Nature and Nurture in the Transmission of Economic Status

Sang Yoon (Tim) Lee
Department of Economics
University of Mannheim

Ananth Seshadri
Department of Economics
University of Wisconsin-Madison

August 29, 2012

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). We find that we can explain a wide range of phenomena including cross sectional inequality and the intergenerational mobility of earnings, consumption, education and wealth. We use the model to isolate the significance of exogenous ability transmission (nature) from other endogenous transmission mechanisms (ability to pay, parent’s human capital) while allowing for natural and environmental forces to interact in a non-linear fashion. Contrary to previous studies, we find that endogenous mechanisms are just as significant as exogenous ones. Consequently, market incompleteness, education subsidies, and social security policies also affect the persistence in economic status across generations.

∗Email: sylee.tim@uni-mannheim.de and aseshadr@ssc.wisc.edu, respectively. We are grateful to Gary Becker, Steven Durlauf, Arthur Goldberger, Nicolas Schutz and Chris Taber for useful discussions.
1. Introduction

The intergenerational elasticity (IGE) of parents’ lifetime earnings with respect to children’s lifetime earnings is 0.6 in the United States.\(^1\) If this degree of correlation is due to natural causes such as the correlation in innate abilities, then there is less room for policy to affect socioeconomic mobility without reduced efficiency. On the other hand, if this level of correlation is due to market imperfections wherein a parent’s ability to earn affects his ability to invest in his child, then economic policy would matter for socioeconomic mobility. Disentangling the contribution of nature from nurture in generating persistence of socioeconomic status across generations has been the subject of much discussion and debate in the social sciences.

*The empirical approach:* Behavioral geneticists have estimated the heritability of IQ and a robust finding in this literature is that while genetic factors explain 50 to 60 percent of the variation in IQ, family background explains little of the variation. While this finding is robust, the interpretation differs widely. Herrnstein and Murray (1994) and Jensen (1972) interpret the small effect of family background to imply that socioeconomic policy aimed at improving the environment of children from poor families will have little effect. Goldberger (1979), amongst others, provides reasons why such conclusions are unwarranted. The methods used for determining heritability typically assume additivity between natural and environmental forces. But family background is endogenous and will typically depend on genes, thus making part of nurture difficult to isolate from nature by simply looking at variations. Also, the determinants of IQ are not the same as the determinants of education or earnings, which are better measures of socioeconomic status. Even if we were to know that much of the existing variance is due to genetic factors, it does not follow that the status of a child is predetermined or that it cannot be altered by socioeconomic policy. To get around such issues, more recent work in the field of epigenetics studies the interaction between nature and nurture. An important finding of this nascent literature is the finding of strong family background effects on heritability coefficients.

There is also a large empirical literature in economics that attempts to disentangle the effect of nature from nurture, especially pertaining to the transmission of educational attainment across generations.\(^2\) Since high income parents also tend to be high ability parents, the main goal of such studies is to separately identify the effect of family income and parental ability on children’s educational outcomes such as years of schooling or college attendance. More recent studies fur-

---

\(^1\)This estimate is from Mazumder (2005). Earlier work estimates an IGE of 0.4, e.g. Solon (1992). The primary interest in most studies is the IGE across lifetime earnings. Direct measurement requires a large panel of a representative sample that spans at least two generations. Since such data sets are scant, various authors have employed different identification techniques to infer the magnitude of the IGE of lifetime earnings from earnings data spanning shorter time periods. The consensus is that the estimated IGE is larger as the time horizon used in the analysis becomes longer, because longer horizons allows one to better average out the transitory components of earnings. Mazumder (2005) finds that when increasing the horizon of measured earnings, the IGE between parent and child increase, up to 0.6 at the 15 year horizon.

\(^2\)Some examples are Behrman et al. (1977); Plug and Vijverberg (2003); Behrman and Rosenzweig (2005).
her emphasize the importance of parental spillovers. 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 and parental spillovers.

While the literature has made progress 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 - as Heckman (2008) points out, “genes and environment cannot be meaningfully parsed by traditional linear models that assign unique variances to each component.” Furthermore, by focusing exclusively on educational attainment, one can measure the relative effects of ability transmission and other parental spillovers, but not the efficacy of education itself. While differences in educational attainment no doubt have a large impact on differences in human capital—which we take to measure socioeconomic status—a better measure of human capital is one that can be inferred from earnings observations. Finally, one cannot evaluate the effectiveness of policies on the basis of empirical studies that hold fixed the degree of transmission. Policies are likely to alter choices that in turn affect the degree of intergenerational transmission, especially given the long time horizon. Hence the equilibrium relationship between the current and future generations’ economic outcomes will vary when policies are altered.

The theoretical approach: The theoretical work in this literature gained traction beginning with the seminal work of Becker and Tomes (1979, 1986). They laid out a simple but powerful two period equilibrium model to derive implications for the intergenerational transmission of economic status. There are several other important advances in the literature that feature models that generate persistence in economic status. 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. More recently, Cunha and Heckman (2007) argue that early childhood investments lead to persistence across generations.

While the Becker-Tomes framework is used widely in the literature, it has been met with some skepticism, most notably by Goldberger (1989). He argued that the economic approach represented little value added relative to the reduced form approach and that education intensities cannot be differentiated from ability transmission. Mulligan (1999) tests some key empirical implications of the Becker-Tomes model and finds limited support. Most importantly, he shows that borrowing constraints, a key feature of Becker and Tomes (1986), seem not to matter, and even if

---

3We define parental spillovers as elements affecting early childhood environment, such as parenting skills. We explain our approach in greater detail below but one important distinction between our approach and prior approaches is that we endogenize the level of such spillovers.

4Loury (1981) was a similar model in a dynastic setting with borrowing constraints.
they do, are empirically irrelevant. Han and Mulligan (2001) further argues that heterogeneous abilities and borrowing constraints are indistinguishable.

Our approach: The purpose of this paper is to use quantitative economic theory à la the pioneering study of Becker and Tomes (1986). Specifically, we combine a fairly standard life-cycle model of human capital accumulation à la Ben-Porath (1967) with an intergenerational transmission mechanism à la Becker and Tomes (1986). So as noted earlier, both the human capital of the child and the parent, as well as their investments, are endogenous. While we are not the first to present a quantitative model, prior formulations were too simplistic to directly compare with the data. Many studies either assume away important model elements emphasized by Becker-Tomes (e.g. Restuccia and Urrutia (2004) abstract away from asset accumulation which is central to the notion that parents might want to equate marginal rates of return to human capital investment with the real interest rate) or have too few empirical counterparts, making identification difficult. Furthermore, almost all models that rely on Becker-Tomes ignore the importance of individual human capital accumulation over the life-cycle - the degree of cross sectional inequality is already determined once the children become young adults. Indeed, Becker himself had pioneered the role of individual human capital accumulation as another source of inequality, e.g. Becker (1975). Since small differences in the initial conditions of young adults can lead to magnified cross sectional inequality over the life-cycle, ignoring the life-cycle dimension of inequality will overestimate the returns to human capital investment. A final element is to carefully consider all forms of government intervention. The data we observe comes from an economy in which the government taxes progressively, subsidizes education, funds social security and assists low income households through welfare payments. These programs may potentially have important effects on various aspects of behavior and in particular on intergenerational persistence, so are included in the model.

Differentiating intra- and intergenerational human capital production is a key element of the model. Augmenting a dynastic version of Becker-Tomes with a standard life-cycle human capital accumulation à la Ben-Porath (1967) yields a wide range of predictions on statistics such as the intergenerational correlations of earnings, consumption, education and wealth, as well as life-cycle earnings profiles and cross sectional inequality. Combining life-cycle and intergenerational elements is nontrivial, especially since the child’s state variables affect the parent’s decision problem. So we do this in a parsimonious way by assuming that children’s human capital production is identical to adults’ in all aspects except one - their production function is augmented by a parameter governing parental spillovers. This is modeled as the parents’ human capital being a productive input in the child’s human capital production function. This spillover (which is not an externality in the traditional sense since the parent internalizes this effect through his altruism to-

---

5The important work of 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 is more parsimonious.
wards the child) is identified by the relative size of investments in the child and the young parent. While the former determines intergenerational persistence, the latter plays a central role in determining the life-cycle profile. Using the model we are able to pin down the component of earnings transmission that is due to the transmission of innate abilities (nature) and parental spillovers and investments (nurture). Finally, since our structural model is cast in a general equilibrium framework, it is an ideal laboratory in which to conduct policy analysis.

As in Becker and Tomes (1986) and subsequent empirical studies, we find that genetics and parental spillovers are important. Eliminating the exogenous ability transmission and spillovers drop the IGE of earnings from 50% to 2%. But the magnitude of the response of the IGE to genetics transmission do depend on imperfect capital markets and government policies. Our results also demonstrate that nature and nurture interact in subtle ways and that it is not easy to describe the complex relationship through a simple variance decomposition. We find that considering both intra- and intergenerational constraints is key to reconciling the Becker-Tomes model with apparently contradictory empirical findings from Mulligan (1999), who concludes that constraints do not affect the IGE, and Mazumder (2005), who concludes that they do. Our model suggests that intergenerational constraints have a smaller impact on the IGE than intragenerational constraints.

We find that policies that affect the life-cycle, and not just intergenerational motives, will have indirect effects on the IGE. For instance, social security programs that apply only at the end of the life-cycle have a quantitatively significant impact that realigns intra- and intergenerational motives, thereby increasing socioeconomic mobility. Our result that welfare subsidy recipient status 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 children of AFDC families in our model simulations have low enough abilities such that conditional on their abilities and parents, education subsidies allow them to obtain relatively more education than children of non-AFDC families.

Finally, although we are primarily interested in intergenerational transmission rather than the life-cycle, we show 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 prior work shows that the standard life-cycle model over-predicts the correlation between lifetime earnings and retirement wealth, we show that accounting for intra- and intergenerational human capital investment and intergenerational transfers can account for these discrepancies.

In sum, we are able to generate substantial persistence in earnings and are able to assess the nature versus nurture breakdown via our structural model. We do this by avoiding the ad-hoc assumptions delineated in Goldberger (1979) that continue to plague the literature on twin studies. We are able to allow for fairly non-linear interactions between genes and the environment and note here that the environment is affected by family choices which in turn is influenced by government
policy. We show that an economic model can be used to understand and explain the observed intergenerational persistence in not just earnings but also consumption, wealth, schooling and welfare program participation. We find that life-cycle incentives interact with intergenerational incentives to produce outcomes different from the Becker-Tomes model. These outcomes, such as the differential degrees of persistence between rich versus poor households, as well as those between bequest-constrained versus non bequest-constrained households, are consistent with the empirical evidence.

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 an extension discussing the relevance of our model for understanding retirement wealth. Section 5 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

There is a unit measure of households, each of whom goes through four periods of life: childhood, young parent, old parent, and retired. Throughout the analysis, primes will denote next period values, and whenever child-parent variables coexist within a same period, the child’s variable is subscripted with a $k$.

Each household gives birth to one child at the beginning of the second period (ages 20-40), which constitutes the child’s first period (ages 0-20). During the child’s first period of life, the parent makes decisions on his behalf. These decisions involve consumption and investment in human capital, which begins at age 6. The child then enters the second period of his life (at age 20) as a young parent and forms his own household. He begins the second period of life with a human capital level that is an outcome of decisions that the parent made on his behalf, as well as bequests, an intergenerational transfer that the parent (now old) chose for him. His earnings at that point in time are also subject to a stochastic “luck shock” $\epsilon \sim F(\epsilon)$ which remains constant throughout his
lifetime. This shock captures idiosyncratic risk associated with human capital accumulation.

A young parent invests \( m \) units of consumption goods and \( n \) units of time in his own human capital accumulation (this can be thought of as higher education and on-the-job accumulation of human capital), and \( m_k \) units of consumption goods in his child’s human capital. The child invests \( n_k \) units of time in his own human capital accumulation (this may be thought of as corresponding to the schooling period), a decision made by the parent.

A child is born with innate ability \( a_k \), which determines his proficiency at accumulating human capital. The child’s ability is stochastic, and its distribution is a function of the parent’s ability \( a \), i.e. \( a_k \sim A(a_k|a) \). This will capture the “genetic” part of the IGE or the aspect of intergenerational transmission that we label "nature".

Human capital production is governed by the production technology

\[
h_k = a_k (n_k \bar{h})^{\gamma_1} (m_k + d)^{\gamma_2} h^{\gamma_3} + (1 - \delta_h) \bar{h}
\]

for the child, where \( d \) denotes a lump-sum education subsidy received from the government, and

\[
h' = a (nh)^{\gamma_1} m^{\gamma_2} + (1 - \delta_h) h
\]

for the parent. The only difference in these accumulation processes is the presence of \( h^{\gamma_3} \) in the child’s function, which we will elaborate on below.\(^6\)

Here \( \bar{h} \) is the initial human capital endowment which is identical across all households. Hence, heterogeneity across households comes from the market and ability shocks. Notice that there are three mechanisms through which human capital is transmitted across generations. First, ability is transmitted across generations, that is, the ability of a child \( a_k \) is drawn from a distribution that depends on the parent’s ability. Second, the resources that a parent invests in his child \( m_k \) will be a function of the parent’s resources. If capital markets are complete, the investment will be a function of the child's ability. However, with capital market imperfections (which we assume), parental resources will matter in determining \( m_k \). 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, as measured by \( \gamma_3 \). This is intended to capture the impact of early childhood environment - children of high human capital parents are more efficient at learning beyond their natural ability. We label the first mechanism nature and the second and third, nurture.

The young parent saves \( s_y \) for the third period (ages 40-60). The old parent leaves bequests \( b \) to his child, who is now a young parent, and saves \( s_o \) for retirement. When retired, she consumes her savings and social security benefits, and dies at age 80. The sequence of events is depicted in Figure 1.

\(^6\)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 the environment.
We assume a standard neoclassical firm that takes physical and human capital as inputs to produce the single consumption good. There is also a government that levies taxes, which is partially used to subsidize the poor and for education subsidies, and runs a balanced-budget social security program. Capital income is taxed at a flat rate $\tau_k$. Earnings in any period, $e$, are taxed at a progressive rate $\tau_e(e)$ plus a flat rate payroll tax $\tau_s$ that is used to fund the social security benefits. Subsidies to the poor, or welfare programs, are described by $T[\cdot]$. We model the welfare program as a negative income tax $t(z) = \max\{0, g - t_{nit}z\}$, where

$$z = (1 - \tau_s - \tau_e(e))e,$$

i.e. $z$ is earnings after the progressive and payroll taxes. So the final after tax, after subsidy earnings of each household is

$$T[z] = z + t(z) = \max\{z, g + (1 - t_{nit})z\}.$$

Social security benefits are also subject to the progressive earnings tax and welfare subsidies (but not the payroll tax). After tax social security benefits are modeled as a function of $\bar{e}$, the average lifetime earnings net of the progressive earnings tax. We model it as an affine function:

$$SS(\bar{e}) = S_0 + S_1\bar{e},$$

where $(S_0, S_1)$ are parameters governing the social security regime. These benefits are net of the progressive earnings tax, but still subject to the welfare subsidy. Hence the pretax benefits and tax revenue generated from these benefits are, respectively,$^8$

$$S_0 + S_1\left[\frac{e_y + e_o}{2}\right], \quad S_1\left[\frac{\tau_e(e_y) e_y + \tau_e(e_o) e_o}{2}\right].$$

The government runs a balanced budget on these benefits by financing them with the payroll tax revenue from before retirement earnings.

We assume that the economy is at a steady state, and that the initial distribution $A_0$ for $a$ is the stationary distribution of $A(a_k|a)$. The steady state equilibrium and government budget balance conditions are explained in detail below.

2.1 Household’s Problem

We solve the model backwards, starting from the retirement period, recursively.

---

$^7$The rest we assume is thrown in the ocean.

$^8$The benefits are paid out based on $\bar{e}$ rather than pretax earnings directly for computational purposes.
2.1.1 Retirement

There are no decisions to be made, and the individual simply consumes his after tax, after transfer social security benefits $e_R$ and savings $(1 + \tilde{r})s_o$:

$$V_R(e_R, s_o) = \max_{c_R} u(c_R)$$

$$c_R = e_R + (1 + \tilde{r})s_o$$

where $\tilde{r} = (1 + \bar{r})^{20} - 1$, the 20 year compounded effective interest rate, and $\bar{r} = (1 - \tau_k)r$ where $r$ is the pretax annual interest rate.

2.1.2 Old adult

The old adult, or old parent, makes consumption, savings and bequest decisions $(c_o, s_o, b_k)$ given the state of his child (now young adult) $(a_k', a_k, e_k, h_k)$, retirement earnings $e_R$ and current period budget $Z$:

$$V_o(a_k'; a_k, e_k, h_k; e_R, Z) = \max_{c_o, s_o, b_k} \left\{ u(c_o) + \beta V_R(e_R, s_o) + \theta V_y(a_k'; a_k, e_k, h_k, b_k) \right\}$$

$$c_o + s_o + b_k = Z$$

$$s_o, b_k \geq 0$$

where $\theta$ is an altruism parameter which the parent uses to discount his child’s lifetime utility. Because the retirement period is deterministic, the old adult’s problem can be written as

$$V_o(a_k'; a_k, e_k, h_k; e_R, Z) = \max_{c_o, s_o, b_k} \left\{ u(c_o) + \beta u(c_R) + \theta V_y(a_k'; a_k, e_k, h_k, b_k) \right\}$$

$$c_o + s_o + b_k = Z, \quad c_R = e_R + (1 + \tilde{r})s_o$$

$$s_o, b_k \geq 0.$$

2.1.3 Young adult

This is the important part of the decision problem. A young adult’s state is characterized by his child’s ability level and his own $(a_k, a)$, his human capital $h$, labor market luck shock $\epsilon$, and bequests from his (now old) parent $b$. His problem is given by

$$V_y(a_k; a, e, h, b) = \max_{c_y, m_k, n_k, s_y} \left\{ u(c_y/q) + \beta \int_{a_k'} \int_{e_k} V_o(a_k'; a_k, e_k, h_k; e_R, Z) dF(e_k) dA(a_k'|a_k) \right\}$$

$$c_y + m_k + s_y = T [(1 - \tau_s - \tau_e(e_y))e_y] + wh(1 - n_k) + b$$

$$e_y = \text{wh}(1 - n) - m, \quad e_o = \text{wh} e$$

$$Z = T [(1 - \tau_s - \tau_e(e_o))e_o] + (1 + \tilde{r})s_y$$
\[ \bar{e} = \frac{(1 - \tau(e_y))e_y + (1 - \tau(e_o))e_o}{2}, \quad e_R = T[SS(\bar{e})] \]
\[ e_y, Z \geq 0, \quad n \in [0, 1], \quad n_k \in [0, \bar{n}] \]

where his own human capital and his child’s human capital evolve according to the production technologies specified above.

The bequests \( b \) that the old parent leaves for the young parent is assumed to be decided after observing the young parent’s state but before the young parent makes any decisions.\(^9\) The young parent chooses current period consumption \( c_y \), human capital production inputs for himself \((m, n)\) and his child \((m_k, n_k)\), and savings for next period \( s_y \), which is subject to a natural borrowing constraint - i.e., the parent can only save against his future (after tax) earnings, but not against his social security benefits nor his children’s future 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. \( q \) is an adult equivalent measure that reflects that adjusts for changes in family size - when young, the household consists of the parent and the child. The education subsidy is only transferred if the parent chooses to invest something in his child.\(^10\)

The child can work or receive education for only a fraction \( \bar{n} < 1 \) of childhood. The parameter \( \gamma_3 \) plays an important role - it captures early childhood environment factors that affect the child during his \( 20(1 - \bar{n}) \) years of childhood. In the calibration we assume \( \bar{n} = 14/20 \), which means that until age 6 the child is an infant during which he is affected by the parent only through \( \gamma_3 \).\(^11\)

We assume that all adult education expenses are paid for by the employer, so is fully reflected in measured earnings when young, \( e_y \). In addition to the borrowing constraints, we assume \( e_y \geq 0 \) so that measured earnings cannot be negative.\(^12\) If \( m \) is interpreted as college expenses, it means that earnings in this period must be large enough to pay for any potential college debt, while the cost of college itself is subsidized by the government (both by reducing taxable earnings and being subject to the welfare subsidies). \( e_o \) is measured earnings for the old parent.

The tax rates on capital income and earnings introduces a clear disincentive to invest. Not as clear is the disincentive created by the welfare program. Absent social security benefits, the natural borrowing constraint creates a disincentive for the parent to work when young whenever he finds himself subject to the welfare subsidy. Similarly, if his earnings when old are subject to the subsidy, there is a disincentive to accumulate his own human capital. However, these disincentives are countered by the social security benefits, which is an increasing function of earnings when both young and old.

Due to the constraints and government policies, the objective functions are non-convex. We

---

\(^9\)This is a restriction imposed by the length of a period being 20 years.

\(^10\)In other words, a child with \( n_k = 0 \) receives no subsidies.

\(^11\)For example, a high \( \gamma_3 \) would imply that the child’s early-age environment is important for future earnings.

\(^12\)In the absence of this constraint, because of the welfare program, poor young adults may invest abnormally large amounts in himself and increase their old age earnings arbitrarily.
explain in the appendix how we solve for the value functions and optimal policies.

2.2 Firm and Stationary Equilibrium

We assume a profit maximizing representative firm with a standard neoclassical technology. It solves

$$
\max_{K, H} F(K, H) - RK - wH,
$$

where $R = (1 + r + \delta)^{20} - 1$ is the competitive rental rate and $(K, H)$ are the aggregate quantities of capital and effective units of labor in the economy, respectively.

Let $X = [X_y, X_o]$ denote the aggregate state spanning two generations, where $(X_y, X_o)$ are the young and old parent’s state vectors, respectively, 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 $(r, w)$ solve

1. Market clearing and stationarity:

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

$$
K = \int_X (s^*_y + s^*_o) \Phi(dX)
$$

$$
H = \int_X \left[ \bar{h} (\bar{n} - n^*_k) + h^*_k (1 - n^*) \epsilon^*_y + h^*_o \epsilon^*_o \right] \Phi(dX).
$$

where $(s^*_y, s^*_o)$ are the optimal savings decisions of the young and old adults, $(h^*_k, h^*_o)$ and $(n^*_k, n^*)$ the optimal human capital and time input decisions of the young adult, and $(\epsilon^*_y, \epsilon^*_o)$ the luck shocks of a parent and child of the same family.

2. Government budget balance:

$$
w_{ratio} \times G = \int_X \left[ t(z^*_y) + t(z^*_o) \right] \Phi(dX)
$$

$$
e_{ratio} \times G = \int_X \mathbb{1}\{n^*_k > 0\} \Phi(dX),
$$

where

$$
G = \tau_k RK + \int_X \left[ \tau_t(e^*_y) e^*_y \left( 1 + \frac{S_1}{2} \right) + \tau_t(e^*_o) e^*_o \left( 1 + \frac{S_1}{2} \right) \right] \Phi(dX)
$$

$$
z_x = (1 - \tau_s - \tau_t(e^*_x)) e^*_x, \quad x \in \{y, o\}
$$

and $w_{ratio}$ and $e_{ratio}$ are the welfare and education expenditures over total government rev-
enue ratio, respectively. The social security regime is also balanced:

$$\tau_s \int_X (e_y^* + e_o^*) \Phi(dX) = S_0 + S_1 \int_X \left[ \frac{e_y^* + e_o^*}{2} \right] \Phi(dX).$$

### 3. Calibration

The model is calibrated to 1990 United States. For those moments where the data points for 1990 are not available, we take data as close as possible to 1990.

#### 3.1 Parametrization

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

$$\log a_k = \mu + \rho \log a + \eta, \quad \eta \sim \mathcal{N}(0, \sigma_\eta^2)$$

where $\mu$ is the mean level of log abilities and $\sigma_\eta$ is the conditional variance of abilities. For luck shocks assume:

$$e \sim \log \mathcal{N}(0, \sigma_e^2).$$

Preferences are modeled as the standard CRRA utility function,

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

and the aggregate production function and capital stock evolution are also standard:

$$F(K, H) = K^\alpha H^{1-\alpha},$$

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

The average tax rate function for earnings that we use follows Gouveia and Strauss (2000) and is given by

$$\tau_e(e) = \tau_e^0 [1 - (\tau_e^2 e^{\tau_e^2} + 1)^{-1/\tau_e}]$$

where $\tau_e^0$ is the asymptotic maximum tax rate, and $\tau_e^2$ is a level variable that normalizes units.

Consequently the model has a total of 29 parameters, of which 14 are related to government policies. Among these 14 parameters, 9 are fixed according to the data while 5 are calibrated. Of the 15 model parameters, 7 are fixed based on existing studies and the remaining 8 are calibrated along with government variables to the moments explained in the following subsection.

Table 1 summarizes the values of the government parameters and variables that are used to calibrate the model. Those in the upper panel are either calibrated or found in equilibrium to
maintain budget balance, while those in the lower panel are fixed. The tax rate on capital income, \(\tau_k\), is the 1990 value from Gravelle (2007)’s study on effective marginal tax rates on capital income. The tax rate on social security benefits, \(\tau_s\), and average earnings tax rate \(\tau_e\), to which \(\tau_e^2\) is calibrated to, is from Wallenius (2009). \((\tau_e^0, \tau_e^1)\) are from Gouveia and Strauss (2000).

The values of \((w_{ratio}, e_{ratio})\) are obtained from the Budget of the United States Government, while \(t_{nii}\) is the AFDC deduction rate for 1968-1981. The values of \((g, d, S_0)\) are relative to equilibrium average earnings. The welfare subsidy \(g\) and education subsidy \(d\) are calibrated in the model so that welfare and education subsidies in the model are exactly \((w_{ratio}, e_{ratio})\), respectively. The social security variables \((S_0, S_1)\) are chosen to match a replacement rate of 40% (Diamond and Gruber (1999)) and balance the social security budget.\(^{13}\)

Of the remaining 15 parameters, 7 are fixed as in Table 2. The values for \(\beta, \sigma, \alpha\) and \(\delta\) are fairly standard values from the literature while \(\bar{h} = 1\) is a normalization. The zero depreciation rate on human capital is also a standard assumption in the literature.

### 3.2 Targets and Accounting

The remaining parameters are calibrated to match the following moments: the annual interest rate, average years of schooling, education expense to GDP (EDU-GDP) ratio, average earnings growth from ages 21-40 to 41-60, Gini coefficient of lifetime earnings, Gini coefficient of wealth at age 60, intergenerational wealth transfers to net worth (WT-NW) ratio, and finally the IGE of lifetime earnings. Below we explain how we obtain these moments from the model, followed by the relevant data sources.

Let \(E[X]\) denote the expectation of \(x\) with respect to the steady state distribution. Our simulated data moments that are targeted to match the 1990 U.S. are:

1. Annual interest rate:

\[ r. \]

\(^{13}\)While not targeted, \(S_1\) is close to the social security replacement rate above $4,288 of 15%.
Table 3: Calibrated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$\mu$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\gamma_3$</th>
<th>$\sigma_x$</th>
<th>$\sigma_\eta$</th>
<th>$\theta$</th>
<th>$\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>18.749</td>
<td>0.040</td>
<td>0.512</td>
<td>0.293</td>
<td>0.572</td>
<td>0.349</td>
<td>0.693</td>
<td>0.600</td>
</tr>
</tbody>
</table>

2. Average years of schooling: $\mathbb{E}[n_k] \times 20.$

3. Aggregate education expenditures over GDP: $\mathbb{E}[m_k + d]/K^\alpha H^{1-a}.$

4. Ratio between average earnings from age 21-40 to 41-60, $\mathbb{E}[e_o]/\mathbb{E}[e_y].$

5. Gini coefficient of lifetime earnings: distribution of $e = e_y + \frac{e_o}{1+\eta}$ computed at $\Phi$.

6. Gini coefficient of wealth at age 60: distribution of $s^*_o$ computed at $\Phi$.

7. Total intergenerational wealth transfers over total net worth: $\mathbb{E}[b_k]/\mathbb{E}[s_y + s_o + b_k].$

8. IGE of lifetime earnings: the regression coefficient of log $e'$ on log $e$.

To compute the equilibrium we fix the interest rate at $r = 4\%$ and iterate on $\mu$ to find the value that clears the capital market. Intuitively, a higher $\mu$ leads to a smaller supply of physical assets, which results in a relative increase in the rental price $R$.\textsuperscript{14} Hence we iterate on $\mu$ until $1 + R = (1 + r + \delta)^{20}$, and calibrate the remaining 7 parameters to match the empirical counterparts of the 7 moments listed above. See Appendix B for computational details. The parameters that give us the best match between the data moments and simulated moments are presented in Table 3. Given the computational complexity involved in solving the model, the fit between model and data moments is fairly good. Clearly changing any one parameter alters all the moments, but some moments are more naturally associated with a select few parameters as explained below and more formally in Appendix C where we discuss the theoretical implications of a simpler model.

**Education efficiency and earnings profile.** All else equal, the coefficient on the time input in the human capital production function, $\gamma_1$, is largely determined by its ability to match an average years of schooling of 12.92 years. The coefficient on expenditures $\gamma_2$ is pinned down in large measure by the ratio of aggregate educational expenditures to GDP of 6.85%. (Both computed from sny (1991).) The sum, $\gamma = \gamma_1 + \gamma_2$, determines the overall return on human capital investment, and hence also affects the average earnings growth profile, which is 1.2 in NLSY79.

Given $\gamma$, the earnings profile is identified through $\gamma_3$, because a larger $\gamma_3$ increases investment in children as opposed to the young parent, which decreases the parent’s earnings when old. In turn, the child grows up into a young adult with larger human capital, which due to the decreasing returns to human capital investment implies less investment in himself. These equilibrium

\textsuperscript{14}Refer to Proposition 1 in Appendix C for a more formal argument.
The transmission of abilities is exogenous throughout generations, while the transmission of human capital, through \( \gamma_3 \), is endogenous, and affects all future generations. Consequently, a grandparent who invests more in the parent will effectively have a grandchild with higher initial human capital. In contrast, self investment is only relevant during the life-cycle, so has a finite horizon.

effects result in a lower earnings profile (as in Proposition 2 in Appendix C).

**Heterogeneity.** The parameters that measure uncertainty and heterogeneity, \( \sigma_\epsilon \) and \( \sigma_\eta \), are calibrated to match the Gini coefficients of lifetime earnings and wealth at age 60. The lifetime earnings Gini of 0.462 we target is larger than the conventional value of 0.32 which is typically computed from the PSID. This value is obtained from Leonesio and Bene (2011) who compute lifetime earnings Gini for all years from 1981 to 2004 using Social Security Administration data for more than 3 million people. This data has the advantage of directly measuring lifetime earnings, while typical datasets will impute lifetime earnings from a limited sample in a limited time span. On the other hand since wealth is a stock variable and is not subject to time horizon issues, consequently we proxy the age 60 wealth Gini by the retirement wealth Gini of 0.62 computed in Hendricks (2007).

Clearly, a larger \((\sigma_\epsilon, \sigma_\eta)\) impacts earnings when both young and old resulting in a larger Gini coefficient of lifetime earnings and wealth. However, while eliminating the i.i.d. \( \epsilon \) shock clearly would increase the IGE, the reverse is true for \( \eta \). Given a certain level of \( \rho \), a larger variance implies a larger IGE and vice versa—earnings inequality across ability groups becomes larger, and the earnings process will become closer to the underlying ability process (Han and Mulligan (2001)). Since the earnings correlation is rather high at 0.5, the abilities process variance \( \sigma_\eta \) must be small enough to match a Gini coefficient of 0.62 for wealth. Similarly, because the luck shock is i.i.d. it should have a smaller impact on the wealth Gini, while the ability shock should interact with \( \rho \) to have a larger impact, since less persistence implies less inequality in wealth accumulation. This is confirmed in our results.

**Intergenerational Transmissions.** Given other parameters, the altruism parameter \( \theta \) affects be-
quests to aggregate wealth ratio. In equilibrium, unconstrained parents invest in the human capital of their children up to the point that the marginal returns to investment equals the marginal returns to savings—if parents care about their children beyond this level, the transfers will occur in the form of physical rather than human capital assets (Proposition 1, Appendix C). Since intergenerational transfers in our model are made in one lump sum, we target the percentage sums of intergenerational transfers listed in Gale and Scholz (1994), which includes college support, inter vivos transfers, and bequests, which add up to 63.8%. Given the variance in abilities, the intergenerational correlation of abilities is a residual used to match the IGE of lifetime earnings.

As discussed in the introduction, the conventional IGE value of 0.4 in the literature is based on short time horizons. Mazumder (2005), among others, demonstrates that when we increase the horizon for which we compute the IGE, the value increases with the horizon, up to 0.6 at 15 years. While in our model, a period is 20 years, even longer than 15 years, we take a conservative approach and target a value of 0.5. Figure 2 illustrates how we separately identify $\rho$, $\gamma$ and $\gamma_3$.

Before we turn to the implications of our model, we also report some other moments in the top panel of Table 4. These moments were not targeted, but assures us that our parsimonious model does not deviate far from important aspects of the data. Some of these moments will be relevant for accounting for the mechanism in our model, as discussed in the next section.

In Table 4, columns “$O-Y(e)$” and “$O(i)-Y(e)$” are the IGE’s when regressing the parent’s old age earnings and income on the child’s young age earnings, respectively. Compare these with results from Mazumder (2005)’s parental income-child earnings IGE using 2-year averages (in contrast to lifetime averages, as we do here and so they are not directly comparable). As expected, family income has a larger effect on child earnings than the does parent’s earnings, and the difference in magnitude is similar to that calculated in Mazumder (2005). The reason why the magnitude itself is much smaller is because we do not control for any life-cycle characteristics — indeed, older IGE estimates that simply regressed the annual earnings of a son on his father’s, in the same year, were as small as 0.1-0.2 (e.g. Becker and Tomes (1986)). This confirms that life-cycle effects can have a large impact on the measured IGE.

**IGE in other variables:** One of the virtues of our model is that it is rich enough to make predictions on the intergenerational transmission of not just earnings, which is the most commonly reported upon measure in the literature, but also consumption, schooling and wealth (to reiterate, these moments were not targeted in our calibration). Mulligan (1999) reports a value of 0.68 for the IGE of consumption. In our model it is 0.71 (column “$con$”). Charles and Hurst (2003) report

<table>
<thead>
<tr>
<th></th>
<th>O-Y(e)</th>
<th>O(i)-Y(e)</th>
<th>con</th>
<th>O-Y(s)</th>
<th>sch</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>0.412</td>
<td>0.518</td>
<td>0.680</td>
<td>0.370</td>
<td>0.340</td>
</tr>
<tr>
<td>Model</td>
<td>0.129</td>
<td>0.219</td>
<td>0.720</td>
<td>0.391</td>
<td>0.334</td>
</tr>
</tbody>
</table>
an IGE of wealth between old parents and young children of 0.37, which is exactly the same as in our benchmark model (column “O-Y(s)”). This is expected in our model—for high ability families, there is more human capital investment in both the parent and the child, so that a rich old parent will tend to have a young parent child who is poor, because he borrows against his old age earnings to accumulate human capital when young. Finally, the correlation of educational attainment between father and son is 0.35 (Hauser (1998)), while in the model it is 0.334 (column “sch”).

We report two other moments before turning to the results. McGarry (1999) reports that the share of the US population that leave zero bequests is approximately 45%. In our quantitative model the share of households with zero bequests is 48.8%. Finally, although not very representative of the population, 29.3% of the PSID sample in Gottschalk (1990) grew up in a household that received AFDC. In our model, 31.5% of young adult households are subject to welfare subsidies. These last two numbers are reported along with the targeted moments for all experiments in the following subsections.

4. Results

To illustrate the quantitative contributions of persistence and education related parameters, we proceed by systematically setting them to zero and simulating a new steady state. We then proceed to isolate the role played by incomplete markets and how they interact with government policies. This will help us understand why the IGE is low for rich groups but not intergenerationally unconstrained groups. Next, we vary government policy parameters, focusing on how it affects intergenerational mobility. The final subsection discusses the relationship between lifetime income and wealth.

4.1 Nature, Nurture, or Education?

The third column in Table 5 shows that the model replicates the moments in the data fairly well (the first 8 moments were targeted, while the last two were not). In the following columns, we

---

15 Although in the calibration we target average years of schooling based only on the first period, it is not clear how to measure schooling in the model since we do not differentiate between on-the-job training and schooling. In the first period, some children will be in early on-the-job training, while others will be in college. This extends to the second period as well, since some individuals will be in college and some on-the-job. So if we assume only the first period is schooling, we underestimate for those who go to college and overestimate for early on-the-job trainees. On the other hand, if we include the second period, we overestimate all those on-the-job. However, it turns out that whichever way we measure, the IGE is basically the same (0.334 vs 0.327).

16 Recall that we do not model the difference between inter vivos and bequests, so our model bequests should encompass both in the data.

17 As in many other quantitative models with incomplete markets, the largest shortcoming of the model is noticeable in the Gini coefficient. However, this is due to the fixed number of grids for the ability and luck shock processes, and we have verified that when increasing the number of grids that we can get an exact match. However the increase in computation time would render the project infeasible. Here we present the benchmark case with 4 grid points for each.
Table 5: Results - Nature and Nurture

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
<th>$\rho=0$ (PE)</th>
<th>$\rho=0$ (GE)</th>
<th>$\gamma_3=0$</th>
<th>$\rho, \gamma_3=0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest rate (%)</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
<td>3.773</td>
<td>5.992</td>
<td>6.016</td>
</tr>
<tr>
<td>earnings prof (%)</td>
<td>120.000</td>
<td>118.302</td>
<td>118.360</td>
<td>119.223</td>
<td>229.618</td>
<td>212.850</td>
</tr>
<tr>
<td>earnings Gini</td>
<td>46.200</td>
<td>48.457</td>
<td>39.809</td>
<td>40.072</td>
<td>35.831</td>
<td>33.142</td>
</tr>
<tr>
<td>wealth Gini</td>
<td>62.000</td>
<td>62.946</td>
<td>56.955</td>
<td>57.715</td>
<td>87.048</td>
<td>86.437</td>
</tr>
<tr>
<td>WT-NW (%)</td>
<td>63.800</td>
<td>62.544</td>
<td>58.643</td>
<td>60.906</td>
<td>71.014</td>
<td>75.215</td>
</tr>
<tr>
<td>IGE (%)</td>
<td>50.000</td>
<td>49.446</td>
<td>18.364</td>
<td>20.606</td>
<td>24.067</td>
<td>1.607</td>
</tr>
<tr>
<td>zero WT (%)</td>
<td>45.000</td>
<td>49.273</td>
<td>46.142</td>
<td>47.313</td>
<td>78.451</td>
<td>75.723</td>
</tr>
<tr>
<td>p1 sub (%)</td>
<td>29.344</td>
<td>31.908</td>
<td>23.576</td>
<td>22.894</td>
<td>99.381</td>
<td>98.194</td>
</tr>
</tbody>
</table>

shut down the persistence parameters ($\rho, \gamma_3$) to measure the role they play in generating intergenerational persistence. For each case, we keep the government policy fixed and compute the new equilibrium values including the interest rate.\(^{18}\)

In the column labeled $\rho = 0$ we present the results when the intergenerational correlation of abilities is set to zero, in partial and general equilibrium. This is our model counterpart to empirical studies that (attempt to) control for all other characteristics so as to isolate the effect of nature. The impact is mitigated in general equilibrium, but only by a quantitatively small margin. We conclude that “natural” forces explain approximately 60% of the observed difference in the IGE, consistent with previous empirical studies. Interestingly, the magnitude of $\rho$ is also around 60%. The changes in most other moments, while small, are in predictable directions. Because of the lower persistence, agents invest more in physical capital, which lowers the interest rate, and consequently education increases slightly as well. The most noticeable changes are in the Gini coefficients - as is usual in these models, lower persistence reduces inequality.

Things are slightly different when we set $\gamma_3 = 0$. In a mechanical sense, parental human capital plays a similar role as ability, but shutting down this channel decreases a parent’s incentive to invest across all generations, so has a much more significant impact. As predicted, this decreases parental investment in the child’s human capital (schooling) and increases self investment (the earnings profile). The IGE drops by approximately 50%, comparable in magnitude to when $\rho = 0$ and consistent with Plug and Vijverberg (2003). However, this also means the economy becomes less productive in terms of human capital and overall the steady state economy becomes much poorer with a much larger portion of the population living on the welfare subsidies. Most individuals are at a point where the marginal returns to investing in human capital are higher than physical capital, so the the demand for physical capital increases, resulting in a higher interest rate.

\(^{18}\)One could instead vary government parameters as well to maintain budget balance. However, Since we are interested in the effect of each model component explaining the aggregate, we hold these variables constant. While not presented here, it should be clear that maintaining government budget balance would mitigate the changes.
Table 6: Results - Education

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
<th>$\gamma_1=0$</th>
<th>$\gamma_2=0$</th>
<th>$\gamma=0$</th>
<th>$\rho, \gamma=0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest rate (%)</td>
<td>4.000</td>
<td>4.000</td>
<td>4.726</td>
<td>4.771</td>
<td>5.244</td>
<td>5.477</td>
</tr>
<tr>
<td>mean sch (years)</td>
<td>12.920</td>
<td>12.835</td>
<td>0.000</td>
<td>5.920</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Edu-GDP (%)</td>
<td>6.850</td>
<td>6.701</td>
<td>12.876</td>
<td>22.658</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>earnings prof (%)</td>
<td>120.000</td>
<td>118.302</td>
<td>156.137</td>
<td>131.251</td>
<td>1975.519</td>
<td>1976.496</td>
</tr>
<tr>
<td>earnings Gini</td>
<td>46.200</td>
<td>48.457</td>
<td>80.741</td>
<td>41.478</td>
<td>37.680</td>
<td>35.470</td>
</tr>
<tr>
<td>wealth Gini</td>
<td>62.000</td>
<td>62.946</td>
<td>90.334</td>
<td>86.046</td>
<td>91.617</td>
<td>90.710</td>
</tr>
<tr>
<td>WT-NW (%)</td>
<td>63.800</td>
<td>62.544</td>
<td>69.551</td>
<td>57.076</td>
<td>73.194</td>
<td>79.839</td>
</tr>
<tr>
<td>IGE (%)</td>
<td>50.000</td>
<td>49.446</td>
<td>-27.530</td>
<td>36.799</td>
<td>19.035</td>
<td>0.533</td>
</tr>
<tr>
<td>zero WT (%)</td>
<td>45.000</td>
<td>49.273</td>
<td>82.378</td>
<td>78.808</td>
<td>88.008</td>
<td>86.797</td>
</tr>
<tr>
<td>p1 sub (%)</td>
<td>29.344</td>
<td>31.908</td>
<td>90.651</td>
<td>86.300</td>
<td>100.000</td>
<td>100.000</td>
</tr>
</tbody>
</table>

rate. The drops in the IGE and earnings Gini is masked by the large increase in the wealth Gini, which is mainly due to households relying only on subsidies rather than own assets or savings. The only significant impact when additionally setting $\rho = 0$ is that the IGE further decreases, suggesting that most of the IGE can potentially be explained by these two parameters.

We next investigate how the returns to education affect the equilibrium, in Table 6. Shutting down the returns parameters leads to a poorer economy with small levels of physical assets, driving up the interest rate. The result is a significant increase in the retirement wealth Gini. This effect is similar to what happens when we set $\gamma_3 = 0$. The earnings profile also becomes steeper in all cases, which may be misleading - in fact, when $\gamma = 0$, the earnings profile explodes! This is mainly due to fact that many young adults choose to stay in poverty and receive welfare subsidies in the first period of their lives with zero earnings. The IGE drops in all cases, which may not be intuitive at first blush. When $\gamma = 0$, human capital is equalized across all individuals so that there is no heterogeneity in parental spillovers. Comparing this with the case when we set $\gamma_3 = 0$, the magnitude is similar, leading us to conjecture that education in itself matters little for explaining intergenerational mobility. However, we cannot separately identify this effect in our model because all children start with the same level of initial human capital.

When $\gamma_1 = 0$, human capital accumulation only requires goods inputs. The earnings Gini is quite large compared to the benchmark. In the benchmark, the time input is cheaper than the goods input - hence the increased variation in goods inputs increases the Gini while also increasing the earnings profile (since it is part of measured earnings). Furthermore, education subsidies play a much larger role with only goods inputs, pushing down the IGE even more than when $\gamma = 0$. In fact, it is pushed to negativity, because poor parents have rich kids thanks to the subsidy, and rich parents cannot afford to educate their kids because it’s too expensive. We will confirm the large impact of subsidies again in Section 4.3. On the other hand, the time input is not only cheaper but also has less returns. So the earnings profile is less affected, while the absence of goods expenses
### Table 7: Results - Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
<th>$\sigma_\epsilon=0$</th>
<th>$\sigma_\eta=0$</th>
<th>$\sigma=0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest rate (%)</td>
<td>4.000</td>
<td>4.000</td>
<td>4.718</td>
<td>4.354</td>
<td>6.677</td>
</tr>
<tr>
<td>Edu-GDP (%)</td>
<td>6.850</td>
<td>6.701</td>
<td>7.420</td>
<td>7.008</td>
<td>8.638</td>
</tr>
<tr>
<td>earnings prof (%)</td>
<td>120.000</td>
<td>118.302</td>
<td>113.495</td>
<td>124.702</td>
<td>120.913</td>
</tr>
<tr>
<td>earnings Gini</td>
<td>46.200</td>
<td>48.457</td>
<td>40.622</td>
<td>32.206</td>
<td>0.000</td>
</tr>
<tr>
<td>wealth Gini</td>
<td>62.000</td>
<td>62.946</td>
<td>64.100</td>
<td>48.887</td>
<td>0.000</td>
</tr>
<tr>
<td>WT-NW (%)</td>
<td>63.800</td>
<td>62.544</td>
<td>58.752</td>
<td>56.075</td>
<td>5.655</td>
</tr>
<tr>
<td>IGE (%)</td>
<td>50.000</td>
<td>49.446</td>
<td>80.283</td>
<td>13.339</td>
<td>100.000</td>
</tr>
<tr>
<td>zero WT (%)</td>
<td>45.000</td>
<td>49.273</td>
<td>39.903</td>
<td>38.663</td>
<td>0.000</td>
</tr>
<tr>
<td>p1 sub (%)</td>
<td>29.344</td>
<td>31.908</td>
<td>17.587</td>
<td>22.453</td>
<td>0.000</td>
</tr>
</tbody>
</table>

A smaller luck shock variance $\sigma_\epsilon$ increases the IGE most significantly, implying the large effect of market incompleteness on intergenerational persistence. Investment in children is shifted from physical capital (wealth transfers) toward human capital (earnings persistence). Cross sectional earnings inequality is reduced, but the wealth Gini is not affected much, which is expected - since $\epsilon$ is an i.i.d. shock that will not persist, wealth accumulation is less affected. Furthermore, higher human capital levels as a young adult decreases the steepness of earnings.

A smaller ability variance $\sigma_\eta$ also decreases cross sectional inequality, but also decreases persistence as earnings becomes less reflective of the ability process. In particular, it has a much larger impact on the wealth Gini, as expected. Furthermore, even as children receive more education from parents so that they begin young adulthood with higher levels of human capital, young adults increase investment in themselves as well, as families are relieved from the grandchild ability shock.

The supply of physical capital (through savings and bequests) decreases in both cases, as there is less uncertainty about the human capital returns. This has two offsetting effects on the interest rate - on the one hand, the increased demand for physical capital pushes the interest rate up, but agents are now richer and able to save more. We find that the first effect always dominates, while it is more dominant for the luck shock, as should be expected - the i.i.d. shock does not affect wealth accumulation as much as the ability shock.
When we shut down uncertainty altogether, the economy behaves as a representative agent, which would result in the values in the last column. With our benchmark parameter values with no uncertainty, the representative parent leaves a positive level of bequests, implying that he is able to efficiently invest in his child. The relatively larger human capital investment, both in himself and his child, increases the interest rate.

To illustrate the importance of wealth and the intergenerational borrowing constraints, we compute other 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 is 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 in our model, as the constrained households are not necessarily the poor ones.

Table 8 shows the IGE of earnings within different groups. The first row shows the IGE when dividing families by whether or not bequests to the child \( (b_k) \) is smaller than the average annual income of old parents. This is roughly equal to $25,000 in 1990. We dub this group the “IG Constrained.” The next row is from Mulligan (1999) who divides families in the PSID by whether the child has received or expects to receive an inheritance of $25,000. The difference between the groups is less than 10 percentage points in our model vs 7 in Mulligan (1999). His numbers are based on four year averages, and the numbers presented here is for his whole sample including all family members. Consequently, our moments are larger in levels, which can partially explain the slightly larger difference. Our model has the potential to reconcile Mazumder’s results with Mulligan’s result that there is not a large difference in the IGE between households that expect to leave a inheritance and do not.

The next row is from when we divide families by whether the young parent has savings below

<table>
<thead>
<tr>
<th></th>
<th>low</th>
<th>high</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low is IG Constrained</td>
<td>0.560</td>
<td>0.462</td>
</tr>
<tr>
<td>Mulligan (PSID)</td>
<td>0.490</td>
<td>0.420</td>
</tr>
<tr>
<td>Low is ( s_y &lt; \text{Median} )</td>
<td>0.629</td>
<td>0.331</td>
</tr>
<tr>
<td>Mazumder (SIPP)</td>
<td>0.458</td>
<td>0.274</td>
</tr>
<tr>
<td>Mazumder (SER)</td>
<td>0.465</td>
<td>0.304</td>
</tr>
</tbody>
</table>

---

19For direct comparison with Mulligan (1999)’s results, we present the IGE of wages \((h\epsilon)\) as opposed to earnings. We also checked the IGE of lifetime earnings, and tried splitting the households according to whether \( b_k = 0 \). Both changes lead to slightly larger gaps (2 percentage points or less) in the IGE between the two groups.

20Note that the consumption IGE’s are also in the directions predicted by theory - consumption mobility is lower than earnings mobility, and the unconstrained group has lower consumption mobility. This is because unconstrained groups are better at smoothing consumption across generations than constrained groups.
Table 9: Human Capital Investment by group

<table>
<thead>
<tr>
<th></th>
<th>IG Constrained</th>
<th>$s_y &lt; \text{Median}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>$EI$</td>
<td>1.008</td>
<td>1.030</td>
</tr>
<tr>
<td>$e_o/e_y$ (%)</td>
<td>117.084</td>
<td>118.869</td>
</tr>
<tr>
<td>$s_y/e$ (%)</td>
<td>-4.708</td>
<td>4.608</td>
</tr>
<tr>
<td>$b_k/e$ (%)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>persistence (%)</td>
<td>67.665</td>
<td>42.326</td>
</tr>
</tbody>
</table>

or above the median $s_y$. The last two rows show the IGE of poor and rich groups in Mazumder (2005). He computes the IGE of earnings by different net worth groups, to refute Mulligan (1999)’s claim that net worth does not seem to matter for the IGE.\(^{21}\) The first row is based on SIPP data and the second on SER data. Note that these numbers are only based on a two-year averages, and the earnings IGE is only for between fathers and sons, as opposed to the 25 year average and parent-child relationship we focus on in our model.\(^{22}\)

As in Mazumder (2005), the IGE’s for the rich and poor are quite different.\(^{23}\) Consistent with his findings, the rich group that has a much smaller IGE than the pooled average, while the poor group’s is much higher. Most importantly, notice that the rift between the poor and rich are much larger in this case than when we divide them by whether they leave bequests.

To understand this, we show several statistics in Table 9 for comparison. In the appendix we show that, given any amount of investment $x_k$ in the child, the interior solution to the child’s human capital accumulation problem is

$$h_k(x_k) = a_k \left( \frac{x_k + d}{\gamma} \right)^{\frac{\gamma_1}{w}} \gamma_1 \gamma_2 \gamma_3 + (1 - \delta_h)\bar{h}.$$  

Hence, if all parent’s are unconstrained, in an interior solution the ratio

$$EI \equiv \frac{h_k - (1 - \delta_h)\bar{h}}{a_k(x_k + d)^{\gamma_2 \gamma_3}},$$

which we define as the “Education Intensity” ratio, should be constant. The $EI$ ratio will be larger for children who are at $n_k = \bar{n}$, meaning that the parent wants the child to spend more time than endowed in school, and smaller for children with $m_k = 0$, meaning that the parent educates the child only up to the amount optimal given the education subsidies. We see in the first two columns that this number is similar across the IG constrained and unconstrained groups. In contrast, the

\(^{21}\)Mazumder compares high and low net worth groups in the SIPP and SER, rather than constrained and unconstrained groups, which was Mulligan’s experiment.

\(^{22}\)The IGE of earnings from parent to child in his analysis is .412 for two year averages, and increases up to .6 when extending the horizon. However, he does not present these numbers for different groups.

\(^{23}\)In Mazumder (2005), the difference is large but statistically insignificant.
earnings profile of the latter group is relatively higher than the former, and savings when young $s_y$ higher. This means that parents in the former group invest less in themselves than parents in the latter group.

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 the child even in the absence of constraints. What is happening here is that constrained parents are in fact investing slightly more in their children than the unconstrained parent, given the child's states - their $EI$ ratio is higher. 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 the parents must also consume in their old age and retirement. In contrast, the unconstrained parents are relatively more able than their children. They have invested more in themselves, and did 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, because these relative ability differences across generations are small and ability regresses to the mean, there is relatively more mobility across these two groups intergenerationally than when dividing families by net worth - a child who received bequests is not much more likely to leave bequests to his own offspring than a child who did not receive any bequests.

What about when we look at families divided by net worth ($s_y$) below and above the median? It turns out that, the reason that these parents have low net worth is because on average, they are relatively more able than their children and needed to invest in themselves first, leaving them little resources to invest in their children. This is reflected in the low $EI$ ratio and steep earnings profile. Clearly, these parents will also not have enough resources to share with their children even after they grow up. Conversely, those with high net worth are relatively less able than their children as seen in the higher $EI$ ratio and flatter earnings profile, so invest much more in their children and are also able to share more assets. The mobility across these two groups is less than when dividing families by bequests, because the children of low net worth parents are still the relatively high ability ones, but because they did not receive much education, will again find it necessary to borrow against their own future earnings to invest in themselves, and not be able to provide for their children. The exact opposite is true for high net worth families.

4.3 The Role of the Government - Welfare or Poverty Traps?

"In fact, this strong emphasis on the “active” family and how family choices might be distorted by all forms of government intervention, both by direct interventions on educational markets and by pure welfare redistribution, is the main contribution of
Table 10: Government Effects - Taxation

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
<th>flat</th>
<th>$\tau_v = 0$</th>
<th>$\tau_k = 0$</th>
<th>$\tau = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest rate (%)</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
<td>4.056</td>
<td>3.020</td>
<td>3.174</td>
</tr>
<tr>
<td>earnings prof (%)</td>
<td>120.000</td>
<td>118.302</td>
<td>117.588</td>
<td>113.942</td>
<td>118.175</td>
<td>111.398</td>
</tr>
<tr>
<td>earnings Gini</td>
<td>46.200</td>
<td>48.457</td>
<td>49.025</td>
<td>51.304</td>
<td>48.661</td>
<td>52.420</td>
</tr>
<tr>
<td>wealth Gini</td>
<td>62.000</td>
<td>62.946</td>
<td>63.434</td>
<td>56.184</td>
<td>57.986</td>
<td>48.421</td>
</tr>
<tr>
<td>WT-NW (%)</td>
<td>63.800</td>
<td>62.544</td>
<td>63.539</td>
<td>66.685</td>
<td>59.358</td>
<td>63.591</td>
</tr>
<tr>
<td>IGE (%)</td>
<td>50.000</td>
<td>49.446</td>
<td>50.850</td>
<td>56.752</td>
<td>49.640</td>
<td>61.744</td>
</tr>
<tr>
<td>zero WT (%)</td>
<td>45.000</td>
<td>49.273</td>
<td>47.658</td>
<td>45.962</td>
<td>43.743</td>
<td>40.513</td>
</tr>
<tr>
<td>p1 sub (%)</td>
<td>29.344</td>
<td>31.908</td>
<td>32.443</td>
<td>19.410</td>
<td>28.207</td>
<td>12.138</td>
</tr>
</tbody>
</table>

Gary Becker and his followers to the study of intergenerational mobility."

[Piketty (2000)]

One advantage of our general equilibrium approach is that we can analyze macro-level government policies within the model, quantify their effects, and perform policy experiments. We illuminate the role of the government by shutting down tax and transfer parameters. In Table 10, we alter only the tax variables while maintaining redistributive policies. Even when tax rates are zero, we assume that the government has a magical budget for subsidies. In such cases, the economy is unambiguously better off. Only in the third column, “flat,” do we maintain revenue equivalence.

The column “flat,” corresponding to the case when the earnings tax is flat, shows that it has barely any impact both in partial and general equilibrium. This is in contrast to conjectures that progressive taxation may help increase mobility. These conjectures usually presume that since there are smaller returns to investing in children for richer households, they will tend to invest less in their children compared to the poor. However, this again ignores the life-cycle effect. There is also a disincentive for the young parent to invest in himself, which countervails the disincentive to invest in his child. This is further mitigated by the fact that the returns to any investment in the child propagates indefinitely, while the returns to investment in himself only lasts during his lifetime.

However, if we were to eliminate earnings taxes altogether, investment in children increase and consequently mobility decreases, as seen in the higher IGE. Interest rates rise as human capital investment becomes uniformly more attractive and fewer people invest in physical capital. Interestingly, the wealth Gini decreases, mainly because more people rise out of welfare subsidies.

Eliminating capital taxes decreases the interest rate as more people invest in physical capital, as can also be seen in the reduced number of households with zero bequests (wealth transfers).

---

24 In fact, the impact is so small that the interest remains at the same level, so there is no general equilibrium effect.
Table 11: Government Effects - Transfers

<table>
<thead>
<tr>
<th>Measure</th>
<th>Data</th>
<th>Model</th>
<th>no edu</th>
<th>no SS</th>
<th>g=0</th>
<th>no gov</th>
</tr>
</thead>
<tbody>
<tr>
<td>interest rate (%)</td>
<td>4.000</td>
<td>4.000</td>
<td>4.740</td>
<td>3.327</td>
<td>3.748</td>
<td>2.625</td>
</tr>
<tr>
<td>earnings prof (%)</td>
<td>120.000</td>
<td>118.302</td>
<td>156.908</td>
<td>116.259</td>
<td>115.708</td>
<td>112.110</td>
</tr>
<tr>
<td>earnings Gini</td>
<td>46.200</td>
<td>48.457</td>
<td>51.151</td>
<td>48.608</td>
<td>60.098</td>
<td></td>
</tr>
<tr>
<td>wealth Gini</td>
<td>62.000</td>
<td>62.946</td>
<td>55.281</td>
<td>57.887</td>
<td>49.214</td>
<td></td>
</tr>
<tr>
<td>WT-NW (%)</td>
<td>63.800</td>
<td>62.544</td>
<td>53.115</td>
<td>62.530</td>
<td>54.662</td>
<td></td>
</tr>
<tr>
<td>IGE (%)</td>
<td>50.000</td>
<td>49.446</td>
<td>58.437</td>
<td>51.550</td>
<td>75.977</td>
<td></td>
</tr>
<tr>
<td>zero WT (%)</td>
<td>45.000</td>
<td>49.273</td>
<td>48.861</td>
<td>43.147</td>
<td>40.400</td>
<td></td>
</tr>
<tr>
<td>p1 sub (%)</td>
<td>29.344</td>
<td>31.908</td>
<td>17.142</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

This also implies that more people are optimally investing in their children, which in turn leads to a slight increase in the earnings Gini and IGE, as rich households are more likely to be affected by capital taxes. Again, the wealth Gini drops as more people rise above the welfare subsidy level, and when taxes are eliminated altogether, we see that these forces combine to generate a significant rise in the IGE but a decline in the wealth Gini. While progressivity does not play a large role, taxation itself has a significant impact on the magnitude of the IGE.

In Table 11, we remove each redistributive policy one by one, while the last column is when we remove the government altogether. The columns “no edu,” “no SS,” and “g = 0” are when we eliminate the education subsidy, social security, and welfare subsidy, respectively. In particular, when we eliminate social security, we also eliminate the social security tax.

As one would expect, the education subsidies have a very large effect. Households now divert more resources to education, but become poorer without the help of the subsidy, resulting in less education and savings in equilibrium. This drives the interest rate up and also increases the earnings profile, as parents invest relatively more in themselves than their children. The poor are relatively worse off than the rich, resulting in large increases in the Gini coefficients and the IGE. This is mainly due to the fact that without it, all households are effectively losing a lump sum transfer which was effectively redistributive.

Social security plays a relatively smaller, but still quantitatively significant role. Remember that in the model and in the real world, agents are not allowed to borrow against social security benefits. Hence adults must prepare for retirement savings, reducing the young adult’s ability to borrow against future income to increase their earnings profile. Furthermore, parents now find it more necessary to invest in their children so they can survive without the benefits, which is reflected in increased education. This further flattens the earnings profile. However, more parents now find themselves constrained, both intra- and intergenerationally, which leads to a decrease in mobility. This also results in a larger earnings Gini, but since more people save for retirement, the
Welfare subsidies have countering effects. Since the poor find no need to save with the subsidy, the wealth Gini is higher in the benchmark. In terms of persistence, on the one hand, it helps families with fewer resources educate their children. On the other hand, low income parents are induced to work even less than they would, and since these young parents have children who are themselves likely to grow up to receive the subsidy, they invest less both in themselves and in their children. Furthermore, these subsidies act as an insurance for the non-recipient households, so that they can safely invest more in their children. All these effects more or less cancel out in equilibrium, so we conduct a separate analysis for clarification.

In Table 12, we show the model simulated probability that a child who grew up with a parent who received welfare subsidies \((t(z_y))\) also receives welfare subsidies when he grows up to be a young adult \((t(z_y'))\). The chances of this happening for such a child are significantly higher than that for the corresponding child whose parent did not receive welfare subsidies, and the magnitudes are consistent with the literature, e.g. 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 the parent not anticipated that the child would receive welfare subsidies. To see whether this is the case, we again look at the measure \(EI\) as defined in the previous subsection. It turns out that this ratio is in fact higher in the subsidized group, meaning that the education subsidy alone was enough to induce education levels higher than average, conditional on the child’s states (refer to the next row). We conclude that as long as the education subsidies are in place, the welfare program has no perverse effects.

Surprisingly, when either redistributive policies or taxation is present, as in the last column of Table 11, either one is sufficient to keep the persistence of earnings at a rather low level, i.e. closer to the calibrated value. When both are eliminated, however, the inequality measures increase to significantly large values, and the persistence of earnings jumps to a very high value of 0.76. Based on the analysis, however, we conclude that this is mainly due to education subsidies.
Table 13: Retirement Wealth and Lifetime Earnings

<table>
<thead>
<tr>
<th></th>
<th>$C_{WE}$</th>
<th>Mean W/E Gap</th>
<th>Mean Gini</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSID</td>
<td>0.61</td>
<td>0.47</td>
<td>0.54</td>
</tr>
<tr>
<td>Hendricks</td>
<td>0.82</td>
<td>0.55</td>
<td>0.39</td>
</tr>
<tr>
<td>Model</td>
<td>0.74</td>
<td>0.35</td>
<td>0.52</td>
</tr>
</tbody>
</table>

4.4 Retirement Wealth

According to Hendricks (2007), the standard life-cycle model cannot match the correlation between lifetime earnings and retirement wealth, even after accounting for inheritance shocks. However, we find in our model that once we account for endogenous bequests, the model performs much better in matching these characteristics.

The problem is that the standard life-cycle has a too strong relationship between lifetime earnings and retirement wealth. This also leads to having too little heterogeneity within groups, and the earnings rich saving abnormally larger amount than the earnings poor, compared to the data. In Table 13, the column $C_{WE}$ is the correlation between lifetime earnings and retirement wealth. In the next column, we compute the average retirement wealth over average lifetime earnings ratio for the first and last earnings deciles, and present the “Mean W/E Gap,” the ratio of these values between the first and last earnings deciles. “Mean Gini” is the average of the retirement wealth Gini’s of each earnings decile. The first row is from the PSID data, and second row is from Hendricks (2007) who constructs a model with inheritance shocks independent of lifetime income. There he shows that inheritance shocks help, but not sufficient quantitatively to account for the data, as shown in the table. Subsequent studies have found that heterogeneous earnings profiles and intergenerational bequest motives may also help.

Our study is not intended to explain this puzzle but has all the ingredients to reconcile the model with the data, as shown in the last row. 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 the bequests will be heterogeneous even across parents with the same profile. Since bequests are not persistent across generations, this also leads to the earnings rich not saving as much as they would in a simpler model.

5. Conclusion

In this paper, we present a rich model of parental investment in children. By adding a life-cycle dimension to a dynastic version of Becker-Tomes and casting it in a general equilibrium framework, we are able to discipline our choice of parameters while at the same time making predictions on
a number of variables for which there are empirical counterparts. Using observations on the IGE of earnings, bequests and schooling across generations, as well as cross sectional variables such as the lifetime earnings and retirement wealth Gini’s, the model is used to pin down the importance of natural factors such as innate ability in the determination of the persistence of earnings across generations and earnings groups. Consistent with empirical studies, we find that both nature and nurture are important in determining intergenerational transmission, there is significant heterogeneity even within similar earnings groups, and that government transfer programs have a large impact on intergenerational transmission. Finally, we also show that life-cycle constraints can have intergenerational effects not found in simple dynastic models.
A. Simplifying the Young Parent’s Value Function

The young parent’s problem is complicated by the number of choice variables and constraints, and the nonlinearity of human capital production and government policies. However, the young parent’s human capital decision is deterministic in most cases, so that it can be solved for separately from the recursive problem. Furthermore, the decision for the child can be expressed as one where the parent only chooses the total cost \( x_k \) and \( h_k \equiv h_k(x_k) \) can be solved for analytically.

We first reformulate the child’s human capital production function in terms of total cost. To be precise, define

\[
h_k(x_k) = \max_{n_k, m_k} \{ a_k(n_k \bar{h})^\gamma_1 (m_k + d)^\gamma_2 \bar{h}^\gamma_3 + (1 - \delta_h) \bar{h} \}
\]

s.t. \( x_k = w \bar{h}n_k + m_k \)

and letting \( \gamma = \gamma_1 + \gamma_2 \), the solution is

\[
h_k(x_k) = a_k \left( \frac{x_k + d}{\bar{h}} \right)^\gamma \left( \frac{\gamma_1}{w} \right)^\gamma_1 (m_k + d)^\gamma_2 \bar{h}^\gamma_3 + (1 - \delta_h) \bar{h}
\]

\[
\bar{h} \bar{h}n_k = \frac{\gamma_1 (x_k + d)}{\gamma}, \quad m_k = \frac{\gamma_2 (x_k + d)}{\gamma} - d
\]

in the interior. There are two possible corner solutions: \( n_k = \bar{n} \) and \( m_k = 0 \). Whenever \( n_k = \bar{n} \),

\[
h_k(x_k) = a_k (\bar{n} \bar{h})^\gamma_1 (x_k + d - w \bar{h} \bar{n})^\gamma_2 \bar{h}^\gamma_3 + (1 - \delta_h) \bar{h}
\]

\[
m_k = x_k - w \bar{h} \bar{n}
\]

and whenever \( m_k = 0 \),

\[
h_k(x_k) = a_k \left( \frac{x_k}{\bar{h}} \right)^\gamma_1 d^\gamma_2 \bar{h}^\gamma_3 + (1 - \delta_h) \bar{h}
\]

\[
n_k = \frac{x_k}{\bar{h}}
\]

Solving the young adult’s own human capital problem is slightly more complicated, as both earnings when young and old are subject to taxation and social security benefits. However, because \( \epsilon \) is constant over the life-cycle, as long as the borrowing constraints are not binding, the problem becomes a deterministic income maximization problem which we can solve for independently of the recursive problem. This problem is

\[
L(h) \equiv \max_{m, n} \left\{ \frac{T [S_0 + S_1 (y(e_y) + y(e_o))]}{(1 + \bar{r})^2} + \frac{T [y(e_o) - \tau_s e_o]}{1 + \bar{r}} + T [y(e_y) - \tau_s e_y] \right\}
\]
where
\[
y(e) = (1 - \tau_e(e))e = \left[1 - \tau_e^0 + \frac{0}{\tau_e^m} \left(\tau_e^m e^{\tau_e} + 1\right)^{-1/\tau_e^m}\right]e
\]
\[
ey_y = \text{whe}(1 - n) - m \geq 0
\]
\[
ey_o = \text{wh}'e
\]
\[
h' = a(nh)^{\gamma_1} m^{\gamma_2} + (1 - \delta)h.
\]

Note that, for any given level of \(h'\), cost minimization implies, for \(n \leq \gamma_1 / \gamma_2\),
\[
\frac{m}{\gamma_1} = \frac{m}{\gamma_2},
\]
while if \(n > \gamma_1 / \gamma\) (by the constraint \(e_y \geq 0\)),
\[
m = \text{whe}(1 - n).
\]

First assume that none of the earnings are subject to the earnings subsidies. We have
\[
\max_{m, n} \left\{ \frac{K_y y(e_o) - \tau_{so} e_o}{1 + \beta} + K_y y(e_y) - \tau_{sy} e_y \right\}
\]
where \((K_y, K_o, \tau_{sy}, \tau_{so})\) are constants:
\[
K_y = 1 + \frac{S_1}{2(1 + \beta)^2}, \quad \tau_{sy} = \tau_s
\]
\[
K_o = 1 + \frac{S_1}{2(1 + \beta)^2}, \quad \tau_{so} = \tau_s
\]

Note that for any level of income \(e\), the after tax function defined from the Gouveia-Strauss average tax rate function preserves monotonicity and concavity:
\[
y'(e) = 1 - \tau_e(e) - \tau_e'(e)e = 1 - \tau_e^0 + \frac{0}{\tau_e^m} \left(\tau_e^m e^{\tau_e} + 1\right)^{-1/\tau_e^m}
\]
\[
y''(e) = -\tau_e^0 (1 + \tau_e^1) \tau_e^m e^{\tau_e} - 1 \left(\tau_e^m e^{\tau_e} + 1\right)^{-1/\tau_e^m - 2}.
\]

Hence the above is a well-defined convex optimization problem, though a closed-form solution does not generally exist. However, there will be cases where it is more beneficial for the parent to subject himself to the welfare subsidies. Let \((p_1, p_2, p_3)\) be indicator functions denoting whether the parent is subject to the welfare subsidy when young, old, and retired, respectively. So we must compare the optimized lifetime incomes of the parent for the following 8 cases:
and the parent simply chooses the largest case. One can see from the values of \((K_y, K_o)\) that the welfare subsidy induces adults to work less. On the other hand, without social security \((K_y, K_o)\) would be zero when subject to welfare subsidies. The social security benefits induce adults to work even when their earnings are below the welfare subsidy floor \(g\).

Denoting all variables that are solved by the parent’s human capital decision with a ‘∗’, the young parent’s problem simplifies to (if not borrowing constrained in both periods)

\[
\begin{align*}
V_y(a_k; a, \epsilon, h, b) &= \max_{c, x_k, Z} \left\{ u(c/q) + \beta \int_{\epsilon_k} \int_{a_k} V_o(a_k'; a_k, \epsilon_k, h_k(x_k); e_k^*, Z) dF(\epsilon_k) dA(a_k'|a_k) \right\} \\
\text{s.t.} \quad &c + x_k + \frac{Z}{1 + r} = T \left[ y(e_y^*) - \tau_s e_y^* \right] + \frac{T [y(e_y^*) - \tau_s e_y^*]}{1 + r} + \bar{w} \bar{h} n + b \\
\quad &Z \geq 0
\end{align*}
\]

so the parent only needs to choose \((x_k, Z)\).

**B. Numerical Algorithm**

**B.1 Solving the Household’s Problem**

By the theorem of Benveniste and Scheinkman, we have

\[
\begin{align*}
V_o(b; a_k, \epsilon_k, h_k; e_R, Z) &= u'(c_o) \\
V_y(b; a, \epsilon, h, b) &= u'(c_y) / q
\end{align*}
\]

where \((c_o, c_y)\) are the optimal consumption decisions when old and young, respectively.

To solve the young parent’s problem we must first solve his lifetime income maximization problem as specified in Appendix A. The state variables for this problem are \((a, \epsilon, h)\) from which we obtain \((n^*, m^*, e_y^*, e_{\epsilon}^*, e_R^*)\). We then include these as states in the young parent’s simplified value function to solve for the rest of his policy functions. The savings decision \(Z\) is solved for by the
Euler equation

\[ u'(c_y/q) = q\beta(1 + \tilde{r})EV_{o6}(a_k'; a_k, \epsilon_k, h_k(x_k); e_R^*, Z) \]

where the integral is suppressed in the expectations operator. We solve this Euler equation for all feasible values of \( x_k \), while \( x_k \) is found by a one dimensional optimization scheme. Note that we are assuming that the borrowing constraints do not bind over the life-cycle. In the computation, we do not resolve the parent’s problem when the borrowing constraints are binding, but in our benchmark specification aggregate debt is less than 2% of GDP, leading us to believe that the true solution will not differ by much.

The old parent’s problem is much simpler. Given \( V_y \), the f.o.c. yields the following Euler equations:

\[
\begin{align*}
    u'(Z - s_o - b_k) & \geq \beta(1 + \tilde{r})u'(e_R + (1 + \tilde{r})s_o), & \text{with equality if } s_o > 0 \\
    u'(Z - s_o - b_k) & \geq \theta V_{y5}(a'_k; a_k, \epsilon_k, h_k, b_k), & \text{with equality if } b_k > 0.
\end{align*}
\]

Given CRRA preferences with parameter \( \sigma \), we have a closed form solution for \( s_o \) for any choice of \( b_k \):

\[
s_o = \max \left\{ \frac{Z - [\beta(1 + \tilde{r})]^{-1/\sigma} e_R - b_k}{1 + (1 + \tilde{r}) [\beta(1 + \tilde{r})]^{-1/\sigma}}, 0 \right\}
\]

so we can solve for \( b_k \) using a one-dimensional equation solver.

### B.2 Calibrating the parameters

The calibration is implemented as follows:

1. Discretize the state variables. In particular, follow Rouwenhorst (1995) for the exogenous AR(1) process for abilities, \( a \). We will assume that \( (a, \epsilon) \) are in fact discrete, while \( (h, b, z) \) are continuous. We choose 4 grid points each for \( (a, \epsilon) \), 24 for \( h \), and 36 each for \( (b, z) \).

2. Guess \( \mu \). Solve for \((V_y, V_{y5}, V_o, V_{o6})\) by iterating on the Euler equations. We solve for the Euler equations by Brent’s method using bilinear interpolation, and \( x_k \) is found by a modified Newton-based optimization method. Caution is taken for kinks in the value function.

3. Starting from an initial distribution, simulate the behavior of 120000 individuals for 200 periods assuming fixed prices. We take the distribution in the 200th period to be our stationary distribution.\(^{25}\) Let \( r' \) denote the interest rate implied by the firm’s first order condition. If

\(^{25}\)Changing this to 300 periods does not alter the results.
\[ r' \simeq r = 4\%, \text{ stop. Otherwise repeat 2 with } \mu' > \mu \text{ as a new guess if } r' < r, \text{ and } \mu' < \mu \text{ otherwise.} \]

The welfare subsidy \( g \), education subsidy \( d \) and social security parameters \( (S_0, S_1) \) are found by numerically solving for the budget balance conditions, along with the calibration of the remaining 7 model parameters plus the level parameter for the progressive earnings tax function \( \tau^2_e \), using a downhill simplex method.

When conducting the experiments, we iterate on \( r \) instead of \( \mu \) to find the price equilibrium. Since the simplex method does not guarantee a global minimizer, we tried starting from many different initial values.

C. A Simple Model

We study a simple version of the model to clarify how the parameters in the quantitave model are identified. Specifically, we focus on providing some intuition on the following:

1. Decomposition of the IGE
2. Intergenerational borrowing constraint
3. Parent’s own human capital accumulation decision

The first two points are analyzed to compare our model with previous studies in the literature - in particular, we emphasize the role of early childhood environment. To the best of our knowledge, we are the first to analyze the third point, as well as welfare subsidies and social security.

The standard empirical model for an earnings regression is

\[
\log e_k = A + \text{IGE} \cdot \log e + \nu, \tag{1}
\]

where \((e, e_k)\) are the parent’s and child’s earnings. We will use a simple variant of Solon (2004), which has been used widely in the literature. We will assume that the wage \( w = 1 \) and that the interest rate \( r \) is at the steady state level. We also ignore the discount factor \( \beta \) and assume log utility. The model is used to derive what implications our quantitative model will have for the IGE as defined in (1). Suppose the parent’s problem is:

\[
\begin{align*}
V(h) &= \max_{c,m_k,b_k} \{u(c) + \theta u(I_k)\} \\
\text{s.t. } c &= I - m_k - b_k
\end{align*}
\]
\[ h_k = a_k(m_k + d)^\gamma h^{\gamma_3} \]
\[ I = T[(1 - \tau_e)e + (1 + r)b, g] \]
\[ I_k = T[(1 - \tau_e)e_k + (1 + r)b_k, g] \]

and that abilities are transmitted across generations according to

\[ \log a_k = \mu + \rho \log a_p + \eta, \quad \eta \sim \mathcal{N}(0, \sigma_a^2). \]

Throughout the analysis, assume that we are in a steady state where the distribution of all variables are stationary across generations. Among other things, this implies the population mean of \( \log a = \log a_k = \mu \).

### C.1 Borrowing constraints, Taxation, and Education subsidies

We begin with the simplest possible case and continue to change the definition of \((e, e_k)\) to clarify our points. The first case is a simple version of Becker and Tomes (1986). The implications are also similar except that we include \( \gamma_3 \). Since we are ignoring prices, assume we measure \((e, e_k)\) as

\[ e = h, \quad e_k = h^k \]

Then we have the following results:

**Proposition 1**

1. An increase in \( \theta (\mu) \) will increase (reduce) average bequests, \( \mathbb{E}[b] \).

2. If all families are unconstrained, 
\[
IGE = \frac{(1 - \gamma)(\gamma_3 + \rho)}{(1 - \gamma)\gamma_3 \rho}.
\]

3. If all families are constrained, 
\[
IGE = \frac{(1 - \tau_e)(\gamma + \gamma_3 + \rho) + (1 + \gamma_3 + \rho)d}{(1 - \tau_e)(1 + \gamma_3 + \rho + d)}, \text{ where } dI = d. \text{ So an increase in tax rates or education subsidies will decrease the IGE.}
\]

4. If there are more families where either the parent or child, or both, are subject to welfare subsidies, the IGE is not affected by the education parameter \( \gamma \), and 
\[
IGE = \frac{\gamma_3 + \rho}{1 + \gamma_3 \rho}.
\]

**Proof:** Assume that no families are subject to the welfare subsidy. If families are not subject to the intergenerational borrowing constraint,

\[ m_k^* = \left[ \frac{(1 - \tau_e)\gamma a_k h^{\gamma_3}}{1 + r} \right]^{\frac{1}{1 - \gamma}} - d \]
\[ h_k^* = \left[ \frac{(1 - \tau_e)\gamma}{1 + r} \right]^{\frac{1}{1 - \gamma}} (a_k h^{\gamma_3})^{\frac{1}{1 - \gamma}}. \]

Since there is no uncertainty and capital markets are perfect, in a steady state \( \theta(1 + r) = 1 \). Hence,
again due to stationarity

\begin{align*}
b^*_k &= (1 - \tau_e) h + b - m^*_k - b^*_k - (1 - \tau_e) h^*_k \\
\bar{b} &= -\bar{m},
\end{align*}

i.e., on average, families will finance all their investment in children by borrowing from the children. Now consider an increase in the altruism parameter \( \theta \). Then it must be the case that the equilibrium interest rate \( r \) decreases (since \( \theta(1 + r) = 1 \)). From (3), it is clear that all families will decrease \( m_k \), leading to an increase in \( \bar{b} \). The converse is true for mean log abilities \( \mu \).

Using (4) we obtain

\begin{equation}
\log e_k = \gamma \log \left( \frac{(1 - \tau_e)\gamma}{1 + \gamma} \right) + \gamma \log(1 + d) + \gamma_3 \log e + \log a_k
\end{equation}

and since we are assuming stationarity across generations, standard time series analysis implies if all families are unconstrained we obtain

\begin{equation}
\text{IGE} = \frac{(1 - \gamma)(\gamma_3 + \rho)}{(1 - \gamma)^2 + \gamma_3 \rho}.
\end{equation}

If families are constrained,

\begin{align*}
m^*_k &= \frac{\theta \gamma I - d}{1 + \theta \gamma} \\
h^*_k &= \left[ \frac{\theta \gamma (I + d)}{1 + \theta \gamma} \right]^\gamma a_k h^{\gamma_3}.
\end{align*}

So we obtain

\begin{equation}
\log e_k = \log \left( \frac{\theta \gamma}{1 + \theta \gamma} \right)^\gamma + \gamma \log(1 + d) + \gamma_3 \log e + \log a_k.
\end{equation}

If all families are constrained, assuming that \( d = \bar{d} I \) we can log-linearize around \( e \) to obtain

\begin{align*}
\log e_k &= \bar{A} + \left[ \frac{(1 - \tau_e)\gamma}{1 - \tau_e + \bar{d}} + \gamma_3 \right] \log e + \log a_k
\end{align*}

where

\begin{align*}
\bar{A} &= \log \left( \frac{\theta \gamma}{1 + \theta \gamma} \right)^\gamma - \frac{(1 - \tau_e)\gamma}{1 - \tau_e + \bar{d}} \log \bar{e} + \gamma \log((1 - \tau_e)\bar{e} + \bar{d}),
\end{align*}

\(^{26}\text{Refer to Solon (2004) for details.}\)
so the measured IGE will be

\[ \text{IGE} = \frac{(1 - \tau_e)(\gamma + \gamma_3 + \rho) + (\gamma_3 + \rho)d}{(1 - \tau_e)(1 + \gamma\rho) + \gamma_3\rho + d}. \]

Depending on whether the parent or child, or both, are subject to welfare subsidies, the composition of the IGE also changes. When the parent is subject to the subsidies but not the child, (6) changes to

\[ \log e_k = \log \left( \frac{\theta \gamma}{1 + \theta \gamma} \right) + \gamma \log(g + d) + \gamma_3 \log e + \log a, \]

and when the child receives subsidies,

\[ \log e_k = \gamma \log d + \gamma_3 \log e + \log a. \]

For these families, it becomes immediate that education is not playing any role in explaining their earnings persistence: in any case, the IGE is

\[ \text{IGE} = \frac{\gamma_3 + \rho}{1 + \gamma_3\rho}. \]

□

The first two results help us identify \((\theta, \mu)\) in the calibration. Since we assume decreasing returns to education while returns to bequests are constant, any parent that is leaving a positive amount of bequests is already investing optimally in the child’s human capital. Hence increasing the altruism factor will merely result in more physical bequests. The second result is also trivial - if average abilities increase, human capital investment becomes more efficient, leading to higher levels of human capital on average. In the calibration, \(\theta\) is chosen to match the aggregate stock of intergenerational wealth transfers, and \(\mu\) is targeted to match the equilibrium interest rate - if \(\mu\) is high (low), the returns to physical capital investment becomes less (more) valuable, so there is a relative increase in human capital investment which pushes up the equilibrium interest rate.

For unconstrained families, if \(\gamma_3 = 0\), \(\text{IGE} = \frac{\rho}{1 - \gamma}\), i.e. the earnings process will reflect the ability process scaled by \((1 - \gamma)\), which is a version of a standard result from Becker and Tomes (1986). For unconstrained families. In this case, a larger \(\rho\) implies a smaller \(\gamma\), and vice versa. However, as long as \(\gamma + \rho < 1\), a model that does not consider the parental spillover term \(\gamma_3\) will attribute too large a role to \(\rho\). Also notice that government intervention has no role in this specification.

For constrained families, if \(\gamma_3 = 0\) and there were no government, the IGE is simply the sum of \(\gamma\) and \(\rho\), again a standard result from Becker and Tomes (1986), but for constrained families. While government parameters have no effect on the IGE for unconstrained families, for constrained
families they interact with the education parameter $\gamma$ and affect how much of the IGE comes from education. The higher the level of government intervention (tax rates and education subsidies), the lower the role of $\gamma$. As with unconstrained families, not accounting for $\gamma_3$ will exaggerate the size of $\rho$.

For parents who are subject to the welfare subsidy, the floor $g$ acts as an income equalizer. Hence there will be no variation in their investment in children beyond the natural components, $\gamma_3$ and $\rho$. Next assume that the child is subject to the welfare subsidy. Then there is no incentive for the parent to invest in the child, so $b_k^* = m_k^* = 0$. Again, since all children receive the same public investment $d$, education has no role to play in the IGE beyond reflecting $\gamma_3$ and $\rho$. Hence when many families are on welfare programs and this is not accounted for, the IGE may exaggerate the role of education, and hence, borrowing constraints.

We saw above that when the government is providing sufficient education subsidies, the IGE would be smaller - this is related to Corak and Heisz (1999), who conjecture that because of subsidies, the IGE may be smaller for the poor. The preceding analysis shows that when the government is also providing welfare subsidies for the poor, the effect of the education subsidy becomes extreme to the extent that poor families could exhibit the highest degree of intergenerational mobility.

### C.2 Own human capital accumulation and Social Security

Now assume agents also make a human capital decision for themselves. We change (2) to

$$e_k \equiv \max_{m_k^*} \frac{h_k'}{1 + r} + \underbrace{h_k - m_k'}_{\text{young age earnings} \equiv e_y}$$

s.t. $h_k' = a_k m_k'^\gamma$,

and ignore taxes and education subsidies but assume that the welfare subsidy is applied to old and young age earnings separately:

$$I_k = T[e_y, g] + T[e_o, g] + (1 + r)b_k,$$

and that there is a social security benefit distributed according to

$$SS = \frac{e_o + e_y}{2}.$$  

---

27 This is also assuming that these families are constrained, since it is not realistic that families leaving bequests are subject to welfare subsidies.
The definitions of \((e, I)\) also change accordingly. We assume that there are no intragenerational borrowing constraints and that the parent’s problem is the same as before. We have

**Proposition 2**

1. If no families are borrowing constrained, an increase in \(\gamma_3\) will decrease the average life-cycle earnings profile.

2. Without social security benefits, if an individual is subject to welfare subsidies when young (old) but not when old (young), the earnings profile will increase (decrease). Investment in children decreases if the child is subject to welfare. If the child is subject to welfare in both periods, his earnings profile is flat and parental investment is zero.

3. With social security benefits, the earnings profile becomes steeper, and investment in children will increase if the parent is life-cycle unconstrained.

**Proof:** When considering the life-cycle, the solution to the simple lifetime income maximization problem is

\[
\begin{align*}
    m_k^* &= \left( \frac{\gamma a_k}{1 + r} \right) \frac{1}{1 - \gamma} \\
    h_k^* &= \frac{\gamma}{1 + r} \left( \frac{a_k}{1 + r} \right) \frac{1}{1 - \gamma} \\
    e_k^* &= \frac{1 - \gamma}{\gamma} \left( \frac{\gamma a_k}{1 + r} \right) \frac{1}{1 - \gamma} + h_k^*.
\end{align*}
\]

In the case of no intergenerational borrowing constraints, \(h_k^*\) is given in (4) above. We also need an expression for \(\dot{h}\), the stationary average of human capital when young. This can be obtained from (5). It is evident that \(h\) is lognormal, so we need only obtain the mean and variance of \(\log h\) to express the mean of any power of \(h\):

\[
(1 - \gamma - \gamma_3) \log h = \gamma \log \gamma + \log a \\
\log h = \frac{\mu + \gamma \log \gamma}{1 - \gamma - \gamma_3}.
\]

The variance is found by expanding (5) into a second-order autoregression:

\[
\log h_k = \frac{(1 - \rho)\gamma \log \gamma + \mu}{1 - \gamma} + \frac{\gamma_3 + \rho}{1 - \gamma} \log h - \frac{\gamma_3 \rho}{(1 - \gamma)^2} \log h - \frac{\log \eta}{1 - \gamma},
\]

where \(h\) is the human capital of the grandparent. We obtain

\[
\mathbb{V} \left[ \log h_k \right] = \frac{[1 + \gamma_3 \rho / (1 - \gamma)^2] \sigma_\eta^2}{[(1 - \gamma)^2 - \gamma_3 \rho] [(1 - \gamma)^2 - \gamma_3] [(1 - \gamma)^2 - \rho^2]}.
\]

38
Note that both the mean and variance are increasing in both $\gamma$ and $\gamma_3$. Hence the mean of any power of $h$ is also increasing in $(\gamma, \gamma_3)$.

The inverse of the child’s earnings profile $1/p = (h_k - m'_k) / h'_k$, is

$$1/p = h^{2\gamma}/\gamma - \gamma.$$ 

Hence for a given level of $\gamma$, the average inverse earnings profile

$$\bar{1}/p = \bar{h}^{2\gamma}/\gamma - \gamma$$

is clearly increasing in $\gamma_3$.

Since $e_y = h_k - m'_k$, if $e_y + b_k < g$ the child will always choose $m'_k = h_k$, so that the earnings profile will increase indefinitely. We have seen above that if the child is subject to the welfare subsidy, his parent will invest zero. Considering the life-cycle, unless $e_o < g$, the parent will now invest a positive amount, but less than when the child is not subject to the subsidies. If $e_o < g$, again investments are zero.

Now consider a social security policy that is distributed according to $(e_y + e_o)/2$. This effectively decreases the importance of old age earnings, so the child will invest more in himself when young:

$$m'_{k^*} = \left[ \frac{(2 + r)\gamma a_k}{2(1 + r)} \right]^{1/\gamma}$$

$$e_{k^*} = 2 \left[ 1 - \gamma \left( \frac{(2 + r)\gamma a_k}{2(1 + r)} \right)^{1/\gamma} + h_k^* \right].$$

Since $m'_{k^*}$ is larger than before, the child’s earnings profile is steeper. The same should be true for the parent. While the parent has more resources to devote to the child (income effect), the child has more resources so there is less need to do so (substitution effect). In net, whether there is more education depends on how much the parent is life-cycle constrained.

The first result helps us identify $\gamma_3$. The intuition is quite clear. Since $\gamma$ matters both for intergenerational persistence and the earnings profile, the difference is captured by $\gamma_3$. When $\gamma_3$ is larger, high (low) human capital parents invest more (less) in the child. Whether or not this also increases the IGE, however, depends on other parameters. In our benchmark calibration, the first effect is much more dominant.

The life-cycle component interacts with the welfare subsidy. We have shown in the preceding subsection that it decreases investment in children but increase mobility among poor families.
This effect is dampened when considering the life-cycle since some individuals only receive the subsidy for one period. The dampening effect is even larger when considering social security benefits.
References


