Explaining Income Inequality and Intergenerational Mobility: The Role of Fertility and Family Transfers

Diego Daruich
New York University

Julian Kozlowski*
New York University

May 10, 2018

Abstract

Poor families have more children and transfer less resources to them. This suggests that family decisions about fertility and transfers increase income inequality and dampen intergenerational mobility. To evaluate the quantitative importance of this mechanism, we extend the standard heterogeneous-agent life-cycle model with earnings risk and credit constraints to allow for endogenous fertility, family transfers, and education. The model, estimated to the US in the 2000s, implies that a counterfactual flat income-fertility profile would—through the equalization of initial conditions—reduce intergenerational persistence and income inequality by about 10%. The impact of a counterfactual constant transfer per child is twice as large.

JEL Classifications: J13, J24, J62, D91.

Keywords: Inequality, Intergenerational mobility, Quantitative model, Fertility.

*ddaruich@nyu.edu; kozjuli@nyu.edu. We are grateful to our advisors, Raquel Fernandez, Ricardo Lagos, and Gianluca Violante. For helpful comments, we thank Jess Benhabib, Violeta Gutkowski, Mi Luo, Virgiliu Midrigan, Simon Mongey, Debraj Ray, Pau Roldan, Edouard Schaal, Laura Veldkamp, and Venky Venkateswaran. We also received valuable feedback from participants at seminars at Brown University, New York University, Di Tella University, ECINEQ 2015, EEA 2015, the 2015 Chicago Summer School on Socioeconomic Inequality, the Fall 2015 Midwest Macro Meeting, and SED 2016. We alone are responsible for any errors and interpretations.

What factors determine intergenerational mobility? Is income inequality mainly due to differences in the opportunities available early in life, or to adult income risk? We study the sources of income inequality and intergenerational mobility with particular interest in the impact of families. Extensive empirical evidence shows that family choices are heterogeneous and correlated with family characteristics: Poor families tend to have more children (e.g., Jones and Tertilt, 2008) and invest fewer resources towards their children than rich families (e.g., Altonji, Hayashi, and Kotlikoff, 1997). First, this heterogeneity in family choices can lead to differences in education outcomes, leading to higher levels of inequality relative to an economy without such heterogeneity. Second, this correlation can lead to lower intergenerational mobility, as the children of richer parents have more resources available for education. To evaluate the quantitative importance of these forces, we build a model in which families face a quantity-quality trade-off between having more children and making larger investments in them. Parents influence children's initial conditions regarding skills and economic resources, both of which shape their education choices and later labor income due to capital market imperfections. The model allows us to study the dynamic interactions between family choices and intergenerational mobility, which is our main contribution.

This paper first explores the empirical evidence on the relation between fertility and income, and its consequences on children's education. We use US Census data to exploit the within-country state variation. We confirm that, on average, poor families tend to have more children than rich families. However, we find that fertility differences between income groups are smaller in richer states. This result is robust to alternative definitions of fertility rates, as well as to different regression specifications. Then, we look at the relation between fertility differentials and inequality. We use older census data to estimate fertility differentials at the time of birth of individuals. More recent census data is used to evaluate education outcomes for individuals born in those states and time. We find that individuals born in states with larger fertility differentials are associated with higher education inequality.

We build a model to quantitatively evaluate whether fertility and child investment differences are important factors driving income inequality and intergenerational mobility. We introduce endogenous fertility, family transfers, and education in the standard heterogeneous agent life-cycle model, with idiosyncratic income risk and credit constraints in the spirit of Huggett, Ventura, and Yaron (2011). A key difference is that in our analysis, initial conditions—defined as the agents' initial state variables—are endogenously related to parental background. This change allows us to study intergenerational mobility. Parent-to-children transfers will be important in the model for children to access higher levels of education. In addition to parental resources, college loans are available to finance higher education. Our model also includes children-to-parent transfers (i.e., old-age support).[1] In addition to altruism, old-age support will provide another incentive to have children.

We estimate the model to the US in the 2000s and use it to analyze the impact of individuals' initial conditions. To estimate the novel elements in our model, we fit moments on the relationship between fertility and income, family transfers, and intergenerational mobility. A set of validation exercises shows that the quantitative model is consistent with both not targeted moments and cross-state evidence on the relationship between average income, fertility, and inequality. The variation in lifetime earnings—a measure of income inequality—can be decomposed into differences in initial conditions and in labor-income shocks. Our model suggests that the variation in lifetime earnings due to differences in initial conditions is 40%. In other words, 40% of lifetime-earnings inequality in the US can be attributed to family background.

We conduct two exercises to analyze the role of fertility and family transfers on inequality and mobility. First, we solve an alternative model in which fertility is exogenous and constant across families. This exercise reveals that in the baseline economy fertility accounts for 4% of the annual income inequality and 13% of the intergenerational mobility observed in the data. Second, we simulate an economy in which fertility is endogenous, but transfers from parents to children are exogenous and constant. These transfers play a major role in the observed inequality and social mobility in the US. According to our model, parental transfers account for 8% of annual income inequality and 29% of intergenerational mobility.

Both of these exercises operate through the distribution of initial conditions, particularly initial

[1]In Section 1 we estimate that the net average lifetime transfer from children to parents is significant, almost 60% of the average annual household income in the US. In the model, old-age support helps explain that fertility differentials are smaller in richer states.

assets (or parental transfers) and human capital. With a counterfactual flat income-fertility profile there are relatively less children born from poor households. As less children are born with low levels of initial human capital and assets, the initial distribution becomes more homogenous. First, an equalized initial distribution of assets leads to an increase in access to education. Since wages depend on education, this implies lower labor-income inequality. Second, a more homogenous initial distribution of human capital directly leads to lower labor-income inequality (independently of education). Note that in this counterfactual there are less children born from poor families and these children have higher levels of initial assets (parental transfers), which improves intergenerational mobility. With counterfactual constant transfers per child the initial distribution also becomes more homogenous, improving inequality and mobility through similar mechanisms. Our findings suggest that to understand income inequality and social mobility one should take fertility differentials and family transfers into account.

Related Literature

This paper relates to two literatures usually studied in isolation: income inequality and intergenerational mobility. However, there is a strong and positive correlation between the two (Corak, 2013). On the one hand, models of inequality typically focus on adult shocks and abstract from endogenous initial conditions (e.g., Keane and Wolpin, 1997; Huggett, Ventura, and Yaron, 2011). On the other hand, models of intergenerational mobility usually focus on initial conditions and abstract from adult income volatility (e.g., Restuccia and Urrutia, 2004; Lee and Seshadri, 2014, 2015). Both initial conditions and labor-income volatility generate income inequality. We contribute by providing a model that combines these two sources and assess their relative importance, which also allows for the joint study of inequality and mobility.[2]

The closest paper to ours is Huggett, Ventura, and Yaron (2011), in which the authors use a Bewley model to study the sources of inequality. They find that most of the income inequality is due to conditions present before entering the labor market.[3] However, these conditions

[2] The literature on quantitative models combining adult uncertainty and endogenous initial conditions is scarce. A relevant exception is Yum (2016) who focuses on parental time investment as the driver of initial conditions.

[3] There is also a small empirical literature that estimate the share of inequality explained by family background, but its estimates tend to be wide. For example, Niehues and Peichl (2014) estimate a lower and upper bound of 16 and 75%, respectively.

are exogenous in their analysis, implying that their results are silent about the forces that determine inequality of opportunity. Parental environment and investments have been shown to be determinant for childhood development and their adult outcomes (Murnane, Willett, and Levy, 1995; Cunha, Heckman, and Schennach, 2010). Hence, modeling family choices is necessary to study the origin of conditions early in life. Our results suggest that one-third of the share of inequality due to initial conditions as defined by Huggett, Ventura, and Yaron is due to education and young labor experience. Even though our paper abstracts from detailed early-childhood human-capital formation, as in Lee and Seshadri (2015), we provide a more precise answer to the relative impact of endogenous initial conditions by also including adult uncertainty.[4]

This paper highlights a quantity-quality trade-off à la Barro and Becker (1989) as a main determinant of initial conditions. There is evidence that poor families have more children than richer ones—i.e., there is a negative elasticity of fertility to income (Jones and Tertilt, 2008). We contribute to this literature by showing that this elasticity is smaller for richer states within the US. We also find that children born in states with larger fertility differentials are associated with higher education inequality. Our quantitative model is consistent with these empirical findings.[5]

Our model combines elements from the literature on adult income, intergenerational mobility, and fertility. We build a full life-cycle heterogeneous agents model—allowing for uninsurable shocks in the spirit of Huggett, Ventura, and Yaron (2011)—that also endogenizes earlier stages of life through choices regarding education, fertility, and family transfers. The calibrated model is able to match several moments on education, fertility, inequality, and intergenerational mobility for the US in 2000. The model is also consistent with the cross-state evidence. First, it is consistent with lower average fertility rates, as well as lower fertility differentials in richer

[4]Note that we focus on labor-income inequality and do not look into wealth inequality. Recent literature also finds a decisive role for family background in explaining wealth inequality (De Nardi and Yang, 2015; Luo, Bisin, and Benhabib, 2015). We also abstract from sorting, another force that has been used to generate inequality through families (e.g., Fernandez and Rogerson, 2001; Fernandez, Guner, and Knowles, 2005).

[5]Quantitative models in the fertility literature include Manuelli and Seshadri (2009) and Roys and Seshadri (2014), which are used to explain differences in average fertility rates across countries and long-term economic growth, respectively. Nevertheless, both abstract from uncertainty, and though heterogeneity is allowed in the second one, it is only in the form of constant skill differences across dynasties.

states. Second, it also captures the relation between education inequality and fertility differentials. To the best of our knowledge, this is the first paper able to explain the aforementioned facts, which are important for the joint study of family choices, income inequality, and intergenerational mobility.

The rest of the paper is organized as follows. Section 1 presents our empirical findings on fertility differentials, inequality, and family transfers. Section 2 introduces the model, and Section 3 explains its estimation and conducts some validation exercises. The model's results on inequality and intergenerational mobility are presented in Section 4. Finally, Section 5 concludes. The Appendix contains additional details.

1 Empirical Findings

Since we are going to analyze social mobility and income inequality through the lens of a model with fertility decisions and family transfers, we analyze the data available on these two. First, we use Current Population Survey (CPS) micro data to study the trends in the US over time. Then, in our preferred analysis, we use census data to exploit the within-country state variation. We find evidence that: (i) at the national level, there is a negative relationship between fertility and income, but it has diminished over time; (ii) states with higher levels of average household income are associated with smaller fertility differentials; (iii) individuals born in states with larger fertility differentials are associated with higher standard deviation of education outcomes; and (iv) old-age support—in the form of both money transfers and time—is sizable.

1.1 Fertility and Income

If children were considered a normal good, we should observe richer people having more children. However, this is not usually the case. Comparing over time, most countries have experienced a decrease in fertility (as they become richer). Comparing across countries, richer countries tend to have fewer children per family. More interestingly for our study, within a country-year it is also the case that richer people tend to have fewer children. Jones and Tertilt (2008) look at US Census data on women born between 1826 and 1960, and find substantial evidence that the relationship between income and fertility was stably negative. Controlling for several factors (for instance, urban versus rural families, location, or race), they suggest that economic factors play a large role in fertility decisions and that the negative relation with income is robust. We update their analysis for the US using micro data from the CPS—between 1968 and 2013—and census data—between 1960 and 2010—from the Minnesota Population Center (IPUMS). In all our analysis, income is defined as annual income at the family level, while our main measure of fertility is the Total Fertility Rate (TFR).[6]

Let $\text{inc}_{q,s,t}$ and $fert_{q,s,t}$ be the mean income and fertility rate, respectively, of income quantile q in region s and year t.[7] We allow for region identifier s for our cross-state analysis. We estimate

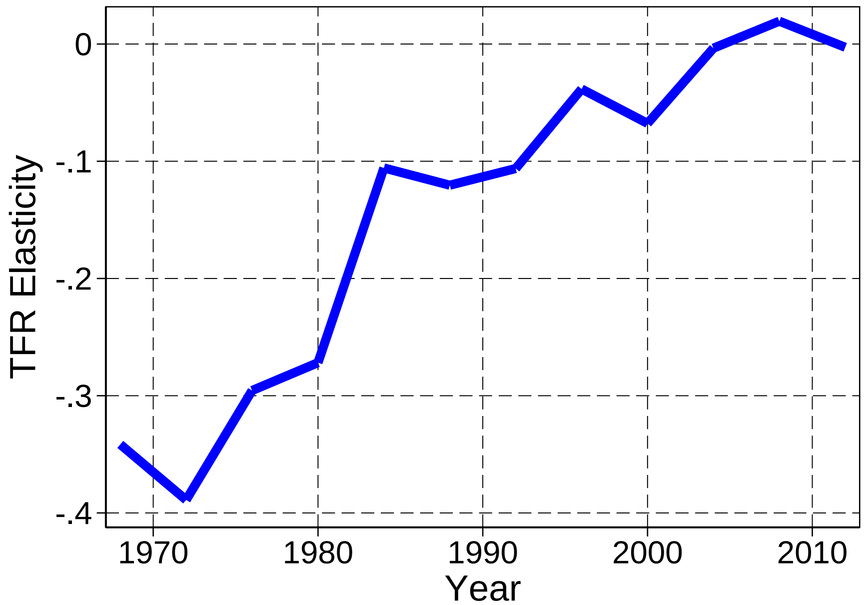
$$\ln\left(fert_{q,s,t}\right) = \alpha_{s,t} + \beta_{s,t}\ln\left(inc_{q,s,t}\right) + \epsilon_{q,s,t}, \tag{1}$$

where $\beta_{s,t}$ will be referred to as the *elasticity of fertility to income* for region s in year t. If this value is negative, richer households tend to have fewer children. Values closer to zero imply that fertility rates are not related to income (at least, according to this specification). Given our main interest on fertility decisions as a function of income, we do not want to mix single-parent households with two-parent households. Hence, we limit our analysis to "marital fertility," i.e., the fertility of those women who, when they answer the survey, indicate that they are married. Appendix A.2 reports details on the sample selection. Figure 1 shows the evolution of the elasticity of fertility to income for the US, with its value on the vertical axis. Figure 1 not only confirms that fertility elasticity has been negative since 1968, but also suggests that it has decreased over time, implying that the difference in the number of children between poor and rich households has become smaller.

[6] We note that our measures of fertility and income differ from Jones and Tertilt, and discuss our choices in detail in Appendix A.1.

[7] Using median income changes the results slightly, but they remain qualitatively the same.

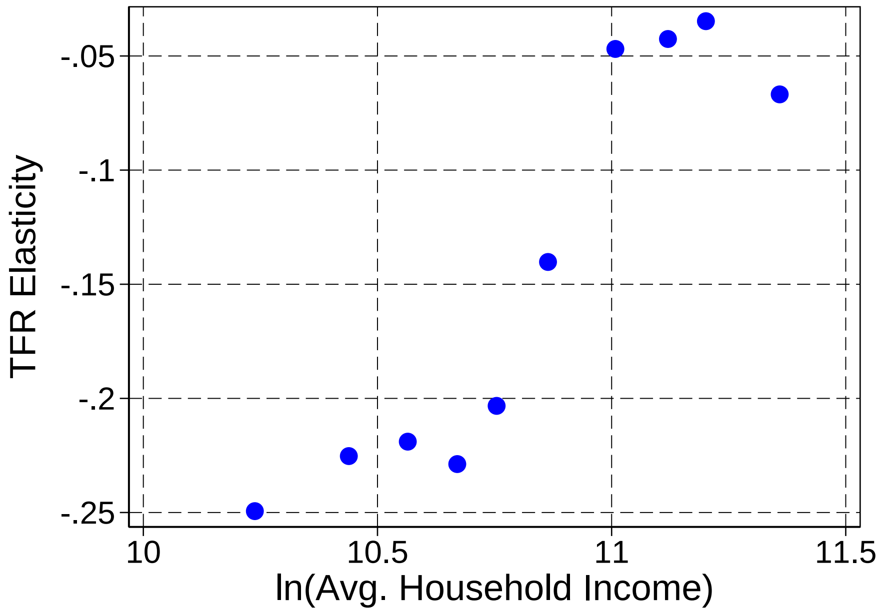
Figure 1: **Elasticity of fertility to income**.



Source: CPS. Years: 1968-2013. Observations are grouped in 3-year windows. Methodology is explained in the main text.

To better understand what is behind this pattern, we extend our analysis to exploit the cross-state variation using US Census micro data from IPUMS, i.e., for each state s and year t we estimate (1). We first analyze the data visually and, then, perform more formal statistical tests. For the visual inspection, we combine all our observations and divide them into deciles according to their levels of real average household income. For each of these groups we calculate the mean household income and fertility elasticity. Figure 2 shows that richer states tend to have smaller elasticities or, in other words, smaller fertility differentials.

Figure 2: **Elasticity of fertility to income by average household income**.

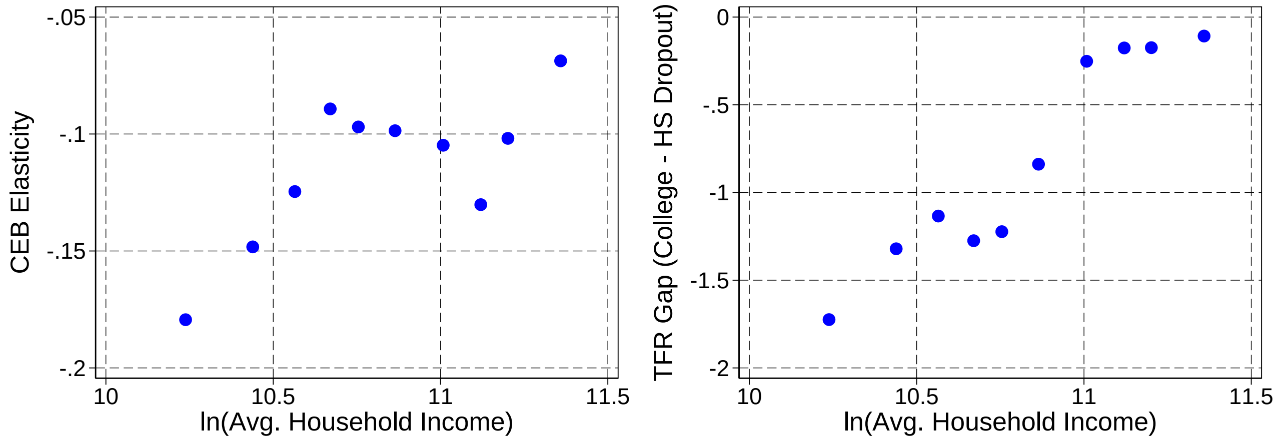


Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. We divide observations into deciles according to average household income. For each decile, we calculate the mean level of household income as well as the mean fertility elasticity of income. Figure A.1 in the Appendix includes all of the data observations before grouping them into deciles. Methodology is explained in the main text.

The pattern of higher fertility differentials being associated with lower average household income is robust to alternative measures of fertility.[8] First, instead of using TFR, we use Children Ever Born (CEB) to compute the fertility elasticity as in (1). The left panel of Figure 3 reports the results. Second, we can use the fertility differences between education groups: We calculate the difference between the TFR of women married to college-graduate men and that of women married to high school dropout men. The right panel of Figure 3 reports the results. Qualitative results in both cases are similar to those using the TFR elasticity to income from Figure 2: Richer states are associated with smaller fertility differentials.

[8] See Appendix A.1 for details on each of these fertility measures.

Figure 3: **Fertility Differentials: Robustness using alternative measures of fertility**.



Source: Census. Years: 1960, 1970, 1980, 1990, 2000 (only TFR Gap), and 2010 (only TFR Gap). We divide observations into deciles according to average household income. For each decile, we calculate the mean level of household income as well as the mean fertility differential measure. Figure A.2 in the Appendix includes all of the data observations before grouping them into deciles. Methodology is explained in the main text.

Various factors related to states' characteristics (e.g., culture) might be driving these results. To attempt to control for these possible concerns, we regress the fertility elasticities on the logarithm of the real average household income, controlling for state and time fixed effects. The regression specification is

$$\text{Fertility Elasticity}_{s,t} = \alpha + \gamma \ln(\text{Avg. Household Income}_{s,t}) + \eta_s + \mu_t + \epsilon_{s,t} \tag{2}$$

where Fertility Elasticity$_{s,t}$ is equal to the estimated $\hat{\beta}_{s,t}$ from (1). Table 1 shows that the elasticity of fertility (TFR) is increasing in real average household income. Once again, this implies that richer states are associated with smaller fertility differentials. This relationship seems stable and robust to controlling for state fixed effects and time fixed effects. This suggests that the pattern seen in Figure 1 may not be due simply to changes over time, but can be thought of as a general relation between average income and fertility differentials.

Table 1: **How the elasticity of fertility to income changes with average income**.

	(1)	(2)	(3)
Ln(Avg. Household Income)	0.228***	0.260***	0.243***
	(0.0169)	(0.0184)	(0.0802)
Observations	300	300	300
R-squared	0.355	0.487	0.582
# of States	51	51	51
State FE	NO	YES	YES
Year FE	NO	NO	YES

*Robust standard errors in parentheses. *, **, *** denote statistical significance at the 10, 5, and 1 percent, respectively. Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. Methodology is explained in the main text.*

How does inequality relate to fertility differentials? Intuitively, we would expect states with larger fertility differentials to present larger inequalities from the earliest stages of life. Larger fertility differentials imply that poor families have relatively more children, which would lead to a larger share of children being born with scarce resources. Assuming this affects their education, we would then expect to observe more inequality in states with larger fertility differentials. To test this hypothesis, we focus on the standard deviation of years of education of individuals born in different states in 1960, 1970, and 1980. For the sake of clarity let's look at how we study children born in 1960. First, we use data from 1960 to calculate the fertility elasticity, average household income and income inequality in each state in their year of birth (i.e., in 1960). Second, we use data from when that generation is 30 years old (i.e., in 1990) to calculate the standard deviation of education for each state.[9] Table 2 shows that individuals born in states with larger fertility differentials are associated with higher levels of inequality in education outcomes. This result is robust to controlling for mean household income and income inequality present in the state and year in which they were born, as well as state or year fixed effects.[10]

[9]The choice of the age and years of birth is limited by the timing of the Census data. To be consistent with the model presented in Section 2, we transform the years of education to groups: 8, 12, and 16 for those who are high school dropouts, high school graduates, and college graduates, respectively. Results are similar without this transformation.

[10]This evidence coincides with findings in Kremer and Chen (2002), who study income inequality and fertility differentials across education groups in different countries.

Table 2: **How education inequality relates to fertility elasticity**.

	(1)	(2)	(3)	(4)	(5)
Fertility Elasticity	-0.303***	-0.310***	-0.252***	-0.237***	-0.260***
	(0.0921)	(0.0921)	(0.0730)	(0.0734)	(0.0735)
Ln(Avg. Household Income)		0.0875**	0.222***	0.289***	-0.356**
		(0.0429)	(0.0420)	(0.0488)	(0.139)
Household Income Gini				2.105**	0.0117
				(0.917)	(0.845)
Observations	151	151	151	151	151
R-squared	0.067	0.098	0.761	0.784	0.816
# of States	51	51	51	51	51
State FE	NO	NO	YES	YES	YES
Year FE	NO	NO	NO	NO	YES

*Robust standard errors in parentheses. *, **, *** denote statistical significance at the 10, 5, and 1 percent, respectively. Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. We use data from 1990 on to calculate the education standard deviation—a measure of inequality—of individuals born 30 years earlier in each state. We then use data from 30 years earlier to calculate the fertility elasticity, average household income and income inequality in each of those states. Methodology is explained in the main text.*

1.2 Family Transfers: Old-age Support

Taking into account the importance in the development literature attached to old-age support when families choose the number of children, as well as the magnitude of the family transfers documented, we include old-age support as a motive for fertility in the model (e.g., Nugent, 1985; Banerjee, Meng, Porzio, and Qian, 2014). Moreover, this motive will explain the decrease in fertility differentials for richer states. We now introduce evidence that old-age support—in the form of both money transfers and time—is sizable.

Cox and Jimenez (1990) summarize the information on private transfers from 9 countries (including the US) and report that between 15 and 50% of people receive family transfers annually. The higher end of that range is dominated by developing countries. However, old-age support is not only monetary transfers. For example, when health problems arise, help from family members can be essential. Lundberg and Pollak (2007) claim that in the US, two thirds of the 5.5 million elderly with disabilities rely on family for help.

We use recent Panel Study of Income Dynamics (PSID) information from the 2013 Rosters and Transfers survey to study transfers between parents and children. As encountered by Abbott, Gallipoli, Meghir, and Violante (2013), data on family transfers are scarce and problematic; more work needs to be done to extract precise information from this type of source.[11] Therefore, we take the evidence presented here as mainly suggestive of one fact: Private transfers, with particular interest on those from children to parents, are substantial. For either direction of transfers, we calculate the average transfer over a lifetime.[12] Data are limited to children who are over 18 years old and do not include "long-term" transfers (for example, tuition or financial help to buy a house). The first column of Table 3 shows transfers going from parents to children, and the second column shows transfers going in the opposite direction. The first thing to notice is that these transfers are sizable, particularly when we include time provided to help either parents or children. Assigning them the mean wage value, we obtain that the average transfer from parents to children is almost 90% of the 2013 average annual household income. In turn, transfers from children to parents are almost 60% larger than those going in the opposite direction.

Table 3: **Family transfers over a lifetime**.

	Parent → Children	**Children → Parent**
Money	$38,589	$37,536
Hours	1,060	3,166
Total	$62,265	$108,222

Source: PSID Rosters and Transfers, 2013. Children are 18+ *years old. Data do not include transfers before that age nor those considered "long-term," such as those for tuition or buying a house. Total is calculated using the average hourly wage. Money and hours include cases of zero transfers.*

[11] Estimates on the size of transfers depend substantially on whether observations with zero transfers are excluded or not.

[12] The procedure is similar to the one used to calculate the TFR in Appendix A.1. We first calculate the average transfer given in the year before the survey (this is the question asked) by age groups. We then multiply this by the age width of the age group to obtain the average transfer given during that age window. Finally, we add across all age groups to obtain the average transfer over a lifetime. Different from Abbott, Gallipoli, Meghir, and Violante (2013), we also include observations with zero transfers. Note that we do not discount transfers to the present value.

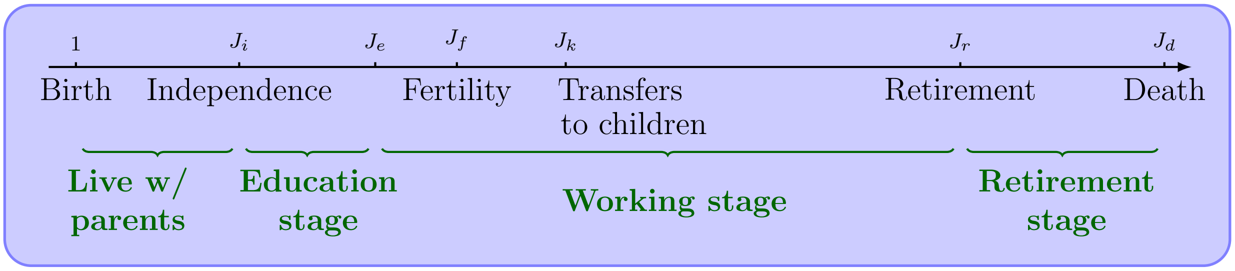
2 Model

We specify a life-cycle economy in a dynastic framework with three main stages. In the first stage, individuals make sequential education decisions: whether to acquire an extra level of education (first high school and then college) or start working. Education increases their human capital and modifies their life cycle of income, as well as the income distribution of their offspring. Once agents exit the education phase, they enter the second stage, which represents their labor market experience. Idiosyncratic uninsurable income risk makes individual earnings stochastic. Throughout their lives, agents choose savings and consumption expenditures. They can borrow only up to a limit, and save through a non-state-contingent asset. During this stage, they also choose how many children to have and how much of their resources to transfer to them. The last stage is retirement. At this time, agents have three sources of income: savings, retirement benefits, and old-age support from their children. We study the partial equilibrium version of this economy (i.e., prices and government policies are exogenous). We now describe the model and discuss the main mechanism.

2.1 The individual problem

Figure 4 shows the life cycle of an agent, in which each period in the model refers to four years. Let j denote age at the beginning of the period. From $j = 1$ until $j = J_i$ the child lives with her parents, who choose the child's consumption. At age $j = J_i$, the child becomes independent. Her initial states are assets, human capital, and school taste (or psychic cost). Initial assets are money transfers from her parents. The initial human capital and school taste are stochastic but correlated with the parents' education and human capital.

Figure 4: **Life cycle**



Agents can only trade risk-free bonds, but interest rates are different for saving and borrowing. Agents with positive savings receive interest rate equal to r, while those borrowing pay interest rate equal to $r^- = r + \iota$, where $\iota \geq 0$. The wedge between interest rates is important to capture the cost of borrowing, which is a form of insurance relevant for the quantitative analysis. Individuals face borrowing limits that vary over the life-cycle. Young workers (i.e., under the age of 20) and retired households cannot borrow. Student loans are explained in detail below. Let $e \in \{1, 2, 3\}$ be the current level of education of the agent, which stand for high school dropout, high school graduate, and college graduate, respectively. Workers with access to borrowing (i.e., after age 20) are subject to credit limit of $\underline{a}(e)$. Estimates of $\underline{a}(e)$ are based on self-reported limits on unsecured credit from the Survey of Consumer Finances.

Education stage: From $j = J_i$ until $j = J_e$, the agent has the option to study. The individual state variables are savings a, human capital h, and psychic cost ϕ. During education, she sequentially chooses whether to continue in school or start working. The education decision is irreversible. All agents become independent as high school dropouts ($e = 1$). If an agent chooses to stay in school, her education increases to $e+1$, while human capital evolves deterministically as $f^s_e(h)$. The monetary cost of education is p_e, but, as is common in the literature (e.g., Heckman, Lochner, and Todd, 2006; Abbott, Gallipoli, Meghir, and Violante, 2013), we also allow for psychic costs $\phi \in [0, 1]$ to affect the total cost of education.[13] Modeling school taste is necessary because resources available to finance schooling and returns to education can only partially account for the observed education patterns. Particularly, we assume that the

[13] We allow education costs and returns to differ between high school and college.

school taste enters as a separate term in the value function. We scale the school taste ϕ by a different constant in each schooling level $\bar{\psi}_e$. After leaving school, the psychic cost is assumed not to affect any adult outcome. While working, human capital evolves stochastically and is distributed by $f^w_{e,j}(h)$. We allow for education- and age-dependent idiosyncratic labor-income shocks. In Section 3, we discuss the estimation of the returns of education and the income process.

Students face borrowing limits $\underline{a}^s(e)$ for subsidized loans. High-school students cannot borrow (i.e., $\underline{a}^s(1) = 0$ since their *current* level of education is $e = 1$). College students have access to subsidized loans at rate $r^s = r + \iota^s$ where $\iota^s < \iota$. To simplify computation, we assume that college student debt is refinanced into a single bond that carries interest rate r^-. $\tilde{a}^s(a')$ is the function performing this transformation. When making this calculation we assume that fixed payments would have been made for 5 periods (i.e., 20 years) following graduation.[14] Borrowing limit $\underline{a}^s(2)$ and wedge ι^s will be based on federal college loans, to be explained in detail in Section 3.

Formally, let V^s_j and V^w_j be the value of an agent of age j in school and working, respectively. Let V^{sw}_j be the value of an agent who can choose between the two

$$V^{sw}_j (a, h, e, \phi) = \max \left\{ V^s_j (a, h, e, \phi), V^w_j (a, h, e) \right\},$$

where V^s_j is defined by

$$V^s_j (a, h, e, \phi) = \max_{c,a'} u(c) - \phi \bar{\psi}_e + \beta V^{sw}_{j+1} \left(\tilde{a}^s(a'), h', e+1, \phi \right) \tag{3}$$

$$c + a' + p_e - h w_e (1 - \tau) = \begin{cases} a(1+r) & \text{if } a \geq 0 \\ a(1+r^s) & \text{if } a < 0 \end{cases}$$

$$a' \geq \underline{a}^s(e), \quad h' = f^s_e(h)$$

[14]Given the fixed payment nature of student loans and the assumption that they are repaid in 5 periods, we can transform college loans into regular bonds using the following formula: $\tilde{a}^s(a') = a' \times \frac{r^s}{1-(1+r^s)^{-5}} \times \frac{1-\left(1+r^-\right)^{-5}}{r^-}$.

The agent is risk averse and her preferences are represented by an increasing, concave, and positive utility function u.[15] She can borrow up to the limit $\underline{a}^s(e)$, and the return on positive savings is $1+r$. However, if the agent is borrowing she pays interest rates $r^s > r$. Future is discounted by β. We denote as w_e the wage for an agent who is currently in school at level e. In particular, we assume that the agent does not work during high school (i.e., $w_1 = 0$), and we allow for (part-time or internship) work while in college (i.e., $w_2 \in [0, w]$).

The value of work V_j^w is defined by

$$V_j^w(a,h,e) = \max_{c,a'} u(c) + \beta \mathbb{E}\left[V_{j+1}^w(a',h',e)\right], \tag{4}$$
$$c + a' - hw(1-\tau) = \begin{cases} a(1+r) & \text{if } a \geq 0 \\ a(1+r^-) & \text{if } a < 0 \end{cases}$$
$$a' \geq \underline{a}(e), \quad h' \sim f_{e,j}^w(h).$$

The agent can borrow up to the limit $\underline{a}(e)$, and the return on positive savings is $1+r$. However, if the agent is borrowing she pays interest rates $r^- > r$. The return from working is the wage w net of taxes τ. There is no disutility from working, and so the labor supply is inelastic. Also recall that the choice to leave education is irreversible.

Working stage: From $j = J_e$ until $j = J_r$, the agent works and her individual problem is equivalent to (4). There are two special periods in which the agent problem will be different, and the number of state variables will change from then on. First, in the exogenously given fertility period $j = J_f$, the agent chooses the number of children. Once the children become independent (at $j = J_k$), the agent chooses the transfer to her offspring. Second, when the agent's parents retire (at $j = J_t$), she provides old-age support, transferring a fraction of her current labor income to her parents.

Fertility: We model altruism à la Barro and Becker (1989), in which parents care about the

[15] The fact that the utility function u is positive is necessary to model altruism. As shown by Jones and Schoonbroodt (2009), the implicit assumption that parents enjoy having children requires that the utility function must be always positive or always negative. If we choose the negative case, we need an extra assumption for the value of having zero children. Therefore, we follow the classic approach of u being always positive and assume that having zero children delivers zero utility.

utility of their children. The problem at the age of fertility $j = J_f$ is

$$
\begin{aligned}
V_j\left(a,h,e\right) = \max_{c,c_k,a',n} \; & u(c) + \beta \mathbb{E}\left[V_{j+1}\left(a',h',e,n\right)\right] + b(n)u(c_k) \qquad (5)\\
& c + nc_k + a' + C(h,n) - hw\left(1-\tau\right) = \begin{cases} a\left(1+r\right) & \text{if } a \geq 0 \\ a\left(1+r^{-}\right) & \text{if } a < 0 \end{cases}\\
& a' \geq \underline{a}(e), \quad h' \sim f^{w}_{e,j}(h), \quad n \in \left\{0,1,\ldots,\overline{N}\right\}.
\end{aligned}
$$

In this period, the agent chooses her consumption c, her children's consumption c_k, savings a', and the number of children n, which is a discrete choice. As usual, the agent derives utility from her own consumption and her continuation utility. Furthermore, similar to Roys and Seshadri (2014), the agent is altruistic and derives utility from her children's consumption. The altruistic discount factor $b(n)$ is positive, increasing, and concave.

Raising children is costly, as is reflected in (5). Parents pay the cost $C(h,n)$ in addition to the money spent on children's consumption and transfers. This cost is assumed to be increasing in the number of children n and in the level of human capital of the parents h. The functional form for C is specified in Section 3. Until the agent's children become independent ($j = J_k$), she chooses the children's consumption and pays the cost C. Hence, the problem is equal to (5), but takes the number of children n as given.

The transfer to each child φ is assumed to be made in the period before the offspring become independent (age $j = J_k$). Moreover, transfers are assumed to be the same for all children.[16] The problem at the age when transfer to children is chosen $j = J_k$ is

$$
\begin{aligned}
V_j\left(a,h,e,n\right) = \max_{c,c_k,a',\varphi} \; & u(c) + \beta \mathbb{E}\left[V_{j+1}\left(a',h',e;\Omega\right)\right] + b(n)\left\{u(c_k) + \beta \mathbb{E}\left[V_{J_i}\left(\varphi,h_k,\phi_k\right)\right]\right\} \qquad (6)\\
& c + nc_k + a' + \frac{n\varphi}{(1+r)} + C(h,n) - hw\left(1-\tau\right) = \begin{cases} a\left(1+r\right) & \text{if } a \geq 0 \\ a\left(1+r^{-}\right) & \text{if } a < 0 \end{cases}\\
& a' \geq \underline{a}, \quad h' \sim f^{w}_{e,j}(h), \quad h_k \sim f^{k}(e,h), \quad \phi_k \sim g^{k}(e)
\end{aligned}
$$

[16] The altruism value derived from children depends on their initial assets. Therefore, we assume that (one period in advance) parents set a fund for their children to receive φ when they become independent.

where Ω contains information about the agent's offspring relevant for old-age support. This is discussed in more detail in the description of the retirement stage. Notice that unlike (5), the value function at this stage now includes the continuation value of the children V_{J_i}. This is the last period in which parents' choices affect their descendants. As the problem is written recursively, this implies that at every period in which parents' choices affect children's outcomes, the value function of their descendants will be taken into account. This embeds the parental altruism motives. The initial human capital and the psychic costs of the children are stochastic, but correlated with the parents' level of education and human capital. The functional form of the altruism, as well at the stochastic processes of human capital, $f^k(e, h)$, and psychic costs, $g^k(e)$, are specified in Section 3.

After the agent's children become independent, the individual problem is equal to (4), except for the following difference: At $j = J_t$ a fraction of the current labor income goes to the agent's parents as old-age support. There are different ways to introduce old-age support. Altig and Davis (1993) allow for double-sided altruism (i.e., children taking into account their parents' utility, as well as parents taking into account that of their children) in a three-period model without heterogeneity. Even in this much simpler model, double-sided altruism brings many difficulties, which lead Altig and Davis to abstract from linkages and strategic behavior from the agents. More recent attempts to include only altruism on the children's side have also been limited to representative agent economies without uncertainty (Boldrin, Nardi, and Jones, 2005). Although not in a family framework, Barczyk and Kredler (2014) solve for the equilibrium transfers in a risk-sharing problem between two agents with double-sided altruism. In our heterogeneous agent model—which includes more than two agents (parents and, potentially, many children); multiple stages of discrete choices (education and fertility); and high dimensionality of the state space—the altruism approach to old-age support is computationally infeasible. Hence, we adopt the simpler rule that children are constrained to transfer an exogenous share $\xi \in [0, 1]$ of their income to their parents, as in Morand (1999). We believe that endogenous old-age support should actually strengthen our proposed channel, since—abstaining from strategic behavior issues—poor parents would expect to receive relatively higher transfers from their children, since their needs are more pressing than those of richer parents.[17] This would make the old-age support motive even greater for poor families.

Retirement stage: At $j = J_r$, the agent retires with three sources of income. On the one hand, she has savings and retirement benefits that depend on her education level and human capital and are progressive. On the other hand, at the first period of retirement the agent receives transfers from her children as old-age support. Parents need to predict how much money they will receive from their children when they get old. An extreme view is that parents know their children's income perfectly, updating it year by year. Independent of the plausibility of this view, in our model this would require extending the state space to include that of each child. Added to the current dimensionality of the model, such a procedure would become computationally infeasible. Hence, we assume that parents have limited information to predict the transfers they will receive. The only information that parents have about their children is $\Omega = \left\{n, \varphi, h_{J_f}\right\}$ and e, i.e., the number of children n, the initial assets of their children φ, and their own education e and human capital when the children were at home h_{J_f}.[18] These state variables remain constant until retirement and are used to predict the old-age support the agent will receive. Note that old-age support is an endogenous random variable whose distribution depends on the children's education choices.[19]

Formally, the problem at the age of retirement is

$$V_j\left(a, h, e, \theta\right) = \max_{c, a'} u\left(c\right) + \beta V^w_{j+1}\left(a', h, e\right), \tag{7}$$

$$c + a' = \theta + \pi(e, h) + a\left(1 + r\right),$$

$$a' \geq 0,$$

[17] The marginal utility of consumption for poor parents would be higher than for richer families, increasing the incentive for children of poor parents to transfer.

[18] The education level and human capital of the parents is essential for the initial draw of human capital and psychic cost of the children. In the solution, we use the human capital from the period before the children become independent to reduce the state space in previous periods. The initial assets help pin down how long children will remain in school. Finally, the number of children n is used to determine, among others, the mean and variance of the transfer to be received.

[19] We remark that the difference with respect to the full-information case is such that uncertainty about children's transfers lasts longer, but at the time of fertility the information available is equal. Hence, in a model with exogenous labor supply like ours, the assumption should only affect savings. As our focus is on labor-income inequality, we believe that our assumption is not harmful.

where θ are the old-age support transfers and π are the retirement benefits, which depend on the education and human capital at the age of retirement.[20]

2.2 Fertility choices

Three main incentives for having children and making transfers to them are at play: (i) altruism, by which parents care about their children's well-being; (ii) old-age support, by which parents care about the help their children will provide them once they grow old; and (iii) the cost of raising children.[21] Altruism implies that parents want to have educated children (because this has a positive effect on their children's income). In the calibrated version of the model, altruism provides incentives for agents to choose a rather constant number of children across income groups. Old-age support implies that children are a private investment for retirement. Due to differences in parents' human capital and education, the opportunity cost of raising children as well as children's distribution of future income vary across income groups. This implies that children have different rates of returns across families. However, children require investment at an inconvenient time in the life cycle, since they are costly at a stage of life in which parents would actually like to borrow.[22] More importantly, for high-income parents, children are particularly costly due to the time they require. Moreover, their return in future transfers is not particularly relevant, since they will have significant other sources of income. The opposite is true for low-income households, which have low time costs and obtain high marginal returns from children. Thus, the old-age support channel is much more valuable for poorer households. In this manner, the altruism channel dominates for richer households, while the investment channel offers an extra incentive for poorer agents to have children.

To understand the qualitative effects, consider a simplified version of the model in which choices

[20] We use education, together with the last level of human capital, as a proxy to approximate average lifetime income with which are the retirement benefits determined. See Section 3 for details.

[21] Intergenerational persistence of human capital and school taste could potentially influence fertility choices. However, Appendix B.2 shows that this channel is not quantitatively relevant for the fertility rate, the fertility elasticity to income, or the average transfers to children.

[22] For example, in the US the current average age of first birth is 27, while the income peak is closer to 50 years old.

of the quantity of children, as well as transfers to them, are contemporaneous and there are no future costs of raising children. Moreover, assume that n is a continuous variable and abstract from the utility derived by the consumption of the children. All these arguments are standard, but complicate the interpretation.[23] Under these assumptions, the first-order condition with respect to n is

$$\frac{\partial u_{J_f}}{\partial c}\left(C_n+\varphi\right)=b_n\left(n\right)\beta\mathbb{E}_{J_f}\left[V_{J_i}\left(h_k,\varphi,\phi\right)|e,h_{J_f}\right]+\beta^{J_r-J_f}\frac{\partial\mathbb{E}_{J_f}\left[V_{J_r}\left(h,a,\theta\right)|n,e,h_{J_f}\right]}{\partial n}. \quad (8)$$

The left-hand side is the marginal cost of an extra child—that is, the transfer φ and the marginal opportunity cost C_n, scaled by the parents' marginal utility at age of fertility $\frac{\partial u_{J_f}}{\partial c}$. The right-hand side is composed of two terms. The first term is the benefit from altruism, and the second is from old-age support. The former channel is standard, with the payoff being the expected value function of the child scaled by the marginal effect of n on the altruism, $b_n(n)$, and discounted by β. The benefit of old-age support is generated by the change in the distribution of the transfers θ. For the sake of clarity, we can decompose the random variable θ into its conditional mean and a martingale shock $\varepsilon\sim g_\varepsilon$. This implies that

$$\frac{\partial\mathbb{E}_{J_f}\left[V_{J_r}|n,e,h_{J_f}\right]}{\partial n}=\mu_\theta\mathbb{E}_{J_f}\left[\frac{\partial u_{J_r}}{\partial c}\right]+\int\left(\int\frac{\partial u_{J_r}}{\partial c}\frac{\partial g_\varepsilon\left(\varepsilon|n,e,h_{J_f}\right)}{\partial n}d\varepsilon\right)df\left(h_{J_r}|e,h_{J_f}\right)$$

where $\mu_\theta=w\left(1-\tau\right)\xi\mathbb{E}\left[h^k_{J_t}|e,h_{J_f}\right]$ is the expected transfer of each child. $h^k_{J_t}$ is the human capital of the child at the age that she provides old-age support, $j=J_t$. The first term is the effect on the conditional mean and shows that as n increases, so does the expected transfer. Once again, note that this is scaled by the expected marginal utility at the age of retirement. The second term reflects the higher-order-moments effects.

With this first-order condition, we can learn why richer agents underweight old-age support and fertility choices are dominated by the altruism channel. Note that both the marginal cost and the old-age support benefits are scaled by the marginal utility of consumption. Moving toward richer agents, the marginal utility diminishes, and so does the old-age support channel. The net effect on the cost depends on the relative magnitudes of the decrease in the marginal

[23]See Appendix B.1 for details about the simplified model.

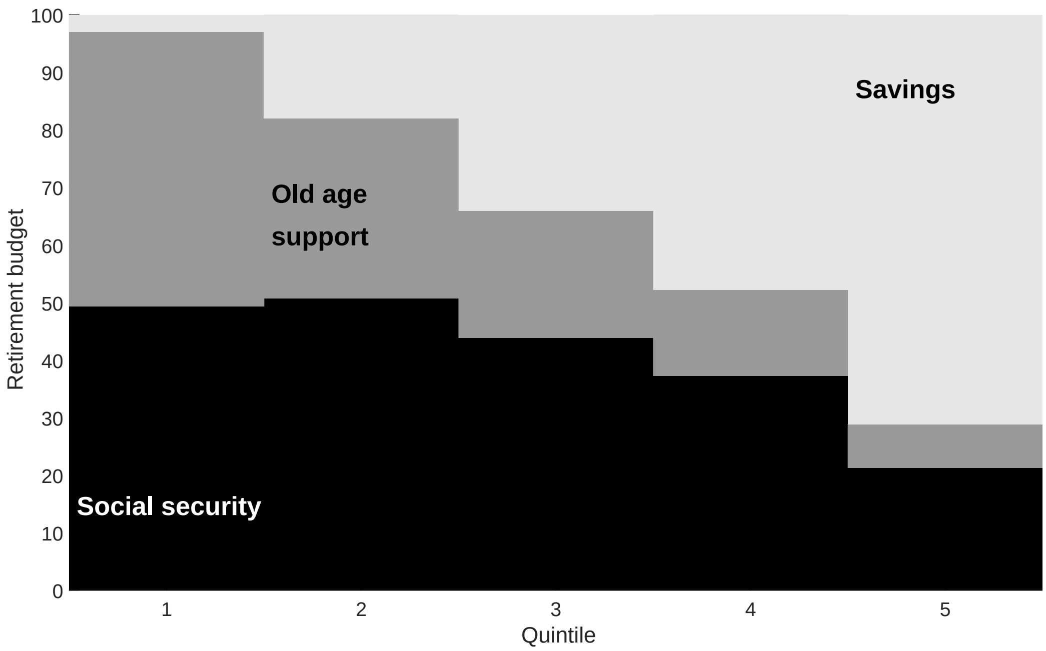
utility and increase in opportunity cost. The altruism benefits are not changed, and therefore dominate the fertility decision over old-age support. Consequently, this implies that as the economy grows—for instance, as w increases—fertility choices will be driven by altruism instead of old-age support. This framework generates the non-homothetic relationship between average income and fertility choices documented in Section 1. Altruism will be the main fertility driver for richer economies in general, as well as for rich agents in poorer ones. In addition to the altruism motive, poorer economies or low-income agents will take into account the old-age support channel, leading them to have more children and generating the negative fertility elasticity. Therefore, the expected return on old-age support is larger for poor families, which reinforces the negative income-fertility relationship.[24]

We quantitatively evaluate the different motives of fertility around the calibrated model. In Appendix B.2 we study the numerical comparative statics of the moments related to fertility with respect to different parameters related to each motive for fertility. In particular, we fix all parameters at their estimated values, and change only one parameter at a time in a neighborhood of the estimated value. First, this exercise shows that altruism significantly influences the transfers to children, while keeping the mean fertility rate rather constant. Second, the opportunity cost is the major driver of the mean fertility rate. Third, old age support leads to large changes on the fertility elasticity to income since it is more important for poorer agents. This is also observed when we look at the resources available during retirement. At that stage there are three sources of income: Social Security, savings, and old-age support transfers. Figure 5 shows the average contribution of each source across income quintiles, under the benchmark calibration described in Section 3.[25] For the poorest quintile, old-age support represents over 45% of resources, but represents about 7% for the richest quintile.

[24]Once an economy is sufficiently rich, it is possible for children to behave as a normal good. Every household prefers educated children, and rich families can afford more of them.

[25]Recall that old-age support transfers occurs only at age J_r but Social Security benefits are received throughout retirement. In Figure 5, we compare the net present value of Social Security with the stock of savings and the old-age support transfers.

Figure 5: **Budget's Assets Distribution by Income Group**.



Average income contribution of social security, savings and old-age support for each income quintile. Results depend on the estimation of the model, shown in Section 3.

3 Estimation

The model is estimated to match household level data. Therefore, an agent in the model corresponds to a household with two adults in the data. Consequently, the number of children n is also in terms of households—i.e., $n = 1$ refers to one household.[26] We use three primary data sources: (i) IPUMS US Census; (ii) CPS Fertility Supplement; and (iii) 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). We select a population for which our model can be taken as a reasonable approximation to household behavior and impose two selection criteria on the data. First, as is standard in the literature (e.g., Huggett, Ventura, and Yaron, 2011), we drop household observations with income below a certain threshold. We choose this threshold as the one that corresponds to one person working 20 hours a week for the minimum wage (approximately $8,000 total annual household income). Second, there is no decision regarding marriage in our model. Given our focus on fertility, we are interested

[26] We set the maximum possible number of children to 6, so $\overline{N} = 3$.

in two-member households. To avoid differences in income and time availability due to single parenthood, we keep only married households. Details about sample selection are reported in Appendix C.1.

We numerically solve the steady state of this economy. Due to the presence of nonlinearities and discrete choices, we implement a global solution method. Some of the computational challenges are that we have up to six state variables and several non-convexities due to the discrete choices in education and fertility. Therefore, we apply a generalized endogenous grid method adapted from Fella (2014). We then compute the ergodic distribution of the economy to match moments from the US in 2000.

We describe below how we parameterize the model economy. Some of the parameters can be estimated "externally," while others must be estimated "internally" from the simulation of the model. Table 4 summarizes all the parameters in the model.

3.1 Preliminaries

Demographics: A period in the model is four years. Individuals become independent at the age of $J_i = 12$, and they start with the equivalent of 7 years of education. They can go to high school (one period) and then to college (another period), and so the maximum age for education is $J_e = 20$. Fertility decisions are made around the average age at first birth, $J_f = 28$. At age $J_k = 36$, one period before the agent's children become independent, she chooses the assets to transfer to her children. Retirement occurs at $J_r = 68$. Therefore, when the agent is $J_t = 40$, the agent's parents retire and she transfers money to them. Death is assumed to occur for all agents at age $J_d = 80$.

Prices: Prices are normalized such that the average income of a high school graduate at age 40 is equal to one in the model. In the data, this income is equal to $\$60,198$. In order to do this adjustment, it is equivalent to move the average initial level of human capital or the wage. We choose to do the first alternative (and, also, fix the wage $w = 1$). We estimate the wage while in college from IPUMS census data. We focus on individuals between the ages of 18 and 22 years old and match the relative earnings of those currently in college relative to those who are not, leading to $w^s = 0.56$. Following Smets and Wouters (2007), we set the annual interest rate to $r = 3\%$. Based on self-reported limits on unsecured credit by family from the Survey of Consumer Finances, we set $\underline{a}(e)$, the borrowing limits for working-age individuals, to $\{-10,000, -24,000, -34,000\}$ for high school dropout, high school graduate, and college graduate, respectively. The payroll tax is $\tau = 0.124$, which is the current rate for Social Security. The yearly price of college is from the Delta Cost Project, where we get $\$6,588$.[27] The yearly price of high school is obtained from the Digest of Education Statistics, using the relative private cost of high school to college. Our estimate of high school cost is about 9% of college cost, which is consistent with the US education system (i.e., relatively low cost of high school when compared to college), leading to a price for high school of $593. Taking into account that education takes one period (4 years) and households contain two members, our normalization leads to $p_1 = 0.08$ and $p_2 = 0.90$.

College Loans: College students have access to subsidized loans at rate $r^s = r + \iota^s$. According to the National Center for Education Statistics report "Student Financing of Undergraduate Education: 1999-2000," among the undergraduates who borrow, nearly all (97%) took out federal student loans—only 13% took out nonfederal loans. Moreover, the average loan value was similar for both federal and nonfederal cases. Since average values were similar but federal loans were significantly more common, we focus on federal loans for our model estimation. Among federal loans, the Stafford loan program was the most common: 96% of the undergraduates who borrowed took out Stafford loans. The second most common loans were the Perkins loans, but they were much smaller: only 11% of borrowers used Perkins loans and average amounts were one quarter of average Stafford amounts. Therefore, we focus particularly on Stafford loans. Stafford offers multiple types of loans so we use the weighted average interest rate to set $\iota^s = 0.009$. The borrowing limit while in college in the model is the set to match the cumulative borrowing limit on Stafford loans ($23,000).

[27] We take into account grants and scholarships, such that only private tuition costs are considered. Prices are in 2000 US dollars.

School Taste: In this class of models, it is difficult to match the high school dropout rate. Previous studies (e.g., Abbott, Gallipoli, Meghir, and Violante, 2013; Krueger and Ludwig, 2015) introduced nonpecuniary (psychic) costs of education. We assume the agent's psychic cost ϕ is between 0 and 1, which will be scaled by different estimated levels according to the education stage ($\bar{\psi}_1$ and $\bar{\psi}_2$). Its distribution is related to parents' education through the parameter ω. Particularly, we assume that the psychic cost for children of high-school-graduate parents is uniformly distributed in that range. On the other hand, we assume that the probability of high psychic costs for children of high school dropouts is increasing in ω, and decreasing for those of college graduates. Hence, the CDF of school taste is

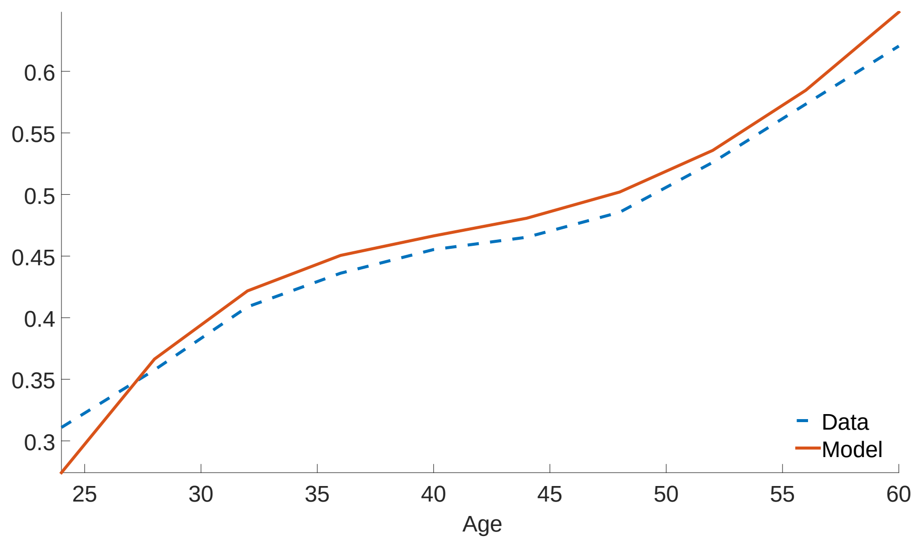
$$G^k(e, \phi) = \begin{cases} \phi^{\omega} & \text{if parents are high school dropouts} \\ \phi & \text{if parents are high school graduates} \\ \phi^{-\omega} & \text{if parents are college graduates} \end{cases}$$

Our estimation suggests that psychic costs are higher for children of less educated parents, which is consistent with previous estimates in the literature.

Education returns: Returns to education are allowed to vary between high school and college as well as between agents, as suggested by Heckman, Lochner, and Todd (2006). Particularly, we specify the human-capital production function to have the nonlinear form $h' = f_e^s(h) \equiv \alpha_e h^{\beta_e}$ for $e \in \{1, 2\}$. Table 4 shows that our estimates for high school are $\alpha_1 = 1.66$ and $\beta_1 = 0.87$, while for college they are $\alpha_2 = 1.72$ and $\beta_2 = 0.53$. If we estimated the return to education at age 40-43, our model would suggest average yearly returns (in the whole population, not just those that do study) of 13%. On the other hand, if we estimated the returns using lifetime earnings and taking into account education costs, the returns would be reduced to an average of 9% per year. These numbers are in line with empirical estimates in the literature summarized in Heckman, Lochner, and Todd (2006).

Labor-income risk: We assume that $h' = f_{e,j}^w(h) \equiv h(1 + \delta)$ where δ is stochastic and can take three values that vary by age j and education e. Together with their probabilities, they are calibrated using the Rouwenhorst method to match the first difference of mean and variance of log earnings between the ages of 24 and 63, by education.[28] Two comments are appropriate. First, income risk is calibrated to include total earnings variation, encompassing what may be considered both wage shocks and hours worked (or effort) differences. Second, even though we propose a simple model of income, we are able to match standard statistics of labor earnings. This is necessary to properly evaluate the impact of initial opportunities on income inequality. Otherwise, the comparison could be favorable for initial opportunities. Figure 6 shows that the variance of log income is well estimated, though slightly higher for older ages. Moreover, in Table 6 we will show that the persistence of the income process is correctly fitted.

Figure 6: **Variance of Log(income) by age**.



Source: Census. The model is simulated at the estimated parameters, and the variance of log(income) is computed for each age.

Opportunity cost of children: The functional form assumed is $C(h,n) = n^{\alpha_n}\beta_n wh(1-\tau)$. This function allows for non-constant returns to scale in the number of children. If an agent has n children, her income is reduced by $n^{\alpha_n}\beta_n$. We estimate the returns to scale $\alpha_n = 0.65$, based on Table 6.4 in Folbre (2009). We internally estimate $\beta_n = 0.10$.

Old-age support: With data from the PSID Rosters and Transfers 2013, we back out the average net transfers to parents over the lifetime (after the age of 18 and not including schooling

[28]We assume that the income-shocks distribution is constant before 24 and after 63, as data are problematic for those ages.

costs). In the model, old-age support occurs only in one period, so we input all these transfers as if they were given only at age J_t. Based on Table 3, we estimate that the fraction of income that goes to parents as old-age support is 15%. Recall that this includes both time and money transfers in the data.[29]

Replacement benefits: The pension replacement rate is based on the Old Age, Survivors, and Disability Insurance federal program. We use education level, as well as the level of human capital at the moment of retirement, to estimate the average lifetime income, on which the replacement benefit is based. See Appendix C.2 for details.

Intergenerational transmission of human capital: We assume that the initial (i.e., at age J_i) level of human capital is stochastic, but correlated with the parents' human capital and education. The initial draw of human capital will be given by

$$log(h_{J_i}) = log(\bar{h}) + \rho \left[log(f(e, h_p)) - log(\bar{f}(e, h_p))\right] + \varepsilon \tag{9}$$

where $\varepsilon \sim N(-0.5\sigma_{h_0}^2, \sigma_{h_0})$ and $f(e, h_p)$ is the mean income of households with education e and income group given by the parents' human capital h_p. Recall that the mean of this distribution is chosen such that the average labor income of a high school graduate at age 40 is normalized to one. This defines f^k, the distribution of the initial draw of human capital in the household problem (6).

Preferences: We specify the period utility over consumption as a CRRA function

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

As discussed in Section 2, the utility function has to be positive, and therefore $\gamma_c \in [0, 1)$. We follow the literature and assume that $\gamma_c = 0.5$ (e.g., Roys and Seshadri, 2014). Other articles, like Manuelli and Seshadri (2009), that have estimated this parameter also obtain roughly this value. As is standard in the literature, the altruism function is assumed to be $b(n) = \lambda_n n^{\gamma_n}$.

[29] Note that in the data, children are restricted to be older than 18 years—but in the model, children become independent at the age of 13. However, based on Table 2 of Hill and Stafford (1980), we estimated that of the time parents devote to children before the age of 18, about 90% is actually spent before the age of 13.

Thirteen parameters of the model are estimated using Simulated Method of Moments. Two parameters, λ_n and γ_n, are related to altruism. σ_{h_0} is the standard deviation of the initial distribution of human capital. ρ relates to the intergenerational persistence of human capital through the initial draw. β_n is the opportunity cost of raising children. α_e and β_e, for $e \in \{1, 2\}$, define the returns to education in high school and college. $\bar{\psi}_e$, for $e \in \{1, 2\}$, defines the distribution (both mean and standard deviation) of the school taste, while ω is related to the correlation with parents' education level. Finally, ι is the wedge in the interest rate between saving and borrowing.

3.2 Simulated Method of Moments: Moments' selection

We internally estimate $K = 13$ parameters in order to match K moments. Although the model is highly nonlinear, so that (almost) all parameters affect all outcomes, the identification of some parameters relies on some key moments in the data. Figure 7 shows the result of the following identification exercise. First, given an hypercube of the parameter space, we draw 200,000 candidate parameter vectors from uniform Sobol (quasi-random) points, and compute the implied moments in the model. Second, for each parameter we associate a relevant target moment. Third, for each parameter, we divide the vector of this particular parameter in 50 quantiles and compute the 25th, 50th, and 75th percentiles of the associated moment in each quantile.[30] Finally, we show these percentiles of the moment along with the value in the data. We claim that a moment is important for a parameter's identification if, as we move across quantiles, the percentiles of the associated moment change and cross the horizontal dashed line (i.e., the value of that moment in the data). The slope of each curve shows how important is that parameter for the associated moment (a steeper curve implies the moment is more informative). The difference between the 25th and 75th percentiles informs about the relative importance of the remaining parameters (other parameters are more important when the 75th and 25th percentiles are further apart).[31]

[30] Notice that for each quantile there are $K - 1$ parameters that are randomly draw from the uniform Sobol points, and, therefore, potentially far away from the estimated parameter.

[31] Similar patterns are observed when we perform a local identification exercise in which we keep all parameters fixed at their estimated value and change only one parameter to study the numerical comparative statics of the

The success of this exercise relies on finding a relevant moment for each parameter. The data on mean fertility, fertility elasticity, and transfers from parents to children identify altruism and opportunity cost parameters, as shown by the first row of Figure 7. More precisely, there is a positive relation between the level of altruism (λ_n) and transfers to children. As parents value more their children (higher λ_n), they increase the transfers to them. Similarly, there is a negative association between the curvature of altruism (γ_n) and fertility elasticity. When $\gamma_n = 0$, the marginal value (due to altruism) of an additional child is equal to zero which implies that all parents have (at most) one child. However, when γ_n is positive the quantity-quality trade-off generates a negative fertility elasticity. Finally, as the opportunity cost increases, it becomes more costly to raise children, which implies a decreasing mean fertility rate.

moment with respect to the parameter in the neighborhood of the estimated parameter. Hence, this exercise is local as it relies on the estimated parameter. Results available upon request.

Figure 7: **Identification**.

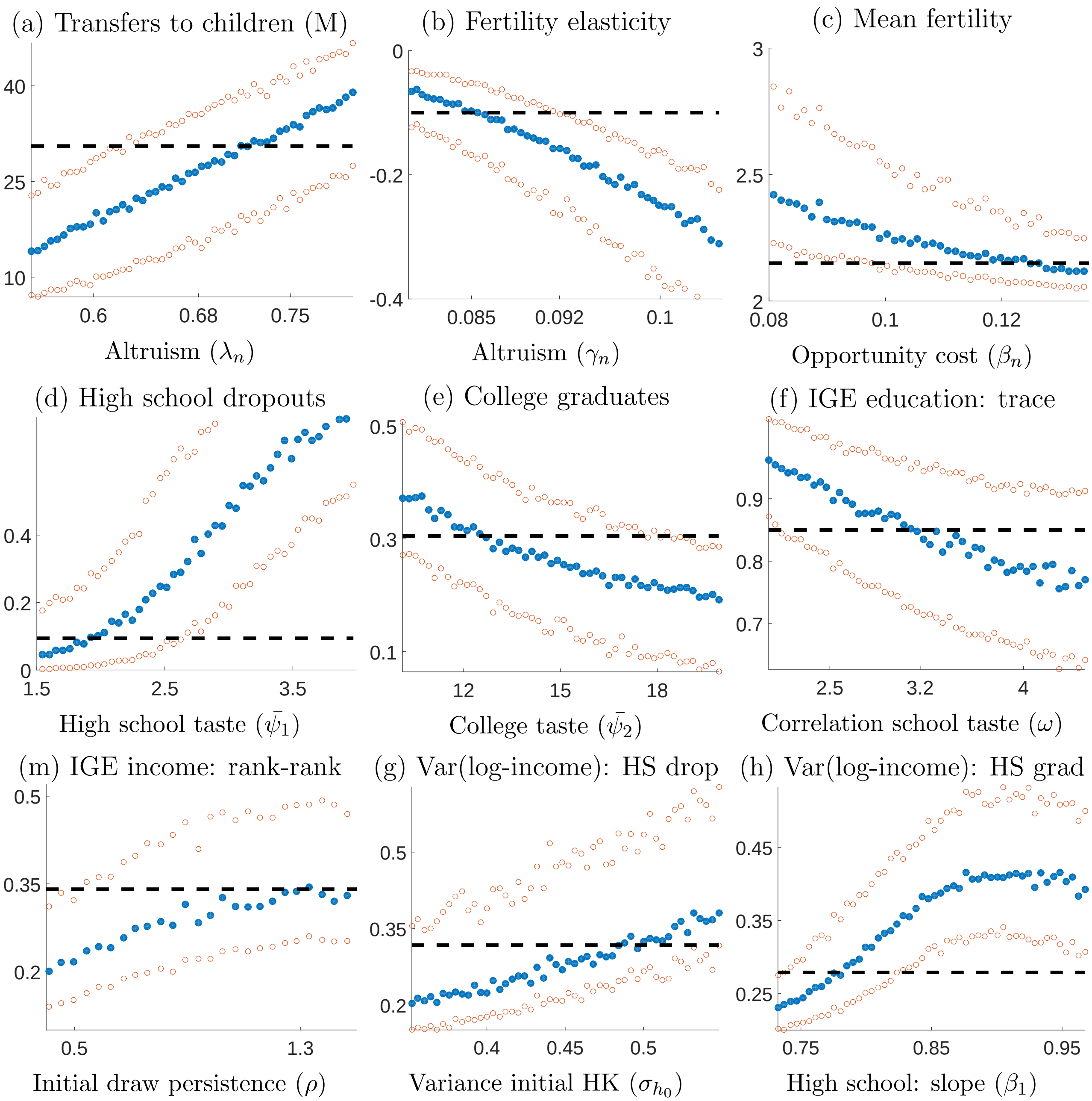
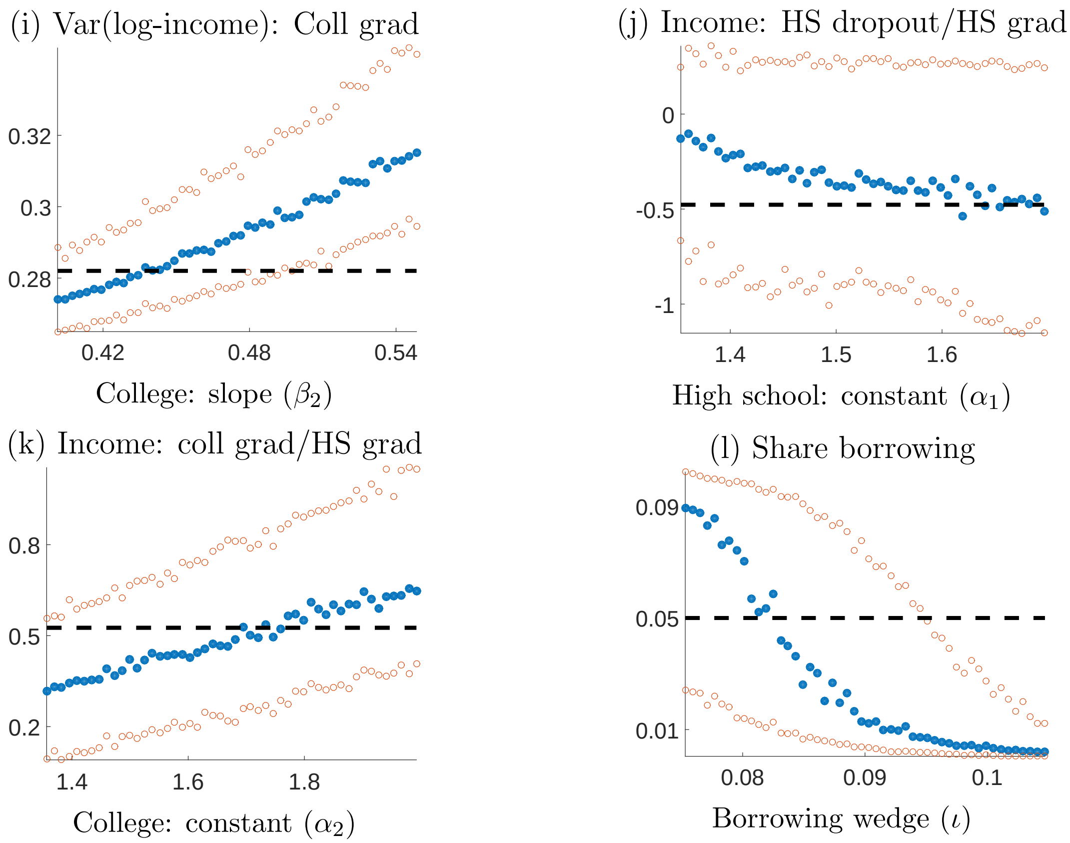


Figure 7 (cont.): **Identification.**



For each parameter's quantile, the (filled) blue dot shows the median while the (empty) red dots show the 25th and 75th percentiles of the assigned moment. The black dashed line shows the value of the moment in the data. Variance of log-income and income ratios are measured at ages 28–31 and 40–43, respectively. Methodology is explained in the main text.

Similarly, the next five plots of Figure 7 show that data on high school dropouts, college graduates, intergenerational persistence (of education and income), and income of high school dropouts identify school taste and initial human capital parameters. In particular, for larger values of high school's (college's) taste shocks, we observe more high school dropouts (less college graduates). In addition, higher correlation between parents' education and child's school taste implies lower intergenerational mobility of education.[32] Similarly, higher persistence in the initial draw lead to a lower intergenerational mobility of income (higher rank-rank coefficient). Finally, note that the variance of log-income of high school dropouts is positively related to the

[32] The trace index of education mobility is defined as $(3 - \text{trace}(P))/2$ where P is the transition matrix of education. Hence, zero mobility would imply an index equal to zero while perfect mobility implies an index equal to one.

variance of the initial draw of human capital. As the initial draw becomes more dispersed, the variance increases.

The next four plots of Figure 7 show that data on relative income between high school dropouts, high school graduates, and college graduates, and variance of log-income of high school and college graduates income identify returns to education.[33] Income ratios identify levels of education returns, while coefficients of variation of income identify curvature parameters. In particular, a higher curvature implies a larger variation of income. Finally, the last figure shows that the share of individuals with negative assets is informative for the interest rate wedge.

3.3 Simulated Methods of Moments: Results

Table 4 shows the estimated parameters while Table 5 shows the moments in the simulated economy. Parent-to-children transfers (average) as well as fertility (average and income elasticity) are successfully matched, which is necessary given their key roles in our model. As for income inequality, the model displays levels similar to the data. Education shares are well matched in the model. We also obtain levels of intergenerational mobility that are close to the empirical evidence.

[33]Recall that education increases human capital by a constant α_i and a slope β_i, where $\ln(h') = \alpha_i + \beta_i \ln(h)$, $i = 1$ for high school, $i = 2$ for college.

Table 4: **Estimated parameters**.

Parameter	Value	Description	Source
Demographics			
	4	Time period	
J_i	12	Independent	
J_e	20	Maximum age for education	
J_f	28	Fertility decisions	
J_k	36	Transfers from parent to children	
J_t	40	Transfers from children to parents	
J_r	68	Retirement	
J_d	80	Death	
Prices			
$\bar{y}_{40}^{HS} = w \times \bar{h}_{40}^{HS}$	$w = 1$, $\bar{h}_{40}^{HS} = 1$	Average Income: HS Graduate, Age 40-43	Normalization
w_2	0.56	College wage	Census
τ	12.4%	Payroll tax	Social Security
p_1	0.08	Price of high school	Digest of Education Statistics
p_2	0.90	Price of college	Delta Cost Project
Financial markets			
r	3%	Interest rate (annual)	Smets and Wouters (2007)
$\underline{a}(1)$	-10	Borrowing limit of HS dropout ($1k)	SCF
$\underline{a}(2)$	-24	Borrowing limit of HS graduate ($1k)	SCF
$\underline{a}(3)$	-34	Borrowing limit of college graduate ($1k)	SCF
ι	10%	Borrowing-saving wedge (annual)	Internally estimated
ι^s	1%	College loan wedge (annual)	NCES
Education			
(α_1, β_1)	(1.66,0.87)	High school return	Internally estimated
(α_2, β_2)	(1.72,0.53)	College return	Internally estimated
$\bar{\psi}_1$	2.59	Maximum high school psychic cost	Internally estimated
$\bar{\psi}_2$	18.39	Maximum college psychic cost	Internally estimated
ω	2.44	Correlation of psychic cost's with parents	Internally estimated
Income process			
Estimated to match mean and variance growth by age and education. See text and Figure 6.			Census
Childcare			
β_n	0.10	Cost	Internally estimated
α_n	0.64	Returns to scale	Folbre (2008)
Old-age support			
χ	15%	Share of labor income transfered to parents	PSID Rosters and Transfers 2013
Retirement benefits			
	$(0.3, 2, 4.1) \times \bar{y}$	Bend points	US Social Security System
	$(0.9, 0.32, 0.15)$	Replacement benefits	US Social Security System
Intergenerational transmission of ability			
ρ	0.57	Intergenerational persistence of initial draw	Internally estimated
σ_{h_0}	0.40	Standard deviation of initial draw	Internally estimated
Preferences			
β	0.96	Discount factor (annual)	
γ_c	0.5	Risk aversion	Roys and Seshadri (2014)
(λ_n, γ_n)	(0.71,0.09)	Altruism	Internally estimated

Note: Prices are normalized using the average income of a high school graduate at age 40, $\bar{y}_{40}^{HS} = \$60,198$*, based on IPUMS.*

Table 5: **Target moments**

Moment	**Data**	**Model**	**Source**
Education			
Dropouts (%)	9.4	9.5	Census
College graduates (%)	30.5	30.3	Census
Log-income ratio: HS Dropout - HS Grad, Age 40-43	-0.48	-0.50	Census
Log-income ratio: College Grad - HS Grad, Age 40-43	0.53	0.56	Census
Family Choices			
Average fertility	2.15	2.11	CPS
Fertility elasticity to income	-0.13	-0.13	Census
Mean transfer to children ($)	30,566	32,167	Abbott et. al. (2013)
Mobility and Inequality			
Intergenerational mobility of income	0.34	0.31	Chetty et. al. (2014)
Intergenerational mobility of education	0.85	0.82	Checchi et. al. (1999)
Variance of log-income: HS Dropout, Age 28–31	0.32	0.32	Census
Variance of log-income: HS Graduates, Age 28–31	0.28	0.28	Census
Variance of log-income: College Graduates, Age 28–31	0.28	0.28	Census
Financial Markets			
Share with negative assets	0.05	0.05	SCF

Source: Census refers to the IPUMS census data in the year 2000. CPS refers to the CPS Fertility Supplement for the most recent years (2010 and 2012), given that we use "Children Ever Born" to estimate these moments. Intergenerational mobility of education is measured using the trace index, as defined in the main text. SCF refers to the Survey of Consumer Finances.

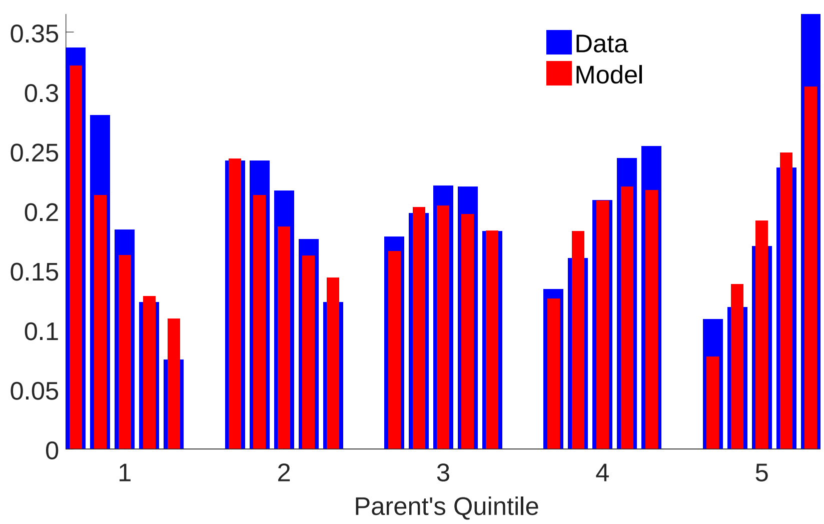
3.4 Validation Exercises

We can test the validity of our estimated model by looking at moments that are not directly targeted. First, we evaluate the model within the estimated steady state given by the parameters in Table 4. The, we look how the model compares with the cross-state evidence reported in Section 1 by moving away from the steady state. Table 6 summarizes the results of these exercises.

First, capturing the correct persistence of income in this type of models is important, as it determines the social mobility within the working lifetime. We estimate an income process for each education group similar to Heathcote, Storesletten, and Violante (2010), but using household income and 4-year periods instead.[34] The first panel of Table 6 shows that the coefficients that determine the persistence of shocks by education level are close to 0.9 in all cases. This is in line with what we find in the data.

Next, we present evidence that the model is able to capture income inequality and intergenerational mobility. The estimation targeted the variance of log-income—a measure of inequality. Table 6 shows that the model is also in line with another measure of inequality, i.e. the Gini coefficient.[35] Regarding mobility, in our estimation we target the rank-rank coefficient but we can compare all the entries of the transition matrix in the model and the data. Figure 8 shows that the model captures the transition probability between generations for all 25 entries of the matrix. Moreover, in the estimation we used the trace of intergenerational transition matrix of education. We check that an alternative measure of education mobility—the determinant of such matrix— is also well fitted by the model. We find that this measure both in the model and in the data is equal to 0.90.

Figure 8: **Intergenerational mobility**.



Source: Chetty, Hendren, Kline, and Saez (2014). For each of the five parent's income quintiles, the bars represent the distribution of children according to their income between ages 28 and 31. Note that, for each set of five bars, the first bar represents the first quintile, while the last one is the fifth quintile. For example, the share of children born to parents in the first quintile who also end up in the first quintile is given by the first bar on the left, i.e. almost 35%.

[34]Details on Appendix C.3.

[35]Other measures of inequality such as the coefficient of variation or the top-bottom are also similar in the estimated model and the data.

Next, we evaluate fertility decisions within different education groups by comparing fertility elasticities to income. The model generates elasticities equal to -0.18, -0.17, and -0.06 for high school dropouts, high school graduates, and college graduates, respectively. In the data these elasticities are similar (-0.22, -0.17, and -0.08). More importantly, both model and data display a decreasing (in absolute terms) relation with education. Overall, the estimated model is quantitatively consistent with fertility choices that were not targeted in our estimation.

Cross-state We do a simple exercise to show that the model is consistent with the cross-state patterns described in Section 1. Recall that in the benchmark calibration, the wage was normalized to one. Consequently, to generate economies with different levels of average household income, as in the data, we move wages, such that the real wage (i.e., in consumption terms) is the main change. The size of wage movements is such that average income in the simulations is in the range of the corresponding empirical estimates. Note that this involves moving the model away from the steady state to which it was estimated. The last panel of Table 6 shows three cross-state estimations, both in the data and the model.

The first two rows in the last panel of Table 6 refer to the relation between average income and fertility. First, the model is able to capture the negative relationship between average income and the fertility rate, as estimated in Appendix Table A.2. The model generates a slope coefficient between the fertility rate and GDP per capita of -0.18, which is close to the empirical estimate of -0.12. More importantly, the model is able to capture the relation between average income and fertility elasticity, as shown in Table 1. The model generates a slope coefficient of 0.26, which is similar to the data value of 0.22. The last row of Table 6 regards inequality of education outcomes and fertility differences across income groups. Table 2 presented evidence that children born in states with larger fertility differences between income groups are associated with larger inequality of education. The slope coefficient was found to be -0.30 in the data, which is at least qualitatively similar to the model outcome of -0.60. We take this as evidence that the model can also capture our main patterns of interest outside of the economy on which the benchmark is estimated.

To summarize, the exercises in this section provide evidence that the model is consistent with many empirical facts that were not used in the estimation procedure. This holds true both for moments within the steady state estimated for the US in 2000 and moments outside such steady state but related to cross-state evidence.

Table 6: **Validation exercises**.

Moment	**Data**	**Model**	**Source**
Income Persistence:			
Dropouts	0.89	0.89	NLSY79
High school graduates	0.88	0.90	NLSY79
College graduates	0.93	0.91	NLSY79
Inequality and Mobility			
Income Gini	0.38	0.39	Census
Determinant of education transition matrix	0.90	0.90	Checchi et al. (1999)
Fertility			
Elasticity within Education Groups			
Dropouts	-0.22	-0.18	Census
High school graduates	-0.17	-0.17	Census
College graduates	-0.08	-0.06	Census
Cross-State Evidence: Regression Coefficients			
Fertility Rate to Avg. Income	-0.12	-0.18	Census
Fertility Elasticity to Avg. Income	0.22	0.26	Census
Education Inequality to Fertility Elasticity	-0.30	-0.60	Census

Source: For the income-persistence estimates, we follow Heathcote, Storesletten, and Violante (2010). Regression coefficients in the data are shown in Tables A.2, 1, and 2.

4 Sources of Inequality and Intergenerational Mobility

We now use the model to estimate the role of initial opportunities on income inequality and intergenerational mobility. We look at two alternative measures of the relevance of initial conditions. First, we follow the inequality of opportunity literature and focus on an entropy decomposition of the income distribution. Second, following Huggett, Ventura, and Yaron (2011), we look at the variation in lifetime earnings that can be attributed to initial conditions (*VLE-IC*). Lastly, we study the role that family choices play in *VLE-IC* as well as inequality and intergenerational mobility. Table 7 summarizes our main results. Focusing on the columns Data and Benchmark, the first panel shows that, as discussed in the previous section, the model generates levels of intergenerational mobility consistent with the data. The second panel looks at different measures of inequality of opportunity. The third panel focuses on the variance decomposition of lifetime earnings. Finally, the third and four columns evaluate the role of fertility and family transfers.

4.1 Inequality of Opportunity

Following the literature on equality of opportunity (EoP), we first define an *outcome* of interest such as average lifetime earnings or income at age 28-31.[36] We then divide the population into different *types* according to their parents' income and education, or to their initial conditions (human capital, parents' transfer and school taste). Finally, we evaluate the effect of types on the outcome of interest by comparing the distribution of outcomes of different types. Equality of opportunity is defined to be achieved when the conditional distributions of outcomes are equalized across types. A suitable way to compare conditional distributions is based on entropy measures that allow us to decompose inequality, both between and within types. As is standard in the literature, we use the Theil-L index, which is explained in Appendix C.4. The relative Theil-L index is the share of total inequality that is due to differences between types. If it is equal to zero, all the variation in outcomes is independent of types.

The empirical problem with the Theil-L index is that many of the variables that potentially determine types and outcomes are rarely observable. This leads to wide estimates in the empirical literature. For instance, Niehues and Peichl (2014) estimate a lower bound of inequality of opportunity of 16% and an upper bound of 75%. The lower-bound estimate focuses on observable initial differences like gender, place of birth, race, or parents' occupation and education. Many of these do not exist in our framework, so we replicate their exercise using the only variables in our model that could potentially be observable. Identifying parents' income and education as

[36] There is a large literature on distributive justice—including equality of opportunity—which has recently been summarized by Roemer and Trannoy (2015).

initial conditions, the third row of Table 7 shows that our estimate is 11.9%, close to but below the lower bound from these empirical studies.[37] This approach has two main disadvantages. First, focusing on income at a given age can be misleading, as lifetime utility is more closely related to lifetime earnings, and temporary shocks could play a bigger role over short periods than longer ones. Furthermore, life-cycle income paths could differ by individuals (for example, due to education groups), making any single year a bad proxy for lifetime earnings, which is not usually available in the data. Second, parents' income can be understood as a proxy for other elements that are closer to initial conditions (e.g., educational investment, time spent with children or job opportunities). Yet these are also not available in the data.

One advantage of having a structural model is that we can look at more interesting measures of outcomes like lifetime earnings instead of only income at a particular age. Using parental characteristics as types, the fourth row shows that the Theil-L index increases from 11.9% to 15.1% when the outcome is changed from income at age 28-31 to lifetime earnings. Moreover, with a structural model we can also identify the true initial conditions in our economy: parents' transfer, initial human capital, and school taste. The fifth and sixth rows of Table 7 show that inequality of opportunity using initial conditions to define types is much higher: 59.7% and 55.9% for income at age 28-31 and lifetime earnings as outcomes, respectively.

Our model provides evidence that it is important to carefully define types when studying inequality of opportunity. By comparing the alternative estimates above we conclude that not choosing the appropriate characteristics that define initial conditions can lead to a large downward bias in the estimation of inequality of opportunity. Hence, our preferred definition of inequality of opportunity identify the model's initial conditions as types and lifetime earnings as outcome, providing an estimated value of 55.9%.

[37] This is not surprising, since our model is focused on married households and does not include initial conditions such as race, gender, or place of birth.

4.2 Variance Decomposition of Lifetime Earnings

An alternative way to evaluate the importance of initial conditions is to decompose the variance of lifetime earnings into variation due to initial conditions and variation due to adult shocks. We can perform this exercise not only for the actual initial conditions (i.e., when the agent becomes independent) but also for the conditions given at different ages (i.e., the state variables at older ages). Define *VFE-CS*(j) as the fraction of the variance of current and future earnings explained by current state variables at age j. To understand *VFE-CS*(j), it is helpful to start at the end of the life-cycle and iterate backwards. In any period, the current state variables determine current labor income. However, future labor income is subject to shocks. Therefore, in the last period of work, all shocks determining labor income have been realized and current state variables explain all future earnings. One period before, the agent knows its current labor income but its next (and last) period's income is subject to an idiosyncratic shock. Iterating backwards towards the initial period, the agent faces more uncertainty about future labor income and as a result the current state provides less information about future earnings. Figure 9 shows that *VFE-CS*(j) is increasing in age and converges to 100%. Interestingly, the curvature of this figure reveals that shocks received between ages 20 and 30 are important to predict future income.

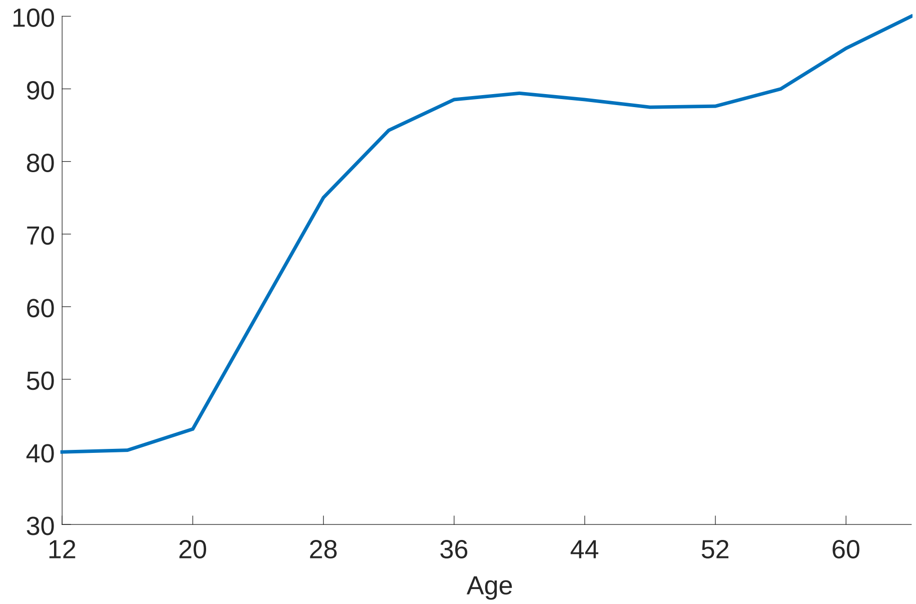
We define initial conditions as the state variables of the agent when become independent (i.e., at age 12). In this case, the variation of lifetime earnings due to initial conditions (hereafter, *VLE-IC*) is equal to 40%. Notice that, given our notation, *VLE-IC* is equal to *VFE-CS*(12). Huggett, Ventura, and Yaron (2011) define initial conditions as of age 23 and found that *VLE-IC* was 61%. In our model, Figure 9 shows that at age 23 current state variables explain 59% of future earnings, an estimate close to Huggett, Ventura, and Yaron. Therefore, we conclude that one third of the of the *VLE-IC* found by Huggett, Ventura, and Yaron can be explained by education choices and young employment experience.

Table 7: **Social mobility and inequality**.

	Benchmark	**Constant fertility**	**Constant transfers**
Social mobility			
Intergenerational Mobility: Rank-Rank	0.31	0.27	0.22
Transition Par Q1-Child Q5 (%)	10.9	11.8	13.0
Inequality of opportunity: Theil-L relative (%)			
Types: parent's income and education			
Outcome: Income, Age 28-31	11.9	9.7	7.1
Outcome: Lifetime Earnings	15.1	12.7	10.6
Types: initial conditions			
Outcome: Income, Age 28-31	59.7	57.0	54.4
Outcome: Lifetime Earnings	55.9	53.4	51.9
Variance decomposition			
CV of Lifetime Earnings	0.71	0.70	0.69
VLE-IC: % expl. by all initial conditions	40	38	37
% expl. by adult income shocks	60	62	63
Total	100	100	100

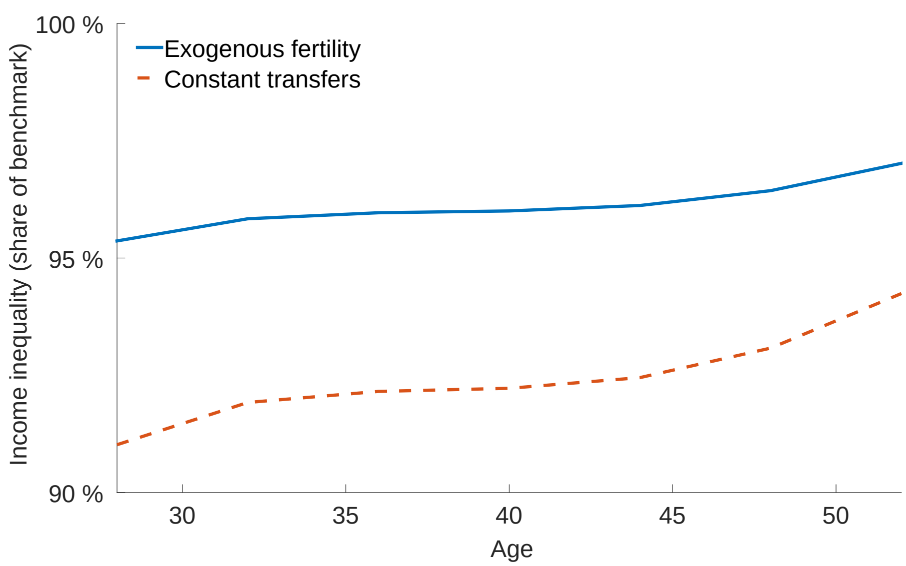
Note: Exogenous fertility solves the benchmark model in which fertility is restricted to be equal to $N = 2$. *Similarly, Constant Transfers refers to the case in which parents-to-children transfers are exogenously given at average value of transfers in the benchmark economy,* $30,566.

Figure 9: **Variance of future earnings explained by current state**.



Note: For each age, we show the share of the variance of future earnings predicted by the current state variables, VFE-CS(j).

Figure 10: **Income Inequality and Family Choices**.



Income inequality is measured using the variance of log income. For each alternative model, we plot the resulting inequality by age as a share of the benchmark economy, i.e., 95% implies a 5% reduction in inequality.

4.3 The Role of Families

The Role of Fertility We can use our model to study the role of endogenous fertility on inequality and intergenerational mobility. For this we keep the same model and estimated parameters from Section 3, but examine the case of constant fertility (assuming fertility es exogenous and each household has two children). Using our measure of choice—variance of log income by age groups—Figure 10 reports that in such a society, inequality would be reduced by around 4%.

Without fertility differentials, Table 7 shows that the rank-rank intergenerational mobility would improve by 13%. This is connected to the reduction of *VLE-IC* to 38%, implying that fertility differences between income groups account for 2 percentage points of the estimated *VLE-IC*.[38,39]

With a counterfactual flat income-fertility profile there are relatively less children born from poor households. As less children are born with low levels of initial human capital and assets, the initial distribution becomes more homogenous. An equalized initial distribution of assets leads to an increase in access to education. Since wages depend on education, this implies lower labor-income inequality. Moreover, a more homogenous initial distribution of human capital directly leads to lower labor-income inequality (independently of education).

Regarding social mobility, in this counterfactual there are less children born from poor families. Note that among this group there is low upward mobility (Figure 8). Holding the initial distribution of these children constant, average intergenerational mobility improves. Moreover, the initial distribution changes since these children now have higher levels of initial assets (parental transfers). This improves their chances to have a higher income as adults, leading to an even higher intergenerational mobility.

[38] Table 7 also shows the role of fertility on inequality of opportunity, with similar conclusions to the ones on *VLE-IC*.

[39] We remark that if we were to remove old-age support—one of the main motives for fertility differentials in our model—we would arrive at a similar conclusion.

The Role of Family Transfers We can also look at the case in which we allow for endogenous fertility, but parents' transfers are exogenously constant at the average value in the data.[40] Similarly to removing endogenous fertility, Figure 10 shows that with constant transfers at the benchmark's average level, the income inequality would be reduced by around 7.5%.

Imposing an exogenous positive transfer to children would improve intergenerational mobility by 29%. With constant family transfers, the large role that initial assets played in education choices is eliminated, and most lifetime earnings are instead due to characteristics less directly related to parents' income. The fact that the initial level of human capital and psychic costs are less strongly correlated to parents' income leads to the improved social mobility. This improvement in mobility is associated with a 20% reduction in the variance of years of education, mostly driven by a decrease in the share of high-school dropouts. Moreover, reducing this major source of initial differences leads to the reduction in the role of initial conditions. Without differences in initial resources, *VLE-IC* can decline to 37%, implying a reduction of 3 percentage points.

Implications Consistent with the early-childhood investment literature, the model suggests that improving children's conditions early in their lives can have a significant impact on their future outcomes (Heckman, Moon, Pinto, Savelyev, and Yavitz, 2010; Gertler, Heckman, Pinto, Zanolini, Vermeersch, Walker, Chang, and Grantham-McGregor, 2013). Moreover, policies that are successful in increasing the resources available to all children earlier in life would reduce inequality and improve intergenerational mobility, through a reduction in inequality of opportunity. Even though fertility differentials play a modest role relative to parental transfers in the US, this may not be the case for other countries with larger fertility differentials. According to our model, this implies that policies that reduce the incentives of poorer households to have children may be successful in improving inequality and social mobility in such countries.

[40] In unreported results we fix the amount transfered to different levels. For instance, we obtained similar results when the transfer is equal to 150% of the cost of high-school and college ($\$43,085$).

5 Conclusion

This paper analyzes the roots of social immobility and income inequality, trying to disentangle the importance of differences in opportunities determined early in life relative to differences in experiences over the working lifetime. We use a standard heterogeneous agent life-cycle model with idiosyncratic risk and incomplete markets extended to account for the role of families (through endogenous fertility, family transfers, and education) in determining initial opportunities. The model also allows for human capital transmission from parents to children. We propose that fertility differentials between rich and poor households can lead to substantial differences in the resources available for children, which can be important for their adult outcomes. The model is able to capture evidence on the relation between fertility differentials, income inequality, and intergenerational mobility. Income risk is calibrated to include total earnings variation, encompassing what may be considered both wage shocks and hours worked differences. Typical statistics on adult income risk are well captured by the model, which is required for an impartial comparison of the importance of adult risk relative to initial conditions.

We find that initial opportunities (as of age 12) accounts for 40% of the lifetime earnings inequality, while adult income risk over the working life accounts for the remaining 60%. Fertility differentials and family transfers, respectively, account for 2 and 3 percentage points of the share of the variation in lifetime earnings due to initial conditions. More importantly, these two forces separately generate 4 and 8% of annual income inequality, as well as 13 and 29% of the intergenerational immobility observed in the data. Doepke and Tertilt (2016) argue that there is a potentially large role for family economics within macroeconomics. Our results are consistent with this: those interested in understanding inequality, intergenerational mobility, or inequality of opportunity may need to take fertility differentials and family transfers into account.

Our results suggest that improving access to education (through education subsidies, for example) might diminish the importance of family transfers, helping, in turn, to reduce income inequality and improve social mobility. Even though our model is silent about the forces that determine the initial level of human capital (though calibrated such that the correlation between parents' and children's human capital holds, as in the data), it can still shed light on its importance for the levels of inequality observed in the data. Research on the determinants of initial human capital (e.g., Cunha, Heckman, and Schennach, 2010) or on ways to improve it (e.g., Gertler, Heckman, Pinto, Zanolini, Vermeersch, Walker, Chang, and Grantham-McGregor, 2013, on early-childhood investment) is needed to understand how this initial distribution could be modified to reduce adult income inequality or promote intergenerational mobility.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2013): "Education Policy and Intergenerational Transfers in Equilibrium," NBER Working Papers 18782, National Bureau of Economic Research, Inc.

Altig, D., and S. J. Davis (1993): "Borrowing constraints and two-sided altruism with an application to social security," *Journal of Economic Dynamics and Control*, 17(3), 467–494.

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1997): "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, 105(6), 1121–1166.

Autor, D. H., L. F. Katz, and M. S. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *The Review of Economics and Statistics*, 90(2), 300–323.

Banerjee, A., X. Meng, T. Porzio, and N. Qian (2014): "Aggregate Fertility and Household Savings: A General Equilibrium Analysis using Micro Data," NBER Working Papers 20050, National Bureau of Economic Research, Inc.

Barczyk, D., and M. Kredler (2014): "Altruistically motivated transfers under uncertainty," *Quantitative Economics*, 5(3), 705–749.

Barro, R. J., and G. S. Becker (1989): "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57(2), 481–501.

Boldrin, M., M. D. Nardi, and L. E. Jones (2005): "Fertility and Social Security," NBER Working Papers 11146, National Bureau of Economic Research, Inc.

Checchi, D., A. Ichino, and A. Rustichini (1999): "More equal but less mobile?: Education financing and intergenerational mobility in Italy and in the US," *Journal of Public Economics*, 74(3), 351–393.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014): "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 129(4), 1553–1623.

Corak, M. (2013): "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27(3), 79–102.

Cox, D., and E. Jimenez (1990): "Achieving Social Objectives through Private Transfers: A Review," *World Bank Research Observer*, 5(2), 205–18.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010): "Estimating the Technology of Cognitive and Noncognitive Skill Formation," *Econometrica*, 78(3), 883–931.

De Nardi, M., and F. Yang (2015): "Wealth Inequality, Family Background, and Estate Taxation," NBER Working Papers 21047, National Bureau of Economic Research, Inc.

Doepke, M., and M. Tertilt (2016): "Families in Macroeconomics," NBER Working Papers 22068, National Bureau of Economic Research, Inc.

Fella, G. (2014): "A generalized endogenous grid method for non-smooth and non-concave problems," *Review of Economic Dynamics*, 17(2), 329–344.

Fernandez, R., N. Guner, and J. Knowles (2005): "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality," *The Quarterly Journal of Economics*, 120(1), 273–344.

Fernandez, R., and R. Rogerson (2001): "Sorting And Long-Run Inequality," *The Quarterly Journal of Economics*, 116(4), 1305–1341.

Flood, S., M. King, S. Ruggles, and J. R. Warren. (2015): Integrated Public Use Microdata Series, Current Population Survey: Version 4.0. [Machine-readable database]. Minneapolis: University of Minnesota.

Folbre, N. (2009): *Valuing children: rethinking the economics of the family*, The Family and Public Policy. Harvard University Press.

Gertler, P., J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2013): "Labor Market Returns to Early Childhood Stimulation: a 20-year Followup to an Experimental Intervention in Jamaica," NBER Working Papers 19185, National Bureau of Economic Research, Inc.

Heathcote, J., K. Storesletten, and G. L. Violante (2010): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

Heckman, J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010): "Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program," *Quantitative Economics*, 1(1), 1–46.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): *Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond.* vol. 1 of *Handbook of the Economics of Education*, chap. 7, pp. 307–458. Elsevier.

Hill, R. C., and F. P. Stafford (1980): "Parental Care of Children: Time Diary Estimates of Quantity, Predictability, and Variety," *The Journal of Human Resources*, 15(2), 219–239.

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of Lifetime Inequality," *American Economic Review*, 101(7), 2923–54.

Jones, L. E., and A. Schoonbroodt (2009): "Complements versus Substitutes and Trends in Fertility Choice in Dynastic Models," Working paper.

Jones, L. E., and M. Tertilt (2008): *An Economic History of Fertility in the U.S.: 1826-1960.* chap. 5, pp. 165–230, Frontiers of Family Economics. Emerald.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), pp. 473–522.

Kremer, M., and D. L. Chen (2002): " Income Distribution Dynamics with Endogenous Fertility," *Journal of Economic Growth*, 7(3), 227–58.

Krueger, D., and A. Ludwig (2015): "Optimal Capital and Progressive Labor Income Taxation with Endogenous Schooling Decisions and Intergenerational Transfers," Working papers.

Lee, S. Y., and A. Seshadri (2014): "On the Intergenerational Transmission of Economic Status," Working paper.

——— (2015): “Economic Policy and Equality of Opportunity,” Working paper.

Lundberg, S., and R. A. Pollak (2007): “The American Family and Family Economics,” *Journal of Economic Perspectives*, 21(2), 3–26.

Luo, M., A. Bisin, and J. Benhabib (2015): “Wealth Inequality and Social Mobility,” 2015 meeting papers, Society for Economic Dynamics.

Luxembourg Income Study (LIS) Database (2014): http://www.lisdatacenter.org (multiple countries; 1981 to 2010). Luxembourg: LIS.

Manuelli, R. E., and A. Seshadri (2009): “Explaining International Fertility Differences,” *The Quarterly Journal of Economics*, 124(2), 771–807.

Minnesota Population Center. (2014): Integrated Public Use Microdata Series, International: Version 6.3 [Machine-readable database]. Minneapolis: University of Minnesota.

Morand, O. F. (1999): “Endogenous Fertility, Income Distribution, and Growth,” *Journal of Economic Growth*, 4(3), 331–49.

Murnane, R. J., J. B. Willett, and F. Levy (1995): “The Growing Importance of Cognitive Skills in Wage Determination,” *The Review of Economics and Statistics*, 77(2), 251–266.

Niehues, J., and A. Peichl (2014): “Upper bounds of inequality of opportunity: theory and evidence for Germany and the US,” *Social Choice and Welfare*, 43(1), 73–99.

Nugent, J. B. (1985): “The Old-Age Security Motive for Fertility,” *Population and Development Review*, 11, 75–97.

Restuccia, D., and C. Urrutia (2004): “Intergenerational Persistence of Earnings: The Role of Early and College Education,” *The American Economic Review*, 94(5), pp. 1354–1378.

Roemer, J. E., and A. Trannoy (2015): “Chapter 4 - Equality of Opportunity,” in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 2, pp. 217 – 300. Elsevier.

Roys, N., and A. Seshadri (2014): "On the Origin and Causes of Economic Growth," Working paper.

Ruggles, S., K. Genadek, R. Goeken, J. Grover, and M. Sobek. (2015): Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]. Minneapolis: University of Minnesota.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

Yum, M. (2016): "Parental Time Investment and Intergenerational Mobility," Working paper.

Appendix

A Empirical Findings: Details

A.1 Fertility and Income

Economic models focus on the decisions made by individual households. Consequently, we would like a measure of fertility decisions at the household level. Probably the closest measure to this is available from the US Census: Children Ever Born (CEB). This variable asks each woman how many children they had had during their lives and allows researchers to compute fertility rates by cohorts. Unfortunately, this variable has some limitations. First, it requires women's fertility period to be over to be of use for our purposes. Even assuming that child bearing age extends only to forty years old, using the most current census possible only women born forty years ago could be used. Notice also that choosing the upper end of the age that determines the sample can bring issues. For example, if we used women up to any age we might get biased measures of fertility if this is correlated with mortality risk. Last but not least, this variable has even been dropped from the US Census after 1990. Hence, we use an alternative measure of fertility for our main analysis, but use CEB to evaluate the robustness of our results.

For the sake of clarity let us introduce the most basic measure of fertility, the Crude Birth Rate (CBR), which is defined as the ratio of births to women alive in one year. A typical issue with the CBR is that it can be too low because of a big share of women who have already completed their child bearing age, but are still pulling the ratio down. The Total Fertility Rate (TFR) attempts to correct some of these issues. It is defined as the sum of the age-specific birth rates over all women alive in a given year. Hence, under the same example, if there is an unusually large number of women outside of the child bearing age, TFR is not affected. Formally, let $f_{a,s,t}$ be the number of children born to women of age a in region s and period t divided by the number of women of age a in region s and period t. Assume that the child bearing age extends between ages a_L and a_H.[41] Then the TFR in region s and period t, $\text{TFR}_{s,t}$, is defined as

$$\text{TFR}_{s,t} = \sum_{a=a_L}^{a=a_H} f_{a,s,t}. \tag{10}$$

Typically these age specific fertility rates are constructed for bands of ages of width 5 years and then summed, with the limits of the sum being $a_L = 15$ and $a_H = 49$.[42] Relative to CEB, the main benefit is that it does not require the data to report how many children has each woman had. Instead, it only needs for the children under the age of one to be associated to theirs mothers within the household—a much more standard requirement. Moreover, TFR does not require for the child bearing age to be complete as it focuses on fertility rates which are not associated with a particular cohort but with the women currently alive. Hence, information on the TFR is more up to date than that of the CEB. For this and other reasons, TFR has been widely used in the literature (Kremer and Chen, 2002; Manuelli and Seshadri, 2009).[43]

In order to connect the fertility rate with income, we define the TFR conditional on the income group. Suppose we divide the mothers according to their household income level in quantiles. Then, let $f_{a,q,s,t}$ be the number of children born to women of age a within quantile q in region s and period t divided by the number of women of age a and income quantile q in region s and in period t. Then, the TFR of income quantile q in region s and period t, $\text{TFR}_{q,s,t}$, is defined as

$$\text{TFR}_{q,s,t} = \sum_{a=15}^{a=49} f_{a,q,s,t}. \tag{11}$$

The appropriate measure of income is not obvious either. Assuming households have perfect foresight of their income, using their lifetime income would probably be the best measure. Jones and Tertilt (2008) use "Occupation Income" as their measure of choice. This is constructed for

[41] Notice that, assuming most women have children only in that period, extending this sample would most likely add only values of zeros to the formula of the TFR.

[42] Notice that when using age bands of width bigger than one year (but having only one year of data), $f_{a,s,t}$ is calculated as the number of children born to women within age band A in region s and in year t divided by the number of women within age band A in region s and in year t, *multiplied by the length of age band A.*

[43] The TFR measure of fertility also has its weaknesses. Since it is computed using data from a given year, it mixes fertility decisions of the different birth cohorts alive at the time. If all of these had the same fertility decisions, both CEB and TFR would be identical. However, if fertility rates are changing from cohort to cohort, then CEB gives the more accurate picture of fertility decision. Given the data limitations, we do our empirical work based on the TFR measure of fertility.

year 1950 by IPUMS and the authors extend it to their whole period of interest by assuming a constant 2% annual increase, equal across all occupations. This assumption does not seem harmless since occupations change their relative importance in the society over time (e.g., Autor, Katz, and Kearney, 2008). Moreover, there is a substantial variation in income across people within a given occupation.[44] Hence, we focus on annual total household income in the year of the sample. In order to get the appropriate quantile groups, we cannot compare the income level of young and old households since, following the typical life cycle of income, young households tend to have lower incomes. Hence, we define quantiles within the appropriate age group used for the TFR calculation.[45] This way the TFR for each quantile-region-year can be estimated.

We consider alternative measures in our robustness analysis reported in Figure 3. To compute the TFR gap between education groups, we calculate the TFR as in (10) but separately for each education group. Then, we calculate the difference in the TFR between women married to college-graduate men and those married to high school dropout men. We use men's education to avoid introducing issues regarding changes in women's educational attainment patterns over time. For the robustness analysis using CEB elasticity we focus on women between 40 and 49 years old and use the same income measure and methodology used for the TFR elasticity.

A.2 Fertility Estimation: Sample Selection

For each year of the US Census, we start with all the women belonging to the main family of each household and with non-missing family income. We drop women outside of the "age of fertility," i.e., 15 to 49 years old. Then, we restrict our attention to those who are either heads or spouse of heads, and report to be married. Finally, we drop those who report to be in school

[44] For example, see the *National Compensation Survey: Occupational Wages in the United States, July 2004, Supplementary Tables* (Bureau of Labor Statistics, August 2005), p. 3; on the Internet at http://www.bls.gov/ncs/ocs/sp/ncbl0728.pdf (visited Jan. 21, 2015).

[45] For example, for households within the age group 15-19 years old, income quantiles are defined among other households in the same age group. Moreover, we use a second-degree polynomial on age within each age-group to approximate each family's income at a fixed constant age within each age group and further reduce this concern. However, results do not change significantly if we omit this last step.

or whose annual household income (in 2000 US$) is less than $4,000. Each entry of Table A.1 shows the number of women after each selection procedure, in the corresponding year.

Table A.1: **Sample Selection**

	1960	1970	1980	1990	2000	2010
Women in Main Family without missing income	4,132,162	3,821,829	5,313,266	5,789,849	6,357,343	6,860,823
Age≥15 & Age≤49	1,946,343	1,806,332	2,675,706	2,903,974	3,087,222	2,996,625
Head or Spouse	1,550,987	1,369,469	2,035,969	2,264,903	2,411,233	2,259,209
Married	1,395,011	1,174,508	1,577,704	1,694,897	1,700,881	1,554,153
Not in School	1,376,347	1,158,518	1,492,430	1,555,541	1,581,212	1,432,147
HH Income≥$4000	1,337,549	1,142,124	1,465,870	1,535,536	1,561,333	1,422,478

Source: Census. Each row reports the number of women in each year after dropping all observations without the characteristics given by that row and those above it. HH Income refers to the annual income at the household level in real terms (2000 US$).

After doing this sample restriction, we estimate the fertility rates and elasticities only in states with samples with more than 1,500 women, to avoid using small noisy estimates in our main analysis of the relation between fertility differentials and average income levels. Moreover, when computing the TFR we require each of the seven age groups (15–19, 20–24,...,45-49) to have at least 50 women and 1.5% of the women in the state's sample. We do this in order to avoid using small age-groups which can add noise to the estimation of the TFR—particularly important for younger age groups since we are focusing on married women. We have tried alternative selection procedures and found results to be qualitatively similar.

A.3 Additional Figures and Tables

Table A.2 reports the results from regressing fertility rate (TFR) on the log of average household income, calculated for each state-year using the main sample selection criteria explained in Appendix A.2.

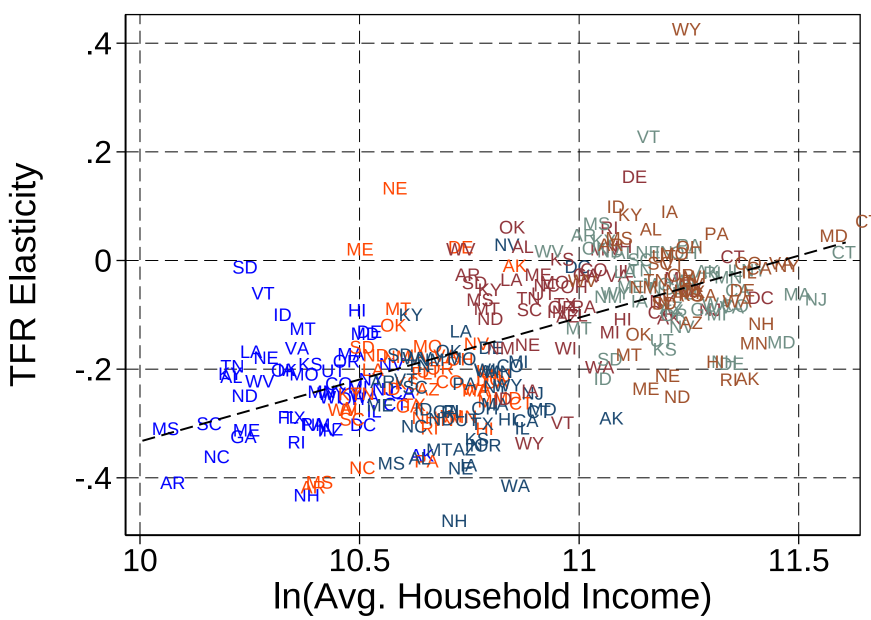
Table A.2: **How the TFR changes with Average Household Income**.

	(1)	(2)	(3)
Ln(Avg. Household Income)	-0.398***	-0.465***	-0.0539
	(0.0319)	(0.0315)	(0.213)
Observations	298	298	298
R-squared	0.281	0.547	0.763
# of States	51	51	51
State FE	NO	YES	YES
Year FE	NO	NO	YES

*Robust standard errors in parentheses. *, **, *** denote statistical significance at the 10, 5, and 1 percent, respectively. Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. Methodology is explained in the main text.*

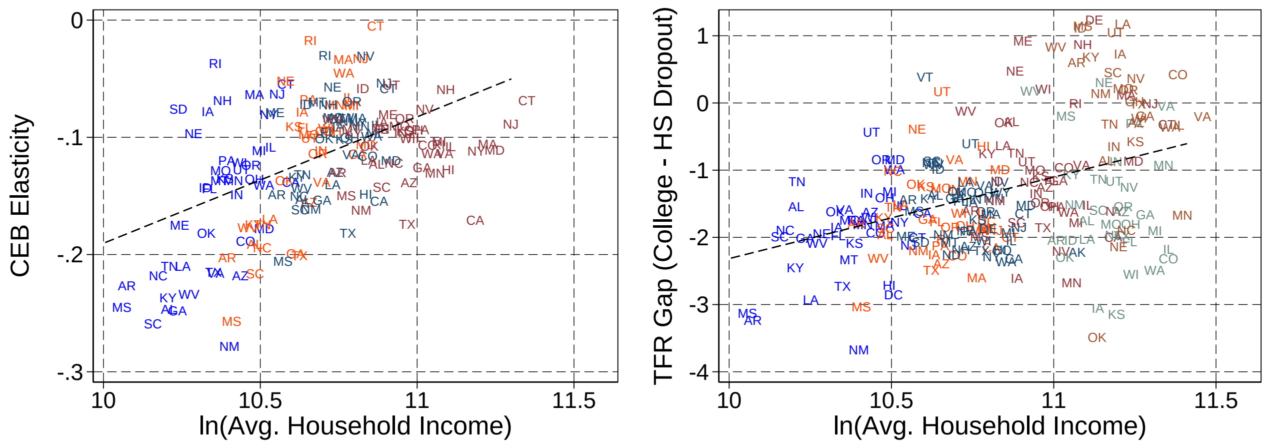
Figures A.1 and A.2 report all the observations used in Figures 2 and 3, respectively.

Figure A.1: **Elasticity of fertility to income and GDP: All observations**.



Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. For each state-year we estimate the elasticity of fertility to income (1). Each census year is represented by a different color. Methodology is explained in the main text.

Figure A.2: **Robustness of fertility differentials; All observations**.



Source: Census. Years: 1960, 1970, 1980, 1990, 2000, and 2010. Each census year is represented by a different color. Methodology is explained in the main text.

B Model: Fertility choices

B.1 Simple model

In this section we consider a simplified version of the model to analyze fertility choices. The life cycle is similar to the benchmark model presented in Section 2 but we change the periods related to family decisions. From $j = J_i$ until $j = J_f - 1$ and from $j = J_k$ until $j = J_d$ the agent's problem is equivalent to the benchmark model of Section 2. From $j = J_f$ until $j = J_k - 1$ the problem will be reduced.

At the age of fertility, $j = J_f$, we assume that the choices of the quantity of children as well as the transfers to them are contemporaneous and there are no future costs of rising children. Also, abstract from the utility derived by the consumption of children. Under these assumptions the problem becomes

$$V_j\left(a,h,e,n\right) = \max_{c,a',n,\varphi} u(c) + \beta\mathbb{E}\left[V_{j+1}\left(a',h',e;\Omega\right)\right] + b(n)\beta\mathbb{E}\left[V_{J_i}\left(\varphi,h_k,\phi_k\right)\right] \tag{12}$$

$$c + a' + \frac{n\varphi}{(1+r)} + C(h,n) = hw\left(1-\tau\right) + a\left(1+r\right),$$

$$a' \geq \underline{a}, \quad h' \sim f^w_{e,j}(h), \quad h_k \sim f^k(e,h), \quad \phi_k \sim g^k(e).$$

From $j = J_f + 1$ until $j = J_k - 1$ the agent's problem is shortened to

$$V_j(a, h, e; \Omega) = \max_{c,a'} u(c) + \beta \mathbb{E}\left[V_{j+1}(a', h', e; \Omega)\right] \tag{13}$$
$$c + a' = hw(1 - \tau) + a(1 + r),$$
$$a' \geq \underline{a}, \quad h' \sim f^w_{e,j}(h).$$

Finally, assume that n is a continuous variable. In this version of the model the first order condition with respect to n reads

$$\frac{\partial u_{J_f}}{\partial c}(C_n + \varphi) = b_n(n)\,\beta \mathbb{E}_{J_f}\left[V_{J_i}(h_k, \varphi, \phi)\,|e, h_{J_f}\right] + \beta^{J_r - J_f}\frac{\partial \mathbb{E}_{J_f}\left[V_{J_r}(h, a, \theta)\,|n, e, h_{J_f}\right]}{\partial n}, \tag{14}$$

which is (8) in the main text.

B.2 Quantitative Analysis

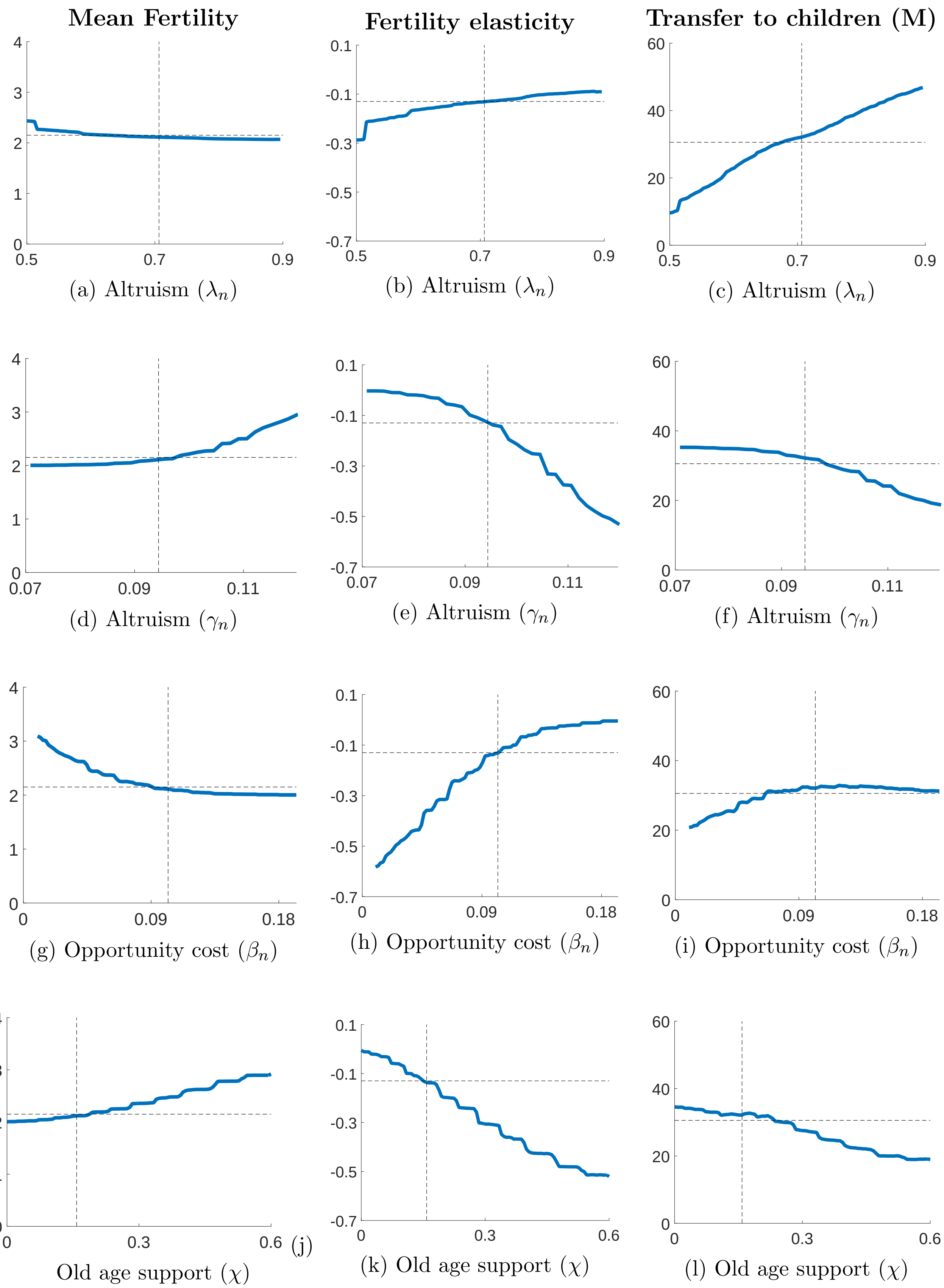
We quantitatively evaluate the four motives for fertility present in the benchmark model of Section 2: (i) altruism; (ii) opportunity cost; (iii) old age support; and (iv) mean reversion across generations. Although the model is highly nonlinear, so that the four motives are interconnected, we can study the numerical comparative statics of the moments related to fertility (mean fertility, fertility elasticity, and mean transfer to children) with respect to different parameters related to each motive for fertility. In particular, we fix all parameters at their estimated values, and change only one parameter in a neighborhood of the estimated values in order to evaluate change in the moments. Hence, we study how different parameters affect mean fertility, fertility elasticity and transfer to children.

Figure B.1 shows the comparative statics of the motives for fertility. First, note that altruism (rows one and two) is relatively less important to identify the mean fertility, with values close to two (sub-figures a and d). However, altruism is important both for the fertility elasticity and for the transfers to children (sub-figures b, c, d, and f). Second, opportunity cost is important both for the mean fertility and the fertility elasticity (sub-figures g and h). However, it is less relevant for the transfer to children (sub-figure i). Third, the fourth row of Figure B.1 shows the comparative statics with respect to old-age support (χ). Sub-figures j, k, and l show that both the mean fertility and the fertility elasticity are sensitive to χ. In particular, when old-age support increases, poor agents have more children than richer ones. This is in line with the intuition highlighted in Section 2.2: Old-age support is a more important motive for poor parents.

Finally, the fifth and sixth rows of Figure B.1 show the comparative statics with respect to mean reversion. In the fifth row we consider alternative transitions between parents and children's initial human capital by changes in the value of ρ in Equation 9. Zero represents draws of human initial capital independent of family characteristics. Effects on fertility choices are quantitatively less important than those of altruism, opportunity cost, and old age support. Qualitatively, as the correlation decreases, mean transfer to children decrease and fertility differentials increase. As parent's and children's human capital are less correlated, the return on having children is larger for poorer families. Finally, the last row shows the comparative statics with respect to the correlation of school taste with parent characteristics, ω. No important effects are observed on the selected moments on fertility choices.[46] Overall, mean reversion does not seem to have large quantitative effects on fertility choices.

[46] Recall that ω is important to identify the intergenerational mobility in education. See Section 3.2.

Figure B.1: **Quantitative analysis of fertility choices**.



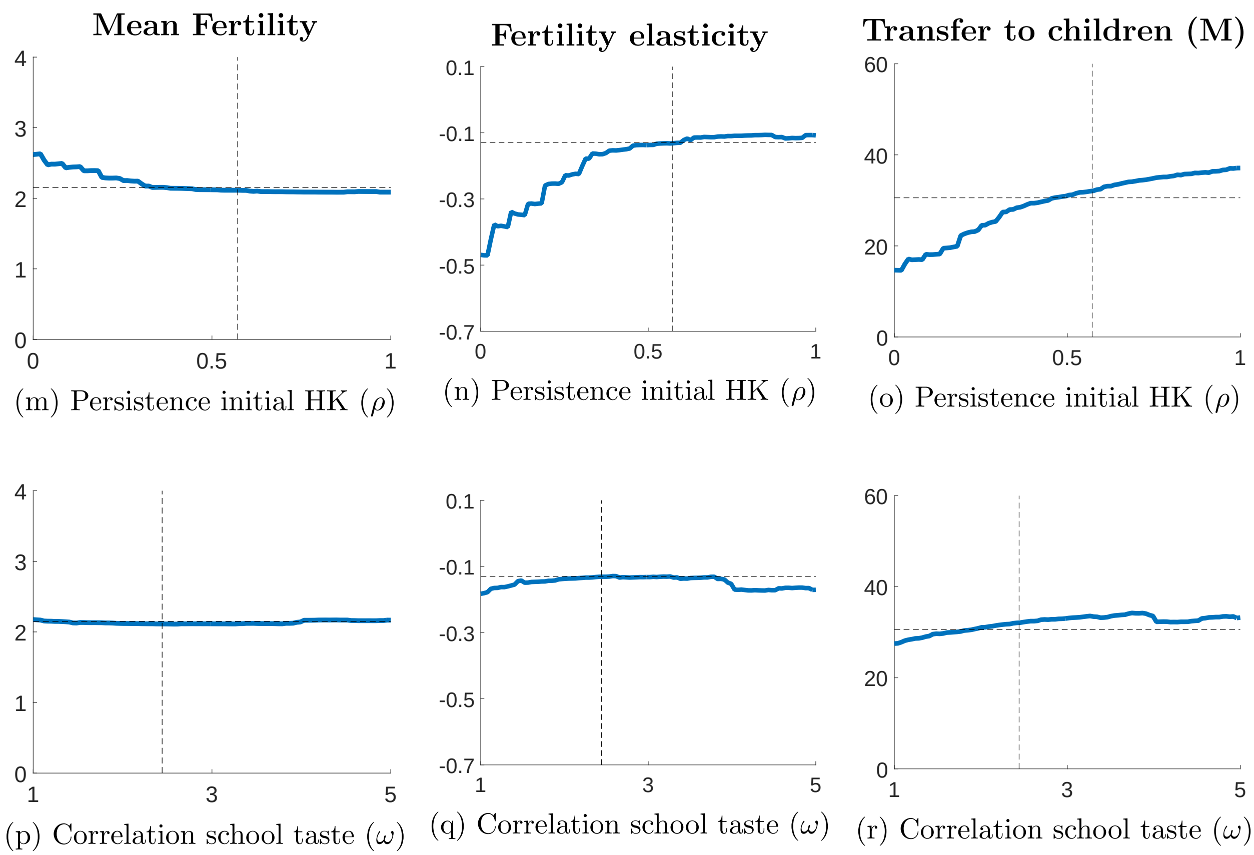
(a) Altruism (λ_n) (b) Altruism (λ_n) (c) Altruism (λ_n)

(d) Altruism (γ_n) (e) Altruism (γ_n) (f) Altruism (γ_n)

(g) Opportunity cost (β_n) (h) Opportunity cost (β_n) (i) Opportunity cost (β_n)

(j) Old age support (χ) (k) Old age support (χ) (l) Old age support (χ)

Figure B.1 (cont.): **Quantitative analysis of fertility choices**.



(m) Persistence initial HK (ρ) (n) Persistence initial HK (ρ) (o) Persistence initial HK (ρ)

(p) Correlation school taste (ω) (q) Correlation school taste (ω) (r) Correlation school taste (ω)

Sub-figures (a) to (r) display the results of numerical comparative statics of mean fertility, fertility elasticity, and transfer to children with respect to altruism, opportunity cost, old-age support, and intergenerational correlation of human capital and school taste. In each sub-figure we change only one parameter, while all other parameters are fixed at their estimated value. The vertical dashed lines show the estimated value of the parameter while the horizontal dashed lines show the value of the moment in the data.

C Estimation and Results: Details

C.1 Income Profile: Sample Selection

We start with 2,900,310 married households from the 2000 US Census data available from IPUMS. We drop the households whose heads are reported to be in school: this reduces the sample to 2,728,958. Dropping households who report yearly household income below $8,000 (equivalent to 50 weeks of 20 hours of work at an hourly wage of $8) further reduces the sample to 2,174,358. Finally, we drop households whose head is outside the age range of 24-63, which gives us a final sample of 1,501,006. When we split this sample in 3 education groups, we get a HS dropouts' sample of 144,935 households, a HS graduates' sample of 891,306 households and a college graduates' sample of 464,765 households.

C.2 Replacement benefits: US Social Security System

The pension replacement rate is obtained from the Old Age Insurance of the US Social Security System. We use education level as well as the level of human capital at the moment of retirement to estimate the average lifetime income, on which the replacement benefit is based. With the last level of human capital before retirement h and the education level e, we estimate the average life time income to be $\widehat{y}\,(h) = \bar{h}(\mathrm{e}) \times h$ with $\bar{h}$ equal to 0.98, 1.17 and 0.98 for high school dropouts, high school graduates and college graduates respectively. Then average annual income $\widehat{y}$ is used in (15) to obtain the replacement benefits.

The pension formula is given by

$$\pi(h) = \begin{cases} 0.9\widehat{y}\,(h) & \text{if } \widehat{y}\,(h) \leq 0.3\bar{y} \\ 0.9\,(0.3\bar{y}) + 0.32\,(\widehat{y}\,(h) - 0.3\bar{y}) & \text{if } 0.3\bar{y} \leq \widehat{y}\,(h) \leq 2\bar{y} \\ 0.9\,(0.3\bar{y}) + 0.32\,(2 - 0.3)\,\bar{y} + 0.15\,(\widehat{y}\,(h) - 2\bar{y}) & \text{if } 2\bar{y} \leq \widehat{y}\,(h) \leq 4.1\bar{y} \\ 0.9\,(0.3\bar{y}) + 0.32\,(2 - 0.3)\,\bar{y} + 0.15\,(4.1 - 2)\,\bar{y} & \text{if } 4.1\bar{y} \leq \widehat{y}\,(h) \end{cases} \tag{15}$$

where $\bar{y}$ is approximately \$70,000.

C.3 Validation of Persistence of Income Process

We estimate an income process similar to Heathcote, Storesletten, and Violante (2010), but using household income and 4 year periods instead. We propose that log earnings of household

i, with age j and education e in period t are represented by

$$\ln(\text{Earnings}_{i,j,e,t}) = \ln(w_{e,t}) + \mu_{j,e} + u_{i,j,e,t}$$

where $w_{e,t}$ is the wage for all labor supplied by those with education e at time t, $\mu_{j,e}$ is the age profile, and $u_{i,j,e,t}$ is the idiosyncratic shock.

From these regressions we obtain the stochastic residual component $u_{i,j,e,t}$. We then model the unobservable shock $u_{i,j,e,t}$ as the sum of two independent components

$$u_{i,j,e,t} = z_{i,j,e,t} + m_{i,j,e,t}$$

where $z_{i,j,e,t}$ is a persistent shock assumed to have an AR(1) structure

$$\begin{aligned} z_{i,j,e,t} =& \rho^e z_{i,j-1,e,t-1} + v_{i,j,e,t} \\ v_{i,j,e,t} \sim& N(0, \sigma_v^e) \end{aligned}$$

and $m_{i,j,e,t} \sim N(0, \sigma_m^e)$ is measurement error (and noise from the point of view of the model). The initial draw is $z_{i,0,e,t} \sim N(0, \sigma_{z_0}^e)$. Note that ρ^e, $\sigma_{z_0}^e)$, σ_v^e and σ_m^e may depend on the education group but are assumed to be independent over time. So we have 12 parameters to estimate, which we will do independently for each education group using a Minimum Distance Estimator.

C.4 Theil-L index: Definition and Decomposition

The Theil-L index is defined as

$$T = \mathbb{E}\left[\ln\left(\frac{y_i}{\mathbb{E}\left[y\right]}\right)\right]$$

where $\mathbb{E}$ denotes the unconditional expectation operator. Suppose there are J types, each of mass m_j. Let $m = \sum_{j=1}^{J} m_j$ and $\mathbb{E}_j$ be the expectation operator conditional on type j. We can decompose the Theil-L index as

$$
\begin{aligned}
T &= \sum_{j=1}^{J} \frac{m_j}{m} \ln \left(\frac{\mathbb{E}_j \left[y \right]}{\mathbb{E} \left[y \right]} \right) + \sum_{j=1}^{J} \frac{m_j}{m} \mathbb{E}_j \left[\ln \left(\frac{y_i}{\mathbb{E}_j \left[y \right]} \right) \right] \\
&= T_b + \sum_{j=1}^{J} \frac{m_j}{m} T_j
\end{aligned}
$$

where T_b is the Theil-L index over the means of each type and T_j is the Theil-L index within each type. Finally, we define the relative Theil-L index as $T_r = \frac{T_b}{T}$. If $T_r = 0$ then all the variation in outcomes is independent of the initial conditions.