_1 \lambda \omega_1} \theta_1^{\alpha_1 \lambda \omega_2} + (\bar{\mu}_2 - \mu_2^1) h_{s,2} I_1^{\alpha_1 \lambda \omega_1} \theta_2^{\alpha_1 \lambda \omega_2}
\]

Our goal is to establish that \( \frac{\partial^2 \tilde{y}}{\partial \mu_1^1 \partial \mu_2^1} < 0 \) which would establish the optimality of perfect assortative matching within that subsample. One can simply do the computations:
\[
\frac{\partial \tilde{y}}{\partial \mu_1} = h_{s,1}^{\alpha_1} \left[ I_1^{\alpha_1} \theta_1^{\alpha_1} - I_2^{\alpha_1} \theta_2^{\alpha_1} \right] \\
+ \frac{\partial \theta_1^{\alpha_1}}{\partial \mu_1} \left[ \mu_1 h_{s,1}^{\alpha_1} + \mu_2 h_{s,2}^{\alpha_1} \right] \\
+ \frac{\partial \theta_2^{\alpha_1}}{\partial \mu_1} \left[ (\bar{\mu}_1 - \mu_1^1) h_{s,1}^{\alpha_1} + (\bar{\mu}_2 - \mu_2^1) h_{s,2}^{\alpha_1} \right]
\]

\[
\frac{\partial \theta_1^{\alpha_1}}{\partial \mu_1} = \frac{\alpha_2 \lambda \omega_2}{\mu_1} s_1 I_1^{\alpha_1} \theta_1^{\alpha_1} \omega_2 \ln \left( \frac{h_{s,1}}{h_{s,2}} \right) \theta_1^{\alpha_1} \\
\frac{\partial \theta_2^{\alpha_1}}{\partial \mu_1} = -\frac{\alpha_2 \lambda \omega_2}{\mu_2} s_2 I_2^{\alpha_1} \theta_2^{\alpha_1} \omega_2 \ln \left( \frac{h_{s,1}}{h_{s,2}} \right) \theta_2^{\alpha_1}
\]

\[
\frac{\partial^2 \tilde{y}}{\partial \mu_1 \partial \mu_2} = -\alpha_2 \lambda \omega_2 \ln \left( \frac{h_{s,1}}{h_{s,2}} \right) \left[ I_1^{\alpha_1} \theta_1^{\alpha_1} - I_2^{\alpha_1} \theta_2^{\alpha_1} \right] \\
+ \alpha_2 \frac{\lambda}{\mu_1} \omega_2 \ln \left( \frac{h_{s,1}}{h_{s,2}} \right) I_1^{\alpha_1} \theta_1^{\alpha_1} \omega_2 s_1 \frac{1}{s_1} h_{s,1}^{\alpha_1} - s_2 h_{s,2}^{\alpha_1} \\
+ \alpha_2 \frac{\lambda}{\mu_2} \omega_2 \ln \left( \frac{h_{s,1}}{h_{s,2}} \right) I_2^{\alpha_1} \theta_2^{\alpha_1} \omega_2 s_2 \frac{1}{s_2} h_{s,1}^{\alpha_1} - s_1 h_{s,2}^{\alpha_1}
\]

This shows that supermodularity holds locally. Hence 'local' output is maximized when students are perfectly match in homogeneous colleges. This ends the proof.
B.2 Closed-Form Solutions for Planner’s Problem

There are two special cases in which this equation offers a closed-form solution for $I(i)$. We maintain the assumption that $\eta \to +\infty$ so that it is optimal to perfectly sort students across homogeneous colleges ability-wise.

B.2.1 Special case (1): two periods problem

The first case is the two-periods model in which case: $I'(i) = 0$ and

$$I(i) = \frac{C}{\gamma} \chi_1 \lambda h_s(i)^{\frac{1 + \alpha_2 \omega_2}{1 - \eta - \mu}}.$$

In this very simple case, educational investment is increasing in abilities if and only if $\lambda \alpha_2 \omega_2 < 1$ which should be true in any sensible parametrization.

B.2.2 Special case (2): $\frac{1+\alpha_2 \omega_2}{1 - \eta - \mu \alpha_2 \omega_2} = \frac{1}{\alpha_1}$

The second case is the edge-case where $\alpha_1 = \frac{\lambda \eta}{\eta - \mu}$ in which case I guess that $I(i) = I_0 h_s(i)^{\varepsilon \eta h_s}$ for time-dependent constant $I_0$ independent of $i$ and applying the guess:

$$I_0 h_s(i)^{\varepsilon \eta h_s} = \frac{C}{\gamma} \left[ \chi_1 t \lambda h_s^{\frac{\lambda \eta}{\eta - \mu}} \right] (I_0 h_s(i)^{\varepsilon \eta h_s}) + \alpha_1 (1 + \alpha_2 \omega_2) I_0 E(h_s(i)^{\varepsilon \eta h_s})$$

Define $\varepsilon = \frac{\lambda \eta}{\eta - \mu} \frac{1 + \alpha_2 \omega_2}{1 - \eta - \mu \alpha_2 \omega_2}$, the terms in $h_s$ cancel out if and only if

$$\varepsilon = \varepsilon \eta h_s = \varepsilon \frac{1}{\alpha_1 \eta - \mu} = \varepsilon$$

When this holds, one obtains the following non-linear recursive equation in $I_0$

$$I_0^{1 - \frac{\lambda \eta}{\eta - \mu \alpha_2 \omega_2}} = \frac{C}{\gamma} \left[ \chi_1 t + \alpha_1 (1 + \alpha_2 \omega_2) I_0 h_s^\varepsilon \right] E((h_s(i)^{\alpha_1 \varepsilon \eta h_s})$$

Instead of solving for the path of $I_0$ I choose the aggregate amount of educational expenditures
defined by:
\[ \bar{I} = \int I(i) di \]

which implies
\[ \bar{I} = I_0 \int h_s(i)^{\bar{I}h_s} di \]
\[ \Rightarrow I_0 = \frac{sY}{\int h_s(i)^{\bar{I}h_s} di} \]

where I define \( s \) the aggregate rate of investment into higher education \( \bar{I} = sY \). The quality an individual of ability \( h_s \) is going to get is therefore
\[ q = Qh_s^{\bar{I}h_s} \]
with
\[ Q = \left( \frac{sY}{\int h_s(i)^{\bar{I}h_s} di} \right)^{\omega_1} \]

The law of accumulation of human capital is:
\[ h' = \kappa \xi_y h_s q^{\alpha_2} \]
\[ = \kappa \xi_y h_s \left( \frac{sY}{\int h_s(i)^{\bar{I}h_s} di} \right)^{\omega_1 \alpha_2} h_s^{\bar{I}h_s} \omega_1 \alpha_2 \]
\[ h' = \kappa \xi_y h_s^{1+\bar{I}h_s} \omega_1 \alpha_2 + \omega_2 \alpha_2 \]

Therefore if the economy starts with a log-normal distribution for human capital, it will stay so all along the equilibrium path. Denote \((m, \Sigma_h^2)\) the mean and variance of the normal distribution of the log of human capital. The law of motion of human capital rewrites:
\[ \ln h' = \ln \kappa + \ln \xi_y + \alpha_p (\ln \xi_b + \ln h) + \omega_1 \alpha_2 \left( \ln \bar{I} - \alpha_1 \bar{I}h_s \left( m_h - \frac{\sigma_b^2}{2} \right) - \frac{(\alpha_1 \bar{I}h_s)^2}{2} \left( \Sigma_h^2 + \sigma_b^2 \right) \right) \]
\[ \alpha_p = \alpha_1 + \alpha_1 \bar{I}h_s \omega_1 \alpha_2 + \omega_2 \alpha_2 \omega_1 \]
\[ \epsilon_{Ih_s} = \frac{\lambda \eta - \alpha_2 \epsilon_2}{1 - \lambda \eta - \alpha_2 \omega_1} \]
\[ m'_h = \ln \kappa - \frac{\sigma_b^2}{2} + \alpha_p \left( -\frac{\sigma_b^2}{2} + m_h \right) + \omega_1 \alpha_2 \left( \ln \bar{I} - \alpha_1 \bar{I}h_s \left( m_h - \frac{\sigma_b^2}{2} \right) - \frac{(\alpha_1 \bar{I}h_s)^2}{2} \left( \Sigma_h^2 + \sigma_b^2 \right) \right) \]
\[ \Sigma_h' = \sigma_b^2 + (\alpha_p)^2 (\sigma_b^2 + \Sigma_h^2) \]

This combined with the property that consumption for all individual is their Pareto weight times the aggregate consumption and the expression for labor supply, gives:
\[
Y = A[\eta(1-s)]^{-\frac{1}{\eta}} \left( \int h(i)^{\lambda \eta^{-1}} di \right)^{\frac{\eta-1}{\eta}}
\]

From the F.O.C for labor I also have:
\[
\ell^\eta = \frac{Ah^\lambda}{\eta C} \Rightarrow \ell^\eta = \int \frac{Ah^\lambda}{\eta C} = \frac{Y}{\eta(1-s)Y} = \frac{1}{\eta(1-s)}
\]

Guessing that the relevant state variables are the mean and standard deviation of the distribution of (the logarithm of) human capital in the economy, the problem rewrites:
\[
\ln \mathcal{U}(m_h, \Sigma_h) = \max_{s, l} (1-\beta) \left[ \ln A(1-s)^{1-\frac{1}{\eta}} \left( \int h(i)^{\lambda \eta^{-1}} di \right)^{\frac{\eta-1}{\eta}} - \frac{1}{\eta(1-s)} \right] + \beta \ln \mathcal{U}(m_h', \Sigma_h')
\]

\[
m_h' = \ln \kappa - \frac{\sigma_y^2}{2} + \alpha_p \left( -\frac{\sigma_b^2}{2} + m_h \right)
\]

\[
+ \omega_1 \alpha_2 \left( \ln \left( sA(1-s)^{-\frac{1}{\eta}} \left( \int h(i)^{\lambda \eta^{-1}} di \right)^{\frac{\eta-1}{\eta}} \right) \right) - \alpha_1 \varepsilon_{Ih_s} \left( m_h - \frac{\sigma_b^2}{2} \right) - \frac{(\alpha_1 \varepsilon_{Ih_s})^2}{2} \left( \Sigma_h^2 + \sigma_b^2 \right)
\]

\[
\Sigma_h' = \sigma_y^2 + \alpha_p^2 \left( \sigma_b^2 + \Sigma_h^2 \right)
\]

I guess that the value function is linear in \((m_h, \Sigma_h^2)\):
\[
\ln \mathcal{U}(m_h, \Sigma_h^2) = \gamma_{0,t} + \gamma_{1,t} m_{h,t} + \gamma_{2,t} \Sigma_{h,t}
\]

Hence:
\[
\gamma_{0,t} + \gamma_{1,t} m_{h,t} + \gamma_{2,t} \Sigma_{h,t} = \max_{s, l} (1-\beta) \left[ \ln A(1-s)^{1-\frac{1}{\eta}} \left( \int h(i)^{\lambda \eta^{-1}} di \right)^{\frac{\eta-1}{\eta}} - \frac{1}{\eta(1-s)} \right]
\]

\[
+ \beta \left( \gamma_{0,t+1} \right)
\]

\[
+ \gamma_{1,t+1} \left( \ln \kappa - \frac{\sigma_y^2}{2} + \alpha_p \left( -\frac{\sigma_b^2}{2} + m_h \right) \right)
\]

\[
+ \omega_1 \alpha_2 \left( \ln \left( sA(1-s)^{-\frac{1}{\eta}} \left( \int h(i)^{\lambda \eta^{-1}} di \right)^{\frac{\eta-1}{\eta}} \right) \right) - \alpha_1 \varepsilon_{Ih_s} \left( m_h - \frac{\sigma_b^2}{2} \right) - \frac{(\alpha_1 \varepsilon_{Ih_s})^2}{2} \left( \Sigma_h^2 + \sigma_b^2 \right)
\]

\[
+ \gamma_{2,t+1} \left( \sigma_y^2 + \alpha_p^2 \left( \sigma_b^2 + \Sigma_h^2 \right) \right)
\]
Identifying all the terms in \( m_h, \Sigma^2_{h_t} \):

\[
\begin{align*}
\gamma^{p}_{1,t} &= (1 - \beta) \lambda + \beta \gamma^{p}_{1,t+1} (\alpha p + \omega_1 \alpha_2 (\lambda - \alpha_1 \varepsilon_{Ih})) \quad (58) \\
\gamma^{p}_{2,t} &= (1 - \beta) \frac{\lambda^2}{2} \frac{\eta}{\eta - 1} + \beta \gamma^{p}_{1,t+1} \omega_1 \alpha_2 \left( \frac{\lambda^2}{2} \frac{\eta}{\eta - 1} - \frac{\alpha_1 \varepsilon_{Ih}^2}{2} \right) + \beta \gamma^{p}_{2,t+1} \alpha_p^2 \\
\gamma^{p}_{0,t} &= (1 - \beta) \left[ \ln A(1 - s)^{\frac{1}{\eta}} - \frac{1}{\eta(1 - s)} \right] \\
&+ \beta \left( \gamma^{p}_{0,t+1} + \gamma^{p}_{1,t+1} \left( \ln \kappa - a^2 - \alpha p \sigma_x^2 + \omega_1 \alpha_2 \left( \ln s A(1 - s)^{-\frac{1}{\eta}} + \alpha_1 \varepsilon_{Ih} \sigma_y^2 - \frac{\alpha_1 \varepsilon_{Ih}^2}{2} \sigma_y^2 \right) \right) \\
&+ \gamma^{p}_{2,t+1} \left( \sigma^2_a + \alpha^2 p \sigma^2_y \right) \right) \quad (59)
\end{align*}
\]

The F.O.C. w.r.t. \( s \) is:

\[
\begin{align*}
\frac{1 - \beta}{1 - s} \frac{\eta - 1}{\eta} - \frac{(1 - \beta) \eta}{(\eta(1 - s))^2} + \beta \gamma^{p}_{1,t+1} \omega_1 \alpha_2 \left( \frac{1}{s} + \frac{1}{\eta(1 - s)} \right) &= 0 \\
-(1 - \beta) \frac{\eta - 1}{\eta} s(1 - s) - \frac{(1 - \beta)}{\eta} s + \beta \gamma^{p}_{1,t+1} \omega_1 \alpha_2 \left( (1 - s)^2 + \frac{s(1 - s)}{\eta} \right) &= 0 \\
\Gamma_1 s(1 - s) - \Gamma_2 s + \Gamma_3 \left( (1 - s)^2 + \frac{s(1 - s)}{\eta} \right) &= 0 \\
\frac{s^2}{\Gamma_1} + \frac{\eta - 1}{\eta} + s \left( \Gamma_1 + \Gamma_3 \frac{\eta - 1}{\eta} \right) + \frac{s}{\Gamma_1 - \Gamma_2 + \Gamma_3 \frac{1 - 2\eta}{\eta}} + \frac{\Gamma_3}{\eta} &= 0
\end{align*}
\]

After simplification one gets:

\[
\begin{align*}
A &= \frac{\eta - 1}{\eta} (1 - \beta + \Gamma_3) > 0 \\
B &= - \left( 1 - \beta + \Gamma_3 \left( 2 - \frac{1}{\eta} \right) \right) < 0 \\
s^+/- &= \frac{-B \pm \sqrt{B^2 - 4AC}}{2A} > 0 \quad (60)
\end{align*}
\]

Both root are positive because \( A > 0 \), its value at 0 is positive and \( \frac{-B}{2A} > 0 \). The lowest one, \( s^- \), is the right one. In particular,

\[
\lim_{\eta \to \infty} s^- = \frac{\beta \omega_1 \alpha_2 \gamma_{1,t+1}}{1 - \beta + \beta \omega_1 \alpha_2 \gamma_{1,t+1}} \quad \text{and} \quad \lim_{\eta \to \infty} s^+ = 1
\]

If all the parameters are constant over time, one can solve for \( \gamma^{p}_{0}, \gamma^{p}_{1}, \gamma^{p}_{2} \):
\[ \gamma_1^p = \frac{(1 - \beta) \lambda}{1 - \beta \left( \alpha^p + \omega_1 \alpha_2 (\lambda - \alpha_1 \varepsilon_{IH}) \right)} \]

\[ \gamma_2^p = \frac{(1 - \beta)^2 \frac{1}{\eta - 1} + \beta \gamma_1^p \omega_1 \alpha_2 \left( \frac{\lambda^2}{2} \frac{\eta}{\eta - 1} - \frac{(\alpha_1 \varepsilon_{IH})^2}{2} \right)}{1 - \beta \alpha_p^2} \]

In this case the saving rate is constant along the equilibrium path. Before summarizing the properties of the first best allocation, let's define \( \mathcal{L} \) as:

\[ Y = A \mathcal{L} \int h(i)^\lambda di \]

\[ \iff \mathcal{L} = \left[ \eta(1 - s) \right]^\frac{1}{\eta} \left( \int h(i)^\lambda \frac{\eta}{\eta - 1} di \right)^\frac{\eta - 1}{\eta} \int h(i)^\lambda di \]

\[ = \left[ \eta(1 - s) \right]^\frac{1}{\eta} e^\frac{\lambda^2 \Sigma_h^2}{\eta - 1} \]

and I will refer to it as 'the aggregate labor supply in the first best'. The equilibrium law of motion for the distribution of human capital in the economy is:

\[ \ln h' = \mathcal{N} \left( m_h', \Sigma_h^2 \right) \]

\[ m_h' = \left( \alpha^p + \omega_1 \alpha_2 (\lambda - \alpha_1 \varepsilon_{IH}) \right) m_h + \omega_1 \alpha_2 \left( \frac{\lambda^2}{2} \frac{\eta}{\eta - 1} - \frac{(\alpha_1 \varepsilon_{IH})^2}{2} \right) \Sigma_h^2 \]

\[ + \ln \kappa - \frac{\sigma_y^2}{2} - \alpha^p \alpha_h^2 + \omega_1 \alpha_2 \left( \ln sa[\eta(1 - s)] - \frac{1}{\eta} + \frac{\alpha_1 \varepsilon_{IH}}{2} (1 - \alpha_1 \varepsilon_{IH}) \sigma_h^2 \right) \]

\[ \Sigma_h^2 = \sigma_y^2 + (\alpha^p)^2 (\sigma_h^2 + \Sigma_h^2) \]

with \( \alpha^p + \omega_1 \alpha_2 (\lambda - \alpha_1 \varepsilon_{IH}) = \alpha_1 + \omega_1 \alpha_2 \lambda + \omega_2 \alpha_1 \omega_2 \) is the same coefficient as the decentralized equilibrium with \( \omega_3 = 0 \).

The following proposition summarizes the properties of the first best allocation:
Proposition B.1.

\[ \forall i \quad \frac{c(i)}{\omega(i)} = C = (1 - s)A_L(\Sigma_h)e^{\lambda m_h + \frac{1}{2}\nu^2_h} \]

\[ q = Q h^s \epsilon_{h_s}^{1+\omega_2} \]

\[ \ln h' = \mathcal{N} \left( \left( m_1', \Sigma_2' \right) \right) \]

\[ m_1' = (\alpha_1 + \omega_1 \alpha_2 \lambda + \alpha_2 \alpha_1 \omega_2)m_h + \omega_1 \alpha_2 \left( \frac{\lambda^2}{2} \right) \frac{\eta}{\eta - 1} - (\alpha_1 \epsilon_{h_s})^2 \right) \Sigma_2_h \]

\[ + \ln \kappa - \frac{\sigma_0^2}{2} - \alpha_p \frac{\sigma_y^2}{2} + \omega_1 \alpha_2 \left( \ln sA[\eta(1-s)]^{-\frac{1}{\eta}} + \frac{\alpha_1 \epsilon_{h_s}}{2} \right) \]

\[ \Sigma_2_h = \sigma_y^2 + (\alpha^p)^2(\sigma_b^2 + \Sigma^2_h) \]

with \( s \) a constant given by (60) and

\[ \gamma_{1,t} = (1 - \beta) \sum_{k=0}^{\infty} \beta^k \lambda_t \prod_{m=0}^{k-1} (\alpha_1 + \omega_1 \alpha_2 \lambda_{t+m} + \alpha_2 \alpha_1 \omega_2) \]

\[ \epsilon_{h_s} = \frac{\lambda \eta}{\eta - 1} \frac{1 + \alpha_2 \epsilon_2}{1 - \lambda \eta - \alpha_2 \omega_1} \]

\[ \alpha^p = \alpha_1 + \alpha_1 \alpha_2 \omega_2 + \omega_1 \alpha_2 \alpha_1 \epsilon_{h_s} \]

\[ Q = \left( \frac{sA_L(\Sigma_h)e^{\lambda m_h + \frac{1}{2}\nu^2_h}}{\epsilon_{h_s}^{1+\omega_2} \left( m_h - \frac{\sigma_0^2}{2} \right) + \frac{\alpha_1 \epsilon_{h_s}}{2} (\Sigma_2^2 + \sigma_0^2)} \right)^{\omega_1} \]

B.3 Competitive Equilibrium Without Financial Constraints

In order to investigate the sources of inefficiencies, let’s restore the completeness of financial markets and see whether is it sufficient to bring back the decentralized equilibrium on the Pareto-frontier. The model is exactly the same as before except that now households can transfer wealth across generations and states.

\[ \ln \mathcal{V}(h_e, \xi_y, \xi_b, a) = \max_{c, q, a} \left\{ (1 - \beta) \left[ \ln c - \ell^n \right] + \beta E \ln \left( \mathcal{V}(h_e', \xi_y', \xi_b', a'(h_e', \xi_y', \xi_b')) \right) \right\} \]

\[ y + a = c + \int a'(\xi_y', \xi_b')Q(\xi_y, \xi_b) + e(q, y, h_s) \quad \lambda_1 \]

\[ y = A \ell^m h^\lambda \]

\[ h_e' = \kappa(\xi_e) \alpha_1 q^{\alpha_2} \quad \lambda_2 \]

\[ h' = \xi_y' h_e' \]

\[ h_s = (\xi_b h)^{\alpha_1} \]
The F.O.C. and E.T. write:

\[
\frac{1 - \beta}{c} = \lambda_1 \tag{61}
\]

\[
(1 - \beta) \frac{\eta}{\mu} \ell^{\eta-\mu} = \lambda_1 Ah^\lambda (1 - e_y) \tag{62}
\]

\[
\alpha_2 h_e' \lambda_2 = \lambda_1 e_e e \tag{63}
\]

\[
\beta \frac{\mathcal{U}_{a'}}{\mathcal{U}} \pi(\xi_b', \xi_y') = \lambda_1 Q(\xi_b', \xi_y') \tag{64}
\]

\[
\beta E_{\xi_y, \xi_b} \left( \frac{\mathcal{U}_{h_e'}}{\mathcal{U}} \right) = \lambda_2 \tag{65}
\]

\[
\frac{\mathcal{U}_a}{\mathcal{U}} = \lambda_1 \tag{66}
\]

\[
\frac{\mathcal{U}_{h_e}}{\mathcal{U}} = \frac{\lambda_1}{h_e} \left( Ah^\lambda \frac{e}{e_y} \lambda y h_e - \varepsilon_{eq} e e_h \frac{h_s}{h_e} \alpha_1 h_s \right) + \frac{h_e'}{h_e} \alpha_1 \lambda_2 \tag{67}
\]

Guess \( e_y = 0 \). Using (61) and (62) one gets:

\[
\frac{\eta}{\mu} \ell^{\eta-\mu} = \frac{Ah^\lambda}{c}
\]

Using (61), (66) and (64) one gets the traditional Euler Equation for consumption:

\[
\beta \lambda_1' (\xi_b', \xi_y') \pi(\xi_b', \xi_y') = \lambda_1 Q(\xi_b', \xi_y')
\]

The ratio of marginal utilities \( \frac{\lambda_1'(\xi_b', \xi_y')}{\lambda_3} = \frac{c}{\mathcal{U}} \) is equalized across agents, which implies that it must also be equal to the ratio of aggregate consumptions: \( \frac{\mathcal{C}}{\mathcal{U}} \) and that each household receives a constant share of aggregate consumption \( c(i) = \omega(i) C \).

Combining (63) and (67) (and applying the guess that \( e_y = 0 \)) one gets for all \((\xi_b, \xi_y)\):

\[
\frac{\mathcal{U}_{h_e}}{\mathcal{U}} = \frac{\lambda_1}{h_e} \left( Ah^\lambda + e \left( \varepsilon_{eq} \frac{\alpha_1}{\alpha_2} - \varepsilon_{ehs} \alpha_1 \right) \right)
\]
Then substituting into (64):

\[
\beta E_{\xi_y, \xi_b} \left( \frac{\mathcal{U}_{h_e'}}{\mathcal{U}} \right) = \lambda_1 \frac{\varepsilon_{eq}}{\alpha_2 h'_e} \\
\iff \beta E_{\xi_y, \xi_b} \left( \frac{\lambda'_1 h'_e}{\lambda_1} \right) \left( A' \ell'h'_{\lambda_1} + \varepsilon' \left( \varepsilon_{eq} \frac{\alpha_1}{\alpha_2} - \varepsilon_{ehs} \alpha_1 \xi'_y \right) \right) = \lambda_1 \frac{\varepsilon_{eq}}{\alpha_2 h'_e} \\
\iff \beta E_{\xi_y, \xi_b} \left( \frac{\lambda'_1 (\xi'_y, \xi'_b)}{\lambda_1} \right) \left( \alpha_2 A' \ell'h'_{\lambda_1} + \alpha_1 e' \left( \varepsilon_{eq} - \alpha_2 \varepsilon_{ehs} \right) \right) = \varepsilon_{eq} \tag{68} \\
\iff \beta E_{\xi_y, \xi_b} \left( \frac{\lambda'_1 (\xi'_y, \xi'_b)}{\lambda_1} \right) \left( \alpha_2 \left( \frac{\mu}{\eta \varepsilon'} \right) \frac{\mu}{\eta} A' h_{\lambda_1} + \alpha_1 e' \left( \varepsilon_{eq} - \alpha_2 \varepsilon_{ehs} \right) \right) = \varepsilon_{eq} \tag{69} \\
\iff \beta E_{\xi_y, \xi_b} \left( \frac{\lambda'_1 (\xi'_y, \xi'_b)}{\lambda_1} \right) \left( \alpha_2 \left( \frac{\mu}{\eta \varepsilon'} \right) A' h_{\lambda_1} + \alpha_1 e' \left( \varepsilon_{eq} - \alpha_2 \varepsilon_{ehs} \right) \right) = \varepsilon_{eq} 
\]

**Proposition B.2.** The competitive equilibrium with complete financial markets is Pareto optimal if and only if \( \omega_3 = 0 \).

**Proof.** Assume \( \omega_3 = 0 \). Since \( \eta \to +\infty \),

\[
I(i) = e(i) \iff \varepsilon_{eq} = \frac{1}{\omega_1} \quad \text{and} \quad \varepsilon_{ehs} = \frac{\omega_2}{\omega_1}
\]

makes (69) and (55) equivalent and is compatible with the F.O.C. and budget constraint of colleges.

\( \omega_3 = 0 \) is a necessary condition for equation (67) to be compatible with the one in the social planner’s problem.

Furthermore, setting \( \omega(i) = \omega(i) \) implies that the consumption share of households are identical in the social planner and the decentralized equilibrium. \( \square \)

### B.4 Efficiency and Peer-effect

**Proof.** 1. The variance of \( \ln h_s \) within a college is given by

\[
\left( \frac{\omega_I}{\omega_A} \right)^2 \frac{\Sigma_h^2}{\Sigma_h^2 + \frac{\omega_A}{\omega_A + \omega_I}} \frac{\sigma_b^2}{\Sigma_b^2} = \left( \alpha_1 x \right)^2 \frac{\Sigma_h^2 \sigma_b^2}{\Sigma_h^2 + \frac{\sigma_b^2}{(x + 1)^2}} \\
= \left( \alpha_1 x \right)^2 \frac{\Sigma_h^2 \sigma_b^2}{(x + 1)^2 \Sigma_h^2 + \sigma_b^2} = f(x)
\]

The derivative w.r.t. \( x \) is:

\[
f'(x) = \left( \alpha_1 x \right)^2 \frac{\Sigma_h^2 \sigma_b^2}{(x + 1)^2 \Sigma_h^2 + \sigma_b^2} \left[ \frac{1}{x} - \frac{(1 + x) \Sigma_h^2 \sigma_b^2}{(x + 1)^2 \Sigma_h^2 + \sigma_b^2} \right]
\]

97
and the condition for $f(x)$ to be decreasing is:

$$\frac{1}{x} - \frac{(1 + x)\Sigma_h^2}{(1 + x)^2\Sigma_h^2 + \sigma_h^2} < 0$$

$$\iff (1 + x)\Sigma_h^2 + \sigma_h^2 < 0$$

which is never true as long as $\varepsilon_I > 0$. So for a given level of inequality of human capital in the economy $\Sigma_h^2$, the variance of $\ln h_s$ is always strictly increasing in the ratio $\frac{\omega_I}{\omega_A}$.

2. When $\omega_1 = 0$, the law of motion of the distribution of human capital both in the competitive equilibrium and in the first best is given by

$$m'_h = \rho m_h - \frac{\sigma_y^2}{2} + \ln \kappa - \frac{\sigma_b^2}{2}$$

$$\Sigma_h'^2 = \frac{\sigma_y^2}{2} + \rho^2 (\sigma_b^2 + \Sigma_h^2)$$

$$\rho = \alpha_1 (1 + \alpha_2 \omega_2)$$

Labor supply however is given by

$$\ell^{CE} = \left(\frac{1 - \tau_y}{\eta}\right)^{\frac{1}{\eta}}$$

$$\ell^{SP}(i) = \left(\frac{Ah(i)^\lambda}{\eta C}\right)^{\frac{1}{\tau - 1}}$$

which converges to 1 because $\eta \to +\infty$, still a necessary assumption to maintain perfect assortative matching in the social planner’s allocation. Hence the allocation of students and resources is exactly the same in the decentralized equilibrium and in the social optimum. The allocations of consumption remain different due to the absence of insurance against idiosyncratic risks in the decentralized economy.

□

**B.5 Second-best and Optimal Higher Education Policy**

I do the proof in the case $\omega_3 = 0$ and $\gamma = \lambda$ without loss of intuition. I also focus on the financial aid part of the optimal tax mix and set the progressive transfers to 0: $\tau_u = 0$. Furthermore it is obvious that $a_u$ and $a_h$ will play the same role so I set $a_u = 0$. The objective of the government is to maximize:
\[
\ln W^g(m_h, \Sigma_h) = \max_{s,l,\tau_m,\tau_y,\tau_n,a_y,a_h} (1-\beta) \left[ \ln \left( \int \left( \frac{(1-s)y}{1+a_c} \right)^{\frac{\sigma-1}{\sigma}} \frac{\sigma}{\sigma-1} \right) + \beta \ln W^g(m_h', \Sigma_h') \right]
\]

\[
m'_h = pm_h + X_1(s, l, \tau_m, \tau_y, \tau_n, a_y, a_h, \Sigma_h') (\mu_1)
\]

\[
\Sigma_h' = (\alpha_h(\tau_y, \tau_n, \tau_m))^2 \Sigma_h + X_2(\tau_m) (\mu_2)
\]

\[
\ell = \left[ (1 - \tau_y) \frac{1}{\eta} \left( 1 + \frac{\beta}{1-\beta} \sigma_2 (1 - \tau_n) - \sigma_3 V' \right) \right] \frac{1}{\eta} (\mu_3)
\]

\[
0 = s_t(1 + a_{u,t})(1 + a_{h,t}) - \frac{a_{y,t}}{1 - a_{y,t}} - \left( a_{c,t} \frac{(1 - s_t)}{1 + a_{c,t}} + s_t \right) (\mu_4)
\]

where \(V'\) is the next generation’s marginal value of human capital. The flow of utility for the government can be rewritten as follows:

\[
\ln \left( \int \left( \frac{(1-s)y}{1+a_c} \right)^{\frac{\sigma-1}{\sigma}} \frac{\sigma}{\sigma-1} \right) - \ell^n = \ln \frac{(1-s)(1-a_y)A^{\ell n}}{(1+a_c)} + \lambda m_h + \frac{(1-\tau_y)^2 \sigma - 1}{\sigma} + \tau_y(2-\tau_y) \frac{\lambda^2}{2} \Sigma_h^2 - \ell^n
\]

\[
= \frac{(1-s)(1-a_y)A^{\ell n}}{(1+a_c)} + \lambda m_h + \frac{\lambda^2}{2} \Sigma_h^2 - (1-\tau_y)^2 \frac{1}{\sigma} \frac{\lambda^2}{2} \Sigma_h^2 - \ell^n
\]

I now derive the optimal set of average consumption tax and subsidies to colleges. The lagrangian is:

\[
\mathcal{L} = -\ln(1-a_c) + ... + \lambda m_h + \gamma \left[ m_h - \rho m_h - \alpha_2 \omega_1 \ln \left( \frac{a_y}{(1-a_y)s} + \frac{a_c}{(1+a_c)s} + 1 \right) + ... \right]
\]

where the constraint is the law of accumulation of \(m_h\) where we dropped the terms that are independent of \(a_c\) and where we have substituted for \((1+a_h)(1+a_u)\) using the government aggregate budget constraint. The F.O.C. w.r.t. \(a_c\) and \(m_h\) write:

\[
- \frac{1}{1+a_c} = \gamma \alpha_2 \omega_1 \frac{(1-s)s}{(1+a_c)^2 s^2} - \frac{a_y}{(1-a_y)s} + \frac{a_c}{(1+a_c)s} + 1
\]

\[
\lambda + \gamma (1-\rho) = 0
\]

\[
\Rightarrow a_c = \frac{\alpha_2 \omega_1 (1-s)(1-a_y)}{1-\rho} - s - a_y
\]

C Estimation - Details

For future reference, I call M1 the benchmark model explored in section 2 and 3 and M2 the version of the model with intergenerational financial transfers and enrollment decision.
C.1 The College Problem in the Quantitative Version

In this section I describe how I keep the college problem tractable despite the loss of closed-form expressions for the distribution of students within the college and equilibrium tuition.

I essentially assume that the problem of the college is given by (42), rewritten below.

$$\max_{\theta, I, r, D, y} I^{\omega_1} \theta^{\omega_2} D^{-\omega_3}$$

$$\ln I \int_0^1 r(\ln h_s, \ln y) d\ln h_s d\ln y = \int_0^1 r(\ln h_s, \ln y) \left( (1 - \tau_y) \ln(e_u)^i + \ln(1 + \alpha_u)T_u/p_I \right) d\ln h_s d\ln y$$

$$\ln \theta \int_0^1 r(\ln h_s, \ln y) d\ln h_s d\ln y = \int_0^1 r(\ln h_s, \ln y) \ln h_s d\ln h_s d\ln y$$

$$\ln D \int_0^1 r(\ln h_s, \ln y) d\ln h_s d\ln y = \int_0^1 r(\ln h_s, \ln y) \ln y d\ln h_s d\ln y$$

In M1, this problem resulted from the combination of the equilibrium tuition schedule, with the budget constraint and the expression for the cost of heterogeneity (41).

In M2, it is no longer possible to derive this problem from the primitive problem (8). This version of the college problem can be reinterpreted as follows: a college seeks to maximize a weighted geometric average of tuition and students abilities and has a social objective. Notice that this reinterpretation of the college’s problem is close to Fu [2014] where colleges maximize a weighted average of student ability and a quadratic function of net tuition. What my formulation entails is a punishment for tuition and ability heterogeneity because averages are geometric and not arithmetic.

Finally, the first order conditions for this problem are the same as in M1, see appendix A.2 for more details.

C.2 External Calibration

**Income Tax Schedule** $a_y, \tau_y$ In order to calibrate $\tau_y$, I take an average between the value estimated by Heathcote et al. [2017a] and the ones needed to match the ratio between the market income and after tax and transfers Gini in the U.S in 2000, $\tau_y = .23$. The latter estimate $\tau_y$ directly using two different datasets over the period 2000-2005: using the data provided by the CBO—itself based on the SOI and the CPS—they obtain $\tau_y = .2$ and using the PSID, they find $\tau_y = .18$. The value needed to rationalize the .12 difference in the Gini coefficient of households market income (.56) and after tax and transfers income (.44) in 2000-2005, within the log-normal framework of M1 is .26. The discrepancy is due on the one hand to the log-normal assumption and on the other to the slightly different different measure of income used—for example, Heathcote et al. [2017a] exclude Medicare benefits from their measure of transfers. I calibrate $a_y$ using average income tax rate provided by CBO: $a_y = .2$.

\[a_{\text{CBO}} \text{(The Distribution of Household Income and Federal Taxes, 2013, CBO and own calculations)}\]

- 41
College Subsidies, $a_h, a_u, \tau_u$  From the IPEDS, I compute $a_u$ by dividing the total amount of public subsidies by the total revenues before public aid. From the NPSAS I obtain $a_h$ by dividing the total amount of public financial aid by the sum of out of pocket payments.

According to the specification for subsidies to university, $\tau_u$ can be estimated in a weighted least-square regression of (log) total revenues per student on (log) revenues before public transfers in the cross-section of colleges, where the weights are given by students enrollment. I run this regression in a companion paper Capelle [2019c] and find $\tau_u = .35$ at the beginning of the 2000s.

Frisch Elasticity, $\eta$  $\eta$ is set to match the Frisch elasticity of labor supply, $\varepsilon_{\ell,w} = \frac{1}{\bar{\eta}}$. Empirical estimates of $\varepsilon_{\ell,w}$ range from .2 to .7 [Chetty et al., 2011]. My preferred estimate in the conservative value $\varepsilon_{\ell,w} = .2$, which implies $\eta = \mu \left(1 + \frac{1}{\bar{\eta}}\right) = 2$.

Ideally we would have elasticities of lifetime household income to wages in order to be consistent with the model. All estimates however are at the individual and yearly level. I argue that they are likely upper bounds for two reasons. First they do not capture intra-household substitution. Second they do take into account the intertemporal substitution stemming from temporary fluctuations in wages.

Calibrating $\bar{\lambda}$  The elasticity of substitution in the educational services sector is set to $\bar{\lambda} = \lambda$ so that the price of educational services is given by $p_I = \frac{A_I}{A}$.

C.3 Calibrating $\lambda$ and $\mu$

$\lambda$ and $\mu$ are calibrated to match the share of income going to raw labor, human capital and capital, $\lambda = .67, \mu = .33$. I also show that microfounding the wage function using a Constant Elasticity of Substitution aggregate production function would lead to a similar parametrization, for reasonable values of the elasticity of substitution across skills. This section also provides details regarding the calibration of the change in the value of $\lambda$ from 2000 to 1980.

Using shares of raw labor, human capital and physical capital.  As in Benabou [2002], assume the primitive production function is a Cobb-Douglas production function with raw labor $\ell$, human capital $h$ and physical capital $k$:

$$y = a\ell^{\phi_{\ell}}h^{\phi_h}k^{1-\phi_{\ell}-\phi_h}$$

If the interest rate is $r$, efficiency requires $\frac{\partial y}{\partial k} = r$ which implies $r = (1 - \phi_{\ell} - \phi_h) \frac{y}{k}$, hence
substituting for $k$ into the production function gives

$$y = A\ell^{\phi_\ell h^{\phi_h}}$$

with

$$A = \left[a \left(\frac{1 - \phi_\ell - \phi_h}{r}\right)^{1 - \phi_\ell - \phi_h} \right]^{\frac{1}{\phi_\ell + \phi_h}}$$

This expression of the production function implies that $\lambda = \frac{\phi_h}{\phi_\ell + \phi_h}$ and $\mu = \frac{\phi_\ell}{\phi_\ell + \phi_h}$. In the U.S., in the beginning of the 2000’s, the share of income going to capital is .4. The share of income going to raw labor has been estimated to be between .05 and .2 [Barro et al., 1995, Krueger, 1999]. This leads to estimates of $\lambda$ between .67 and .92 and estimates of $\mu$ between .33 and .08.

Our favorite value is the lower bound $\lambda = .67, \mu = .33$. One should think of it as a very conservative calibration: choosing a higher value of $\lambda$ would only magnify the results regarding the amplification role of colleges for inequality and social immobility by making the income-sorting channel only stronger.

**Mapping into an aggregate CES production function** Building on Benabou [2002], it is possible to show that if the aggregate production is CES with elasticity of substitution $\sigma$

$$Y = a \left[ \int x(s) \frac{\sigma - 1}{\sigma} ds \right]^{\frac{\sigma}{\sigma - 1}}$$

where $s$ stands for skills, if an individual with human capital $h$ and labor $\ell$ can produce $h\ell$ units of intermediary inputs and the distribution of human capital is log-normally distributed, then the individual market income function is given by

$$y_m = Ah^{\frac{\sigma - 1}{\sigma}} \ell^{\frac{\sigma - 1}{\sigma}}$$

with

$$A = a^{\frac{\sigma - 1}{\sigma}} Y^{\frac{1}{\sigma}}$$

This implies a direct mapping between the elasticity of substitution and $\lambda = 1 - \frac{1}{\sigma}$. When skills become more substitutable, the returns to human capital increase.

The range of estimates for $\sigma$ is large: Borjas et al. [2012] find estimates ranging from 2 to 20, implying a value of $\lambda$ between .5 and .95. Because the main counterfactual exercise consists in increasing $\lambda$ over time, I rerun the benchmark regression by Autor et al. [2008] of college/high school log wage gap over college/high school relative supply allowing for a change in the coefficient before and after the middle of the sample period, i.e. allowing for the one coefficient for the 1963-1983 period to be different from the one for the 1984-2005 period. As in their paper, I include a time trend allowing for a break in 1992, as well as log minimum wage and male prime-age unemployment rate. The results are displayed in table (A1). Column 1 replicates column 3 of Autor et al. [2008].

---

[42]Krueger [1999] shows how to relate the share of raw labor to the intercept in a Mincer regression in the U.S. Barro et al. [1995] uses estimates from income per capita cross-country regressions.
Column 2 allows for an interaction of the relative supply of college graduates with an indicator function equal to one for years greater than 1984. Column 3 adds two additional controls: minimum wage and the unemployment rate. The coefficient for the post-1984 period on the relative supply of graduates in the third column implies \( \lambda = 1 - .328 = .672 \), which is very close to the conservative calibration of \( \lambda \) discussed above. This estimated coefficient is also likely to be a conservative estimate of the elasticity of substitution since individuals are partitioned in only two large skills groups, college and high school graduates. A more granular partition would likely gives higher estimates for \( \lambda \).


<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLG/HS relative supply</td>
<td>-0.619***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0657)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLG/HS relative supply ×</td>
<td>-0.563***</td>
<td>-0.526***</td>
<td></td>
</tr>
<tr>
<td>pre-1983</td>
<td>(0.0799)</td>
<td>(0.124)</td>
<td></td>
</tr>
<tr>
<td>CLG/HS relative supply ×</td>
<td>-0.410**</td>
<td>-0.328</td>
<td></td>
</tr>
<tr>
<td>post-1984</td>
<td>(0.183)</td>
<td>(0.233)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>0.0258***</td>
<td>0.0235***</td>
<td>0.0215***</td>
</tr>
<tr>
<td></td>
<td>(0.00229)</td>
<td>(0.00294)</td>
<td>(0.00447)</td>
</tr>
<tr>
<td>Time × post-1992</td>
<td>-0.00798***</td>
<td>-0.00977***</td>
<td>-0.00891***</td>
</tr>
<tr>
<td></td>
<td>(0.00202)</td>
<td>(0.00248)</td>
<td>(0.00260)</td>
</tr>
<tr>
<td>Log real minimum wage</td>
<td>-0.0657</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0529)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male prime-age unemp. rate</td>
<td>0.000823</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00363)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.146**</td>
<td>-0.0931</td>
<td>0.0589</td>
</tr>
<tr>
<td></td>
<td>(0.0573)</td>
<td>(0.0714)</td>
<td>(0.188)</td>
</tr>
<tr>
<td>Observations</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.953</td>
<td>0.955</td>
<td>0.957</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Calibrating the change in \( \lambda \) from 1980 to 2000 Both calibration strategy leads to an increase over time of \( \lambda \). According to the first strategy, since the share of capital has gone up from .36 to .4,\(^{43}\) and Krueger [1999] provides evidence that the share of aggregate income going to

\(^{43}\)BLS
raw labor has declined by about 7 percentage points over the same period, it must be that \( \lambda \) has increased over time, with \( \lambda_{1980} = .578 \).

According to the second strategy, column 3 of table (A1) implies that for the earlier sample period, \( \lambda_{1980} = 1 - .526 = .474 \). This direct empirical evidence of a rise in the elasticity of substitution across skills is corroborated by a large body of research. Kremer and Maskin [1996] provide evidence that the recent growth in wage inequality has been accompanied by greater segregation of high- and low-skill workers into separate firms. More recently, Song et al. [2018] show that one-third of the rise in the variance of (log) earnings occurred within firms, whereas two-thirds of the rise occurred between firms due to a widening gap in the composition of their workers. The theoretical literature has also provided mechanisms through which the aggregate elasticity of substitution could rise in an economy experiencing skill-biased technological change, increasing international trade and adoption of modern information and communication technologies. Kremer and Maskin [1996] show in an assignment model that skill-biased technological change can account for increasing segregation and increase in inequality. Grossman and Maggi [2000] show that international trade leads the country with the most diverse labor force to specialize in the production of goods that require more substitutability between factors. Garicano and Rossi-Hansberg [2006] show that better communication technology may lead to organization changes implying higher substitutability across workers and increasing inequality.

Left with a range of \( \lambda \in [.474, .578] \), in the final calibration, I pick \( \lambda = .55 \).

### C.4 Internal Calibration

In this section, I provide details about the way the empirical moments are constructed, how I construct the model-consistent variables and more importantly discusses the sources of identification. In M1, the log-linearity of the equilibrium relationship makes very transparent which moments inform which parameters since one can map directly moments into combinations of parameters. This analytical mapping is lost in M2, but the intuitions go through.

#### Assumption of Steady-State

I assume that the economy is at steady-state at the beginning of the 2000s. The assumption of steady-state really matters only for the identification of \( \sigma_y^2 \). Moreover, the endogenous convergence speed is quite fast—the half life of the AR(1) for the variance of income is given by \( \frac{-\ln(2)}{\ln(\alpha_h^2)} \simeq .5 \), which corresponds to half a generation. But one could still be concerned that the movement of exogenous variables—especially persistent increase in \( \lambda \)—have not reached a steady-state. However, the future increase in \( \lambda \) feedbacks into the current allocation only through the aggregate investment rate into higher education—at least in the version of the model for which I have closed-forms—and through labor supply, so would not change the cross-sectional and inequality implications, on which most of the identification is based.
Identification of $\sigma^2_b$ and Construction of Model-Consistent Abilities  At this point of the procedure, I construct a grid on $\alpha_1$. I estimate $\sigma^2_b(\alpha_1)$ on this grid as follows: according to the assumed relationship between children’ high school abilities and parental human capital, one has

$$\ln h_{s,i} = \ln(\xi_{h,i} h_i)^{ \alpha_1} = \frac{\alpha_1}{\lambda} \ln y_{m,i} + \alpha_1 \ln \xi_{h,i}$$

with $V(\ln \xi_{h,i}) = \sigma^2_b$

In NLSY97, I observe gross parental income $y_{m,i}$ and the ranking in test scores $\text{rank}(h_{s,i})$. For a given $(\alpha_1, \lambda)$, there exists a unique $\sigma^2_b$ that matches the correlation between parental income and the rank of the child at the test, $\rho(y_{m,i}, \text{rank}(h_{s,i}))$. This identifies $\sigma^2_b(\alpha_1)$.

I can then construct a model-consistent measure of abilities $\{\ln h_{s,i}\}(\alpha_1, \varepsilon)$. I generate $\{\ln h_{s,i}\}$ consistent with $(\alpha_1, \lambda, \sigma^2_b(\alpha_1))$ and $\text{rank}(h_{s,i})$ to use in subsequent steps. This last step implies a random draw that introduces some noise in the measure of ability, i.e. $\{\ln h_{s,i}\}$ are not deterministic function of $\alpha_1, \lambda, \sigma^2_b(\alpha_1)$ and $\text{rank}(h_{s,i})$. I provide more details in appendix C.7 about the identification of $\sigma^2_b$ and the construction of abilities.

The Progressivity of Financial Aid, $\tau_n, \tau_m$ and $\omega_3$. According to the assumed functional form for government financial aid and the equilibrium tuition schedule, the following is true

$$\ln \frac{e^h_i}{e^u_i} = \tau_n (1 - \tau_y) \ln y_{m,i} - \tau_m \ln h_{s,i} + c_0$$

$$\ln e^h_{i,j} = \gamma_j + \left( \frac{\omega_3}{\omega_1} (1 - \tau_u) + \tau_n \right) (1 - \tau_y) \ln y_{m,i} - \left( \frac{\omega_2}{\omega_1 (1 - \tau_u)} + \tau_m \right) \ln h_{s,i} + c_1$$

where $\gamma_j$ are college fixed effects and the second equation is true only in M1 (the first one is true in M1 and M2).

In the NCES-NPSAS dataset, one observes parental income $y_{m,i}$, test score, institutional aid, government aid as well as out of pocket payment. Regressing the (log) ratio of after-government aid payment on before-government aid payment over parental income and student ability gives us $\tau_n$ and $\tau_m$ (first equation). This equation also holds true in M2.

The second equation relates between before-government aid payment and college fixed effect, parental income and student ability. It tells us that the elasticity of before-government aid tuition to parental income $\varepsilon_{e,y_m}(\alpha_1)$ identifies the progressivity of institutional financial aid $\omega_3$, what I called the social objective parameter. In the data I find that the fit of the second equation to the data is very high, $R^2 = 80\%$. In the model M2, I run the second regression on a simulated population.

Constructing a Measure of Quality, $q_i$. At this point of the procedure, I need to define a grid on $\frac{\omega_2}{\omega_1}$. At each point of the grid $(\alpha_1, \frac{\omega_2}{\omega_1})$, I construct a measure of annual quality delivered by all colleges, indexed by $j$. I collect average real spending per student by college using IPEDS and median test score within a college. The median test score is defined as the mean between the
bottom and top quartile—the only available data in the IPEDS—which is exactly the median if the
distribution is symmetric. When test scores data are not available for 2000, I either use years up
to 2004, or impute them based on a regression of test scores on spending per student and other
characteristics. This measure based on ACT or SAT is converted into a model-consistent measure
of abilities using a scale based on publicly available quantiles of the distribution of these scores and
using the guess on \( \frac{\omega_2}{\omega_1} \). I construct the annual quality delivered by a college consistent with the
functional form for the production function given by 9 and the guess for \( \frac{\omega_2}{\omega_1} \).44

I then construct a measure of quality received by an individual \( i \). \( q_i \) is a weighted average of the
\( q_j \) where the weights depend on the time spent in each college and whether they have graduated at
any point. I have checked the robustness of the results to alternative ways of aggregating annual
college qualities. More details on this may be found in appendix C.6.

**Estimating the Elasticity of Quality to Ability, \( \varepsilon_{q,h_s}(\alpha_1, \frac{\omega_2}{\omega_1}) \) and Identifying \( \frac{\omega_2}{\omega_1} \).**

According to the equilibrium sorting rule in M1, the elasticity of quality to ability identifies \( \frac{\omega_2}{\omega_1} \):

\[
\ln q_i = c + h(\Sigma_h) \left[ \left( \left( 1 - \tau_u \right)(1 - \tau_n) - \frac{\omega_3}{\omega_1} \right) (1 - \tau_y) \ln y_{m,i} + \left( \frac{\omega_2}{\omega_1} + \tau_m (1 - \tau_u) \right) \ln h_{s,i} \right] \tag{71}
\]

with

\[
h(\Sigma_h) = \frac{\omega_1}{1 - \nu(\Sigma_h)\omega_3}
\]

Since the measure of quality is based on a guess on \( \frac{\omega_2}{\omega_1} \), finding \( \frac{\omega_2}{\omega_1} \) is a fixed point problem. Empirically, the elasticity of quality to ability doesn’t change a lot with the guess on \( \frac{\omega_2}{\omega_1} \), the latter
is therefore tightly identified. Although this relationship should hold perfectly in the model, the
\( R^2 \) in the data associated with this regression is 23%. Finally notice that even in M1, it is not
possible to recover \( \frac{\omega_2}{\omega_1} \) directly from the elasticity of quality to abilities without numerically solving
the model, since it depends on some endogenous aggregate variable, \( h(\Sigma_h) \).

**The Human Capital Accumulation Function \( \alpha_2 \left( \alpha_1, \frac{\omega_2}{\omega_1} \right), \alpha_3 \left( \alpha_1, \frac{\omega_2}{\omega_1} \right) \).** According to the
law of motion for human capital

\[
\ln y_{m,i} = c_y + \lambda \ln h_{s,i} + \alpha_2 \omega_1 \lambda \ln q_i + \alpha_3 \ln y_{m,i} + \ln \xi_{y,i}
\]

running a regression of children’ income on parental income, abilities and quality identifies \( \alpha_2 \)
and \( \alpha_3 \). I run the regression using the NLSY97, where I observe parental income, children ability,
children college quality and children earnings.

At this point, I gather all the moments that depend on \( \left( \alpha_1, \frac{\omega_2}{\omega_1} \right), \left( \sigma_b^2, \tau_n, \tau_m, \varepsilon_{e,y_m}, \varepsilon_{q,h}, \varepsilon_{q,h_s} \right) \alpha_2 \alpha_3 \)

The remaining targeted moments are independent of \( \left( \alpha_1, \frac{\omega_2}{\omega_1} \right) \).

---

44I do not need to take into account the within college heterogeneity, corresponding to \( \sigma_a^2 \) because, at least
in the model, it is common to all colleges and will therefore factor out and leave our regression coefficients
unchanged.
Using the Gini Coefficient for Income to Identify $\sigma^2_y$. I target the Gini coefficient for income. In M1, from the steady-state variance of human capital, one has

$$
\text{Gini}(y_m) = 2\Phi\left(\lambda \sqrt{\frac{(\Sigma^S)^2}{2}}\right) - 1
$$

$$
(\Sigma^S)^2 = \frac{\sigma^2_y + \left(\alpha_1[1 + \alpha_2(\Sigma^S) + \tau_m(1 - \tau_u)\epsilon_1(\Sigma^S)]\right)^2 \sigma^2_h}{1 - (\alpha_h)^2}
$$

where $\Phi$ is the c.d.f. of a standard normal. The best estimate for the Gini coefficient of lifetime labor earnings is from Kopczuk et al. [2010] who have access to administrative data. There would be two issues with the NLSY97: first children labor earnings are observed only up to 2015, i.e. their first years of labor market experience and a lot of them are not in a households yet. Secondly, top income are censored. Kopczuk et al. [2010] finds that the eleven-year Gini coefficient is between .45 and .50. This is slightly lower than the annual Gini coefficient, which is between .49 and .57—depending on the exact measure of gross income used—in 2000 according to the CBO, probably because of transitory income shock. I keep a Gini of lifetime labor earnings of .45 as a target.

The Intergenerational Elasticity identifies $\alpha_1$. From the steady-state equilibrium law of motion of human capital in M1, the children income elasticity to their parents income is the IGE:

$$
\ln y_{m,i}' = c + \alpha_h \ln y_{m,i} + \varepsilon_i
$$

with $\alpha_h = \alpha_1 + \alpha_3 + \alpha_1\alpha_2\varepsilon_2 + \tau_m(1 - \tau_u)\varepsilon_1 + \alpha_2(-\varepsilon_1(1 - \tau_u)(1 - \tau_n) - \varepsilon_3)(1 - \tau_y)\lambda$

As shown in the expression in the second line, the IGE directly informs $\alpha_1$. There is some disagreement in the literature regarding the magnitude of the IGE, with estimates ranging from .3 to .6. Even the recent literature that uses administrative data is not immune to the short-panel and lifecycle biases.\textsuperscript{45} I take the intergenerational elasticity from Mazumder [2015] who provides the most compelling and robust estimation, and target $\alpha_h = .5$.

$\beta$ and the Share of Private Spending for Higher Education in GDP. I calibrate the intergenerational discount factor, $\beta$, to match the average private spending on higher education in GDP. In M1, the latter is:

$$
s(1 - a_y) = \frac{\beta\alpha_2\omega_1(1 - \tau_u)V(1 - a_y)}{1 - \beta + \beta\alpha_2\omega_1(1 - \tau_u)V}
$$

The OECD reports that share of private spending for higher education in GDP in the U.S. over

\textsuperscript{45}Given that only the first years of the children earnings are observed while the parents are observed when they are already older, the estimate of the intergenerational elasticity from the NLSY97 is biased downward—I indeed find $\hat{\alpha}_h = .3$. 

107
the period 2000-2004 is 1.3%.\footnote{For reference, they also report that the share of public spending is 1%, making spending in higher education 2.3% of GDP.}

C.5 Calibrating $q, a, \bar{a}$ and $r$ in M2

In M2, there are two additional parameters to calibrate: $q$ and $r$.

$q$ and the enrollment rate. To calibrate $q$—the outside option to college—it is natural to target the enrollment rate: the lower $q$, the stronger the incentives to go to college. The immediate enrollment rate, provided by the NCES, in the U.S. in the 2000s is between 67 and 70%.

Calibrating limits to intergenerational financial transfers, $a, \bar{a}$. There is no limit to how much individuals can bequeath, $\bar{a} = +\infty$. For the lower bound, I target the official limit on student loan, as a percentage of lifetime GDP per capita, which amounts to 3%.

$r$ and the income-sorting channel. First I consider a small open economy, and do not try to find the interest rate that ensures market clearing of the financial asset market. Instead, I target the elasticity of quality to parental income, $\varepsilon_{q,y}$ in the regression (71). It turns out that in M1, the elasticity resulting from the estimation is too high compared to what is in the data—.4 instead of .2. By increasing $r$, one gives incentives to individuals to avoid debt, which relaxes the borrowing constraint and decreases the dependence of college quality to parental income. I find a generational net interest rate of 180% which corresponds to an annual interest rate of 3.5% for a generation length $H = 30$ years.

C.6 Measuring individual-level quality of higher education

In this subsection, I describe in details how I construct $q$ for each individual. The general formula is as follows. Define $\mathcal{C}$ the set of all colleges, $N$ is the number of colleges. Define an arbitrary ordering onto this set from 1 to $N$. $\mathcal{C}_i \in \mathcal{C}$ the subset of colleges visited by $i$ at some point, and $C_i$ the vector whose first element is the index of the first college visited, the second element the index of the second etc... I restrict the size of $C_i$ to be no larger than 5 which is also the maximum number of college visited in the NLSY. Define $y_i$ the vector of years spent in each college, ordered as in $C_i$, $s_i$ the vector of degree sought in each college (measured in years of higher education, 2 years, 4 years etc...), $d_i$ the vector of dummy equal to 0 if the previous degree has not been obtained and 1 otherwise and $k_i$ the vector of yearly quality of each college visited. The general formula for the quality of higher education received by individual $i$ is:

$$q_i = \sum_{n=1}^{\text{size}(C_i)} w_n k_i(n)$$
with \( w_n^* = n(n)(1 + d_i) \) which means that if the student has graduated from a college they started at, I multiply the yearly quality by the number of years the program should theoretically last and multiply it by half the time if they dropped out and/or didn’t graduate. For example, if a student has started a four-year degree and has graduated, the weight for this college is \( w = 4 \), if they have dropped out, the weight is only 2.

I deal with students transferring to a 4-year college after graduation in a 2-year college in the following intuitive way: the weight on the 2-year college quality is 2, and the weight on the 4-year college quality is 2 if they obtain the 4-year degree and 1 otherwise. I have tried other weighting schedules:

1. Select the college a student has spent the most time at. Multiply the quality at this college by the theoretical number of years from high school until graduation but divided this weight by two if the student hasn’t graduated. This methodology is the closest to the way Chetty et al. [2017] assign students to college.

2. Select the college a student has obtained the highest degree from (and if has obtained no degree, take the highest degree sought) and multiply the yearly quality by the same weight as in the previous bullet point.

3. Select the college with highest quality a student has visited. Apply the same weight as in the previous bullet points.

I find very high correlation between these 4 measures and virtually identical results in the estimation of the model.

C.7 Estimating \( \sigma^2_b \) and constructing abilities

In the data I observe a measure of abilities and a measure of parental income. I want to construct a measure of abilities which is consistent with the model. I start with values of \( \alpha_1, \lambda_\sigma^2_b \). To this end, I generate a random sample of abilities from the true distribution in the model and the empirical distribution of parental income - I use our guess for \( \lambda, \alpha_1 \). By definition, this sample has the correct mean and variance as in the model. I reorder this sample in the exact same order as the original ordering of the children’s abilities. I call this reordered variable.

I do the procedure described in the previous paragraph for a lot of different \( \sigma^2_y \). It turns out that the correlation between our constructed measure of abilities and the empirical measure of parental income identifies uniquely \( \sigma^2_b \) (for a given \( \alpha_1, \lambda \)).

I have checked that this methodology works (see numerical_test.m.) well. Test-scores as measures of ability raise some issues. The Armed Services Vocational Aptitude Battery (ASVAB) scores reported in the 1997 wage of the National Longitudinal Study of Youth has the advantage of being a rather comprehensive tests including different topic, verbal and math and design to last approximately three hours. The scaling of the scores do not matter for our analysis since I use only the correlation between the parental income and the rank in scores. Although the ASVAB
scores are constructed as posterior modes of an individual latent ability, and is therefore potentially
bias towards the mean, this doesn’t raise a concern for us since I use only the ranks. Moreover,
the ASVAB does not use conditioning variables which could be a source of noise in the ranking.
Although the ASVAB is a rather broad and comprehensive test, one might still be concerned by
classical and non-classical measurement error. To check the robustness of our methods to classical
measurement error, I run simulations. I find that classical measurement error would bias our
estimate of the variance of the birth shock—σ_b—upwards and our estimate of α_1 downward. The
bias is negligible as long as the standard deviation of the (normally distributed) noise is an order of
magnitude lower than \sqrt{σ_b^2}, i.e. as long as it is smaller than \sqrt{σ_b^2}/10.

In theory, I would need to have a measure of A and ℓ but the estimates of \sigma_b^2 is completely
independent of it, as well as the implied moments in the following regression. It just shifts the
constant in the regressions.

C.8 Miscellaneous issues

Dealing with top-coded income Income data in the NLSY are top-coded such that the top
2% all have the same value. I have tried censored regression and non-censored regression, it doesn’t
affect our results.

C.9 Analytical Evidence of Identification in M1

The analytical expressions for the equilibrium relationships in M1 allows us to derive, not only
a heuristic argument regarding the role of each moments in the identification of each parameter,
but also a formal identification result of the set of parameters \theta_1 given the set of chosen moments
specified in table (1). This uniqueness result, formalized in 8, is derived under two restrictive
assumptions: individuals’ ability are directly observable and there is no social objective, \omega_3 = 0.

The data doesn’t satisfy the first assumption, because we observe only the rank of abilities and
not their level, a problem that the literature in education has dealt with for a long time. This is the
reason why I need to generate, at some point in the estimation procedure, levels of abilities that are
both consistent with the model (in particular the correlation structure imposed by it) and with the
empirical ranking of abilities.

The identification result therefore helps build confidence in the uniqueness of the global solution
found by our estimation procedure, but doesn’t completely ensure it. That is the reason why, I
provide further numerical evidence of identification in appendix C.10.

Lemma 8. In M1, given our set of targets (detailed in column 4 of table (1)), if \omega_3 = 0 and
{h_{si}} are observables, there exists a unique set of parameters \theta_1 consistent with the model-implied
restrictions.
Proof. When \( \omega_3 \), the set of restrictions for the estimation is:

\[
\ln h_{s,i} = \ln(\xi_{b,i}h_i)^{\alpha_1} = \frac{\alpha_1}{\lambda} \ln y_{m,i} + \alpha_1 \ln \xi_{b,i} \tag{72}
\]

with \( V(\ln \xi_{b,i}) = \sigma_y^2 \)

\[
\ln \frac{e_i^h}{e_i} = \tau_n(1 - \tau_y) \ln y_{m,i} - \tau_m \ln h_{s,i} + c_0 \tag{73}
\]

\[
\ln e_{i,j}^h = \gamma_j + \left( \frac{\omega_3}{\omega_1}(1 - \tau_u) + \tau_n \right) (1 - \tau_y) \ln y_{m,i} - \left( \frac{\omega_2}{\omega_1(1 - \tau_u)} + \tau_m \right) \ln h_{s,i} + c_1 \tag{74}
\]

\[
\ln q_i = c + h \left[ (1 - \tau_u)(1 - \tau_n) - \frac{\omega_3}{\omega_1} \right] (1 - \tau_y) \ln y_{m,i} + \left( \frac{\omega_2}{\omega_1} + \tau_m(1 - \tau_u) \right) \ln h_{s,i} \tag{75}
\]

\[
\ln y'_{m,i} = c_y + \lambda \ln h_{s,i} + \omega_2 \omega_1 \lambda \ln q_i + \alpha_3 \ln y_{m,i} + \ln \xi_{i,j} \tag{76}
\]

\[
\ln y'_{m,i} = c + \alpha_h \ln y_{m,i} + \varepsilon_i \tag{77}
\]

\[
\text{Gini}(y_m) = 2 \Phi \left( \lambda \sqrt{\frac{(\Sigma_h^2)^{SS}}{2}} \right) - 1 \tag{78}
\]

\[
(\Sigma_h^2)^{SS} = \frac{\sigma_y^2 + \left[ \alpha_1[1 + \alpha_2(\Sigma_h^{SS}) + \tau_m(1 - \tau_u)\varepsilon_1(\Sigma_h^{SS})] \right]^2 \sigma_b^2}{1 - (\alpha_h)^2} \tag{79}
\]

\[
s(1 - a_y) = \frac{\beta \alpha_2 \omega_1(1 - \tau_u)\varepsilon_1(1 - a_y)}{1 - \beta + \beta \alpha_2 \omega_1(1 - \tau_u)} \tag{80}
\]

From equations (73) and (74) one immediately get \( \tau_n, \tau_m, \omega_3/\omega_1 \). From the coefficient in front of \( h_s \) in equation (75), one identifies \( \omega_3/\omega_1 \). From (76), one identifies \( \alpha_2 \omega_1 \), and \( \alpha_3 \). The IGE is obtained with (77) (although I use another, more reliable source, as explained in the main text), which together with the previously obtained \( \alpha_2 \omega_1, \alpha_3, \omega_2/\omega_1, \omega_3/\omega_1, \tau_u, \tau_n, \tau_m, \tau_y, \lambda \) gives \( \alpha_1 \). From equation (72) one gets \( \sigma_b^2 \). The computation of the steady-state Gini identifies \( \sigma_y^2 \) (equation (78) and (79)). Finally, targeting the LHS of (80) and computing the steady-state value of \( V \) gives us \( \beta \).

The assumption that \( \omega_3 = 0 \) implies that \( h(\Sigma_h) = h \) a constant independent of the model, in equation (75). The assumption that one observes \( h_s \) is maybe the most worrying one of the two: there could be two completely disjoint part of the state-space of parameters compatible with this set of restrictions when one doesn’t observe abilities. The next appendix provides numerical evidence that this is not the case and that the parameters are well-identified.

C.10 Numerical Evidence of Identification in M1

The following figures display respectively the model-generated moments, empirical moments, parameters and value of the criterion for the five points in the grid with the lowest value of the criterion. The third table suggests that all parameters are tightly identified because none of the set
of parameters significantly depart from the set of parameters that minimizes the criterion.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Moments1</th>
<th>Moments2</th>
<th>Moments3</th>
<th>Moments4</th>
<th>Moments5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ah</td>
<td>0.49992</td>
<td>0.51191</td>
<td>0.50147</td>
<td>0.49133</td>
<td>0.50262</td>
</tr>
<tr>
<td>Giniy</td>
<td>0.45127</td>
<td>0.45707</td>
<td>0.44398</td>
<td>0.44074</td>
<td>0.45258</td>
</tr>
<tr>
<td>GiniI</td>
<td>0.18611</td>
<td>0.19013</td>
<td>0.18097</td>
<td>0.17701</td>
<td>0.18538</td>
</tr>
<tr>
<td>YpOn3Hs</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>YpOn3Q</td>
<td>0.13589</td>
<td>0.13556</td>
<td>0.13555</td>
<td>0.13528</td>
<td>0.13391</td>
</tr>
<tr>
<td>YpOn3Y</td>
<td>0.19664</td>
<td>0.19683</td>
<td>0.19638</td>
<td>0.19673</td>
<td>0.19664</td>
</tr>
<tr>
<td>QOnHsYY</td>
<td>0.40297</td>
<td>0.40296</td>
<td>0.40297</td>
<td>0.40296</td>
<td>0.40297</td>
</tr>
<tr>
<td>QOnHsYHs</td>
<td>0.83563</td>
<td>0.83238</td>
<td>0.883</td>
<td>0.93128</td>
<td>0.93036</td>
</tr>
<tr>
<td>EOOnHsYY</td>
<td>0.15004</td>
<td>0.15006</td>
<td>0.15004</td>
<td>0.15008</td>
<td>0.15004</td>
</tr>
<tr>
<td>EarnPrem</td>
<td>1.8877</td>
<td>1.9206</td>
<td>1.8489</td>
<td>1.8321</td>
<td>1.8984</td>
</tr>
</tbody>
</table>

Model-generated moments

<table>
<thead>
<tr>
<th>Parameters</th>
<th>MomentsE1</th>
<th>MomentsE2</th>
<th>MomentsE3</th>
<th>MomentsE4</th>
<th>MomentsE5</th>
</tr>
</thead>
<tbody>
<tr>
<td>ahE</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>GiniyE</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>GiniIE</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td>YpOn3HsE</td>
<td>0.17929</td>
<td>0.17201</td>
<td>0.18594</td>
<td>0.19068</td>
<td>0.17854</td>
</tr>
<tr>
<td>YpOn3QE</td>
<td>0.13589</td>
<td>0.13556</td>
<td>0.13555</td>
<td>0.13528</td>
<td>0.13391</td>
</tr>
<tr>
<td>YpOn3YE</td>
<td>0.19664</td>
<td>0.19683</td>
<td>0.19638</td>
<td>0.19673</td>
<td>0.19664</td>
</tr>
<tr>
<td>QOnHsYYE</td>
<td>0.18232</td>
<td>0.18308</td>
<td>0.18239</td>
<td>0.18378</td>
<td>0.18542</td>
</tr>
<tr>
<td>QOnHsYHsE</td>
<td>0.8785</td>
<td>0.85484</td>
<td>0.9095</td>
<td>0.94351</td>
<td>0.89406</td>
</tr>
<tr>
<td>EOOnHsYYE</td>
<td>0.14988</td>
<td>0.14933</td>
<td>0.14988</td>
<td>0.15008</td>
<td>0.14988</td>
</tr>
<tr>
<td>EarnPrem</td>
<td>1.6964</td>
<td>1.6964</td>
<td>1.6964</td>
<td>1.6964</td>
<td>1.6786</td>
</tr>
</tbody>
</table>

Empirical moments

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Solutions1</th>
<th>Solutions2</th>
<th>Solutions3</th>
<th>Solutions4</th>
<th>Solutions5</th>
</tr>
</thead>
<tbody>
<tr>
<td>lambda</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
<td>0.67</td>
</tr>
<tr>
<td>a1</td>
<td>0.2125</td>
<td>0.22292</td>
<td>0.2125</td>
<td>0.20208</td>
<td>0.2125</td>
</tr>
<tr>
<td>o2o1</td>
<td>0.78947</td>
<td>0.78947</td>
<td>0.83684</td>
<td>0.88421</td>
<td>0.88421</td>
</tr>
<tr>
<td>a2o1</td>
<td>0.20283</td>
<td>0.20232</td>
<td>0.20232</td>
<td>0.20191</td>
<td>0.19987</td>
</tr>
<tr>
<td>o3o1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.2875e-05</td>
<td>0</td>
</tr>
<tr>
<td>sigma2y</td>
<td>0.74359</td>
<td>0.74359</td>
<td>0.74359</td>
<td>0.74359</td>
<td>0.74359</td>
</tr>
<tr>
<td>sigma2b</td>
<td>7.4177</td>
<td>6.9779</td>
<td>6.5751</td>
<td>7.0937</td>
<td>7.2691</td>
</tr>
<tr>
<td>a3</td>
<td>0.19664</td>
<td>0.19683</td>
<td>0.19638</td>
<td>0.19673</td>
<td>0.19664</td>
</tr>
<tr>
<td>tn1</td>
<td>0.19486</td>
<td>0.19489</td>
<td>0.19486</td>
<td>0.19484</td>
<td>0.19486</td>
</tr>
<tr>
<td>tm1</td>
<td>0.071006</td>
<td>0.066014</td>
<td>0.071006</td>
<td>0.072364</td>
<td>0.071006</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Criterion1</th>
<th>Criterion2</th>
<th>Criterion3</th>
<th>Criterion4</th>
<th>Criterion5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of the criterion</td>
<td>0.051785</td>
<td>0.065799</td>
<td>0.045447</td>
<td>0.050882</td>
<td>0.051577</td>
</tr>
</tbody>
</table>

112
D Counterfactuals: Details

In this appendix I provide more details regarding the derivation of the counterfactuals.

D.1 Policy Counterfactuals (INCOMPLETE)

In this appendix, I display the results for both M1 and M2. This allows to see the extent to which the quantitative results depend on the specification for the borrowing constraint and for the outside option.

D.2 Counterfactuals with respect to the returns to education, $\lambda$

Counterfactual 1 in M2  When changing $\lambda$ in the first counterfactual, one changes not only the returns to education ("the slope") but also "the intercept" of the production function. Although this has no meaningful implications in M1 for the moments we are interested in—the intercept affects only averages of logs, not the measures of inequality or of intergenerational mobility—it does matter in M2 because of the enrollment choice. Lowering $\lambda$ and increasing $\mu$ does significantly lower GDP and results in a very sharp decrease in the enrollment rate for a given $q$, to a completely counterfactual level. To address this issue, I consider three possible assumptions.

1. Keep $q$ unchanged.
2. Adjusting $q$ to target an enrollment rate of 50% which corresponds to its level in 1980.
3. Adjusting $q$ to target an enrollment rate of 70% which corresponds to its level in the original calibration.

The results for the main variables of interest are reported in table (A2):

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Enrollment</th>
<th>s</th>
<th>Gini y_m</th>
<th>IGE</th>
<th>Endog. Amplif. of Gini y_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda = .67$</td>
<td>70%</td>
<td>1.3%</td>
<td>.45</td>
<td>.5</td>
<td>-</td>
</tr>
<tr>
<td>1</td>
<td>5%</td>
<td>-92%</td>
<td>-23%</td>
<td>-14%</td>
<td>+ 35%</td>
</tr>
<tr>
<td>2</td>
<td>50%</td>
<td>-36%</td>
<td>-17%</td>
<td>-2.8%</td>
<td>+ 6.3%</td>
</tr>
<tr>
<td>3</td>
<td>70%</td>
<td>-18%</td>
<td>-16%</td>
<td>-.6%</td>
<td>+ 2.4%</td>
</tr>
</tbody>
</table>

Legend: The three assumptions are as follows: 1- Keep $q$ unchanged. 2- Adjusting $q$ to target an enrollment rate of 50% which corresponds to its level in 1980. 3- Adjusting $q$ to target an enrollment rate of 70% which corresponds to its level in the original calibration.

Fixing Tuition and Spending within Counterfactual 1: Details  In this appendix, I document how I run our counterfactual illustrating and quantifying how endogenous reaction of colleges has amplified the initial increase in inequality and decreased social mobility.
Start from the price schedule in equilibrium faced by HH in 1980:

\[ e(q, h_s, y) = h_s^{-\tau_m}y^{\tau_n} \frac{T_e}{(1 + a_h)} \left( \frac{p_I}{(1 + a_u)T_u} q \frac{1}{\tau_1} \left( \frac{y}{\kappa_2} \right)^{\frac{\varepsilon_3}{\varepsilon_1}} \right)^{\frac{1}{1 - \tau_u}} \]

Consider the marginal distributions of high school ability and income, denoted \( F_{h_s,1980}(h_s), F_{y,1980}(y) \). As well as the distribution of colleges quality \( F_{q,1980}(q) \). Denote \( F_{h_s,1980}^{-1}(\cdot), F_{y,1980}^{-1}(\cdot), F_{q,1980}^{-1}(\cdot) \) the respective quantile function. The object I fix is the following function

\[ e(rk_q, rk_{h_s}, rk_y) = C \left( F_{q,1980}^{-1}(rk_q) \right)^{\frac{1}{\varepsilon_1(1 - \tau_u)}} \left( F_{h_s,1980}^{-1}(rk_{h_s}) \right)^{-\frac{\varepsilon_2}{\varepsilon_1(1 - \tau_u)}} - \tau_m \left( F_{y,1980}^{-1}(rk_y) \right)^{\frac{\varepsilon_3}{\varepsilon_1(1 - \tau_u)}} \]

where \( rk \) denotes rank in the distribution and

\[ C = \frac{T_e}{(1 + a_h)} \left( \frac{p_I}{(1 + a_u)T_u} \left( \frac{1}{\kappa_2} \right)^{\frac{\varepsilon_3}{\varepsilon_1}} \right)^{\frac{1}{1 - \tau_u}} \]

as well as \( \varepsilon_1, \varepsilon_2, \varepsilon_3, \tau_u \) are fixed at their 1980 value.

Families buy a college rank, not a given quality and the tuition depends only on their ranks in the distribution. It is not possible to fix the schedule in terms of quality, because this is an endogenous object that depends on the peer effects, hence on the composition of the college. The rank is a well defined object. As for the rank of income and abilities, I haven’t found a good enough reason which this is better than in levels. I also show the counterfactual in terms of level in the next section. One good reason to fix in rank rather than in level is that, if I leave tuition as a function of the level of income and abilities, it implies that an increase in inequalities, will increase tuition for the top 1 percent, because their income is now higher. But my goal is to consider a world in which rich people do not face an increase in tuition...

From the constant saving rate across all households at a given time \( t \), I get the following relationship:

\[ sy = C \left( F_{q,1980}^{-1}(rk_q) \right)^{\frac{1}{\varepsilon_1(1 - \tau_u)}} \left( F_{h_s,1980}^{-1}(rk_{h_s}) \right)^{-\frac{\varepsilon_2}{\varepsilon_1(1 - \tau_u)}} - \tau_m \left( F_{y,1980}^{-1}(rk_y) \right)^{\frac{\varepsilon_3}{\varepsilon_1(1 - \tau_u)}} + \tau_n \]

\[ rk_q = F_{q,1980} \left[ \left( \frac{sy}{C} \right)^{\varepsilon_1(1 - \tau_u)} \left( F_{h_s,1980}(rk_{h_s}) \right)^{\varepsilon_2 + \tau_m \varepsilon_1(1 - \tau_u)} \left( F_{y,1980}(rk_y) \right)^{-\varepsilon_3(1 - \tau_u) \varepsilon_1 \tau_n} \right] \]

In the counter-factual I fix the saving rate to the one in the final steady-state of the non-counterfactual economy, the idea being to imagine an economy in which I fix the saving and labor choice of individual and simply change the price schedule, I want to identify the partial/marginal effect of colleges’ reactions to market forces keeping the policy rules of families constant.

The next step consists in mapping \( rk_q \) to an actual level of quality. Here I do two experiments. One in which, the level of quality remains constant at what it was in 1980, and one in which I allow
for change in peer effects but not in spending. For this, I create a grid on \( r_{k,q} = [0, 1] \), put people in bins according to their choice of rank, and take the geometric average. This gives us a mapping \( \tilde{\theta}(r_{k,q}) : [0, 1] \to \mathbb{R}^+ \). I then combine it with our fixed mapping of investment per student to get quality \( q \).
Figure A1: Quality, $\theta$ and $I$ along the Quality Ladder in Counterfactuals

Legend: the blue lines represent qualities (top-left panel), spending per student (bottom-left panel) and the average ability (top-right panel) in the original steady-state in 1980, the red in the final steady-state in the beginning of the 2000s, and the green in the counterfactual, for each percentile of the quality distribution where tuition schedule and spending per students have been fixed at what they were in the original steady-state. The bottom-right panel represents the Gini coefficient of human capital in the 3 different allocations, on the left for the original steady-state, on the right for the counterfactual (green) and final steady-state (red).
D.3 Discussion on joint maximization of $\tau_n$ and $\tau_u$

When simultaneously maximizing with respect to $\tau_n$ and $\tau_u$, one obtains the surprising result that $\tau_u$ should be .62 while the optimal $\tau_n$ should be 0. This is shown on figure A2. The optimal $\tau_u$ is therefore not very different from the case where $\tau_n$ is held fixed. Since $\tau_n$ and $\tau_u$ are partial substitute, it is not surprising that $\tau_n$ should be lower when allowed to maximize with respect to both at the same time. What is more surprising is that the optimal $\tau_u$ is lower than in the constrained case when $\tau_n$ is higher. This is because $\tau_u$ trades off two forces: on the one hand it redistribute resources, lessen inequality across colleges which decreases inequality; and decreases the misallocation coming from the income-sorting channel. But on the other hand, the planner would also like to direct resources to the smartest students, to increase output, which incentivizes them to decrease $\tau_u$. When $\tau_n$ is small, the latter effect is strongest, while it is weaker when $\tau_n$ is high. This explains why $\tau_u$ is higher when $\tau_n$ is equal to 0 than when it is kept equal to .195.

The second surprising result is that the optimal $\tau_u$ should decrease when $\lambda$ increases. This can be understood along the same lines: when the planner already cares a lot about matching resources with the smartest students, it should care even more when human capital becomes more productive (higher $\lambda$), therefore $\tau_u$ should decrease when $\lambda$ increases.

![Figure A2: Optimal $\tau_n$ and $\tau_u$](image)

D.4 Discussions on $\tau_m$

Maximizing simultaneously with respect to merit-based, need-based aid and transfers to colleges gives rise to the following unrealistic solution: college subsidies should relatively and realistically high, but both need-based and merit-based should be extremely high, at levels that are too high and should be seen as mathematical curiosity telling us something about the forces at play in the model more than anything else. Maximizing with respect to need and merit-based aid keeping subsidies to college subsidies constant leads to the same issue. For very low, and counterfactual, level of $\tau_u$, the optimal $\tau_n$ and $\tau_m$ are reasonable.
The issue is that whenever \( \tau_u \) is sufficiently high, the mean of log human capital \( m \) and the variance \( \Sigma_h^2 \) becomes both increasing in \( \tau_m \) such that the optimal is very high, potentially infinite. This is because increasing \( \tau_m \) takes full advantage of the complementarities with the abilities of children, and the implied inequality can be addressed through an increase in \( \tau_n \) or \( \tau_u \). In figure A3 I plot the optimal \( \tau_n, \tau_m \) keeping \( \tau_u = .35 \) on the left panel and the joint optimal \( \tau_n, \tau_m, \tau_u \) on the right panel, forcing \( \tau_n \) and \( \tau_u \) to be between 0 and 1 and \( \tau_m \) to be smaller than 2.8. It appears clearly that whatever values of \( \lambda \) and \( \sigma \), \( \tau_m \) binds at 2.8. To offset the increase in inequality that this policy implies, the optimal \( \tau_n \) should be very high (left panel). When the government maximizes with respect to \( \tau_u \) as well, one obtains the same result as described in the previous section D.3, a very high \( \tau_m \), a high \( \tau_u \) but no need-based policy.

![Figure A3](image)

(a) Optimal \( \tau_n, \tau_m \) for \( \tau_u = .35 \)

(b) Optimal \( \tau_u, \tau_n, \tau_m \) (joint)

Figure A3: Optimal \( \tau_n, \tau_u \) and \( \tau_m \) as a function of the preference for redistribution parameter
E  Additional Figures

Figure A4: Before-Aid Tuition Fees (real)

Legend: In 1980 constant Dollars. The series has been normalized to 1 in 1980. The series for tuition fees and CPI is from the Bureau of Labor Statistics.