

# On the Intergenerational Transmission of Economic Status

Sang Yoon (Tim) Lee

Department of Economics

University of Mannheim

Ananth Seshadri

Department of Economics

University of Wisconsin-Madison\*

January 2014

## Abstract

We present a model of human capital investment within and across generations, with incomplete markets and government transfer programs. Our model combines a fairly standard life-cycle model of human capital with an intergenerational model à la [Becker and Tomes \(1986\)](#). The human capital technology features multiple stages of investment during childhood, a college decision, and on-the-job accumulation. The model can jointly explain a wide range of intergenerational relationships, such as the intergenerational elasticities (IGE) of lifetime earnings, education, poverty and wealth, while remaining empirically consistent with cross-sectional inequality. Unlike previous models, intergenerationally constrained families have similar IGE's as the unconstrained families. Exogenous ability transmission has a modest impact on the IGE in partial equilibrium, but this effect disappears in general equilibrium. On the other hand, investment in children and parents' human capital have a large impact on the equilibrium IGE. Education subsidies and progressive taxation can significantly reduce the persistence in economic status across generations.

---

\*Email: [sylee.tim@uni-mannheim.de](mailto:sylee.tim@uni-mannheim.de) and [aseshadri@ssc.wisc.edu](mailto:aseshadri@ssc.wisc.edu), respectively. We are very grateful for constructive comments by Gary Becker, Steven Durlauf, Arthur Goldberger, John Knowles, Nicolas Schutz, Chris Taber, seminar participants at the 2012 MIE conference, the 2012 Vienna Macro Workshop, the University of Mannheim, and the University of Chicago. We also thank James Heckman and four anonymous referees for their detailed suggestions. All remaining errors are our own.

# 1. Introduction

The intergenerational elasticity (IGE) of parents' lifetime earnings with respect to children's lifetime earnings is 0.6 in the United States.<sup>12</sup> Wealth, consumption, schooling and poverty are also persistent across generations. Understanding this degree of persistence has been a longstanding goal of social science. Disentangling the respective contributions of the transmission of innate abilities (nature), family background (nurture), or economic policy in generating persistence has also been the subject of much discussion and debate. However, most such analyses are statistical models, and economic models as of yet have come short of rationalizing various patterns we see in the data. The goal of this paper is to provide such a framework to better understand these complex empirical relationships, and to analyze the impact of various policies.

The theoretical work in this literature gained traction beginning with the seminal work of [Becker and Tomes \(1979, 1986\)](#).<sup>3</sup> They laid out a simple but powerful two-period equilibrium model to derive implications for the intergenerational transmission of lifetime earnings and wealth.<sup>4</sup> While the Becker-Tomes framework is used widely in the literature, it has been met with some empirical skepticism, most notably by [Goldberger \(1989\)](#). He argued that an economic approach represented little value added relative to mechanical approaches, i.e. those that do not rely on household optimizing behavior. [Mulligan \(1999\)](#) tests some key empirical implications of the Becker-Tomes model and finds limited support. Most importantly, he shows that the inability of a parent to borrow against the future income of his child (hereafter "intergenerational borrowing constraint"), a key feature of [Becker and Tomes \(1986\)](#), seems not to matter, and even if it does, is empirically irrelevant. [Han and Mulligan \(2001\)](#) further argue that heterogeneous abilities and such borrowing constraints are indistinguishable.

One reason for this apparent invalidation of the Becker-Tomes model is that entire lifetimes are condensed into two periods. In these models, parents make a once-and-for-all investment in children who grow up to earn a once-and-for-all income. Human and physical capital investments

---

<sup>1</sup>There is a large empirical literature that attempts to measure the magnitude of the IGE of lifetime earnings, and the number we cite here from [Mazumder \(2005\)](#) is different from the ubiquitously used number of 0.4 in the literature largely due to [Solon \(1992\)](#). [Mazumder \(2005\)](#) finds that when increasing the horizon of measured earnings, the IGE between parent and child increase, up to .6 at the 15 year horizon. This is not so unusual—many other authors, including [Solon \(1992\)](#), find a number larger than 0.4 using different specifications as well.

<sup>2</sup>Throughout the text, we will refer to the IGE of lifetime earnings as simply "the IGE." For other IGE's we explain to which IGE we are referring to.

<sup>3</sup>[Loury \(1981\)](#) was a similar model in a dynastic setting with borrowing constraints.

<sup>4</sup>There are several other important papers in the theoretical literature that focus on intergenerational persistence. [Benabou \(1993\)](#) and [Durlauf \(1996\)](#) present models of segregation, [Galor and Zeira \(1993\)](#) focus on poverty traps while [Banerjee and Newman \(1993\)](#) present a model featuring mobility traps.

(bequests) are decided upon simultaneously. The only moments predicted by the model are childhood investments, lifetime earnings and bequests, with each period comprising as many as 20-30 years. Hence, when bringing the model to the data, empirical moments are averaged over extensive time periods, with no consideration for less than perfectly substitutable child investments across periods.<sup>5</sup> Any decisions over the life-cycle other than child investments are ignored, which could affect bequest decisions that happen much later in life. Furthermore, such data spanning the entire lifetimes of multiple generations are scant at best, making it difficult to validate the key mechanisms of the model.

The main purpose of this paper is to fill this gap by using a richer model that overcomes these shortcomings. Specifically, we combine a fairly standard life-cycle model of human capital accumulation à la [Ben-Porath \(1967\)](#)<sup>6</sup> with an intergenerational transmission mechanism à la [Becker and Tomes \(1986\)](#). While we are certainly not the first to present a quantitative model, prior theoretical formulations were too simplistic to directly compare with available data. Many studies either assume away important model elements emphasized by Becker-Tomes (e.g. [Restuccia and Urrutia \(2004\)](#) abstract away from asset accumulation) while others ignore the empirical skepticism raised against it. In our model, bequests occur long after the parent completes child investments. Hence, consistent with the empirical findings of [Mulligan \(1999\)](#), whether or not the parent leaves bequests is not much of an indicator of differential child investments. In contrast, whether or not the parent's wealth level at a younger age is above the median predicts child investments, consistent with the findings in [Mazumder \(2005\)](#). The inclusion of the parent's own human capital accumulation over the life-cycle, which happens simultaneously with child investments but prior to bequest decisions, is central to generating this result.

Parents' anticipation of the life-cycle decisions their children make as grown-ups also affect child investments. In most models that rely on Becker-Tomes, parental states are fixed exogenously, and cross-sectional inequality among children is pre-determined once they become adults. That is, there are only two generations in the model and they do not take into account the effect that subsequent members of the dynasty have on current decisions. We cast our model in an overlapping generations framework with infinitely-lived dynasties who are altruistic.<sup>7</sup> The initial

---

<sup>5</sup>This criticism also forms part of the bases for studies such as [Cunha and Heckman \(2007\)](#); [Cunha et al. \(2010\)](#); [Caucutt and Lochner \(2012\)](#) who argue that there is strong complementarity across periods.

<sup>6</sup>We also explore sensitivity by redoing the analysis using a learning-by-doing specification.

<sup>7</sup>Becker-Tomes as well as many of the ensuing papers used two period models in which parents care about their own consumption and the income of their children. A dynastic formulation in which parents care directly about descendants' utility is more parsimonious.

distribution over the next generation's human capital and assets are determined by investments in children and bequest decisions, while the terminal distributions of human capital and assets over the parent generation are determined by own human capital investment and life-cycle savings decisions. Since parents make these decisions, we require individual behavior to be consistent with observing these distributions in a stationary equilibrium. In particular, rationalizing the distribution of human capital jointly with assets is critical, since a framework that can account for earnings persistence but inconsistent with wealth inequality and its transmission would be unconvincing.

A large literature in macroeconomics and public finance takes the earnings process as given and attempts to match wealth inequality (e.g. [Aiyagari \(1994\)](#); [Hubbard et al. \(1995\)](#); [Castañeda et al. \(2003\)](#)). In particular, [Castañeda et al. \(2003\)](#) demonstrate that accounting for intergenerational relationships is crucial to account for cross-sectional inequality, but take these relationships as exogenous. A separate line of work takes initial conditions around age 20 as given and attempts to match the distribution of earnings (e.g. [Huggett et al. \(2011\)](#)). In most of these papers, a joint distribution of initial human capital and assets is estimated to match life-cycle earnings dynamics and distributions later in life. They find that small differences in initial conditions can lead to large differences in earnings and wealth over the life-cycle. In contrast, we take neither the earnings process nor initial conditions as given, and require the intergenerational mechanism of our model to endogenously generate empirically valid distributions of earnings and wealth within and across generations. Hence, not only does the life-cycle component help explain intergenerational data, but the intergenerational investments also help explain cross-sectional data.

Our human capital formation process<sup>8</sup> is rather complex, but remains a parsimonious combination of features that are separately well understood. [Cunha et al. \(2010\)](#) show that there is complementarity between early and later child investments, while [Del Boca et al. \(2013\)](#) show that parental time is an important input in the production of a child's human capital, particularly when the child is young. We capture these relationships by assuming complementarity between early investments in the form of parental time, and later investments in the form of goods. We model the final stage of human capital formation as a college enrollment decision coupled with skill acquisition at college age, as much of inequality can be attributed to differences in educational attainment. The college choice also allows us separate cross-sectional inequality into differential skill prices and different levels of skill, a focus of many recent studies. Post-schooling life-cycle wage growth follows a Ben-Porath human capital accumulation specification.

---

<sup>8</sup> By "formation," we mean skill acquisition prior to entering the labor market.

Combining childhood, life-cycle, and intergenerational elements is not trivial, especially because the child's state variables affect the parent's decision problem. Adding to the complexity is the asset and labor market equilibrium with two skills. Nonetheless, it is precisely these two features (multiple time periods and equilibrium) that make comparison with empirical counterparts feasible. In the data, typically, we observe parent and child variables at different ages, which we can compare to the end and beginning of life-cycles in our model. The stationary equilibrium setup allows us to discipline parent's behavior early in life, or children's behavior later in life, which is also not simultaneously observed in the data.

To the best of our knowledge, ours is the first attempt at incorporating an endogenous human capital accumulation process into an intergenerational setting where ability is transmitted across generations and pre-labor market initial conditions are determined by endogenous parental investments. In addition, we also model most of the conventional forms of government intervention. The data we observe comes from an economy in which the government taxes income progressively to subsidize education, fund social security and assist low income households through welfare payments. Our results demonstrate that the interactions between the parent's investment in his own human capital and his child's human capital are important not only for explaining empirical facts, such as the intergenerational persistence of earnings, wealth, and educational attainment as well as life-cycle earnings profiles and cross-sectional inequality, but also for counterfactual policy analyses. We bolster the robustness of our model by showing its empirical consistency with several non-targeted moments, inspection of an influence matrix à la [Gentzkow and Shapiro \(2013\)](#), and equilibrium responses to changes in moments.

In addition to rationalizing apparently conflicting findings from [Mulligan \(1999\)](#) and [Mazumder \(2005\)](#), we find that augmenting the intergenerational component into a life-cycle model helps improve the performance of the latter in explaining the relationship between lifetime earnings and retirement wealth. While [Hendricks \(2007\)](#) shows that the standard life-cycle model over-predicts the correlation between lifetime earnings and retirement wealth, we show that life-cycle and intergenerational human capital investments, together with intergenerational transfers, can account for these discrepancies. In fact, the weak correlation between lifetime earnings and retirement wealth, together the weak correlation between bequests by parents and child's lifetime earnings, is the same mechanism that generates the Mulligan-Mazumder result.

The removal of education subsidies or progressive taxation (move to a flat tax regime) has the largest adverse impact on the IGE. Relatedly, our result that subsistence on lumpsum trans-

fers is persistent across generations is in line with studies that confirm the persistence of AFDC participation, e.g. [Gottschalk \(1990\)](#). At the same time, however, it turns out that the education transfers, which are also distributed lumpsum (but tied to education), are large compared to the innate abilities of the children of subsistent families in our model simulations.

Finally, our model is related to a large literature focusing on disentangling the effect of nature and nurture on educational attainment.<sup>9</sup> These studies typically exploit natural experiments by contrasting identical and fraternal twins, their offspring, relatives and siblings, or biological and adopted children, and most conclude that more than half of observed differences in educational attainment can be attributed to differences in genes. While the literature has made great strides over the past few decades, the empirical approach to disentangling nature from nurture is not without its critics. It may well be unrealistic to assume that twins differ in terms of schooling but not in terms of any other characteristic or experience that may affect the education of their children, e.g. [Griliches \(1979\)](#). The linearity assumptions employed in the nature-nurture decomposition have also been criticized.<sup>10</sup> Furthermore, while differences in educational attainment no doubt have a large impact on differences in human capital, we believe a better measure of human capital is observed earnings. By only focusing exclusively on educational attainment, one can measure the relative effects of ability transmission and environment, but not the efficacy of education itself.

Our framework allows for nature and nurture to interact in a non-linear fashion. And our structural model allows us to shut down the ability transmission process and run counterfactual experiments.<sup>11</sup> We find that indeed, removing ability transmission drops the IGE by up to 10 log points. But our results indicate that human capital formation parameters in childhood have a larger impact on the IGE, and that imperfect capital markets and government policies also play a large role.

While the benchmark adulthood human capital accumulation is modeled as Ben-Porath, in a sensitivity analysis we also consider a simple learning-by-doing (LBD) formulation. Most of our results still go through qualitatively, but the quantitative results are not as strong. We argue that what is more important than how human capital accumulation happens over the life-cycle is

---

<sup>9</sup>Examples are [Behrman et al. \(1977\)](#); [Plug and Vijverberg \(2003\)](#).

<sup>10</sup>According to [Heckman \(2008\)](#), “genes and environment cannot be meaningfully parsed by traditional linear models that assign unique variances to each component.”

<sup>11</sup>As a caveat, we should emphasize that we do not use test scores to measure ability. Ability in our model is different from measured ability, which itself is a combination of nature and nurture. Test scores are affected not only by genes but also the environment. If we wanted to use test scores, we would simply incorporate a test score function that maps model ability to measured test scores, which in turn maps into measured earnings, resulting in the additional parameters and moments. Instead, we only assume a relationship between model ability and measured earnings.

whether or not there is a trade-off between investments in the human capital of oneself and one's child. Our LBD specification still includes such a trade-off in early adulthood, although weaker than our Ben-Porath benchmark, leading to such results.

We also conduct a number of additional sensitivity analyses on model parameters, in particular the elasticity of substitution between time and good inputs during childhood, and the demand elasticity of substitution between high school and college labor. The former has a large, non-monotonic impact on the IGE, consistently with [Cunha and Heckman \(2007\)](#), while the latter has little impact (except at extreme values). Hence, it is early rather than later human capital accumulation that matters for intergenerational persistence.

The rest of the paper is organized as follows. Section 2 lays out our model, and section 3 explains the calibration strategy. In section 4, we present our results, and also discuss the relevance of our model for understanding retirement wealth and welfare subsidies. Section 5 includes robustness checks and sensitivity analyses. Section 6 concludes. The appendix details the numerical implementation and a simpler version of the model from which we derive implications that help us identify the model and analyze government policy effects.

## 2. Model

A period is 6 years. There is a unit measure of households whose members live through a life-cycle, where an individual is born at age 0. So each individual goes through 13 periods of life ( $j = 0, \dots, 12$ ) or 4 stages:

1. Child/College:  $j = 0, 1, 2, 3$ . Attached to parent, choice to enter the labor market or attain college education (human capital accumulation) at  $j = 3$ .
2. Young parent:  $j = 4, 5, 6, 7$ . Independence and bears child at  $j = 4$ , and continues own human capital accumulation. Simultaneously makes early childhood and subsequent investments in child. Joint college decision for child at  $j = 7$ .
3. Grandparent:  $j = 8, 9, 10$ . Continues to accumulate own human capital, and saves for retirement and bequests to child.
4. Retired:  $j = 11, 12$ . Consumes social security payouts, dies *after*  $j = 12$ .

Throughout the analysis, primes will denote next generation values.

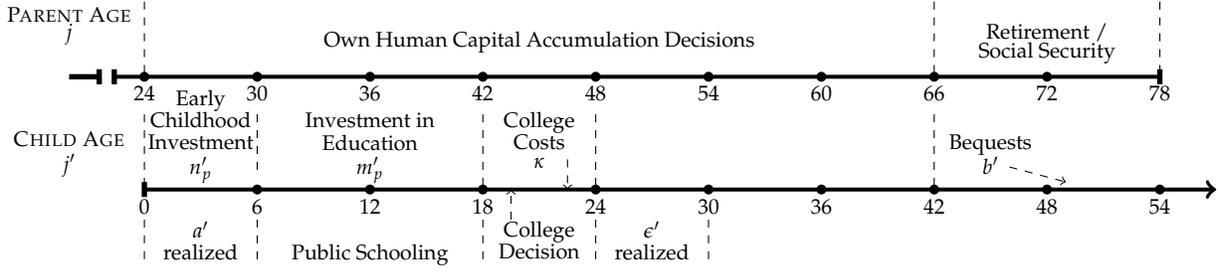


Figure 1: Sequence of Events

Fertility is exogenous, and one parent is assumed to give birth to one child at  $j = 4$ , so  $j = j' + 4$  is fixed throughout the lifetimes of a parent-child pair.<sup>12</sup> Until the child's full independence at  $j' = 4$ , the parent-child pair solve a Pareto problem. These decisions involve consumption, savings, investment in human capital, and college. The family decides whether the child should attend college when he becomes 18 years old ( $j' = 3$ ). The direct cost of college is fixed at  $\kappa$ , but there is also an implicit cost of the child's forgone earnings by delaying entry into the labor market. In retirement, the grandparent consumes his savings and social security benefits, and dies at age 78. He also sets aside a bequest  $b'$  for his child at  $j = 11$ , which is transferred to his child at  $j = 12$ . The sequence of events is depicted in Figure 1.

## 2.1 Human Capital Formation

A child is born with innate ability  $a'$ , which determines his proficiency at accumulating human capital. The child's ability is stochastic, and depends on the parent's ability. This will capture the "genetic" part of the IGE or the aspect of intergenerational transmission that we label "nature." An adult's earnings, or returns to human capital, are subject to a stochastic "market luck shock"  $\epsilon$  that depends on his own ability, and remains constant throughout his lifetime. This shock captures idiosyncratic risk associated with human capital accumulation, while the dependence on abilities captures both the potential persistence in luck across generations, and a human capital returns premium for high ability individuals. In addition to intergenerational transmission, we include three important aspects on child human capital formation that are well appreciated in the literature: i) parental time and good investments, ii) complementarity between inputs, and iii) the

<sup>12</sup>The only papers with investments in children that also feature endogenous fertility are [Aizer and Cunha \(2012\)](#) and [Gayle et al. \(2013\)](#). Both papers are partial equilibrium two generation models, with a fixed initial distribution for parents that leads to an endogenously determined distribution of children, which is estimated to the data. In our paper, dynasties live infinitely, capital and labor markets must clear, and government tax-transfer policies must satisfy a budget-balance condition. With endogenous fertility then, the steps that these decisions have to be aggregated over becomes exponentially costly.

child's own investments.

Formally, denoting a young parent's ability as  $a$ , at  $j = 4$  (age 24) he draws  $(a', \epsilon)$  from a joint probability distribution conditional on  $a$ :

$$(a', \epsilon) \sim F(a', \epsilon | a).$$

During the early childhood of his offspring (ages 0-5,  $j' = 0$ ), a young parent invests  $n_4$  units of time in his own human capital accumulation (on-the-job accumulation of human capital), and spends  $n_p$  units of time with his child. When the child goes to secondary school ( $j = 3$ ), he supports his education with  $m_p$  units of consumption goods.<sup>13</sup> The education production function (later childhood human capital formation) is CES in time and goods inputs:

$$h'_3 = \left[ \gamma_k^{\frac{1}{\phi}} (m_p + d)^{\frac{\phi-1}{\phi}} + (1 - \gamma_k)^{\frac{1}{\phi}} (n_p h_4)^{\frac{\phi-1}{\phi}} \right]^{\frac{\phi}{\phi-1}}$$

where  $d$  is a government subsidy, and  $\gamma_k$  captures the expenditure share of education. We have assumed that the time investment in the child,  $n_p$ , interacts with the human capital level of his parent,  $h_4$ . This is to capture the fact that more educated households (mothers) spend more time with their children (Behrman et al. (1999)).

The CES parameter  $\phi$  is the elasticity between parental time and good inputs, which can also be interpreted as complementarity between early and later childhood investments. Cunha et al. (2010), estimate this parameter in a model with multiple stages of child investments but only in terms of goods, while Del Boca et al. (2013) provide evidence that time is the dominant input when the child is young with its importance declining with age, and that the reverse is true for goods. From ages 18-23 ( $j = 3$ ), a child chooses whether or not to attend college, and invests own time  $n'_3$  into human capital production according to:

$$h'_4 = a' n_3'^{\gamma_{S'}} h_3'^{\gamma_p} = a' n_3'^{\gamma_{S'}} \left[ \gamma_k^{\frac{1}{\phi}} (m_p + d)^{\frac{\phi-1}{\phi}} + (1 - \gamma_k)^{\frac{1}{\phi}} (n_p h_4)^{\frac{\phi-1}{\phi}} \right]^{\frac{\phi \gamma_p}{\phi-1}},$$

where  $\gamma_{S'}$  depends on whether the child is in college ( $S' = 1$ ) or not ( $S' = 0$ ), and  $\gamma_p$  captures the returns to parental inputs. This composite function is intended to capture the fact the child's own time investment becomes more important in later years (Del Boca et al. (2012)).

<sup>13</sup>It is not the timing of time and good investments that is germane to our model, but the complementarity between them.

From age 24 until retirement, we assume a Ben-Porath type human capital accumulation function,

$$h_{j+1} = a(nh_j)^{\gamma_S} + h_j, \quad \text{for } j = 4, \dots, 10,$$

where  $h_{j+1}$  is human capital tomorrow and  $n_j$  is the time investment in own human capital accumulation. The parameter  $\gamma_S$  again depends on whether the adult is college educated ( $S = 1$ ) or not ( $S = 0$ ).

Heterogeneity across households comes from the ability and market shocks,  $(a, \epsilon)$ , respectively. Notice that there are four mechanisms through which human capital is transmitted across generations. First, the ability to *learn* is transmitted across generations, that is, the child's  $a'$  is drawn from a distribution that depends on the parent's  $a$ . Second, there is partial inheritance of market luck through the correlation of  $(a, \epsilon)$ . This can be viewed as heritable disabilities that affect the ability to *earn*.<sup>14</sup> Third, the resources devoted to child human capital formation,  $m_p$ , will be a function of the parent's resources. If capital markets are complete, the investment will be independent of parental resources. However, with capital market imperfections (which we assume), parental resources will matter. Finally, we assume that a higher human capital parent (large  $h$ ) is better at transmitting human capital beyond his ability to pay for more resources per unit of  $n_p$ . We regard the first and second mechanisms as nature, and the third and fourth as nurture.

The interaction between child and adult human capital accumulation is what differentiates us the most from other models. Overlapping generation models that focus on childhood typically take adulthood wage processes as exogenous, e.g. [Caucutt and Lochner \(2012\)](#); [Del Boca et al. \(2013\)](#). The focus of these papers were not on the IGE, but our analysis in section 5, where we consider a LBD formulation in place of the Ben-Porath specification, suggests that including an *earnings motive* (invest in own human capital) is important to generate our results, in particular the result that intergenerationally constrained families have similar IGE's to unconstrained ones. In other words, it is not so much that the parent needs to invest goods in the child that is important. He must also sacrifice some of his own *time*, either in the form of working time or own human capital formation (learning time). A pure *savings motive* (life-cycle physical capital accumulation) is not enough.

---

<sup>14</sup>In general,  $\epsilon' \sim F(\epsilon'|a', \epsilon)$ , but we ignore the direct dependence of  $\epsilon'$  on  $\epsilon$  because i) it is not clear what a persistence of "luck" means conceptually, and ii) such persistence would not be identifiable without multiple generations of data. The current formulation leaves only one pure channel for genes.

The assumption that the human capital level of the parent affects both the parent *and* the child is also important. In the data, wages (or *measured* ability) may be thought of as reflecting both *innate* ability as well as acquired human capital. Viewed this way, measured ability will be affected by unobservable parental human capital in addition to observables such as parental education or earnings. So we need a model when evaluating the effect of innate ability on the IGE. Our results suggest that this channel is indeed important, somewhat in contrast to empirical studies that emphasize innate abilities, such as [Plug and Vijverberg \(2003\)](#). In particular, the non-linear human capital formation process plays a role. For example, we argue in section 5 that in a counterfactual world with stronger complementarity between parental time and goods investments, innate abilities play no role at all.

## 2.2 Firm and Government

We assume that there is a government that levies taxes. The proceeds are partially used for a lumpsum transfer  $g$  and an education subsidy  $d$ . The government also runs a balanced-budget social security program.<sup>15</sup>

We assume that earnings are taxed progressively. Denote by  $e_j$  the earnings in any working period  $j$ . Then,  $e_j$  is taxed at a progressive rate  $\tau_e(e_j)$ , and also subject to a flat rate payroll tax  $\tau_s$  that is used to fund the social security benefits. The final after tax, after subsidy net earnings of a working adult is

$$f(e_j) = [1 - \tau_s - \tau_e(e_j)] e_j.$$

Capital income is taxed at a flat rate  $\tau_k$ , so we can define  $\tilde{r} = [1 + (1 - \tau_k)r]^6 - 1$ , the 6 year compounded effective interest rate faced by a household, where  $r$  is the pretax annual interest rate.

Social security benefits  $p$  are modeled as a function of  $\bar{e}$ , the average lifetime earnings from ages 24-65 ( $j = 5$  to 10).<sup>16</sup> We model it as an affine function:

$$p = p(\bar{e}) = p_0 + p_1 \bar{e}, \quad \bar{e} = \frac{\sum_{j=5}^{10} e_j}{6}$$

where  $(p_0, p_1)$  are parameters governing the social security regime. Social security benefits are not

<sup>15</sup>The rest we assume is thrown in the ocean.

<sup>16</sup> Social security benefits in the U.S. are based on the 35 years of an individual's highest earnings.

subject to any tax. The government runs a balanced budget on these benefits by financing them with the payroll tax revenue.

We assume that the initial distribution  $F_0$  for  $(a', a, \epsilon)$  is the stationary distribution of  $F(a', \epsilon|a)$ . The steady state equilibrium and government budget balance conditions are explained in detail below.

### 2.3 Household's Problem

We assume that an adult faces natural borrowing constraints for all savings and education decisions (but not bequests). Given the timing of shocks, we can split the life-cycle into two sub-periods: young and old. For any  $j$ , it is useful to define the state vector

$$\mathbf{x}_j = [S, a, \epsilon; h_j],$$

i.e. the educational attainment (college or not,  $S \in \{1, 0\}$ ), ability, market luck shock, and current human capital level of an individual in period  $j$  of life. Let  $(y, o) = (4, 8)$ , i.e., the first periods in the young parent and grandparent stages, respectively. We present the recursive representation of the problem.

**Young parent ( $y = 4$ )** This is the most important part of the decision problem. Note that there are no realizations of shocks between  $j = 4$  to  $7$ , which effectively collapses the college choice into the young parent's problem. Symmetrically, there are no realizations of shocks for the grandparent throughout the same period. The timing of the model is such that a young parent anticipates the bequests  $b$  that will be decided by his now old parent (the grandparent) when he becomes 42, or  $j = 7$ . Hence, we are assuming that first, the grandparent makes a bequest decision  $b$ , followed by a joint decision between the young parent and child on whether the child should go to college,  $S' \in \{0, 1\}$ .

Suppose the family chooses education level  $S'$  for the child. The young parent recognizes that any earnings made in the previous period and today ( $j = 3, 4$ ) will not matter for social security. Hence his state is his own state vector  $\mathbf{x}_y$ , the anticipated bequests  $b$ , and the ability of his child, born in the beginning of this period,  $a'$ . Given this state, he chooses consumption streams for himself and his child,  $c = \{c_{y+l}, c'_{y+l-4}\}_{l=0}^2$ , own human capital investments and savings  $(n, s) = \{n_{y+l}, s_{y+l+1}\}_{l=0}^2$ , and investments in his child's human capital formation,  $(n_p, m_p)$ . Furthermore,

the family jointly decides whether the child should attend college at  $j = 7$ , and how much the child should contribute toward joint family earnings versus the child's human capital accumulation at age  $j' = 3$ . The value function is

$$W_y(S', a'; \mathbf{x}_y; b) = \max_{c, s, n, n_p, m_p} \left\{ \sum_{l=0}^3 \beta^l \left[ u(c_{y+l}) + \theta u(c'_{y+l-4}) \right] + \beta^4 \int_{a'', \epsilon'} V_o(a''; \mathbf{x}'_y; \mathbf{x}_o, \bar{e}_o; s_o + b) dF(a'', \epsilon' | a') \right\}$$

subject to consumption-savings decisions

$$\begin{aligned} c_{y+l} + c'_{y+l-4} + s_{y+l+1} &= f(e_{y+l}) + (1 + \tilde{r})s_{y+l} + q_A \cdot g, & \text{for } l = 0, 1, \\ c_{y+2} + c'_{y-2} + s_{y+2+1} + m_p &= f(e_{y+2}) + (1 + \tilde{r})s_{y+2} + q_A \cdot g, \\ c_{y+3} + c'_{y-1} + s_{y+4} &= f(e_{y+3}) + (1 + \tilde{r})s_{y+3} + f(e'_{y-1}) - S' \cdot \kappa + 2g \end{aligned}$$

and human capital accumulation decisions

$$\begin{aligned} e_y &= w_S h_y \epsilon (1 - n_y - n_p) \quad \text{and} \quad n_y \in [0, 1], \quad n_p \in [0, n_y], \\ e_{y+l} &= w_S h_{y+l} \epsilon (1 - n_{y+l}) \quad \text{and} \quad n_{y+l} \in [0, 1], \quad \text{for } l = 1, 2, 3, \\ e'_{y-1} &= w_S h'_{y-1} (1 - n'_{y-1}) \quad \text{and} \quad n'_{y-1} \in [0, 1], \\ \bar{e}_o &= \frac{e_{y+1} + e_{y+2} + e_{y+3}}{3}. \end{aligned}$$

where his own and child's human capital production, and earnings net of taxes  $f(\cdot)$ , are specified above. The transfer  $g$  is multiplied by an adult equivalence scale  $q_A$  when the child is young, while a college aged child ( $j' = 3$ ) receives the full subsidy. The parent and child are effectively solving a Pareto problem, where the child has a Pareto weight of  $\frac{\theta}{1+\theta}$ , as determined by the parent's altruism.<sup>17</sup> The parameters  $(\gamma_p, \phi)$  play an important role, as they capture how much parental investments matter for education. If  $\gamma_p$  is high, the parent may need to sacrifice own human capital accumulation to educate the child, and even more so when time and goods are less substitutable ( $\phi \rightarrow 0$ ). Given this, the household chooses  $S'$ :

$$V_y(a'; \mathbf{x}_y; b) = \max_{S' \in \{0,1\}} \{W_y(S', a'; \mathbf{x}_y; b)\}.$$

<sup>17</sup>Additionally, we could also assume that the parent makes college transfers at  $j = 3$ , and the child saves for independence,  $s_{y-1}$ , but the timing of the problem would imply both are zero.

We assume a fixed cost for college, as our focus is more on parental influences on earlier stages of education rather than college.<sup>18</sup>

**Grandparent** ( $o = 8$ ) From the young parent's problem, it is apparent that the grandparent's states are the ability of his grandchild  $a''$ , the state of the young parent,  $\mathbf{x}'_y$ , and his own state  $(\mathbf{x}_o, s_o)$  and  $(\bar{e}_o, b)$ . In particular,  $b$  is the bequest his retired parent (the great-grandparent) decided when the grandparent was still a young parent, but only received when old. Given this, he makes consumption, own human capital investment, and savings decisions,  $(c, s, n) = \{c_{o+l}, n_{o+l}, s_{o+l+1}\}_{l=0}^4$ . In period  $j + 3 = 11$  he leaves bequests,  $b'$ , for the young parent, which is in turn received in the next period. His value function is

$$V_o(a''; \mathbf{x}'_y; \mathbf{x}_o, \bar{e}_o; s_o, b) = \max_{c, n, s, b'} \left\{ \sum_{l=0}^4 \beta^l u(c_{o+l}) + \theta V_y(a''; \mathbf{x}'_y; b') \right\}$$

$$c_o + s_{o+1} = f(e_o) + (1 + \tilde{r})(s_o + b) + g,$$

$$c_{o+l} + s_{o+l+1} = f(e_{o+l}) + (1 + \tilde{r})s_{o+l} + g, \quad \text{for } l = 1, 2,$$

$$c_{o+3} + s_{o+4} + b' = p(\bar{e}) + (1 + \tilde{r})s_{o+3} + g, \quad b' \geq 0,$$

$$c_{o+4} + s_{o+5} = p(\bar{e}) + (1 + \tilde{r})s_{o+4} + g, \quad s_{o+5} \geq 0$$

$$e_{o+l} = w_S h_{o+l} \epsilon (1 - n_{o+l}) \quad \text{and} \quad n_{o+l} \in [0, 1], \quad \text{for } l = 0, 1, 2,$$

$$\bar{e} = \frac{\bar{e}_o}{2} + \frac{\sum_{l=0}^2 e_{o+l}}{6}.$$

The parent faces only a natural borrowing constraint on own savings ( $s_{o+5} \geq 0$ ), but bequests must be non-negative ( $b' \geq 0$ ); i.e the parent cannot borrow against their child's income. This is in the spirit of [Loury \(1981\)](#), where parents cannot borrow against their children's future income to invest in their children's human capital.

Again, notice the timing of the bequests: we assume that the bequests are *decided upon* after the grandchild's college decisions  $S''$ , but before the realization of  $(a''', \epsilon'')$ , i.e. the great-grandchild's ability shock and grandchild's market luck shock. However, the transfer *occurs* after their realizations. This keeps the problem convex, and given the time horizon, does not seem to be an unreasonable assumption.

We explain in the appendix how the shock structure helps reduce the state space, allowing us

---

<sup>18</sup>While there is heterogeneity in actual college costs, in particular when considering financial aid, our results regarding the IGE are insensitive to assuming some degree of heterogeneity (e.g. financial aid for the poor). In addition, our results indicate that the college margin does not play a major role in explaining the IGE.

to simplify the problem, and solve for the value functions and optimal policies. The simplification also shows that parents with more resources leave more bequests, but all else equal, if their children have high ability and/or human capital, they leave *less* bequests.

## 2.4 Firm and Stationary Equilibrium

We assume a standard neoclassical firm that takes physical and human capital as inputs to produce the single consumption good. It solves

$$\max_{K, H_h, H_l} F(K, H_0, H_1) - RK - w_0 H_1 - w_1 H_1,$$

where  $R = (1 + r + \delta)^6 - 1$  is the competitive rental rate and  $(K, H_0, H_1)$  are the aggregate quantities of capital and effective units of labor by skill in the economy, respectively.

Let  $X$  denote the aggregate state spanning all generations, and denote its stationary distribution by  $\Phi(X)$ . Let  $\Gamma(\cdot)$  denote the law of motion for  $\Phi$ , which is derived from the agents' policy functions. In a stationary equilibrium, prices  $(r, w_0, w_1)$  solve

1. Market clearing and stationarity:

$$\Phi = \Gamma(\Phi)$$

$$K = \int_X (s_j^* + b^*) \Phi(dX) \tag{1}$$

$$w_S H_S = \int_X e_j^* \Phi(dX, S, 3 \leq j \leq 10). \tag{2}$$

where  $s_j^*$  are the optimal savings decisions in period  $j$ ,  $b^*$  the bequest decision,  $e_j^*$  the earnings in period  $j$  resulting from the optimal human capital and time investment decisions of adults, and  $\epsilon$  the luck shock of the decision making household.

2. Government budget balance:

$$g = \pi_g \bar{e}^*, \quad d = \pi_d \bar{e}^*,$$

i.e. the subsidies  $(g, d)$  are fixed fractions  $(\pi_g, \pi_d)$  of equilibrium average earnings  $\bar{e}^*$ :

$$\bar{e}^* = \int_X e_j^* \Phi(dX). \tag{3}$$

The social security regime is also balanced:

$$\bar{e}^* = 2 \left[ p_0 + p_1 \int_X \bar{e}^* \Phi(dX) \right].$$

### 3. Calibration

The model is calibrated to 1990 United States. There are several reasons we choose 1990 as the benchmark. First of all, many of our target moments need to be based on data spanning the working lives of at least two generations. If we were to target data collected on a parent-child pair beyond this time, most children will not have yet reached the peak of their working lives. Hence, many of the intergenerational statistics we use as target moments are also from studies based on observations in this period. Second, we want to look at a period of relative stability, so that the earlier and later generations are not living in environments that differ by temporal events (such as the oil crisis of the 70s or Great Recession of recent). For those moments where the data points for 1990 are not available, we take the closest available data to 1990, as we explain below.

#### 3.1 Parametrization

**Exogenous processes** We assume an AR(1) process for ability shocks:

$$\log a' = -\frac{(1 - \rho_a)\sigma_a^2}{2} + \rho_a \log a + \eta, \quad \eta \sim \mathcal{N}(0, (1 - \rho_a^2)\sigma_a^2) \quad (4)$$

where  $\sigma_a^2$  is the unconditional variance of abilities. For luck shocks assume:

$$\log \epsilon' = -\sigma_v^2/2 + \rho_\epsilon (\log a' + \sigma_a^2/2) + v, \quad v \sim \mathcal{N}(0, \sigma_v^2),$$

so both abilities and luck are assumed to have a mean of 1.

**Preferences and Technology** Preferences are modeled as a standard CRRA utility function,

$$u(c) = \frac{c^{1-\chi}}{1-\chi},$$

for all individuals, including children and retirees. The aggregate production function and capital stock evolution are parameterized as:

$$F(K, H_0, H_1) = K^\alpha H^{1-\alpha}, \quad H = \left[ v^{\frac{1}{\sigma}} H_1^{\frac{\sigma-1}{\sigma}} + (1-v)^{\frac{1}{\sigma}} H_0^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

$$K' = (1 - \delta)K + I.$$

The parameterization of the aggregate production function is consistent with [Heckman et al. \(1998\)](#), who present evidence that the elasticity of substitution between capital and aggregate labor are not significantly different from unity. Their point estimate for the elasticity between skilled and unskilled labor is  $\sigma = 1.441$ , in the range of many studies cited in the same paper. In addition, we show in [section 5](#) that key results are insensitive to the assumed value of  $\sigma$ .

The parameters for human capital production are calibrated within the model, except for  $\phi$ . We set the elasticity of substitution between age 6 and 18 human capital investments  $\phi = 2.569$  as estimated in [Cunha et al. \(2010\)](#). We analyze the sensitivity of our results to different values of  $\phi$  in [section 5](#).

**Policy Parameters** Tax rates should capture marginal tax rates, not average tax rates. Hence,  $\tau_k$  is set as 0.31, the 1990 value from [Gravelle \(2007\)](#)'s study on effective marginal tax rates on capital income. For labor income taxes, we follow the log specification of average taxes in [Guner et al. \(2013\)](#), which they show retains attractive properties for the marginal tax rate:

$$\tau_e(e) = \tau_0 + \tau_1 \log(e/\bar{e}),$$

where again,  $\bar{e}$  is the average earnings in the economy. We set  $(\tau_0, \tau_1)$  to the values estimated in their study, and calibrate  $\bar{e}$  to the average earnings in our model economy.<sup>19</sup>

We then use  $\bar{e}$  to compute transfers and college costs. Parameters for the social security system,  $(p_0, p_1, \tau_s)$ , are set as follows. First, we fix  $p_1 = 0.32$ , the median replacement rate for social security payments. Given  $p_1$ , we set  $(p_0, \tau_s)$  to balance the social security budget and match a replacement rate of 40%, as reported in [Diamond and Gruber \(1999\)](#). Instead of adding two additional nested loops to our numerical problem, we assume that earnings in model periods 3 – 4 (ages 18-29) are negligible in the aggregate so that  $\bar{e} \approx \int_X \bar{e} \Phi(dX)$ , where  $\bar{e}$  was individual av-

<sup>19</sup>Although their estimates are based on 2000 IRS tax returns, the 1990s displayed significantly modest changes to tax policies compared to other decades.

verage earnings from ages 30-65. Given this assumption,  $\tau_s = 0.1$  and  $p = \pi_p \bar{e}$  with  $\pi_p = 0.133$ . Transfers as a fraction of average earnings,  $(\pi_g, \pi_d)$ , are set to  $(2\%, 5/(1 - \alpha)\%)$ , respectively. We view lumpsum transfers mainly as welfare for the poor. The size of welfare transfers in the U.S. is approximately 1-2% of total GDP throughout the late 1980s to mid 1990s, of which we take the upper-bound for two reasons: the model transfers captures more than welfare transfers, and is modeled as a fraction of average earnings which is smaller than GDP/capita. Education transfers in the model are assumed to be public spending on secondary education and below in the data, which is obtained from the 1990 Digest of Education Statistics. As a fraction of GDP, this value is 5%, and the division follows from the labor income share of total output, since  $\pi_d$  is a fraction of average earnings. Annual college costs are set to equal  $\pi_\kappa = 30\%$  of average earnings, which is roughly equal to the ratio of college costs from the National Center of Education Statistics and average earnings from the CPS throughout the 1980s-1990s, so that

$$\kappa = \frac{1 - \beta^4}{1 - \beta^6} \cdot \pi_\kappa \bar{e}.$$

Since all policy parameters are expressed as a fraction of  $\bar{e}$ , we solve the fixed point problem (3) such that in equilibrium, individual choices given the implied values of tax, subsidies and college costs leads to the average earnings used to compute those values. Lastly,  $q_A$ , the adult equivalence scale for households with children, is set to 1.7, which is the adjustment factor for two-adult households with two children versus those without children, used by the OECD.

**Setting Prices** The parameters  $(\beta, v)$  are found in equilibrium as follows. The firm's profit maximization implies that

$$RK = \alpha F, \quad WH = (1 - \alpha)F, \quad W = (1 - \alpha) \left( \frac{\alpha}{R} \right)^{\frac{\alpha}{1-\alpha}}$$

where

$$W = \left[ v w_1^{1-\sigma} + (1 - v) w_0^{1-\sigma} \right]^{\frac{1}{1-\sigma}}. \quad (5)$$

In the benchmark calibration, we choose the discount factor  $\beta$  to imply an annual equilibrium interest rate of  $\bar{r} = 4\%$ , while in experiments we fix  $\beta$  to its calibrated value and choose  $r$  to clear the asset market. Since we do not directly observe skill prices in the data, we set  $(w_0, w_1)$  to match

Parameter	Value	Description
$\chi$	2	CRRA coefficient, $u(c) = \frac{c^{1-\chi}}{1-\chi}$
$\phi$	2.569	CES age 6 vs 18 human capital investments, <a href="#">Cunha et al. (2010)</a>
$\sigma$	1.441	CES high vs low skill labor, <a href="#">Heckman et al. (1998)</a>
$(\alpha, \delta)$	(0.322, 0.067)	capital income share / depreciation rate, <a href="#">Huggett et al. (2011)</a>
$(\tau_0, \tau_1)$	(0.099, 0.035)	earnings tax constants $\tau(e) = \tau_0 + \tau_1 \log \frac{e}{\bar{e}}$ , <a href="#">Guner et al. (2013)</a>
$(q_A, \pi_g)$	(1.7, 0.02)	adult equivalence scale, lumpsum subsidies as fraction of average earnings $\bar{e}$
$(\pi_d, \pi_k)$	(0.05, 0.3)	education subsidies and cost of college, as fraction of average earnings $\bar{e}$
$(p_1, \pi_p, \tau_s)$	(0.32, 0.08, 0.12)	social security parameters implied by balanced social security budget and median replacement rate of 40% ( <a href="#">Diamond and Gruber (1999)</a> ), refer to text.
$(\bar{r}, \bar{EP})$	(4%, 46.8%)	rate of return on capital / earnings premium in 1990 IPUMS CPS

Table 1: Fixed Parameters

the observed earnings premium. The firm's cost minimization implies the equilibrium wage ratio

$$w \equiv \frac{w_0}{w_1} = \left[ \frac{(1-v)H_1}{vH_0} \right]^{\frac{1}{\sigma}},$$

and since we observe  $EP$ , the college earnings premium in the data, wages must satisfy

$$\begin{aligned} \frac{H_1}{wH_0} &= \left( \frac{v}{1-v} \right)^{\frac{1}{\sigma}} \left( \frac{H_1}{H_0} \right)^{\frac{\sigma-1}{\sigma}} = EP &\Rightarrow & \frac{H_1}{H_0} = \left( \frac{1-v}{v} \cdot EP^\sigma \right)^{\frac{1}{\sigma-1}}, \\ & &\Rightarrow & w = \left( \frac{1-v}{v} \cdot EP \right)^{\frac{1}{\sigma-1}}. \end{aligned} \quad (6)$$

Using  $(w, W)$  from equations (5) and (6), we can set  $(w_0, w_1)$  for any given  $(r, v, EP)$ . In the calibration, we choose  $v$  to match an earnings premium of  $\bar{EP} = 46.8\%$ , which we compute from the 1990 IPUMS CPS. In experiments, we fix  $v$  to its calibrated value and find the wage ratio  $w$  that clears the labor market by skill.

### 3.2 Simulated Method of Moments

In sum, 16 parameters are taken from the literature or exogenously fixed by the data, see Table 1. As explained above, three parameters,  $(\beta, v, \bar{e})$ , are found in equilibrium to satisfy equations (1)-(3), respectively. The remaining 9 parameters are chosen by using the model to simulate 9 equilibrium moments to match an exactly identified number of empirical moments. Specifically, the parameter vector is

$$\Theta = [\gamma_0 \ \gamma_1 \ \gamma_p \ \gamma_k \ \rho_a \ \sigma_a^2 \ \rho_\epsilon \ \sigma_v^2 \ \theta]'$$

$M_{s,i}$	Description	value
$i = 1.$	Average wage ratio at ages 48-53 vs. 30-35, $S = 0$	141.0%
2.	Average wage ratio at ages 48-53 vs. 30-35, $S = 1$	153.7%
3.	College enrollment rate	41.3%
4.	(Secondary and below) education expenses-GDP ratio	6.9%
5.	Education persistence $P(S' = 1 S = 1)$ (refer to text)	63.4 %
6.	IGE of lifetime earnings	0.6
7.	Gini coefficient of lifetime earnings	0.488
8.	Gini coefficient of retirement wealth	0.62
9.	Fraction of population receiving zero bequests	45%

Table 2: SMM Moments. The probability measure  $P(\cdot)$  is taken over the equilibrium stationary distribution.

and the vector of empirical moments,  $M_s$ , is summarized in Table 2. Moments 1-3 and 5 are computed from the NLSY79. College,  $S = 1$ , is defined as at least one year of education beyond high school graduation, and high school graduates or below are defined as  $S = 0$ .<sup>20</sup> To obtain the persistence of education we refer to the Children of NLSY79, from which we can compute  $\mathbb{P}(S' = 1|S = 1)$  in the sample.

The education expense ratio is from the 1990 Digest of Education Statistics. Note that this is the aggregate of both public and private expenses. The IGE of lifetime earnings is from Mazumder (2005) using data from the SER. The number 0.6 is the coefficient obtained by running a regression on the logarithm of fathers' earnings averaged over the years 1970-1985, where the dependent variable is the logarithm of children's earnings (both sons and daughters) averaged over the years 1995-1998.<sup>21</sup> The Gini coefficient of lifetime earnings is from Leonesio and Bene (2011). Their data set is based on Social Security Administration data from 1980-2004, for more than 3 million people. The number we take is the Gini coefficient of a 12-year present discounted value of earnings using data on men aged 31-50 in 1993. The retirement wealth Gini is computed in Hendricks (2007) using data from the PSID. McGarry (1999) documents detailed information on intergenerational transfers using data from the HRS and AHEAD. The AHEAD survey elicits the subjective probability of leaving a bequest, of which the sample average is 0.55.<sup>22</sup>

<sup>20</sup> Hence, we make no distinction between high school dropouts and high school graduates, nor college dropouts, some college, college graduates and beyond. Increasing the education categories would perhaps make the study more interesting, however the numerical analysis becomes exponentially costly.

<sup>21</sup> In a sensitivity analysis, we also consider a value of 0.4. Our main results are robust to the exact number we target for the IGE.

<sup>22</sup> In comparison, 43% of respondents in the HRS respond that affirmatively to the question on whether they "expect to leave a sizable inheritance." Although this is a qualitative question, it does not seem to be at odds with the AHEAD respondents.

We find the point estimate  $\hat{\Theta}$  by

$$\hat{\Theta} = \arg \min_{\Theta} [M(\Theta) - M_s]' W_s [M(\Theta) - M_s], \quad (7)$$

where  $M(\Theta)$  are the simulated model moments and  $W_s$  is a weighting matrix. Most of the empirical moments come from different data sources, so we set  $W_s = \mathcal{I}$ , the identity matrix, as it is not clear how to obtain an optimal weighting matrix  $\hat{W}_s$ .<sup>23</sup> Furthermore, note that this is a nested fixed point problem, since for every  $\Theta$  we compute the 2 equilibrium conditions (1), (2) and the budget balance condition (3).<sup>24</sup> The resulting benchmark parameters are summarized in the top panel of Table 3.

**Moment Influence on Parameters** In the next section, we will discuss the role of each parameter and how they influence moments or mechanisms of interest. Before doing so, we do the reverse as a robustness check: measure the influence of moments on parameters. Let  $J$  denote the Jacobian matrix of  $M(\Theta)$ . Since our model is an exactly identified system,

$$\Lambda = \hat{J}^{-1}, \quad \hat{J} = \nabla_{\theta} M(\hat{\Theta})$$

is the influence matrix, or what [Gentzkow and Shapiro \(2013\)](#) call “sensitivity.” The influence of moment  $j$  on parameter  $i$  is measured by an element of  $\Lambda$ ,  $\lambda_{ij} = \frac{\partial \hat{\theta}_i}{\partial M_j} |_{M=\hat{M}}$ . The matrix  $\Lambda$  is shown in the bottom panel of Table 3.

The matrix  $J$  is approximated by computing the impact of a percentage point change in  $\Theta$  on equilibrium moments. Specifically, we vary each parameter by  $\pm 0.005$  of their benchmark values, and compute the new market clearing prices  $(r, w)$  and budget balance condition (3) holding  $(\beta, v)$  fixed at their calibrated values. Note that all parameters except  $(\sigma_a, \sigma_v)$  are assumed to lie in the interval  $[0, 1]$  in absolute terms. However, since  $(\sigma_a, \sigma_v)$  are the standard deviations of a logarithm, these can also be interpreted as percentage point deviations. Furthermore, all moments except 6-8 are expressed in percentage point ratios, and elasticities and Gini coefficients are easily interpreted on a  $[0, 1]$  scale as well. Hence  $\Lambda$  is easily interpreted as the percentage point change in parameters from a 1 percentage point change in moments.

It is clear from the table that the non-linearity of the model renders no single moment identi-

<sup>23</sup>Since the parameters are recovered using an identity weighting matrix, it makes no sense to present standard errors.

<sup>24</sup>Or, it could be formulated as an MPEC problem. However, we solve the numerical problem as a nested fixed point.

Parameter	$\gamma_0$	$\gamma_1$	$\gamma_p$	$\gamma_k$	$\rho_a$	$\rho_\epsilon$	$\sigma_a$	$\sigma_v$	$\theta$
Value	0.63	0.74	0.97	0.64	0.35	-0.55	0.17	0.52	0.54
Moment Influence									
1. Age-wage ratio, $S = 0$	0.108	0.487	0.168	0.099	0.126	0.518	-0.079	0.437	-0.117
2. Age-wage ratio, $S = 1$	-0.038	0.043	-0.006	0.048	-0.084	-0.005	0.012	0.023	0.010
3. Enrollment rate	-0.027	0.215	0.065	0.247	0.158	0.020	0.057	0.198	-0.047
4. Educ. exp./GDP	-1.364	0.608	0.421	1.840	-4.129	2.601	-0.219	0.391	-1.074
5. $P(S' = 1 S = 1)$	-0.019	-0.326	-0.140	-0.051	-0.065	-0.344	0.068	-0.282	0.135
6. IGE lifetime earnings	0.103	0.107	0.057	0.000	0.122	0.142	0.009	-0.088	-0.054
7. Lifetime earnings Gini	-0.940	-1.216	-0.634	0.217	-1.172	-1.336	-0.418	-0.563	0.457
8. Retirement wealth Gini	0.660	0.886	0.328	0.141	0.340	0.533	0.163	0.931	-0.102
9. $P(b' = 0)$	0.320	-0.667	-0.109	-0.263	0.209	-0.526	0.594	-0.542	-0.130
Chosen in Equilibrium	$\beta$ 0.99	$v$ 0.60							

Table 3: Benchmark Parameters and Moment Influence. The probability measure  $P(\cdot)$  is taken over the equilibrium stationary distribution. Units are in percentage points.

fying a single parameter. Two moments have a uniformly large influence on all parameters: the GDP share of education expenditures and the lifetime earnings Gini coefficient. In generic models with stochastic persistence, persistence and variance jointly determine cross-sectional inequality and intergenerational mobility (e.g., [Han and Mulligan \(2001\)](#)). In our model, overall the cross-sectional inequality moments have greater influence on all four parameters ( $\rho_a, \rho_\epsilon, \sigma_a, \sigma_v$ ) than the intergenerational moments. Note that a negative  $\rho_\epsilon$  is needed to simultaneously match the high degree of persistence and relatively low Gini(s). In the next section, we also show that conversely, persistence related moments are more sensitive to  $(\gamma_p, \gamma_k)$  and  $(\sigma_a, \sigma_v)$  than  $(\rho_a, \rho_\epsilon)$ .

In the NLSY79, wage profiles are noisy before age 30, steeply increase until approximately age 55, and flatten out thereafter, which is why we defined moments 1 and 2 as the average steepness of wage profiles from ages 30 to 54. However, note that college related moments 1-3 have little influence over all parameters, indicating that the college distinction has less influence on the human capital parameters  $(\gamma_{S=0,1}, \gamma_p, \gamma_k)$  than secondary education expenditures and cross-sectional inequality. Since the parameters  $\gamma_{S=0,1}$  determine life-cycle profiles, it is intuitive that they are strongly influenced by moments 7 and 8. But why do moments 1-3 have a small influence?

The benefits to college enrollment in the model are i) higher wages, and ii) faster human capital accumulation. Explicit college costs are exogenously fixed by  $\kappa$ , but there is also the implicit cost of foregone earnings. In partial equilibrium, the two Ben-Porath parameters  $\gamma_S$  govern the average wage ratios of these two age groups, for the two education categories  $S \in \{0, 1\}$ . For any given

skill price ratio, a larger  $\gamma_1$  in comparison to  $\gamma_0$  would make the college age-wage profile steeper, ignoring equilibrium effects. But the equilibrium effect is the opposite—a larger  $\gamma_1$  induces less selection, and would result in a flatter wage profile for college graduates. Since the labor market clears by skill, these countervailing effects are washed out in the equilibrium. Relatedly, table 3 also shows that the college profile ( $S = 1$ ) has even less influence on the parameters, once the high school age-wage profile is taken into account.<sup>25</sup>

The parameter  $\gamma_k$  is the goods share of pre-college education, which would be 0 if  $\gamma_k$  were zero. A larger value lead to larger shares of education expenses, and it is visually clear in Table 3 that  $\gamma_k$  is largely pinned down by the GDP share of education expenses. Given this,  $\gamma_p$ , the returns to parental investments, is strongly influenced by secondary education, and even more so by lifetime earnings inequality. For all  $\gamma_k < 1$ , our model displays positive selection, so that rich parents not only tend to be wealthier, but also have higher ability and market luck shocks. Since there is positive persistence in abilities, high ability parents have higher ability children on average. To match a larger degree of inequality in a stationary distribution, therefore, either i) high ability parents must spend less time with their children and more on their own human capital accumulation, increasing the gap between earnings profiles or ii) pre-labor market investments must have a smaller effect on lifetime earnings, so there is no “catching up,” especially given the decreasing returns ( $\gamma_p < 1$ ). Both effects lead to a smaller value of  $\gamma_p$ .

A key difference in our model compared to the previous literature is that parents care both about their own future and their children’s, discounted at the rates of  $\beta$  and  $\beta\theta$ , respectively.<sup>26</sup> Nonetheless, the intuition from [Becker and Tomes \(1986\)](#) still applies: the level of  $\theta$  determines how much parents invest in their children’s physical capital as opposed to human capital. As is the case for most other parameters, the moments that have the largest influence on  $\theta$  are education expenditures and the lifetime earnings Gini. Recall that  $\beta$  is chosen to match a general equilibrium interest rate of 4%, separately from  $\theta$ . So generally, a lower value of  $\theta$  is associated with a smaller value of  $\beta$ . The fact that  $\theta$  responds negatively to education expenditures is not because parents care less about their children, but because the calibrated  $\beta$  becomes larger, so that parents shift more resources into their children’s human capital rather than their consumption or

<sup>25</sup>Our estimates for  $\gamma_0, \gamma_1$  are consistent with the value of 0.7 found in [Huggett et al. \(2011\)](#), who did not model a college enrollment choice. It is somewhat lower than the values found in [Heckman et al. \(1998\)](#), who estimate  $\gamma_0$  as high as 0.945 and  $\gamma_1$  as high as 0.939, but consistent with the small differences between the two parameters. Also note that both these papers take initial conditions at college age as given, which are endogenous in our model.

<sup>26</sup>For example, in 2-period OLG models such as [Becker and Tomes \(1986\)](#); [Mulligan \(1999\)](#), parents die immediately when their children grow up, so the discount factor and altruism parameter are one and the same.

physical capital (bequests).

### 3.3 Non-targeted moments

In Table 4, we compare non-targeted moments with moments that are available from other datasets. In the first two columns, to compare with previous empirical studies that use data from parent and children at different ages, we show IGE's of earnings and wealth when using numbers from different periods of the parent and child's life-cycle. The first column shows the IGE of earnings using parental earnings in period 6 and child earnings in period 5. This is notably lower than our 0.6 in the benchmark using lifetime earnings, and remarkably similar to the magnitude found in Mazumder (2005) when using 6-year averages of earnings.<sup>27</sup> Since our model has no idiosyncratic earnings shocks following the once-and-for-all luck shock in period 4, the different number comes purely from life-cycle effects. Next, we compare wealth holdings at parent period 9 and child period 6. Again, the magnitude is very close to what is found in Charles and Hurst (2003) who compute the IGE of wealth for parents around the age of 50 and children around the age of 37 years old. While not shown in the text, we have verified that for both earnings and wealth, the IGE's become larger when the age of the parent and child are closer to each other.

In the last column, we compare the fraction of households who participate in welfare programs ( $WPP = 1$ ) in the model to the fraction of families who receive AFDC in the NLSY79, as tabulated by Gottschalk (1990)). In the Gottschalk sample, AFDC receipts comprise approximately 10% of the participating family's total income. Our model does not have any welfare transfers per se, so instead we define  $WPP = 1$  households as those for whom the lumpsum transfer,  $g$ , exceeds 10% of the parent's earnings in period 5, when the child is 6-11 years old. The fraction of  $WPP$  households thus defined is very much in line with Gottschalk's numbers. Moreover, as we will see in subsection 4.3, the persistence of  $WPP$  is virtually identical.

Hence, the calibrated parameters are virtually unchanged when we use these moments instead as targets. The fact that we match more moments than parameters is a virtue. Recall that one of the critiques of the original Becker-Tomes framework was that the model was not identified—there were too few empirical counterparts. In summary, we believe that the model not only matches the moments that we are explicitly targeting, but also moments of interest that are not directly targeted by the calibration procedure.

---

<sup>27</sup>Most children in his study were in their early 30s, so period 5 is a reasonable proxy. For parents, he looks back 10-15 years, so that the children are of high school or college entry age. Although period 6 is slightly lower than the average age in his sample, this is when the child is of high school/college age.

	IGE of $x(j_p, j_k)$ :		$WPP = 1$
	$e(6, 5)$	$s(9, 6)$	
Data	0.528	0.365	0.293
Model	0.513	0.341	0.326

Table 4: Non-targeted Moments

## 4. Results

Given our benchmark parametrization, we conduct a number of comparative static exercises (comparisons across different steady states). First, we investigate the effect of deep parameters on intergenerational persistence through the lense of our model. This will help us understand why the IGE is low for rich groups but not intergenerationally unconstrained groups. Next, we vary government policy parameters to investigate the counterfactual implications for early childhood human capital formation and long-run intergenerational mobility. We also show that our model can match the observed weak correlation between lifetime income and retirement wealth.

### 4.1 Nature, Parental Investments, and Education

Most empirical studies of intergenerational persistence attribute a large portion to nature. In this subsection, we show that even in the absence of exogenous parameters that govern natural transmission, we would still observe a significant degree of persistence, due to earlier investments made before individuals enter the labor market. Consequently, it is inequality in earlier investments, especially time investments, that have a large impact on mobility.

**Exogenous Transmission** In Table 5, we shut down the exogenous transmission parameters,  $(\rho_a, \rho_\epsilon)$ , and tabulate the comparative static response. The columns in Table 5 are self-explanatory except for column i, where we only compute a partial equilibrium satisfying government budget balance, and in column ii take into account the change in prices in a general equilibrium. In a partial equilibrium, we fix the prices  $(r, w)$  at their benchmark values and solve the individual’s problem. Then we compute the aggregate quantities implied at the stationary distribution, and in the last two rows show the difference in the prices implied by the firm’s factor demand schedule at these quantities, compared to the benchmark. In a general equilibrium, we show how much the market clearing prices differ from the benchmark.

Clearly, in the partial equilibrium, shutting down intergenerational ability persistence increases

	Benchmark	i. (PE) $\rho_a = 0$	ii. (GE) $\rho_a = 0$	iii. $\rho_\epsilon = 0$	iv. $\rho = 0$
1. Age-wage ratio, $S = 0$ (%)	141.0	-3.554	+1.815	-5.232	-7.864
2. Age-wage ratio, $S = 1$ (%)	153.7	+11.684	+11.080	-3.242	-2.699
3. Enrollment rate (%)	41.3	-12.605	-1.792	-3.188	-1.140
4. Educ. exp./GDP (%)	6.9	+1.021	+0.031	+0.014	+0.131
5. $P(S' = 1 S = 1)$ (%)	63.4	<b>-23.437</b>	<b>-9.704</b>	<b>-1.133</b>	<b>-7.794</b>
6. IGE lifetime earnings ( $\times 100$ )	60.0	<b>-10.918</b>	<b>-1.052</b>	<b>+10.008</b>	<b>-7.170</b>
7. Lifetime earnings Gini ( $\times 100$ )	48.8	-3.918	-0.607	+3.649	-0.030
8. Retirement wealth Gini ( $\times 100$ )	62.0	-0.630	-0.831	+3.034	+1.694
9. $P(b' = 0)$ (%)	45.0	+3.450	+1.397	+0.520	-1.756
- Annual $r$ (%)	4.0	+0.864	+0.077	-0.191	-0.139
- Skill price ratio $w$ (%)	93.1	-40.224	-1.373	-1.659	-0.944

Table 5: Effect of Persistence Parameters. Columns i-iv are the change in moments from their benchmark values. The last column is when both  $(\rho_a, \rho_\epsilon)$  are set to zero.

mobility, as shown in the large drops in both education and lifetime earnings persistence. However, notice that it does not drop to zero, and furthermore, the implied skill prices diverge significantly away from their calibrated values. Consequently, in GE, much of this change in mobility is washed out. A comparison between columns i and ii shows that nature can be overestimated when recovering ability persistence from a sample without taking into account market forces. In any case, there is still a large degree of persistence even without any persistence in the underlying model.

When we set  $\rho_\epsilon = 0$ , there is a large increase in the IGE of lifetime earnings (row 6). Remember that the benchmark value for  $\rho_\epsilon$  is negative, pushing lifetime earnings toward the mean. Without the mean reverting luck shock conditional on ability  $a$ , and since the persistence in educational attainment does not change much, both earnings persistence and cross-sectional inequality increase. When both  $(\rho_a, \rho_\epsilon) = 0$  (column iv), education and earnings persistence both decline. It may be puzzling that the IGE of lifetime earnings declines in contrast to the increase in column iii. In column iii, the absence of  $\rho_\epsilon$  amplifies the reflection of ability in lifetime earnings, but in column iv. both ability and luck are purely i.i.d. shocks. Hence the change in earnings persistence closely reflects the change in the persistence of educational attainment.

The overall message of this exercise is threefold: equilibrium effects may mitigate the relevance of ability persistence in explaining earnings persistence, educational attainment alone does not explain earnings persistence, and most importantly, the persistence in innate abilities falls far short of explaining the degree of intergenerational persistence. This last result is in stark contrast to the

	Benchmark	i. $\gamma_k/2$	ii. $\phi/2$	iii. $\gamma_p = 1$
1. Age-wage ratio, $S = 0$ (%)	141.0	+25.525	-3.693	-19.826
2. Age-wage ratio, $S = 1$ (%)	153.7	+37.981	-7.931	-21.672
3. Enrollment rate (%)	41.3	-1.890	-6.306	+15.192
4. Educ. exp./GDP (%)	6.9	-4.395	+2.800	-0.886
5. $P(S' = 1 S = 1)$ (%)	63.4	<b>-5.684</b>	<b>+8.713</b>	<b>-13.267</b>
6. IGE lifetime earnings ( $\times 100$ )	60.0	<b>+20.604</b>	<b>-1.169</b>	<b>-41.932</b>
7. Lifetime earnings Gini ( $\times 100$ )	48.8	+3.566	+2.488	-11.292
8. Retirement wealth Gini ( $\times 100$ )	62.0	-12.328	+3.930	-9.815
9. $P(b' = 0)$ (%)	45.0	+2.560	+1.199	-13.967
- Annual $r$ (%)	4.0	+0.115	-0.173	-0.140
- Skill price ratio $w$ (%)	93.1	-2.496	+1.958	-4.855

Table 6: Effect of Childhood Human Capital Parameters. Columns i-iii are the change in moments from their benchmark values.

conclusion of many empirical studies.

**Childhood and Education** Next, we vary childhood and pre-college education parameters, namely  $(\gamma_k, \phi, \gamma_p)$ , and tabulate their impact on target moments, focusing on intergenerational persistence. Columns i-iii in Table 6 show the change in moments when  $\gamma_k$  and  $\phi$  are set to half their calibrated values, and  $\gamma_p$  set to 1. We do so because the equilibrium is unstable with extreme values of  $\gamma_k = 0$  or  $1$ ,  $\phi = 0$  or  $\infty$ , and  $\gamma_p = 0$ . When  $\gamma_k$  is small, parental time rather than goods investments becomes more important in human capital formation, and when  $\phi$  is small time and good investments become more substitutable— $\phi/2 \approx 1.285$ , close to Cobb-Douglas. When  $\gamma_p = 1$ , there is no decreasing returns to parental investments in children.

Observe that in column i, as time investments become more important, there is less persistence in educational attainment, and much more in lifetime earnings. As time investments become more important, parents spend more time on their children than on themselves. This increases the wage profile in adulthood (rows 1 and 2). But in particular, high human capital parents spend disproportionately more time with their children, resulting in an even steeper wage profile. In contrast, low human capital parents find themselves even more constrained in child's investments, for several reasons. First, parental time investments are not subsidized, unlike goods investments. Second, the income effect dominates the substitution effect for low human capital parents, who earn less per time unit. Furthermore, the returns to time investment in children is increasing in parental human capital. Consequently low human capital parents cannot afford to spend enough

time with their children, resulting in the large increase in the IGE.

When time and goods investment become less substitutable, parents shift away their resources from time to goods, for the same reasons just mentioned above but in the opposite direction. As a result, the GDP share of education expenditures increases, and fewer children make it to college. Educational attainment becomes more persistent as only richer parents are able to send their children to college. But this has a small impact on lifetime earnings persistence. This is because the larger goods investments made by parents are offset by the decline in time investments, which plays a more direct role in children's human capital formation (since the returns to parental time investment is increasing in the parent's human capital, which determines earnings). However, the comparative static in  $\phi$  is by no means monotone, which we cover in more detail in section 5.

The case when  $\gamma_p = 1$  is when parents have the largest impact on children's lifetime earnings. All children reach higher levels of human capital and more make it to college. Not only do more individuals start their working lives with high levels of human capital, parents also find it optimal to invest in children rather than themselves. These two effects make life-cycle human capital accumulation decline, so the profiles are less steep. This is the ideal world in which persistence and cross-sectional inequality both decrease and overall the economy more productive.

It may be puzzling that a larger  $\gamma_p$  leads to less, not more, persistence. Indeed, for the individual's problem, this would favor the rich and high human capital parents. But this is countered by the fact that more individuals make it to college, and that there is an education subsidy  $d$  and an upper bound on  $n_p$ , parents' time investments in children. Poor parents increase time investments to match the level of  $d$ , while additional goods inputs by rich parents are still subject to decreasing returns. This is evidenced by the drop in education expenses. We elaborate more on this in section 5 when we discuss the relationship between  $\gamma_p$  and  $\rho_a$ .

The points to take away are that pre-working age investments have a large impact on both mobility and inequality, and that much of intergenerational mobility stems from poor parents' lack of time investments.

## 4.2 Market Incompleteness and Intergenerational Persistence

Having analyzed the impact of innate ability persistence and early human capital formation, we turn to analyzing how the environment impacts mobility in a stationary equilibrium. Since we have established that endogenous investments made by parents play a large role in determining persistence, this exercise is of interest as those investments will change in response to changes in

	Benchmark	i. $\sigma_a/2$	ii. $\sigma_v/2$	iii. $\sigma/2$
1. Age-wage ratio, $S = 0$ (%)	+141.0	-6.993	+15.557	-3.909
2. Age-wage ratio, $S = 1$ (%)	+153.7	+2.158	+10.158	+6.167
3. Enrollment rate (%)	41.3	+7.052	-9.815	+2.388
4. Educ. exp./GDP (%)	6.9	-0.069	-0.902	+0.015
5. $P(S' = 1 S = 1)$ (%)	63.4	-6.291	-6.150	-6.206
6. IGE lifetime earnings ( $\times 100$ )	60.0	-0.854	+30.584	+8.421
7. Lifetime earnings Gini ( $\times 100$ )	48.8	-4.691	+6.744	-7.881
8. Retirement wealth Gini ( $\times 100$ )	62.0	-4.471	-2.972	-3.265
9. $P(b' = 0)$ (%)	45.0	-5.309	+3.560	-0.882
- Annual $r$ (%)	4.0	+0.171	+0.458	+0.425
- Skill price ratio $w$ (%)	93.1	-4.362	+1.908	-2.823

Table 7: Effect of Idiosyncratic Shocks. Columns i-iii are the change in moments from their benchmark values. The last column is when  $(\sigma_a, \sigma_v)$  are each set to  $1/\sqrt{2}$  of their benchmark values.

the environment. Specifically, we halve the level of uncertainty faced by individuals (measured in log standard deviations) and compute the comparative statics. Then we illustrate that our model can explain why the IGE's of intergenerationally constrained and unconstrained families can be similar.

**Idiosyncratic Uncertainty** As is apparent in Table 7, the largest change in the lifetime earnings IGE occurs when  $\sigma_v$  is halved. This is partly due to the fact that the benchmark value of  $\sigma_v$  is already large to begin with, but this change is still pronounced in comparison to changes in other aggregate moments. In column i, as individuals vary less in learning abilities, more children make it to college and there is a smaller difference in educational attainment across parental groups  $S = 0$  and 1. This translates into a slightly smaller lifetime earnings IGE, but more notably a drop in cross-sectional inequality and in particular, more parents leaving positive bequests.

To interpret the comparative static of a smaller  $\sigma_v$ , recall that the benchmark value of  $\rho_\epsilon$  is negative. Since individuals are hit with the luck shock  $\epsilon$  at the beginning of working life, and this shock is negatively correlated with innate ability  $a$ , differences coming from innate abilities and early human capital formation become dampened. Hence when the variance of this shock is reduced, not only do lifetime earnings across generations become more persistent but cross-sectional inequality also increases. Parents become more intergenerationally constrained, as seen in row 9. When both variances are reduced, the stationary equilibrium in column iii is more or less an intermediate outcome of columns i and ii.

	low	high
<hr/>		
Low is IG Constrained		
Model	0.476	0.401
Mulligan (PSID)	0.490	0.420
<hr/>		
Low is NW<Median		
Model	0.591	0.409
Mazumder (SIPP)	0.458	0.274
Mazumder (SER)	0.465	0.304

Table 8: IGE's by group, data and model.

Investments in children are made after parents observe their children's ability and their own luck shock, but before they observe their grandchildren's ability and children's luck shock. Bequests are made after all of the aforementioned shocks are observed. Hence, the relative magnitudes of these shocks determine the relative investments in the physical or human capital of the child. Bequests, or physical capital investments, will depend more on the luck shock while human capital investments will depend more on the ability shock. The differential timing of these shocks also affect investment decisions, which is only present via the inclusion of the life-cycle component. We explain this in greater detail below.

**IGE's by group** To illustrate the importance of wealth and intergenerational borrowing constraints, we compute group-specific statistics that the model is not specifically calibrated to. While earlier theories such as [Becker and Tomes \(1986\)](#) predict that poorer, constrained families would show less mobility, [Mulligan \(1999\)](#) shows that dividing households in the PSID by those who anticipate significant bequests and those who do not results in no significant difference between their IGE's. More recently, however, [Mazumder \(2005\)](#) has shown that when dividing households by net worth into poor and rich, poorer households tend to have a much higher IGE. These results are not at odds with our quantitative model, as we show in [Table 8](#).

The idea behind [Mulligan \(1999\)](#) is that children expecting bequests are the families for which the parent's intergenerational borrowing constraint is not binding. Accordingly, we measure the parent's earnings in the period of his life-cycle when he makes the bequest decisions (period 8) and divide the simulated sample into whether the parent leaves zero bequests. We call this group the "IG Constrained." We then measure the IGE for both groups, but instead of lifetime earnings

use the (6-year) parental earnings in period 8 and child earnings in period 5.<sup>28</sup> The results are summarized in the top panel of Table 8. The difference in the model-implied IGE's of the two groups are in line with Mulligan (1999).

Analogously, the reasoning behind Mazumder (2005) is that if constraints have an effect on intergenerational persistence, it would show when the parent is actually making investments in the child. In our model, the parent makes direct goods investments in the child in periods 6 and 7. For comparison with his results, we divide families into whether parental wealth was above or below the median in period 6, and again measure the 6-year IGE for these two groups.<sup>29</sup> We call the low wealth group "NW<Median." The results are summarized in the bottom panel of Table 8. Again, the difference in the model-implied IGE's of the two groups are similar to Mazumder (2005)'s, both for the SIPP and SER data.<sup>30</sup>

The table shows that, both in the model and the data, bequests are not so much of an indicator of inefficient investments, while earlier investments are.<sup>31</sup> This is consistent with Haider and Solon (2006)'s results that annual earnings are more correlated with lifetime earnings in an individual's ages from early 30s to 40s, but then tends to drop with age. Relatedly, in our model the correlation between parent's lifetime earnings and the bequests he leaves for his child is 0.59, far from 1. The correlation between the parent's bequest and the child's lifetime earnings is very low at 0.06. Hence, bequests are only a weak indicator of the parent's human capital, and barely related to the child's.

To further understand this result, we show several other statistics in Table 9 for comparison. In the first two rows, we show the fraction of parents and children, in each group, that are college educated. In the model, if college were free, everyone would attend college. Hence, if the difference in the parent-child generations' educational attainment is large, it means there is more mobility, and vice versa if it were small. In the table, it is not surprising that there is upward educational mobility in the poorer group and downward mobility in the richer one, since the underlying stochastic process displays regression toward the mean. What is notable here is the

---

<sup>28</sup>In both studies, child earnings are measured in their early thirties.

<sup>29</sup>Specifically, we divide groups by whether  $s_6 + b/(1 + \bar{r})$  is above or below the median.

<sup>30</sup>Perhaps surprisingly, even the levels of the IGE's are in line for the Mulligan (1999) comparison, who used earnings averaged over 4 years to compute IGE's. This is not the case for the Mazumder (2005) comparison, who used 2 year averages. Remember, however, that the 6-year IGE in our model is similar to what he finds in the same study (Table 4). Both in our model and in his data, the overall IGE increases as one increases the number of years that earnings are averaged over, so the difference in levels may just be a life-cycle noise effect.

<sup>31</sup>Of course, the results are different based on which periods of life, both for the parent and the child, we use to compute the IGE. But the underlying effect that, what matters is the time at when the parent is making investments in the child, not leaving bequests, still remains.

	IG Constrained		NW<Median	
	yes	no	yes	no
$P(S = 1)$	0.288	0.512	0.366	0.457
$P(S' = 1)$	0.311	0.497	0.373	0.453
NW<Median	0.761	0.287	-	-
IG Constrained	-	-	0.685	0.215
persistence	0.676	0.734	0.782	0.692

Table 9: Human Capital Investment by group. The probability measure  $P(\cdot)$  is taken over the equilibrium stationary distribution.

magnitude of the differences, which is much larger when the sample is split according to whether families are IG constrained. In other words, IG constrained parents send more of their children to college than the parents themselves, compared to low net worth parents. The reverse is true for the unconstrained vs high net worth parents.

This is very different from [Mulligan \(1999\)](#)'s conjecture that constrained families are not affected by constraints, because those parents would not have invested in their children even in the absence of constraints. What is happening here is that constrained parents have already done all their investing by the time they are leaving bequests. These parents are relatively less able than their children, so they have not invested much in themselves and borrowed against their own future income to invest in their children. But when the children grow up, these parents find that they do not have enough assets to share with them, as parents must also consume in their old age and retirement. In contrast, the unconstrained parents are relatively more able than their children. They invest more in themselves, and do not find much of a need to invest in their children. Because they have more assets when they become old, both due to the fact that they have more lifetime earnings and did not have to invest in their children in the previous period, they are able to leave more bequests for them.

Furthermore, whether or not the parent leaves bequests depends not only on how much education the child received, but also on the market luck shock which hits the child when he becomes an adult. This shock is negatively correlated with ability, which determines the level of education, and has large variance. Because bequests depend more on the luck shock, children receiving zero bequests include both efficiently educated children and less educated children with good luck shocks. Additionally, there is relatively more upward mobility across these two groups than when dividing families by net worth, as can be seen in the last row. A child who does not re-

ceive bequests is less likely to find himself IG constrained, compared to a child growing up to be categorized as low net worth if his parent were categorized as low net worth.

What if we were to divide families by whether or not their net worth was above the median? It turns out that these parents are indeed constrained in educating their children, as can be seen in the small differences in educational attainment rates. The increase and decrease in educational attainment is smaller for the low and high net worth groups, respectively. Clearly, low net worth parents will also not have sufficient resources to share with their children in the form of bequests, in contrast to high net worth parents. The upward mobility across these two groups is less than when dividing families by bequests, which is contributing to the high IGE among low net worth families. The exact opposite is true for high net worth parents. Note that this does not mean that the IG constrained and low net worth families are very different—76.1% of the IG constrained families are low net worth parents. But there is enough mobility out of the low net worth state among IG constrained children so that the IGE is not as high as among low net worth children.

### 4.3 Government Interventions

As we have shown above, human capital formation, and the environment that determines parental incentives to invest in child's human capital, is important in understanding intergenerational persistence and related empirical facts. Our benchmark model and comparative static exercises were conducted under the assumption that government subsidies and taxes were neutral in terms of average earnings. Policies realign how incentives respond to the environment, and we show that all government tax/transfer systems in our model contribute to reducing intergenerational persistence, except for capital taxation. In particular, it should not be surprising that education subsidies, followed by earnings taxation, have the largest counterfactual effect on the IGE, given that they directly affect human capital investments.

**Transfers** We begin by computing the new equilibria in the three counterfactual scenarios where there are no lumpsum transfers, which mainly affects the poor; no education transfers, which again primarily constrains poor parents' investment in children; no (PAYGO) social security system, which does not affect young parents' investment in children directly, but nonetheless has life-cycle and equilibrium feedback effects.

Eliminating lumpsum transfers has an equalizing effect on college enrollment, as the aggregate enrollment rate increases while the conditional enrollment rate decreases. Nonetheless, this leads

	Benchmark	i. $g = 0$	ii. $d = 0$	iii. No SS
1. Age-wage ratio, $S = 0$ (%)	141.0	-1.977	-4.566	-16.379
2. Age-wage ratio, $S = 1$ (%)	153.7	+0.492	-4.852	-21.247
3. Enrollment rate (%)	41.3	+2.107	-4.383	+10.715
4. Educ. exp./GDP (%)	6.9	-0.261	+0.030	+1.186
5. $P(S' = 1 S = 1)$ (%)	63.4	-8.171	-6.718	-8.599
6. IGE lifetime earnings ( $\times 100$ )	60.0	+4.620	+40.000	+1.702
7. Lifetime earnings Gini ( $\times 100$ )	48.8	+0.113	+7.553	-6.156
8. Retirement wealth Gini ( $\times 100$ )	62.0	-1.250	+3.024	-8.868
9. $P(b' = 0)$ (%)	45.0	-2.226	+2.412	-0.095
- Annual $r$ (%)	4.0	+0.011	-0.077	-0.516
- Skill price ratio $w$ (%)	93.1	-2.271	+0.979	-4.544

Table 10: Government Transfers.

	$WPP = 1$	0
Gottschalk (1990)	0.553	0.218
Model	0.547	0.219
$n_k = 1$	0.910	0.942
$d/(m_p + d)$	0.143	0.062

Table 11: Intergenerational Persistence of Welfare Program Participation

to an increase in the IGE, indicating that poor parents become more constrained regardless of college enrollment. High school parents spend less time with their children (row 1) compared to college parents, and have to save more to prepare for retirement, resulting in a decline in the retirement wealth Gini (row 8).

The lack of education transfers causes the largest increase in the IGE, despite the drop in the conditional enrollment rate; in fact, the resulting elasticity is 1.<sup>32</sup> This means that there is almost perfect persistence in earnings, which also leads to large increases in cross-sectional inequality, moments 7 and 8. Of course, some efficiency can be lost by tying transfers to secondary education (in-kind transfers) rather than distributing them lumpsum in aggregate. But in equilibrium, parents who anticipate that their children will receive lumpsum transfers when they grow up may have less incentive to invest in their education. Table 10 indicates that the lumpsum transfers lead to greater mobility. We now turn to the question of whether these transfers play a negative role in children's education when there are in-kind education transfers in place.

<sup>32</sup>In our stationary equilibrium, the IGE of lifetime earnings is equivalent to the IGC.

**Welfare Program Participation** In Table 11, we contrast two groups—those groups who live on welfare and those who don't, as defined in subsection 3.3. The table shows the model simulated probability that a child who grew up with a parent who was on welfare, ( $WPP = 1$ ), also receives welfare subsidies when he grows up to be a young parent ( $WPP' = 1$ ). The chances of this happening for such a child are significantly higher than that for the corresponding child whose parent was not on welfare, and the magnitudes are consistent with [Gottschalk \(1990\)](#). The question is whether the children were also just born with low ability, or whether they could have grown up to attain higher human capital levels had they received more external support from outside the family or had no lumpsum transfers to look forward to in adulthood. To see whether this is the case, in the next two rows we compare the amount of time the child spends in own human capital accumulation at college age ( $n_k$ ), and how much the parent pays out of his own pocket ( $m_p$ ) compared to the education subsidies ( $d$ ). It turns out that the former is lower, and the latter higher, in the  $WPP$  group. This means that given a level of investment,  $WPP$  children found it less worthwhile to continue to increase their human capital, and that the education subsidy alone was sufficient to induce desirable education levels conditional on the child's states. We conclude that as long as the education subsidies are in place, the welfare program has no perverse effects.

**Lifetime Income and Retirement Wealth** The model social security program has only a small impact on the IGE, as shown in column iii of Table 10. However, it has a large impact on individual wage profiles and college enrollment rates. The retirement wealth Gini declines, as individuals must save on their own to prepare for retirement. Additionally, with social security, work becomes more important later in life because earnings are tied to social security payments, and since social security payments are heavily discounted earlier in life. Consequently the wage profiles decline because there is less incentive to concentrate working hours toward later periods, and also because the incentive to work rather than invest in own human capital increases in the absence of the flat rate payroll tax. This implies that endogenous wage profiles and social security payments are strongly related to retirement wealth.

According to [Hendricks \(2007\)](#), the standard life-cycle model overshoots the correlation between lifetime earnings and retirement wealth, even after accounting for inheritance shocks. The strong correlation leads to having too little heterogeneity in wealth within earnings deciles, and too much across deciles, compared to the data. In the textbook model, the earnings poor do not save at all, while the earnings rich save abnormally larger amounts than the earnings poor. We

	$C_{WE}$	Mean W/E Gap	Mean Gini
PSID	0.61	0.25	0.54
Hendricks	0.82	0.55	0.39
Model	0.55	-0.09	0.59

Table 12: Retirement Wealth and Lifetime Earnings

summarize his findings in Table 12.

The column  $C_{WE}$  is the correlation between lifetime earnings and retirement wealth. The next column, “Mean W/E Gap,” is defined in Hendricks (2007) as follows. First, we compute average retirement wealth over average lifetime earnings ratio for each earnings decile. Then, we compute the slope of a fitted line across the deciles. The last column, “Mean Gini,” is the average of the retirement wealth Gini’s of each earnings decile. We present these statistics from the PSID data in the first row, Hendricks (2007)’s exogenous inheritance shock model in the second, and our model in the last row. If savings behavior is monotonically determined by earnings, we should see a higher Mean W/E Gap and a smaller Mean Gini, respectively, than in the data.

The benchmark model of Hendricks (2007) falls short of explaining the data, as inheritance shocks that hit individuals in later periods of life do not generate enough within group heterogeneity in savings behavior. However, we find that our model with anticipated bequests from one’s parent, intentional bequest motives toward one’s child, and perhaps most importantly, endogenously heterogeneous earnings profiles create a great degree of within group heterogeneity. In fact, it *undershoots* the correlation, and consequently *understates* the Mean W/E Gap. Furthermore, the Mean Gini is close to the population Gini, meaning there is almost as much heterogeneity in each decile as there is in the population.

In short, our study is not intended to explain this puzzle but has all the ingredients to reconcile the model with the data. Because of heterogeneous earnings profiles, retirement wealth can be very different even within the same earnings deciles. Furthermore, the earnings profile of the child can again be very different from the parent, so bequests will be heterogeneous even across parents with the same profile. Since bequests are not persistent across generations, this also leads to very heterogeneous savings behavior across the earnings rich.

**Taxation** Table 13 reports the change in equilibrium moments when the progressive earnings tax function is made flat, and earnings or capital taxation is eliminated altogether. The flat earnings tax rate used to generate column i is set equal to the ratio of total earnings tax revenues over total

	Benchmark	i. flat	ii. $\tau_e = 0$	iii. $\tau_k = 0$
1. Age-wage ratio, $S = 0$ (%)	141.0	+50.598	-13.693	-10.944
2. Age-wage ratio, $S = 1$ (%)	153.7	-2.596	-16.902	-10.232
3. Enrollment rate (%)	41.3	-15.793	+8.386	+2.553
4. Educ. exp./GDP (%)	6.9	-0.698	+1.995	+0.306
5. $P(S' = 1 S = 1)$ (%)	63.4	-12.669	-15.320	-2.883
6. IGE lifetime earnings ( $\times 100$ )	60.0	+29.285	+11.827	-5.883
7. Lifetime earnings Gini ( $\times 100$ )	48.8	+24.664	-1.458	-3.693
8. Retirement wealth Gini ( $\times 100$ )	62.0	-7.109	+1.415	-1.716
9. $P(b' = 0)$ (%)	45.0	+13.191	-1.105	-6.641
- Annual $r$ (%)	4.000	-0.199	-0.278	-1.285
- Skill price ratio $w$ (%)	93.061	-3.355	-6.700	-0.498

Table 13: The Effect of Taxation.

earnings in the benchmark equilibrium, which is 26.1% including the payroll tax.

The flat tax implementation has several noticeable effects. In the benchmark model, the progressive tax function essentially plays the role of further diminishing the returns to human capital accumulation, since every additional unit of human capital investment is not only less effective but also subject to higher taxation. When this is eliminated, the aggregate college enrollment rate declines, and both the lifetime earnings IGE and Gini increase substantially. It is easy to see that in the absence of progressive taxation, there would be an increase in cross-sectional inequality in our model which features endogenous human capital accumulation throughout the life-cycle. This mechanism works intergenerationally as well, since high ability/human capital/income parents face a larger returns to investment in children. Basically, in terms of child investment, progressive taxation has similar implications as education transfers by favoring disadvantaged parents.

Eliminating earnings taxes all together (column ii) has a similar qualitative impact on the IGE as does the move to a flat tax regime (column i), but there is a quantitatively small impact on cross-sectional inequality while college enrollment increases. The difference between columns i and ii can be viewed as an endowment effect—remember that in column i, the average tax rate was set equal to the benchmark. In column ii, everyone is better off, not only the rich. The endowment effect shows up in the larger enrollment rate (more parents can afford college), which partially mitigates the jump in the IGE seen in column i.

Reducing the capital income tax rate in our model is equivalent to increasing the interest rate paid on savings. As physical capital investments become more attractive compared to the bench-

mark, life-cycle accumulation declines (rows 1 and 2). It also relaxes the intergenerational borrowing constraint, seen in the decline in the fraction of parents leaving zero bequests (row 9) and a reduction in the IGE and Gini of lifetime earnings. It is important to note that capital income taxation has different implications from labor income taxation, and that this is due to the stationary equilibrium set up. Although it shifts away incentives from own human capital accumulation, which would seem to increase inequality and persistence, in a world with intergenerational credit constraints it helps the poor that the equilibrium results in a higher aggregate capital stock and a lower interest rate, which is more beneficial for the poor than the rich.

In sum, we find that educational subsidies have the largest impact in reducing the IGE, surpassing any other policy instrument present in our model. Greater progressivity in labor income taxation leads to less inequality and higher mobility. Indeed, our model suggests that US-Europe differences in inequality and mobility may be partially explained by differences in these two policies.

## 5. Sensitivity Analysis

In subsection 3.3, we showed that the model is consistent with other moments we did not target in the benchmark calibration. In the previous section, we demonstrated that the model implied moments are also consistent with i) IGE's by group, where groups are defined as low/high net worth or IG (un-)constrained, ii) the intergenerational persistence of welfare program participation status, and iii) the correlation between lifetime earnings and retirement wealth. Recovering the parameters using these moments rather than our targeted moments barely change our quantitative results. In this section, we consider the following additional robustness checks.

To validate that the life-cycle structure embedded into an intergenerational framework with child investment can reconcile empirical irregularities, our model incorporates various features that the recent literature has emphasized (e.g. parental time inputs and complementarity between early and later investments in childhood, college choice and human capital accumulation post-childhood). However, there is no consensus on whether human capital accumulation in adulthood is better specified as Ben-Porath or learning-by-doing, see [Heckman et al. \(2002\)](#). So we first report how the main results from our model change with this different specification.

Then, we consider the results from when varying key parameters: i)  $\phi$ , the elasticity of substitution between time and good investments in children, ii)  $\gamma_p$ , the returns to parental investments

in children, iii)  $\sigma$ , the demand elasticity of substitution between high school and college workers, and iv)  $\chi$ , the risk aversion (or intertemporal elasticity of substitution) parameter.

## 5.1 Learning-by-Doing (LBD)

We consider the LBD formulation

$$h_5 = a [h_4(1 - n_p)]^{\tilde{\gamma}^S} + h_4$$

$$h_{j+1} = ah_j^{\tilde{\gamma}^S} + h_j, \quad j = 5, \dots, 10$$

for  $S = 0, 1$ , so  $(\hat{\gamma}_0, \hat{\gamma}_1)$  govern the growth of wages. In [Heckman et al. \(2002\)](#) and other models featuring LBD, there is a trade-off between leisure and human capital accumulation (and foregone earnings). The exclusion of a leisure choice is compensated by the fact that in this specification, there is still the time cost from investing in children, which comes at the cost of sacrificing own human capital accumulation. With this specification, we recalibrate the model to match the same benchmark moments in [Table 2](#). The resulting parameters are mostly similar except that  $(\hat{\gamma}_0, \hat{\gamma}_1) = (0.21, 0.44)$  compared to  $(\gamma_0, \gamma_1) = (0.63, 0.74)$  in the benchmark. The absolute magnitudes are smaller as expected, since wages grow mechanically in the LBD model. That  $\gamma_0$  is relatively smaller compared to  $\gamma_1$  should also be expected: with Ben-Porath, low (learning) ability individuals spend less time accumulating human capital early in life. So without this choice margin,  $\gamma_0$  must drop relatively more than  $\gamma_1$  compared to the Ben-Porath model, to account for the differential in time spent accumulating human capital.

Despite this difference, the role of ability persistence does not change much in magnitude. Compared to a 0.07 drop in the benchmark, the IGE of lifetime earnings drops by 0.06 when both  $(\rho_a, \rho_\epsilon)$  are set to zero. The persistence of education,  $P(S' = 1 | S = 1)$  drops by approximately 10%, compared to 8% in the benchmark. However, there is a difference in the relative roles of these two parameters. While the implied changes in education persistence are similar to the benchmark, the IGE of lifetime earnings responds more strongly to  $\rho_a$ , dropping by 0.05 compared to 0.01 in the benchmark.

The reason, of course, has again to do with differential abilities and the allocation of time. In the benchmark Ben-Porath specification,  $\gamma_0$  and  $\gamma_1$  are closer in magnitude. So at the margin, individuals with similar ability allocate similar amounts of time in human capital accumulation regardless of whether they went to college. Hence, the relationship between the discrete col-

lege choice and lifetime earnings is weaker with the Ben-Porath specification. Since the values of  $(\gamma_0, \gamma_1)$  diverge more with LBD, lifetime earnings become more reflective of college choices, and the levels of persistence also move together.

For analogous reasons, parents with low (high) net worth when young become more identical to the IG (un-)constrained parents when old, which was the focus of our analysis in subsection 4.2. Qualitatively, the difference between the IGE's of the low/high groups are still larger when split by net worth when parents are young, but only by 0.05. The lack of a time allocation decision in the LBD model post-early childhood raises the correlation between the parents' lifetime earnings and the bequests they leave for their children to 0.70, compared to 0.59 in the benchmark. More importantly, the correlation between the child's lifetime earnings and bequests received is 0.13, compared to 0.06 in the benchmark. In the benchmark with Ben-Porath, parents "make-up" for the resources spent on their children when they were young by accumulating their own human capital and subsequently leaving less bequests. With LBD this margin does not exist, so parents who had more resources early on also have more resources later on. However, the early childhood trade-off still exists, so qualitatively our results still go through.

While our LBD formulation is too simplistic to conclude that the Ben-Porath model is more consistent with the data, we do believe this is evidence that some form of effort must be present for human capital accumulation, at least in early adulthood, to rationalize some of our results.

## 5.2 Child Investment Complementarity $\phi$

In Table 6 we showed the change in moments when  $\phi$  is half its benchmark value ( $\phi/2 \approx 1.285$ ). There, the most notable difference in moments was that while education persistence increased by 8 percentage points, while the IGE of lifetime earnings barely changed. The difference between the IGE's of the low/high groups are larger when split by net worth when parents are young, than by whether or not parents are constrained (0.011), but much smaller than in the benchmark (0.107). In aggregate, investments are shifted away from time to goods. But at the household level, not only can the rich parents afford these good investments but they can also match it with increased time investments. This counterfactual world uniformly favors rich parents, and with a lifetime earnings-bequests correlation of 0.69, poor parents when young also leave fewer bequests.

But this impact is not uniform. In Figure 2, we plot the equilibrium IGE of lifetime earnings as a function of  $\phi$ . We vary  $\phi$  from 1, corresponding to a Cobb-Douglas specification, to 5, approaching perfect substitution. As  $\phi \rightarrow 1$ , the complementarity effect dominates. Goods investments,

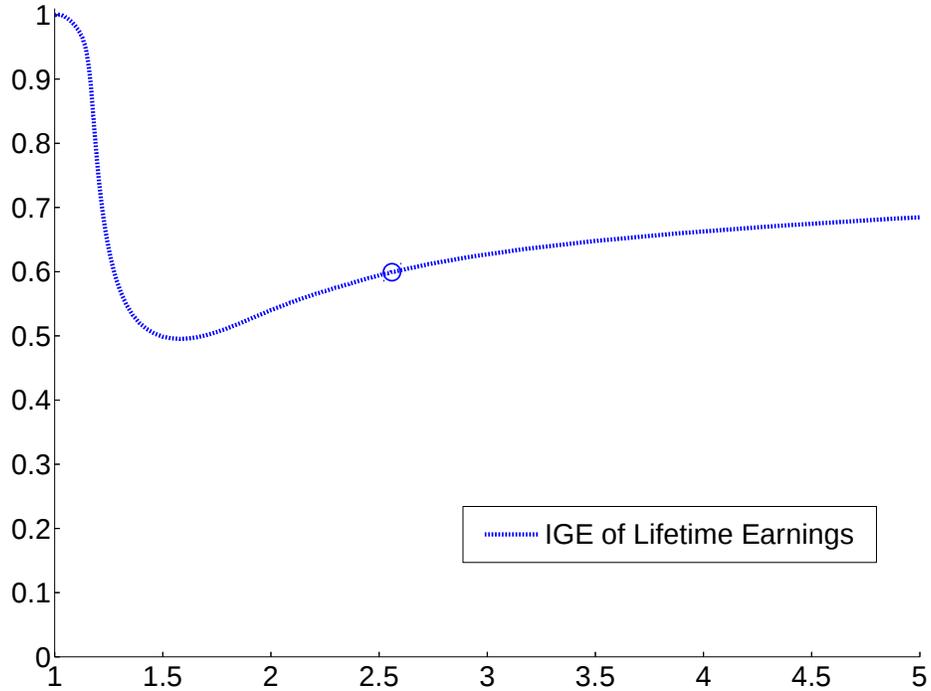


Figure 2: IGE of lifetime earnings and  $\phi$

The graph plots values of  $\phi$  on the  $x$ -axis and the steady state equilibrium IGE of lifetime earnings on the  $y$ -axis. The benchmark value from Cunha et al. (2010) is highlighted in a circle.

which are subsidized but which poor parents cannot match with time investments, become less important and the IGE approaches 1. This should remind the reader of the importance of the education subsidies in subsection 4.3. On the other hand, the IGE is also increasing as time and good investments become more substitutable, but reaches a quantitative upper-bound. Although richer parents have more resources in both time and goods, poor households still benefit from the education subsidy. The figure illustrates that we need a moderate level of substitution for our results, and also that complementarity plays a quantitatively important role in explaining the IGE. Indeed, the sensitivity of the IGE in response to  $\phi$  approaches zero as  $\phi \rightarrow 1$ .

### 5.3 Returns to Parental Investments $\gamma_p$

We showed in section 4.1 that  $\gamma_p$  has large impacts on model moments when increased from its benchmark value of 0.97 to 1. The region of  $\gamma_p$  in which the model is operational is rather narrow—the model collapses for values less than 0.9 or larger than 1. In this region, the IGE monotonically declines with  $\gamma_p$  as early human capital investments increase; at the same time, the IGE becomes more sensitive to the correlation parameters  $(\rho_a, \rho_\epsilon)$ , which is more closely reflected

in the intergenerational earnings process.

To investigate this interaction between  $\gamma_p$  and intergenerational persistence more deeply, we also recalibrated the model to match an IGE of 0.4 instead of 0.6, the conventional number typically attributed to Solon (1992). In this case, the calibrated value of  $\gamma_p$  is larger, approximately 0.98 compared to 0.97 in the benchmark. Given the upper-bound of 1, this increase is not trivial. The calibrated values of  $(\gamma_0, \gamma_1)$  are (0.68, 0.77), compared to the benchmark (0.63, 0.74), an increase which is needed to reconcile age-wage profiles with increased human capital levels at the beginning of working age. Most interestingly, *the value of  $\rho_a$  is close to zero*, in the range of 0 to 0.1.

In terms of our results, two facts are of interest. Because human capital investments during both childhood and adulthood play a larger role due to the larger values of  $(\gamma_p, \gamma_0, \gamma_1)$ , the difference in the IGE's of the low/high groups are now much larger when split by net worth when parents are young, than by whether or not parents are constrained. The difference is as high as 0.130 compared to a difference of 0.107 in the benchmark.

Since the calibrated value of  $\rho_a$  is small, there is almost no quantitative impact when we set it to  $\rho_a = 0$  for most cases when  $\rho_a \leq 0.1$ . However, if we do the reverse exercise of *increasing*  $\rho_a$  back to its benchmark value of 0.35, the IGE jumps back to values close to 0.6. Hence, the fact that the IGE is more sensitive to correlation parameters at larger values of  $\gamma_p$  remains, even in the recalibrated model.

## 5.4 Further Robustness Checks

In our benchmark model we set the college-high school labor demand elasticity of substitution parameter,  $\sigma$ , to 1.441, the estimate from Heckman et al. (1998). Varying  $\sigma$  to the values of 0.5, 1 and 2 (where  $\sigma = 1$  corresponds to the Cobb-Douglas case) does little to change aggregate moments, mostly because the differential wage effects are washed out in general equilibrium. This is similar to why  $(\gamma_0, \gamma_1)$  was not influenced much by moments in section 3.2. This is also in line with sensitivity results from Heckman et al. (1998), where key outcomes were unchanged qualitatively. Other results in our model are also insensitive to the value of  $\sigma$ , and even the quantitative differences are minor, except when  $\sigma = 0.5$ . In that case, larger high school wages compensate for the fact that  $\gamma_0 < \gamma_1$ , resulting in less heterogeneity both in the cross-section and across generations.

Our results go through qualitatively when the risk aversion (or intertemporal elasticity of substitution) parameter,  $\chi$ , is set to 1.1 (close to log utility) or 3 as well, but both cases result in physical and human capital investments becoming more correlated quantitatively. In our model, physical

capital investments are safe while human capital investments are risky. Hence bequests serve as a device for parents to insure their children against shocks not realized when child investments are being made. When  $\chi$  is small, this insurance matters less, so parents would rather invest in their children’s human capital rather than wait and leave bequests later in life. Conversely, when  $\chi = 3$ , risky human capital investments are shifted toward savings. Rich parents who are able to invest more in their children also leave larger bequests. Similarly, the IGE of lifetime earnings is larger by 0.1, and more sensitive to the values of  $(\rho_a, \rho_\epsilon)$ , when  $\chi = 1.1$ . The exact opposite is true when and  $\chi = 3$ , with the IGE declining by 0.04.

Relatedly, we also considered the case where parents do not know the exact values of their children’s innate ability, so that they integrate over dispersed values of  $a'$  when making decisions.<sup>33</sup> Qualitatively, this is similar to increasing risk aversion—since human capital investments become more risky from the parents point of view, there is some shifting of resources toward the safer physical capital asset. Quantitatively, however, this has a negligible impact on the IGE. This might be expected, since larger risk aversion did not have a very large impact, in addition to the fact that the moments we compute from the simulated model are still based on the true values.

## 6. Conclusion

In this paper, we present a rich model of parental investments in children by presenting an amalgam of a life-cycle and an intergenerational model of human capital. Our model is a full-fledged dynastic framework which features multiple generations co-existing. Parents make investments in their children’s human capital and also decide how much in bequests to pass onto them. The human capital formation process is rather complex, but remains a parsimonious combination of features that are separately well understood. Consistent with prior work, we assume complementarity between early investments in the form of parental time, and later investments in the form of goods. We also model the college enrollment decision, and model life-cycle wage growth via a standard neoclassical human capital accumulation equation. We cast this environment in an equilibrium setting with various government policies in place. We demonstrate that the resulting framework has stark implications for various moments and demonstrate that we can explain various intergenerational relationships using our environment.

We can match the intergenerational persistence of earnings, wealth, educational attainment

---

<sup>33</sup>Specifically, we assume that parents integrate over a discretized version of equation (4), but that the conditional variance is  $0.4\sigma_\eta^2$ , where  $\sigma_\eta^2$  is the variance of the idiosyncratic shock in our benchmark.

and welfare recipient status as well as inequality in these variables at a point in time. Our economy is consistent with evidence on the persistence of earnings when the sample is split on the basis of whether parents leave bequests or not. While Mulligan views this as evidence as inconsistent with the basic Becker Tomes framework, we find that these relationships are indeed consistent with our model. When we shut down the ability transmission, the IGE falls from 0.6 to 0.5. If we were to categorize three different elements that result in the IGE, nature does not play a very important role, while nurture and economic policy play a substantial role. This is in contrast to many other studies that suggest that nature plays a very important role. We find that educational subsidies are the most effective tool with which to reduce the IGE, surpassing any other policy instrument. We also find that greater progressivity in labor income taxation leads to less inequality and higher intergenerational mobility. While more work needs to be done, our model has the potential to account for US-Europe differences in inequality and mobility by assigning a first-order role to policy differences.

# Appendices

## A. Simplifying the Household's problem

Our assumptions allow the reduction of choice variables and states as follows.

### A.1 Consumption Smoothing problem

Denote the young parent household's total consumption in period  $j$  as  $C_j = c_j + c'_{j-4}$ ,  $j = 4, \dots, 7$ .

Since the young household solves a Pareto problem, we have

$$c_j = \frac{C_j}{1 + \theta^{\frac{1}{\chi}}}, \quad c'_{j-4} = \frac{\theta^{\frac{1}{\chi}} C_j}{1 + \theta^{\frac{1}{\chi}}}$$

which allows us to define a household utility function

$$U(C_j) = qu(C_j), \quad q = \left(1 + \theta^{\frac{1}{\chi}}\right)^\chi.$$

Similarly, consumption is smoothed in all periods where there is no realization of a new shock, namely throughout the young household period  $j = 4, 5, 6, 7$  and old adult period  $j = 8, \dots, 12$ . For both stages we can define a utility function in terms of the PDV of total consumption within each stage:

$$U_y(C_y) = q_y u(C_y), \quad q_y = q \left[ \sum_{j=4}^7 \left( \beta^{\frac{1}{\chi}} (1 + \tilde{r})^{\frac{1}{\chi} - 1} \right)^{j-4} \right]^\chi, \quad C_y = \sum_{j=4}^7 \frac{C_j}{(1 + \tilde{r})^{j-4}} = \left( \frac{q_y}{q} \right)^{\frac{1}{\chi}} C_4$$

$$U_o(C_o) = q_o u(C_o), \quad q_o = \left[ \sum_{j=8}^{12} \left( \beta^{\frac{1}{\chi}} (1 + \tilde{r})^{\frac{1}{\chi} - 1} \right)^{j-8} \right]^\chi, \quad C_o = \sum_{j=8}^{12} \frac{c_j}{(1 + \tilde{r})^{j-8}} = q_o^{\frac{1}{\chi}} c_8.$$

### A.2 Lifetime Income Maximization problem

Given the timing of shocks and choices, the parent's own human capital decision problem is deterministic and can be solved as a lifetime income maximization problem at age  $y + 1 = 5$ . Given  $\mathbf{x}_{y+1}$ , the PDV of lifetime earnings is

$$z_{h,y+1}(\mathbf{x}_{y+1}) = \max_{\{n_{y+l}\}_{l=1}^6} \left\{ \sum_{l=1}^6 \frac{1}{(1 + \tilde{r})^{l-1}} \cdot \left[ 1 - \tau_e(e_{y+l}) - \tau_s + \frac{(2 + \tilde{r})p_1}{6(1 + \tilde{r})^{8-l}} \right] e_{y+l} \right\},$$

where

$$e_{y+l} = w_s h_{y+l} \epsilon (1 - n_{y+l}),$$

$$h_{y+l+1} = a(n_{y+l} h_{y+l})^{\gamma_s} + h_{y+l}.$$

This problem can be easily solved recursively. Let

$$\tilde{y}_{y+l} = \left[ 1 - \tau_e(e_{y+l}) - \tau_s + \frac{(2 + \tilde{r})p_1}{6(1 + \tilde{r})^{8-l}} \right] e_{y+l},$$

then given  $n_{10} = 0, z_{h,10} = \tilde{y}_{10}$ ,

$$z_{h,j}(\mathbf{x}_j) = \max_{n_j} \left\{ \tilde{y}_j + \frac{1}{1 + \tilde{r}} \cdot z_{h,j+1}(\mathbf{x}_{j+1}) \right\}$$

for  $j = 5, \dots, 9$ . Hence the young parent can take  $z_{h,y+1}$  as given.

### A.3 Reformulation of the value functions

We now reformulate the young and old's value functions. The grandparent of stage  $o = 8$  solves

$$V_o(a''; S', a', \epsilon', h'_y; z_o) = \max_{C_o, z'_y} \left\{ q_o u(C_o) + \theta V_y(a''; S', a', \epsilon', h'_y; z'_y) \right\}$$

$$C_o + z'_y = z_o$$

$$z'_y \geq 0,$$

where the new state  $z_o$  summarizes the lifetime PDV of savings and earnings carried over from when young, and  $b' = (1 + r)^3 z'_y$  are the implied bequests. The young household of stage  $y = 4$  decides the child's skill level

$$V_y(a'; S, a, \epsilon, h_y; z_y) = \max_{S' \in \{0,1\}} \left\{ W_y(S', a'; S, a, \epsilon, h_y; z_y) \right\},$$

where

$$W_y(S', a'; S, a, \epsilon, h_y; z_y) =$$

$$\max_{C_y, z, n_y, n_p, m_p, n_k} \left\{ q_y u(C_y) + \beta^4 \int_{a'', \epsilon'} V_o(a''; S', a', \epsilon', h'_y; z_o) dF(a'', \epsilon' | a') \right\}$$

$$\begin{aligned}
C_y + \frac{z_o}{(1+\tilde{r})^4} + \frac{m_p}{(1+\tilde{r})^2} &= f(e_y) + \frac{z_{h,y+1}(S, a, \epsilon, h_{y+1})}{1+\tilde{r}} + \frac{f(e_k) - S' \cdot \kappa}{(1+\tilde{r})^3} + z_y + G, \\
e_y &= w_S h_y \epsilon (1 - n_y - n_p) \quad \text{and} \quad n_y, n_p \geq 0, \quad n_y + n_p \leq 1 \\
e_k &= w_{S'} h_k (1 - n_k) \quad \text{and} \quad n_k \in \left[ \frac{2S'}{3}, 1 \right], \\
h_{y+1} &= a(n_y h_y)^{\gamma_S} + h_y, \quad h'_y = a' n_k^{\gamma_{S'}} h_k^{\gamma_P}, \\
h_k &= \left[ \gamma_k^{\frac{1}{\phi}} (m_p + d)^{\frac{\phi-1}{\phi}} + (1 - \gamma_k)^{\frac{1}{\phi}} (n_p h_y)^{\frac{\phi-1}{\phi}} \right]^{\frac{\phi}{\phi-1}},
\end{aligned}$$

and  $(e_k, h_k, n_k)$  denote, respectively, the child's earnings, human capital level, and time spent in human capital accumulation in college (period 3). The last term in the budget constraint  $G$  is

$$G = g \cdot \left[ \sum_{l=y}^{y+2} \frac{A}{(1+\tilde{r})^{l-y}} + \frac{2}{(1+\tilde{r})^3} + \sum_{l=0}^{o+4} \frac{1}{(1+\tilde{r})^{l-y}} \right] + \frac{p_0(2+\tilde{r})}{(1+\tilde{r})^8},$$

is the lifetime PDV of subsidies.

## B. Solving the Household's problem

### B.1 Grandparent's problem

When solving the grandparent's problem, notice that  $V_y$  will have kinks according to the grandchild's college decision. However, we know exactly where those kinks are, i.e. we know the  $\hat{z}_y(a''; \mathbf{x}'_y) \geq 0$  s.t.

$$W_y(0, a''; \mathbf{x}'_y; \hat{z}_y(a''; \mathbf{x}'_y)) = W_y(1, a''; \mathbf{x}'_y; \hat{z}_y(a''; \mathbf{x}'_y)),$$

so we can search for the solution before and after the kink and choose the max. The necessary condition for the grandparent's optimization problem is

$$q_o u'(C_o) \geq \theta V_{y,6}(a''; S', a', \epsilon', h'_y; z'_y), \quad \text{with equality if } z'_y > 0, \tag{8}$$

and by the envelope theorem

$$V_{o,6}(a''; S', a', \epsilon', h'_y; z_o) = q_o u'(C_o).$$

The simplification also makes clear why parents who leave bequests are not necessarily those who were not able to invest efficiently in their children. In fact, inspection of the grandparent's simplified problem reveals that all else equal, parents with high ability and/or human capital children will leave *less* bequests (as long as the value function is monotone).

## B.2 Young parent's problem

Note that due to standard Inada conditions, in an optimum  $[z_o, n_y, n_p, n_k] > 0$ . The possible corner solutions are when  $n_y + n_p = 1$  and/or  $n'_k$  hitting either boundary; and/or  $m_p = 0$ , which happens when the subsidy  $d$  is larger than the private optimum.

**Euler Equations** The optimality condition for  $z_o$  implies

$$q_y u'(C_y) = [\beta(1 + \tilde{r})]^4 J_4 \quad (9)$$

where

$$J(S', a', h'_y; z_o) = \int_{a'', \epsilon'} V_o(a''; S', a', \epsilon', h'_y; z_o) dF(a'', \epsilon' | a'),$$

while the envelope theorem implies

$$W_{y,7}(S', a'; S, a, \epsilon, h_y; z_y) = q_y u'(C_y).$$

The first order conditions w.r.t.  $n_k$  in an interior are

$$q_y u'(C_y) \cdot y'(e_k) w_{S'} h_k = \beta^4 (1 + \tilde{r})^3 J_3 \cdot \frac{\gamma_{S'} h'_y}{n_k}.$$

It may well be the case that  $n_k = 1$ , or  $\frac{2}{3}$  for college kids, in which case the equality does not hold.

**Human Capital Formation** The parent is forward-looking and takes the child's choices into consideration. The optimality conditions when  $n_y + n_p < 1$  are (note that  $n_y, n_p > 0$  strictly due to Inada conditions):

$$n_y : y'(e_y) w_{S'} \epsilon = \frac{z_{h,y+1,A}}{1 + \tilde{r}} \cdot \frac{\gamma_{S'} a (n_y h_y)^{\gamma_S}}{n_y h_y}, \quad (10)$$

$$n_p : q_y u'(C_y) \cdot y'(e_y) w_S \epsilon = \beta^4 J_3 \cdot \frac{\gamma_p h'_y}{h_k} \cdot \left[ \frac{(1 - \gamma_k) h_k}{n_p h_y} \right]^{\frac{1}{\phi}},$$

$$m_p : q_y u'(C_y) \geq \beta^4 (1 + \tilde{r})^2 J_3 \cdot \frac{\gamma_p h'_y}{h_k} \cdot \left( \frac{\gamma_k h_k}{m_p + d} \right)^{\frac{1}{\phi}}, \quad \text{with equality if } m_p > 0.$$

If  $n_y + n_p = 1$ , the parent divides time between himself and his child so that the returns are equal:

$$q_y u'(C_y) \cdot \frac{z_{h,y+1A}}{1 + \tilde{r}} \cdot \frac{\gamma_S a(n_y h_y)^{\gamma_S}}{n_y h_y} = \beta^4 J_3 \cdot \frac{\gamma_p h'_y}{h_k} \cdot \left[ \frac{(1 - \gamma_k) h_k}{n_p h_y} \right]^{\frac{1}{\phi}}. \quad (11)$$

The choice for  $(m_p, n_k)$  can be simplified as:

$$z_{h,y+1A} \cdot \frac{\gamma_S a(n_y h_y)^{\gamma_S}}{n_y h_y} = \frac{1}{1 + \tilde{r}} \left[ \frac{1 - \gamma_k}{\gamma_k} \cdot \frac{m_p + d}{n_p h_y} \right]^{\frac{1}{\phi}} \quad \text{if } m_p \geq 0, \quad (12)$$

$$= \frac{\gamma_p}{\gamma_S} \cdot \frac{y'(e_k) w_S n_k}{(1 + \tilde{r})^2} \cdot \left[ \frac{(1 - \gamma_k) h_k}{n_p h_y} \right]^{\frac{1}{\phi}} \quad \text{if } n'_k \leq 1. \quad (13)$$

Basically, all this says is that in an interior solution, the optimal solution for  $m_p$  is such that the time-goods investment ratio in the child's human capital formation is a power of the costs; and the college-age student's time cost per unit of stage  $y$  human capital must equal the parent's.

## C. Solution Method

Our SMM is a nested fixed point problem. For each  $\Theta$ , we obtain individual decision rules, the stationary distribution, and find  $(\beta, \mu, \bar{\epsilon})$  that solve (1)-(3). For given  $\Theta$ , we compute:

1. Guess  $(\beta, \mu, \bar{\epsilon})$ .
2. Obtain individual choices.
  - (a) Outside value function iteration, solve (A.2). For each state, store all values of  $\{h_j, n_j\}_{j=5}^{10}$ . Obtain  $z_{h,y+1}, z_{h,y+1A}$ .
  - (b) Guess  $J, J_4$ , and solve for  $W_y, W_{y,7}$  as follows. In the most outer-loop, solve for  $(n_y, n_p)$  using Brent's method:
    - i. Assuming  $n < 1$ , solve for  $n_p$ . Note that  $n_y$  can be solved for without the value function by (10).
    - ii. Optimize over  $n_p$  by (11)—(12) and (13) give closed form solutions for  $(m_p, n_k)$ .

- iii. Solve for  $z_o$  using the Euler Equation (9).
  - iv. Now resolve the problem assuming  $n_y + n_p = 1$  ( $n_y = 1 - n_p$ ), and solve for  $n_p$ .  
Check whether or not  $n = 1$  by comparing values from each case.
- (c) Obtain  $\hat{z}_y(a'; \mathbf{x}_j)$ : interpolate to find where  $W_y(0, a'; \mathbf{x}_j; \hat{z}_y) = W_y(1, a'; \mathbf{x}_j; \hat{z}_y)$ .
  - (d) Solve grandparent's problem (choose  $z'_y$ ) using Euler Equation (8).
  - (e) Integrate over Store  $V_o, V_{o,6}$  w.r.t.  $(a'', \epsilon')$  to obtain  $J', J'_4$ .
  - (f) Repeat (2b) until  $(J', J'_4) \approx (J, J_4)$ .
3. Due to the size of the state space, we obtain the stationary distribution via Monte-Carlo simulation rather than directly approximating the distribution.
- (a) Simulate  $N = 120,000$  households for  $T = 200$  generations, using decision rules obtained above.
  - (b) Obtain implied life-cycle values of  $\{n_j, h_j\}_{j=4}^{10}$  from (2a). From this, we can compute each individual's consumption, savings and earnings decisions.
  - (c) Aggregate to obtain implied  $(r, EP, \bar{\mathbf{e}}')$ .
4. Iterate from (1) until  $(r, EP, \bar{\mathbf{e}}) \approx (4\%, 46.8\%, \bar{\mathbf{e}}')$ .

For each  $\Theta$ , we solve (7) with equal weights on all moments. In each iteration, the equilibrium vector  $(\beta, \mu, \bar{\mathbf{e}})$  is solved for by numerical iteration. The choice of  $(N, T)$  is arbitrary, but increasing to  $N = 240,000, T = 300$  had no effects. Further details are available upon request.

## References

- Aiyagari, S. R. (1994, August). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics* 109(3), 659–684.
- Aizer, A. and F. Cunha (2012). The production of human capital: Endowments, investment, and fertility. Technical report. NBER Working Paper No. 18429.
- Banerjee, A. V. and A. F. Newman (1993, April). Occupational choice and the process of development. *Journal of Political Economy* 101(2), 274–98.
- Becker, G. S. and N. Tomes (1979, December). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87(6), 1153–89.
- Becker, G. S. and N. Tomes (1986, July). Human capital and the rise and fall of families. *Journal of Labor Economics* 4(3), S1–39.
- Behrman, J. R., A. D. Foster, M. R. Rosenzweig, and P. Vashishtha (1999). Women’s schooling, home teaching, and economic growth. *Journal of Political Economy* 107(4), pp. 682–714.
- Behrman, J. R., P. Taubman, and T. J. Wales (1977). Controlling for and measuring the effects of genetics and family environment in equations for schooling and labor market success. In P. Taubman (Ed.), *Kinometrics: Determinants of Socioeconomic Success within and between Families*. North-Holland.
- Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy* 75, 352.
- Benabou, R. (1993, August). Workings of a city: Location, education, and production. *The Quarterly Journal of Economics* 108(3), 619–52.
- Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003). Accounting for the u.s. earnings and wealth inequality. *Journal of Political Economy* 111, 818–857.
- Caucutt, E. M. and L. Lochner (2012, October). Early and late human capital investments, borrowing constraints, and the family. Working Paper 18493, National Bureau of Economic Research.
- Charles, K. K. and E. Hurst (2003, December). The correlation of wealth across generations. *Journal of Political Economy* 111(6), 1155–1182.

- Cunha, F. and J. J. Heckman (2007, May). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- Cunha, F., J. J. Heckman, and S. M. Schennach (2010, 05). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3), 883–931.
- Del Boca, D., C. Flinn, and M. Wiswall (2013, April). Transfers to households with children and child development. Working Papers 2013-001, Human Capital and Economic Opportunity Working Group.
- Del Boca, D., C. Monfardini, and C. Nicoletti (2012, March). Children’s and parents’ time-use choices and cognitive development during adolescence. Working Papers 2012-006, Human Capital and Economic Opportunity Working Group.
- Diamond, P. and J. Gruber (1999). Social security and retirement in the united states. In *Social Security and Retirement around the World*, NBER Chapters, pp. 437–473. National Bureau of Economic Research, Inc.
- Durlauf, S. N. (1996, March). A theory of persistent income inequality. *Journal of Economic Growth* 1(1), 75–93.
- Galor, O. and J. Zeira (1993, January). Income distribution and macroeconomics. *Review of Economic Studies* 60(1), 35–52.
- Gayle, G.-L., L. Golan, and M. A. Soytaş (2013). What accounts for the racial gap in time allocation and intergenerational transmission of human capital? Technical report.
- Gentzkow, M. and J. M. Shapiro (2013). Measuring sources of identification in nonlinear econometric models. Technical report, Chicago Booth.
- Goldberger, A. S. (1989, June). Economic and mechanical models of intergenerational transmission. *American Economic Review* 79(3), 504–13.
- Gottschalk, P. (1990, May). Afdc participation across generations. *American Economic Review* 80(2), 367–71.
- Gravelle, J. G. (2007). Historical effective marginal tax rates on capital income. Technical report, Congressional Research Service Report for Congress.

- Griliches, Z. (1979, October). Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy* 87(5), S37–64.
- Guner, N., R. Kaygusuz, and G. Ventura (2013, July). Income taxation of u.s. households: Facts and parametric estimates. Working Papers 705, Barcelona Graduate School of Economics.
- Haider, S. and G. Solon (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review* 96(4), 1308–1320.
- Han, S. and C. B. Mulligan (2001, April). Human capital, heterogeneity and estimated degrees of intergenerational mobility. *Economic Journal* 111(470), 207–43.
- Heckman, J., L. Lochner, and C. Taber (1998, January). Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics* 1(1), 1–58.
- Heckman, J. J. (2008, 07). Schools, skills, and synapses. *Economic Inquiry* 46(3), 289–324.
- Heckman, J. J., L. Lochner, and R. Cossa (2002). Learning-by-doing vs. on-the-job training: Using variation induced by the eitc to distinguish between models of skill formation. Technical report.
- Hendricks, L. (2007, 05). Retirement wealth and lifetime earnings. *International Economic Review* 48(2), 421–456.
- Hubbard, R. G., J. Skinner, and S. P. Zeldes (1995, April). Precautionary saving and social insurance. *Journal of Political Economy* 103(2), 360–99.
- Huggett, M., G. Ventura, and A. Yaron (2011, December). Sources of lifetime inequality. *American Economic Review* 101(7), 2923–54.
- Leonesio, M. V. and L. D. Bene (2011). The distribution of annual and long-run us earnings, 1981-2004. *Social Security Bulletin* 71(1).
- Loury, G. C. (1981, June). Intergenerational transfers and the distribution of earnings. *Econometrica* 49(4), 843–67.
- Mazumder, B. (2005, July). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *The Review of Economics and Statistics* 87(2), 235–255.

- McGarry, K. (1999, September). Inter vivos transfers and intended bequests. *Journal of Public Economics* 73(3), 321–351.
- Mulligan, C. B. (1999, December). Galton versus the human capital approach to inheritance. *Journal of Political Economy* 107(S6), S184–S224.
- Plug, E. and W. Vijverberg (2003, June). Schooling, family background, and adoption: Is it nature or is it nurture? *Journal of Political Economy* 111(3), 611–641.
- Restuccia, D. and C. Urrutia (2004). Intergenerational persistence of earnings: The role of early and college education. *American Economic Review* 94, 1354–1378.
- Solon, G. (1992, June). Intergenerational income mobility in the united states. *American Economic Review* 82(3), 393–408.