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

Diego Daruich
New York University

Julian Kozlowski*
New York University

February 14, 2016

Abstract

How much of income inequality is due to initial opportunities relative to adult income risk? What factors determine intergenerational mobility? We study these questions with a particular interest in the impact of family choices: fertility and transfers. Fertility rates, which are higher for low-income than high-income families, are associated with differences in the level of resources available for children's education. To evaluate the quantitative importance of fertility and family transfers we extend the standard heterogeneous agent life-cycle model with idiosyncratic income shocks and incomplete markets to allow for endogenous fertility, family transfers and education. Initial conditions are defined as the agents' initial state variables, which are endogenously related to parental background. We find that initial conditions account for 50% of the variation in lifetime-earnings. Moreover, fertility and parents-to-children transfers generate 20 and 39% of the intergenerational mobility respectively.

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 Ricardo Lagos and Gianluca Violante. For helpful comments, we thank Jess Benhabib, Raquel Fernandez, 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 in Brown University, New York University, Di Tella University, ECINEQ 2015, EEA 2015, the 2015 Chicago Summer School on Socioeconomic Inequality and the fall 2015 Midwest Macro Meeting. We alone are responsible for all errors and interpretations.

1 Introduction

Is income inequality mainly due to differences in the opportunities available early in life, or to adult income risk? What factors determine intergenerational mobility? We study the sources of income inequality and intergenerational mobility with a particular interest in the impact of families. In our main analysis, families 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. Families face a quantity-quality trade-off between having more children and making larger investments in them. Extensive empirical evidence shows that these 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 than rich ones (e.g., [Altonji, Hayashi, and Kotlikoff, 1997](#)). First, the 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 are the ones with more resources available for education.

We quantitatively evaluate whether fertility and child investment differences are important factors in 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. Parent-to-children transfers will be important in the model for children to access higher levels of education. We also introduce children-to-parent transfers (i.e., old-age support) for the model to be consistent with our finding that fertility differences between income groups are smaller in richer countries.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, children-to-parent transfers, and the relationship between children’s test scores and parental background. A set of validation exercises shows that the quantitative model is consistent with both non-targeted US moments and cross-country evidence on the relationship between development, fertility and intergenerational mobility. 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 (hereafter *VLE-IC*) is 50%. Therefore, half of lifetime-earnings inequality in the US can

---

¹The net average lifetime transfer from children to parents is almost 60% of the average annual household income in the US. See Section 2.

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. In this case, *VLE-IC* reduces to 45%, suggesting that fertility differentials account for 5 percentage points of *VLE-IC*. Moreover, fertility also accounts for 5% of the annual income inequality and 20% 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. In this economy, *VLE-IC* reduces to 34%, suggesting that parental transfers explain 16 percentage points of the estimated *VLE-IC*. These transfers play a major role in the observed inequality and social mobility in the US. According to our model, parental transfers account for 7% of annual income inequality and 39% of intergenerational mobility. These results suggest that to understand social mobility 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).² On the one hand, models on 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 on 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, also allowing for the joint study of inequality and mobility.

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

---

²The cross-country correlation between pre-tax labor income Gini and intergenerational mobility is 0.52. See Appendix Figure B.1.

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

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 countries. Consistent with [Kremer and Chen \(2002\)](#), we also find that fertility differentials are larger in more unequal countries. 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 of 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-country evidence. First, it captures the relation between income inequality and social mobility as well as the reduction of fertility differentials in more equal countries. Secondly, it is also consistent with lower average fertility rates as well as lower fertility differentials in more developed countries. 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 2 shows our empirical findings on fertility differentials and family transfers. Section 3 introduces the model, while Section 4 explains its estimation and shows some validation exercises. The model's results on inequality and intergenerational mobility are presented in Section 5. Finally, Section 6 concludes. The Appendix contains details on the empirical measures used to study fertility and inequality of opportunity, as well as additional tables and figures.

---

⁴Note 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](#)).

⁵Quantitative 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 of them abstract from uncertainty and, though heterogeneity is allowed in the second one, it is only in the form of constant skill differences across dynasties.

2 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 first analyze the data available on these two. Using both US and international micro data, we find evidence that: (i) in the US there is a negative relationship between fertility and income, but it has diminished over time; (ii) countries with higher levels of GDP per capita are associated with smaller fertility differentials; (iii) countries with higher levels of income inequality are associated with larger fertility differentials; (iv) old-age support - in the form of both money transfers and time - is sizable.

2.1 Fertility and Income

If children were considered a normal good, we should observe richer people having more children. However, this is not usually found in the data. Comparing over time, most countries have experienced a decrease in fertility (as they become richer). Comparing across countries, richer countries tend to have a smaller number of children per family. More interestingly for our study, within a country-year it is also the case that richer people tend to have a lower number of children. [Jones and Tertilt \(2008\)](#) look at Census data for United States on women born between 1826 and 1960 and provide 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 big 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 Current Population Survey (CPS) between 1968 and 2013. For most of our empirical work we use CPS and census data from Minnesota Population Center (IPUMS). For the international analysis we also use household survey data from Luxembourg Income Study Database (LIS). Table [B.1](#) in the Appendix details the countries and years used as well as the source for each case. We remark that our measures of fertility and income differ from [Jones and Tertilt](#). We discuss in detail our choices in Appendix [A](#).

Let $\text{inc}_{q,c,t}$ and $\text{fert}_{q,c,t}$ be the median income and fertility rate (TFR), respectively, of income quantile q in country c and year t .⁶ We estimate

$$\ln(\text{fert}_{q,c,t}) = \alpha_{c,t} + \beta_{c,t} \ln(\text{inc}_{q,c,t}) + \epsilon_{q,c,t}, \quad (1)$$

where $\beta_{c,t}$ will be referred to as the *elasticity of fertility to income* for country c in year t . If this value is negative, richer households tend to have a lower number of children. Values closer to zero imply that fertility rates are not related to income (at least according to this specification). Figure [1](#) shows

---

⁶Using mean income changes the results slightly, but they qualitatively remain the same.

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 the 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.

Figure 1: **Elasticity of fertility to income (US, 1968-2013).**

The elasticity of fertility to income in the US has been negative and it becomes modest over time. Source: CPS. Methodology is explained in the main text. Income is computed at family level (within the household).

To better understand what is behind this pattern we extend our analysis to include 28 countries (for a total of 116 country-year observations) using micro data from both IPUMS and LIS. We combine all our observations and divide them into deciles according to their levels of real GDP per capita. We use the logarithm of the real GDP per capita (PPP) of each country-year, as a proxy for development. Then, for each of these groups we calculate the median logarithm of real GDP per capita (PPP) and the median elasticity of fertility to income. Figure 2 shows that richer or more developed countries tend to have smaller elasticities or, in other words, smaller fertility differentials.⁷

Various factors related to countries' cultures or infrastructure might be driving those results. In order to try to control for these possible issues we run a regression of fertility elasticities to income on the logarithm of the real GDP per Capita (PPP) controlling for either country or time fixed effects.⁸ The

⁷Income is computed at the household level since we do not have family identifiers in all the samples. It is possible, particularly among the poorer, that multiple families live within a household. We would like to separate their incomes between the families but are not able to for all the samples. However, when focusing on the US in the estimation section of the model as well as in Figure 1, we can identify families and hence compute income at the family level (within the household).

⁸For time fixed effects we limit to decade fixed effects since very rarely do we have observations for different countries in the same year, as census or surveys take place in different times for each country. Moreover, we are not able to control for both country and time fixed effects since we do not have enough observations or variation in the data.

Figure 2: **Elasticity of fertility to income by GDP Decile.**

Source: IPUMS International and LIS. GDP is obtained from Total Economy Database. We divide observations into deciles according to GDP per capita. For each decile we calculate the median level of GDP per capita as well as the median fertility elasticity of income. Figure B.2 in the Appendix includes all the data observations before grouping them into deciles.

regression specification is

$$\text{Fertility Elasticity}_{c,t} = \alpha + \beta \ln(\text{GDP}_{c,t}) + \eta_c + \mu_t + \epsilon_{c,t} \quad (2)$$

where Fertility Elasticity_{c,t} is equal to the estimated $\hat{\beta}_{c,t}$ from (1). Table 1 shows that the elasticity of fertility is increasing in real GDP per Capita. Once again, this implies that more developed or richer countries are associated with smaller fertility differentials.⁹ This relationship seems stable and robust to controls for country fixed effects as well as decade fixed effects. This suggests that the pattern from Figure 1 may not be exclusive to the US nor simply due to changes over time but can be thought as a general relation between development and fertility differentials.

How does income inequality relate to fertility differentials? Intuitively, we would expect countries with bigger fertility differentials to present stronger inequalities from the earliest stages of life. As poor families have more children and fewer resources to split than richer ones, many children are born with scarce resources while a smaller group is born with a bigger pie to split among fewer hands. Assuming

---

⁹For some countries this relationship may actually be positive, meaning that richer households tend to have more children than poorer ones. This seems to be the case for countries where the population fertility rate is below the replacement ratio, like Finland or Denmark.

VARIABLES	(1)	(2)	(3)
ln(GDPpc)	0.207*** (0.0423)	0.137*** (0.0449)	0.203*** (0.0424)
Observations	116	116	116
R-squared	0.160	0.871	0.167
# of countries	28	28	28
Country FE	NO	YES	NO
Decade FE	NO	NO	YES
Robust standard errors in parentheses			
*** p<0.01, ** p<0.05, * p<0.1			

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

Source: IPUMS International and LIS. Methodology is explained in the main text.

this affects their education or business opportunities, we would expect to observe more inequality in countries with bigger fertility differentials. Figure 3 shows that unequal countries are associated with higher fertility differentials. This evidence coincides with findings in [Kremer and Chen \(2002\)](#) who look at fertility differentials across education groups.

2.2 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 countries. 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. They also present evidence that more than one-third of the elderly in developing countries receive transfers from their children. However, old-age support is not only monetary transfers. For example, when health problems arise help from family 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 Trans-

Figure 3: **Income Elasticity of Fertility by Income Gini Decile.**

Inequality is defined over total household income. Source: IPUMS International and LIS. We divide observations into deciles according to income inequality levels. For each decile we calculate the median level of inequality as well as the median fertility elasticity of income. Figure B.3 in the Appendix includes all the data observations before grouping them in deciles.

fers survey to study transfers between parents and children. As encountered by [Abbott, Gallipoli, Meghir, and Violante \(2013\)](#), data on family transfers is scarce and problematic so more work needs to be done to extract precise information from this type of sources.¹⁰ Therefore, we take the evidence presented here as suggestive of mainly 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.¹¹ The data is limited to children who are over 18 years old and does not include “long-term” transfers (for example, tuition or financial help to buy a house). The first column of Table 2 shows transfers going from parents to children, while the second column shows transfers going in the opposite direction. The first thing to notice is that these transfers are sizable, particularly once 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

¹⁰Estimates on the size of transfers depend substantially on whether observations with zero transfers are excluded or not.

¹¹The procedure is similar to the one used to calculate the TFR in Appendix A. We first calculate the average transfer given on 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.

household income. In turn, transfers from children to parents are almost 60% bigger than those going in the opposite direction.

	Parent $\rightarrow$ Children	Children $\rightarrow$ Parent
Money	\$38,589	\$37,536
Hours	1060	3166
Total	\$62,265	\$108,222

Table 2: **Family transfers over a lifetime.**

Source: PSID Rosters and Transfers, 2013. Children are 18+ years old. Data does 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.

3 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 life 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. Finally, the last stage consists of 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.

3.1 The individual problem

Figure 4 shows the life-cycle of an agent, where 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 psychic cost. Initial assets are money transfers from her parents. The initial human capital as well as school taste (or psychic cost) are stochastic but correlated with the parents’ education and human capital.

Figure 4: Life cycle

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 . While on education, she sequentially chooses whether to continue in school or to start working. The education decision is irreversible. 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. 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 $f_e^s(h)$. The 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 school taste or psychic costs $\phi \in [0, 1]$ to affect the cost of education.¹² Particularly, we assume that these costs are perceived by the agent as increasing the financial cost of education.¹³ 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 $f_{e,j}^w(h)$. We allow for education- and age-dependent idiosyncratic labor income shocks. In Section 4 we discuss the estimation of the returns of education and the income process.

Formally, let V_j^s and V_j^w be the value of an agent of age j in school and working, respectively. Let V_j^{sw} be the value of an agent who can choose between the two

$$V_j^{sw}(a, h, e, \phi) = \max \{V_j^s(s, h, e, \phi), V_j^w(s, h, e)\},$$

where V_j^s is defined by

$$V_j^s(a, h, e, \phi) = \max_{c, a'} u(c) + \beta V_{j+1}^{sw}(a', h', e + 1, \phi) \quad (3)$$

¹²We allow education costs and returns to depend on the age of the agent in order to highlight the difference between high school and college in the estimation.

¹³Note that these costs have a marginal value in consumption terms, which can be an alternative way to interpret them. The benefit of this specification (relative to having the cost in utility terms directly) is that it simplifies the calibration without altering the interpretation.

$$\begin{aligned}
c + a' + p_e + \phi \bar{\psi}_e &= hw_e(1 - \tau) + a(1 + r) \\
a' &\geq \underline{a}, \quad h' = f_e^s(h)
\end{aligned}$$

The agent is risk averse and her preferences are represented by an increasing, concave and positive utility function u .¹⁴ She can borrow up to the limit $\underline{a}$ and the return on savings is $1 + r$. Future is discounted by β . Note that in this particular problem there is no uncertainty. We denote as w^s the wage for an agent that is currently in school at level s . 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

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

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 of leaving 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 where 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 \(1989\)](#) where parents care about the utility of their children. The problem at the age of fertility $j = J_f$ is

$$\begin{aligned}
V_j(a, h, e) &= \max_{c, c_k, a', n, T} u(c) + \beta \mathbb{E} [V_{j+1}(a', h', e, n)] + b(n)u(c_k) \\
c + nc_k + a' + C(h, n, T) &= hw(1 - \tau) + a(1 + r), \\
a' &\geq \underline{a}, \quad h' \sim f_{e,j}^w(h).
\end{aligned} \tag{5}$$

In this period the agent chooses her consumption c , her children's consumption c_k , savings a' and

---

¹⁴The fact that the utility function u is positive is needed to model altruism. As shown by [Jones and Schoonbroodt \(2009\)](#) the implicit assumption that parents like 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.

the number of children n . As usual, the agent derives utility from her own consumption and her continuation utility, whose states are explained below. 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). The parents pay the cost $C(h, n, T)$ on top of 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 . T is an indicator variable, $T \in \{\text{hire childcare, stay at home}\}$. This captures the idea that childcare is a good that can be either produced at home or hired in the market. Childcare reduces the time needed to raise children. Therefore, parents with higher human capital will optimally choose to hire childcare in the market whereas parents with lower human capital will choose to stay at home and pay this cost exclusively with a reduction in time at work.¹⁵ The functional form for C is specified in Section 4. Until the agent's children become independent ($j = J_k$), she chooses the children consumption and pays the cost C . Hence, the problem is equal to (5) but taking 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 are chosen $j = J_k$ is

$$\begin{aligned} V_j(a, h, e, n) = & \max_{c, c_k, a', \varphi, T} u(c) + \beta \mathbb{E} [V_{j+1}(a', h', e; \Omega)] + b(n) \{u(c_k) + \beta \mathbb{E} [V_{J_i}(\varphi, h_k, \phi_k)]\} \quad (6) \\ & c + nc_k + a' + \frac{n\varphi}{(1+r)} + C(h, n, T) = hw(1-\tau) + a(1+r), \\ & a' \geq \underline{a}, \quad h' \sim f_{e,j}^w(h), \quad h_k \sim f^k(e, h), \quad \phi_k \sim g^k(e) \end{aligned}$$

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 differently to (5), the value function at this stage now includes the continuation value of the children V_{J_i} . This is the last period where parents' choices affect their descendants. As the problem is written recursively this implies that at every period where 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 level of education and human

---

¹⁵We allow for childcare for quantitative reasons. We explored different costs functions and we found that a model that allows for childcare fits the data on the elasticity of fertility the best. Without childcare, fertility differentials are too big relative to the data and do not get smaller with higher levels of income as reported in Figure 2. Finally, note that child care has no effect on the children's human capital.

¹⁶The 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. Moreover, the utility and value functions are specified at the household level (with two individuals). Hence, for example, when parents choose to have two children, they create one household so $n = 1$.

capital of the parents. The functional form of the altruism as well as the stochastic processes of human capital, $f^k(e, h)$, and psychic costs, $g^k(e)$, are specified in Section 4.

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 (children taking into account the utility of their parents as well as parents taking into account that of their children) in a 3 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 the 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 as their needs are much more pressing than those from richer parents.¹⁷ This would make the old-age support motive even bigger 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 the 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 $\Omega = \{n, \varphi, h_{J_f}\}$, i.e. the number of children n , the initial assets of their children φ and their own education and human capital when the children were at home h_{J_f} .¹⁸ These

---

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

¹⁸The 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 helps pin down how long the children will remain in school. Finally, the number of children n is used to determine, among others, the

state variables remain constant until retirement and are used to predict the old-age support that the agent will receive. Note that the old-age support is an endogenous random variable whose distribution depends on the education choices of the children.¹⁹

Formally, the problem at the age of retirement is

$$\begin{aligned} V_j(a, h, e, \theta) &= \max_{c, a'} u(c) + \beta V_{j+1}^w(a', h, e), \\ c + a' &= \theta + \pi(e, h) + a(1 + r), \\ a' &\geq \underline{a}, \end{aligned} \tag{7}$$

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

3.2 Fertility choices

Two incentives for having children are at play: (i) Altruism, from which parents care about their children's well-being; and (ii) Old-age support, from which parents care about the help their children provide them once they grow old. Altruism implies that parents want to have educated children (as 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, children vary both in terms of raising costs and distribution of future incomes. 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 in a stage of life where parents would actually like to borrow.²¹ More importantly, for high income parents children are particularly costly due to the time they require. Moreover, their return in future transfers are not particularly relevant since they will have significant other sources of income. The opposite is true for low income households, who have low time costs and obtain high marginal returns from children. Thus, the old-age support channel is much more valuable for poorer households. This way, the altruism channel dominates for richer households while the investment one gives an extra incentive

---

mean and variance of the transfer to be received.

¹⁹We remark that the difference with respect to the full information case is such that uncertainty on children's transfers last 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 our assumption is not harmful.

²⁰We 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 4 for details.

²¹For example, in the US the current average age of first birth is 27 while the income peak is closer to 50 years old.

for poorer agents to have children.

To understand the qualitative effects consider a simplified version of the model where 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. 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.²² Under these assumptions, the first order condition with respect to n is

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

The left-hand side is the marginal cost of an extra child, that is the transfer φ and the marginal childcare 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, while the second one 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

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

where $h_{J_t}^k$ is the human capital of the child at the age that they provide old-age support, $j = J_t$. The expected transfer of each child is $\mu_\theta = w(1 - \tau) \xi \mathbb{E} [h_{J_t}^k | e, h_{J_f}]$. 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 towards richer agents, the marginal utility diminishes and both the cost and the old-age support channels become less important. On the other hand, the altruism benefits do not decrease and therefore dominate the fertility decision. 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 development and fertility choices that we documented in Section 2. In the calibrated model, altruism motives will always be present

---

²²See Appendix C for details about the simplified model.

and lead families to have few children. This will be the main fertility driver for developed economies in general as well as for rich agents in poorer ones. On top of 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. Moreover, the return of the investment on old-age support differs among families due to mean reversion of income across generations. On the one hand, poor families expect to have relatively richer children and therefore expect transfers to be larger. On the other hand, richer families expect to have relatively poorer children with low transfers. Hence the expected return of old-age support is larger for poor families which reinforces the negative income-fertility relationship.²³

We can evaluate the effects of old-age support under the benchmark calibration described in Section 4. At the age of retirement 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.²⁴ For the poorest quintile old-age support represents 16% of resources while it represents less than 3% for the richest quintile. There are two forces at play. First, due to the negative income-fertility relationship poor households choose to have more children and therefore receive a larger number of transfers. Secondly, richer households accumulate more savings so the transfers from their children are relatively smaller. This coincides with the intuition obtained from equation (8): old-age support is more relevant for poor agents in our model.

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 4.

²³Once 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.

²⁴Recall 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.

4 Estimation

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 \(2014\)](#). Then we compute the ergodic distribution of the economy to match moments from United States in 2000. Some of the parameters can be estimated “externally”, while others need to be estimated “internally” from the simulation of the model. Table 3 summarizes all the parameters in the model.

The model is estimated based on three primary data sources. We complement data from IPUMS Census with data from the CPS Fertility Supplement as well as from the NLSY79. We also take some estimates from other studies. Details are presented below. We select a population for which our model can be taken as a reasonable approximation to household behavior. We 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 corresponding to one person working 20 hours a week for the minimum wage (approximately \$8,000 of total annual household income). Secondly, there is no decision regarding marriage in our model. Given our interest on fertility, we are interested in two-member households. In order to steer clear of differences in income and time availability due to single parenthood, we keep only married households. Details about our sample selection are reported in Appendix D.1.

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 of 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: We normalize the model using the average wage of a high school graduate at age 40, $\bar{y}_{40}^{HS} = \$60,198$ (adjusting the average initial level of human capital such that we can hold 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 [Roys and Seshadri \(2014\)](#) we set the annual interest rate to $r = 5.5\%$. We assume unsecured borrowing is not possible, $\underline{a} = 0$. The payroll tax is $\tau = 0.124$, as in the current social security. The yearly price of college is from Delta Cost

Project, where we get \$6588.²⁵ The yearly price of high school is obtained from 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 (relatively low cost of high school when compared to college), leading to a price of 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 $p_2 = 0.90$.

School Taste: In this class of models it is difficult to match the high school dropout rate. Previous studies such as [Abbott, Gallipoli, Meghir, and Violante \(2013\)](#) introduced nonpecuniary (psychic) costs of education. Our specification is closer to [Krueger and Ludwig \(2015\)](#) which introduces heterogeneous costs of education. We estimate this cost as entering the budget constraint. This simplifies the estimation and has many interpretations such as extra expenses required to learn or marginal utility through cost in consumption terms. We assume agent's psychic cost ϕ are 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 of children of high school graduates 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

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

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: The returns to education are allowed 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 non-linear form $h' = f_e^s(h) \equiv \alpha_e h^{\beta_e}$ for $e \in \{1, 2\}$. Table 3 shows that our estimates for high school are $\alpha_1 = 1.48$ and $\beta_1 = 0.66$, while for college they are $\alpha_2 = 1.38$ and $\beta_2 = 0.90$.

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 14%. 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 8% per year. These numbers are in line with empirical estimates in the literature summarized in [Heckman, Lochner, and Todd \(2006\)](#).

---

²⁵We take into account grants and scholarships, such that only private tuition costs are considered. Prices are in 2000 US dollars.

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.²⁶ Two comments are appropriate. First, income risk is calibrated to include total earnings variation, encompassing what may be considered both wage shocks as well as 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 coefficient of variation of log income is well estimated. Moreover, in Table 5 we will show that the persistence of the income process is correctly fitted.

Figure 6: CV of log(income) by Age.

Childcare: We propose the following functional form for the cost of childcare

$$C(h, n, T) = \begin{cases} n^{\alpha_n} (wh(1 - \tau)\beta_c\beta_n + \beta_n(1 - \beta_c)p_c) & \text{hire childcare} \\ n^{\alpha_n} (wh(1 - \tau)\beta_n) & \text{stay at home} \end{cases}.$$

This function allows for non-constant returns to scale in the number of children. Moreover, this functional form also allows for a reduction in the time needed to take care of children if childcare is hired in the market, which is done by around 35% of the population (NLSY79). If an agent stays at home, her income is reduced by $n^{\alpha_n}\beta_n$. On the other hand, hiring a nanny implies a reduction in this time cost of raising children of $(1 - \beta_c)$. This option has the extra cost of hiring a nanny $n^{\alpha_n}\beta_n(1 - \beta_c)p_c$.

We estimate the returns to scale $\alpha_n = 0.41$ which is in line with Table 6.4 in Folbre (2009). We estimate that if the agent stays at home, her labor income is reduced by $\beta_n = 0.12$ but if she hires childcare her labor income is reduced only by $\beta_n\beta_c = 0.02$. Angrist and Evans (1998) estimate the effect of children on labor supply for households with different levels of education. If we assume

²⁶We assume that the income shocks distribution is constant before 24 and after 63 as data is problematic for those ages.

that low-educated households stay at home, while high-educated households hire nannies (as in our model), our estimates are similar to the ones from their empirical estimates. [Folbre \(2009\)](#) provides an estimate of p_c based on the average wage of nannies. This leads to an estimated cost of childcare $n^{\alpha_n} \beta_n (1 - \beta_c) p_c = n^{\alpha_n} \times 0.25$.

Old-age support: With data from 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 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 2, 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.²⁷

Replacement benefits: The pension replacement rate is based on 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. See Appendix D.2 for details.

Intergenerational transmission of ability: We assume that the initial level of human capital is stochastic but correlated with the human capital and education of the parent. From NLSY79 we back out the transition matrix between parents' education and income and children's ability (measured according to the reading test) by quintile, shown in Figure 7. We highlight that the difference between education groups as well as the difference between income groups (within education) is remarkable. For example, children born in the bottom tercile of low educated parents have around 50% probability of being in the bottom quintile of test scores. On the other hand, the same probability for those born of high educated parents is always below 10% (and even below 5% for those with high income). In our model, each child first draws her initial human capital quintile from the distribution given in Figure 7 corresponding to her parents' education and income tercile (within education group). Then, she draws h_{J_i} uniformly within the given quintile of (positive) truncated normal distribution with standard deviation σ_{h_0} to be estimated. This defines f^k , the distribution of the initial draw of human capital in the household problem (6). Recall that another source of intergenerational persistence is that we also allow for psychic costs to be correlated with parents' education level (as explained above).

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 3 the utility function has to be positive and therefore $\gamma_c \in [0, 1)$. We follow

---

²⁷Note 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 from the time parents devote to children before the age of 18, about 90% is actually spent before the age of 13.

Figure 7: **Correlation of parents' background and children ability (%)**.

Source: NLSY79. Each cell reports the children's test score quintile probability conditional on mother's education and income group. Probabilities are detailed in Appendix Table D.2.

the literature and assume $\gamma_c = 0.5$ (e.g., Roys and Seshadri, 2014). Other papers, like Manuelli and Seshadri (2009), that have estimated this parameter are also around this value. As is standard in the literature, the altruism function is assumed to be $b(n) = \lambda_n n^{\gamma_n}$.

Thirteen parameters of the model are estimated “internally”. 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. Three parameters $\{\beta_n, \beta_c, \alpha_n\}$ are related to child cost. α_e and β_e , for $e \in \{1, 2\}$ define the heterogeneous returns to education in high school and college. Finally, $\bar{\psi}_e$ defines the distribution (both mean and standard deviation) of the psychic costs, while ω is related to the correlation with parents' education level.

In order to pin down the value of these parameters we aim to match eighteen moments from the data, shown in Table 4. First, given the importance of education for our questions of interest, we target five moments related to education. These include education distributions and income ratios between education groups. Second, five moments on fertility are included. Average fertility - both across the entire population and by education groups - as well as the fertility elasticity discussed in Section 2 are targeted. Finally, given our interest on social mobility and income inequality, we target moments on the rank-rank correlation and various coefficients of variations of income for different ages and education groups.²⁸

²⁸Given the different values of the moments, we adjust the weighting matrix such that the difference used in the objective function of the minimization is the percentage deviation. More precisely, moment j is weighted by $\frac{1}{\max\{\underline{m}, m_j^2\}}$ where m_j is the value of moment j in the data. Since some moments could potentially have the value of 0, we limit the denominator to be above $\underline{m} = 0.1$. We have tried other alternatives to solve for this issue, but we have not found significant differences.

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	IPUMS
r	5.5%	Interest rate (annual)	Roys and Seshadri (2014)
$\underline{a}$	0	Borrowing limit	
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
$\bar{\psi}_1$	0.77	Maximum of high school psychic cost	Internally estimated
$\bar{\psi}_2$	1.81	Maximum of college psychic cost	Internally estimated
ω	2.10	Correlation of psychic cost's with parents	Internally estimated
Return to education			
(α_1, β_1)	(1.48, 0.66)	High school return	Internally estimated
(α_2, β_2)	(1.38, 0.90)	College return	Internally estimated
Income process			
<i>Estimated to match mean and variance growth by age and education. See text and Figure 6.</i>			Census
Childcare			
β_n	0.12	Total cost	Internally estimated
β_c	0.20	Reduction by hiring childcare	Internally estimated
p_c	2.60	Wage of childcare	Folbre (2009)
α_n	0.41	Returns to scale	Internally estimated
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			
<i>Correlation of initial draw of human capital by parent's education and income. See Figure 7.</i>			NLSY79
σ_{h_0}	0.34	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)	(0.29, 0.05)	Altruism	Internally estimated

Table 3: **Estimated parameters**

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.

Moment	Data	Model	Source
Education			
Dropouts	9.4%	9.3%	Census
High school graduates	60.1%	62.5%	Census
College graduates	30.5%	28.2%	Census
Income Ratio: HS Dropout to HS Grad, Age 40-43	0.643	0.596	Census
Income Ratio: College Grad to HS Grad, Age 40-43	1.787	1.802	Census
Fertility			
Average fertility	2.150	2.092	CPS
Average fertility: HS Dropout	2.910	2.620	CPS
Average fertility: HS Graduate	2.160	2.055	CPS
Average fertility: College Graduate	2.050	2.001	CPS
Fertility elasticity	-0.100	-0.101	Census
Mobility and Inequality			
Intergenerational Mobility: Rank-Rank	0.341	0.321	Chetty et al. (2014)
CV income: All ages and educations	0.862	0.888	Census
CV income: HS Dropout, Age 24-27	0.918	0.935	Census
CV income: HS Graduates, Age 24-27	0.609	0.600	Census
CV income: College Graduates, Age 24-27	0.598	0.480	Census
CV income: Age 24-27	0.666	0.604	Census
CV income: Age 40-43	0.849	0.836	Census
CV income: Age 56-59	0.940	0.968	Census

Table 4: **Targeted moments.**

Source: Census refers to IPUMS Census data in the year 2000. CPS refers to IPUMS CPS Fertility Supplement for the most recent years (2010 and 2012) given that we use “Children Ever Born” to estimate this moments.

The model performs generally well. Fertility levels and its income elasticity are successfully matched, which is necessary given the key role of fertility in our model. Regarding education, the three levels are well matched in the model. As for inequality, the model displays levels similar to the data. We also obtain levels of intergenerational mobility which are close to the empirical evidence.

4.1 Validation Exercises

We can test the validity of our estimated model by looking at moments which are not directly targeted. For this, we will start by evaluating the model within the estimated steady state given by the parameters in Table 3. Then we will look at how the model compares with the cross-country evidence reported in Section 2 by moving away from the steady state. Table 5 summarizes the results from

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 similar to [Abbott, Gallipoli, Meghir, and Violante \(2013\)](#), but using household income and 4 year periods instead. See [Appendix D.4](#) for details. Estimating this process (by education groups), the first panel of [Table 5](#) shows that the coefficients determining the persistence of the shocks by education level to be 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 intergeneration mobility of income and education. [Figure 8](#) opens the transition matrix on which the rank-rank estimate is based. Looking at children’s income (ages 28-31) and parent’s income (ages 40-43), we verify that the transition probability between the generations for all twenty-five entries is well captured. Moreover, we estimate the intergeneration mobility of education in the model. We evaluate two measures of education persistence for which we have data ([Checchi, Ichino, and Rustichini, 1999](#)): the trace and the determinant of the intergenerational transition matrix of education. We find that the trace in the model is equal to 0.79 while it is equal to 0.85 in the data. Similarly, the model generates a determinant of 0.87 which is close to the empirical value of 0.90.

Figure 8: **Intergenerational mobility**

Regarding family choices, we can look at the size of transfers from parents to children as well as fertility decisions within the different education groups. First, the model generates an average transfer of \$24,985 in 2000 USD, with a coefficient of variation of 51.5%. Once we restrict to strictly positive transfers, the average increases to \$27,068 which is close to the empirical estimate of \$30,566 from [Abbott, Gallipoli, Meghir, and Violante \(2013\)](#). Secondly, we can compare the fertility elasticities within education groups. The model generates elasticities equal to $(-0.262, -0.007, -0.002)$ for high school dropouts, high school graduates and college graduates. In the data these are $(-0.220, -0.170, -0.080)$. We highlight that the values in the model and the data are similar. More importantly, both model

and data display a decreasing (in absolute terms) relation with education. Finally, we can compare the utilization of childcare services. In the model 44% of the households hire nannies whereas in the data 35% does so. Overall, the estimated model is quantitatively consistent with family choices that were not targeted in our estimation.

Cross-country We do a simple exercise in order to show that the model is consistent with the cross-country patterns described in Section 2. Recall that in the benchmark calibration the wage was normalized to one. Consequently, in order to generate economies with different levels of GDP per capita as in the data, we move wages, such that the real wage (i.e., in consumption terms) is the main change. Note that this involves moving the model away from the steady state to which it was estimated. The last panel of Table 5 shows four cross-country correlations both in the data and the model.

The first two rows in the last panel of Table 5 are on the relation between development (proxied by GDP per capita) and fertility. First, the model is able to capture the negative relationship between development and the fertility rate. The model generates a correlation between the fertility rate and GDP per capita of -0.997, which is close to the empirical estimate of -0.984. More importantly, the model is able to capture the relation between development and fertility elasticity shown in Figure 2. The model generates a correlation of 0.943 which is similar to the data value of 0.827.

The last two rows of Table 5 regard income inequality and intergenerational mobility as well as fertility differences across income groups. Corak (2013) showed that more unequal countries are associated with lower levels of intergenerational mobility (higher mobility coefficient). The model is able to capture this as it generates a correlation of 0.503 which is close to the data estimate of 0.519. Finally, Figure 3 showed evidence - consistent with Kremer and Chen (2002) - that more unequal countries display bigger fertility differences between income groups. The correlation between Gini and fertility elasticity was found to be -0.835 in the data, which is close to the model outcome of -0.967. 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 as well as moments outside such steady state but related to cross-country evidence.

Moment	Data	Model	Source
Income Persistence			
Dropouts	0.89	0.89	PSID & NLSY79
High school graduates	0.88	0.90	PSID & NLSY79
College graduates	0.93	0.91	PSID & NLSY79
Intergenerational Mobility: Education			
Trace of transition matrix	0.850	0.792	Checchi et al. (1999)
Determinant of transition matrix	0.900	0.867	Checchi et al. (1999)
Family Transfers: Parents to Children			
Average transfer		\$24,985	
CV transfer		51.5%	
Average transfer: conditional on > 0	\$30,566	\$27,068	Abbot et al. (2013)
Fertility: Elasticity within Education Groups			
Dropouts	-0.220	-0.262	Census
High school graduates	-0.170	-0.007	Census
College graduates	-0.080	-0.002	Census
Child costs			
Hire childcare	35.3%	43.8%	NLSY79
Cross-Country Evidence			
corr(GDP,Mean Fertility)	-0.984	-0.997	WDI
corr(GDP,Fertility Elasticity)	0.827	0.943	IPUMS & LIS
corr(Gini,Intergenerational Mobility)	0.519	0.503	Corak (2013)
corr(Gini,Fertility Elasticity)	-0.835	-0.967	IPUMS & LIS

Table 5: **Validation exercises.**

Source: For the income persistence estimates, we follow *Abbott, Gallipoli, Meghir, and Violante (2013)*. Correlations in the data are computed using the values in Figures B.1, 2 and 3. The same procedure was used for average fertility and GDP using data from the World Development Indicators.

5 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 in-

come distribution. Secondly, following [Huggett, Ventura, and Yaron \(2011\)](#), we look at the variation of lifetime-earnings that can be attributed to initial conditions (*VLE-IC*). Lastly, we study the role that family choices have in *VLE-IC* as well as inequality and intergenerational mobility. [Table 6](#) summarizes our main results.

5.1 Inequality of Opportunity

Following the literature on equality of opportunity (EoP), we first define an *outcome* of interest such as lifetime income or educational attainment.²⁹ We then divide the population into different *types* according to their initial conditions (e.g. parents' income, initial assets or human capital). Finally, we evaluate the effect of initial conditions 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. We initially do an informal visual comparison by plotting the conditional distributions. For a more formal evaluation, we define and compare measures based on the relative entropy of the distributions.

Visual Inspection We start by looking at the income of young adults and defining the different types according to parents' income. The advantage of this definition is that we can do this exercise both in the model and in the data (using NLSY79).³⁰ We split parents in two groups according to their incomes: bottom 50% and top 50%. [Figure 9](#) shows the distribution of children's income conditional on parents' types in the data (dashed lines) and in the model (solid lines). The model is able to match the data relatively well. For example, the degree of inequality of opportunity in the model, captured by the distance between the two solid lines, is similar to the data (distance between dashed lines). Focusing on the median, the top group has an income almost one-third higher than the poor parents group. More generally, the distribution of income of children with high income parents (first-order stochastic) dominates the one of children with poor parents. This suggests that parents' income is a relevant factor determining children's income and equality of opportunity does not seem to hold neither in the data nor in the model.

---

²⁹There is a large literature on distributive justice - including equality of opportunity - which has recently been summarized by [Roemer and Trannoy \(2015\)](#).

³⁰In the model we use the average income of children between the ages of 28-31 and average of parents' income between the ages 40-43. The data used is from NLSY79 and, as usual, is much noisier than the model, so we adapt it as follows. For parents' income we use total household income and average it between the ages of 40-45. Data on children's income is more problematic so we extend the window of years used in the data to the average between the ages 26-32, using only individuals who are not enrolled in school. Finally, since the model and data refer to different years (and ages) we rescale the data so that the median income is the same as in the model. Note that this procedure does not correct for anything related to the effect of parents' income.

Figure 9: **Child's Income Distribution by Parents' Income Group.**

Average income of children (parents) is defined between the ages of 28-31 (40-43). See main text and footnote 30 for details.

However revealing Figure 9 may be, 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 in short periods than in 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. Secondly, parents' income can be understood as a proxy for other elements which 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.

Formal Evaluation We now move towards a quantitative evaluation of EoP. A suitable way to compare conditional distributions is based on entropy measures that allow to decompose inequality both between and within types. As it is standard literature we use the Theil-L index, which is explained in Appendix E. 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 the initial conditions.

The empirical problem with the Theil-L index is that many variables that potentially determine initial conditions (or 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 initial conditions, the fifth row of Table 6 shows that our estimates are close but below to the lower bound from these empirical studies. 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.

The advantage of having a structural model is that we can identify the true initial conditions in our economy: parents' transfer, initial human capital and school taste. Moreover, we can also look at more interesting measures of outcomes like lifetime-earnings. In this case, the eighth row of Table 6 shows that our actual level of inequality of opportunity (as measured by the relative Theil-L index) is 72%.

Our model provides evidence that it is important to carefully define both types and outcomes when studying inequality of opportunity. We can evaluate this by comparing the two alternative estimates above (i.e., 7.1 vs 72.3%). Using a short period of income during young age and not choosing the appropriate characteristics that define initial conditions can lead to a large downward bias in the estimation of inequality of opportunity.

5.2 Variance Decomposition of Lifetime Earnings

An alternative way to evaluate the importance of initial conditions is the one followed by Huggett, Ventura, and Yaron (2011), which is shown in the second part of the fourth panel of Table 6. We decompose the variance of lifetime-earnings into variation due to initial conditions (*VLE-IC*) and variation due to adult shocks. Defining initial conditions as of age 23, Huggett, Ventura, and Yaron (2011) found that *VLE-IC* was approximately 61%. Taking into account the role of families, we find that most of this figure is actually determined even before high school. Table 6 reports that in our model, using initial conditions as of age 12, *VLE-IC* is equal to 50%.

Furthermore, we can also look at the relative importance of each of the initial conditions. We find that most is determined by the level of human capital, which is correlated to parents' characteristics. However, the third panel of Table 6 shows that parents' transfers play a crucial role in education, explaining about 38% of its total variance. This is consistent with the results from Belley and Lochner (2007) who find that family income and wealth are relevant for educational choices.

	Data	Benchmark	Exogenous fertility	Constant transfers
Income inequality				
CV income: All ages	0.862	0.888	0.847	0.830
CV income: Age 24-27	0.666	0.604	0.556	0.521
CV income: Age 40-43	0.849	0.836	0.790	0.765
Theil-L index: Age 28-31	0.247	0.255	0.207	0.164
Social mobility				
Intergenerational Mobility: Rank-Rank	0.341	0.321	0.201	0.131
Transition Par Q1-Child Q5 (%)	7.5	9.1	13.7	15.5
Inequality of opportunity				
Theil-L relative (%): Income, Age 28-31 (Types: parent's income and education)	16-36	7.1	1.4	0.5
Theil-L relative (%): Lifetime Earnings (Types: initial conditions)	30-75	72.3	65.0	52.7
Variance decomposition				
CV of Lifetime-Earnings		0.78	0.73	0.71
% expl. by initial HK		36.1	33.2	26.0
% expl. by transfers		3.2	2.5	0.0
% expl. by school taste		4.4	3.7	3.3
% expl. by interaction		6.7	6.0	4.4
<i>VLE-IC</i> : % expl. by all initial conditions		50.4	45.4	33.7
Variance years of education	5.7	5.4	4.1	3.4
% expl. by initial HK		6.6	4.9	8.6
% expl. by transfers		38.3	6.2	0.0
% expl. by school taste		51.2	67.6	75.2
% expl. by interaction		3.9	21.3	16.2
Total: % expl. by all initial conditions		100.0	100.0	100.0

Table 6: **Social mobility and income inequality**

Note: Exogenous fertility solves the benchmark model where fertility is restricted to be equal to $N = 2$. Similarly, Constant Transfers refers to the case where parents to children transfers are exogenously given. We choose the value of this transfers such that high school and college are feasible.

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 4, but look at the case of exogenous fertility (assuming each household has two children). Using our measure of choice - CV of income for all ages - column 3 of Table 6 reports that in such a

society inequality would be reduced by 5%. If we focused on smaller age groups or used other measures of inequality, this reduction may increase to 19%.

Without fertility differentials the rank-rank intergenerational mobility would improve by 20%. This is connected to the reduction of *VLE-IC* to 45%, implying that fertility differences between income groups account for 5 percentage points of the estimated *VLE-IC*.³¹

The Role of Family Transfers We can also look at the case where we allow for endogenous fertility but parents' transfers are exogenously constant.³² Similarly to removing endogenous fertility, the fourth column shows that without endogenous transfers the CV of income for all ages would be reduced by 7%. If we focused on some other inequality measures reported in the table, the reduction could be as high as 36%.

Imposing an exogenous positive transfer to children would improve intergenerational mobility by 39%. With constant family transfers, the big role that initial assets played in education choices is eliminated and most of lifetime-earnings is 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. 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* declines to 34%, implying a reduction of 16 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 have a modest role relative to parental transfers in the US, this may not be the case for other countries. In Section 2 we showed that developing countries have bigger fertility differentials. According to our model, this implies that policies which reduce the incentives of poorer households to have children may be successful in improving inequality and social mobility in such countries.³³

---

³¹We remark that if we were to remove old-age support - one of the main motives for fertility differentials in our model - we would arrive to a similar conclusion.

³²We can fix the amount transfered to different levels. Forcing zero transfers is not useful because in such a case no one would be able to study, which would move us to a completely different economy. For this reason, we choose a level such that children are able to go to high school, and those with high initial ability are able to go to college.

³³An older version of this paper was estimated using the US in 1960 and results were in that direction.

6 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, calibrated to match that in the data. 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 both US and cross-country 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 as well as 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.

There are different ways to analyze the impact of initial conditions but we broadly find that initial opportunities (as of age 12) and adult income risk over the working life account for equal parts of income inequality. Our model suggests that 50% of the variation in lifetime-earnings is due to differences in conditions determined even before high school. Fertility differentials and family transfers respectively account for 5 and 16 percentage points of the share of the variation in lifetime-earnings due to initial conditions. Moreover, these two forces generate 5 and 7% of annual income inequality as well as 20 and 39% of the intergenerational immobility observed in the data. These results imply that those interested in understanding inequality, intergenerational mobility or inequality of opportunity may need to take these two forces into account.

Our results suggest that improving access to education (through education subsidies for example) may diminish the importance of family transfers, helping reduce income inequality and improve social mobility. Even though our model is silent about the forces determining 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 may 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.
- ANGRIST, J. D., AND W. N. EVANS (1998): “Children and Their Parents’ Labor Supply: Evidence from Exogenous Variation in Family Size,” *American Economic Review*, 88(3), 450–77.
- 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.
- BELLEY, P., AND L. LOCHNER (2007): “The Changing Role of Family Income and Ability in Determining Educational Achievement,” *Journal of Human Capital*, 1(1), 37–89.
- 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.
- 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.
- 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.

A 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 is only available for a small set of countries and has even been dropped from the US Census after 1990. Hence, we use an alternative measure of fertility.

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 $f_{a,c,t}$ be the number of children born to women of age a in country c and period t divided by the number of women of age a in country c and period t . Assume that the child bearing age extends between ages a_L and a_H .³⁴ Then the TFR in country c and period t , $\text{TFR}_{c,t}$, is defined as

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

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 $a_H = 49$.³⁵ Relative to CEB, the main benefit is that it does not require the surveys 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 their 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

---

³⁴Notice 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.

³⁵Notice that when using age bands of width bigger than one year (but having only one year of data), $f_{a,t}$ is calculated as the number of children born to women within age band A in country c and in year t divided by the number of women within age band A in country c and in year t , *multiplied by the length of age band A* .

and Seshadri, 2009).³⁶

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 $f_{a,q,c,t}$ be the number of children born to women of age a within quantile q in country c and period t divided by the number of women of age a and income quantile q in country c and in period t . Then, the TFR of income quantile q in country c and period t , $\text{TFR}_{q,c,t}$, is defined as

$$\text{TFR}_{q,c,t} = \sum_{a=15}^{a=49} f_{a,q,c,t}. \quad (9)$$

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 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 (Autor, Katz, and Kearney, 2008). Moreover, there is a substantial variation in income across people within a given occupation.³⁷ Finally, this Occupation Income measure is not available for the panel study we are interested in. Hence, for this section 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.³⁸ This way the TFR for each quantile-country-year can be estimated.

---

³⁶The 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.

³⁷For 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).

³⁸For example, for households within the age group 15-19 years old, income quantiles are defined among other households in the same age group.

B Additional Figures and Tables

Country	Years	Source
Austria	1987	LIS
Brazil	1970, 1980, 1991, 2000, 2010	IPUMS
Canada	1991, 1994, 2000, 2004, 2007, 2010	LIS
China	2002	LIS
Colombia	2004, 2007, 2010	LIS
Czech Republic	1992	LIS
Denmark	1987, 1992, 1995, 2000, 2004, 2007, 2010	LIS
Dominican Republic	1981, 2002	IPUMS
Finland	1987, 1991, 1995, 2000, 2004, 2007	LIS
France	1978, 1984, 1989, 1994, 2000	LIS
Germany	1973, 1978, 1983, 2000, 2004	LIS
Guatemala	2006	LIS
India	2004	LIS
Italy	1986, 1987, 1989, 1991, 1993, 1995, 1998, 2000	LIS
Mexico	1984, 1989, 1992, 1994, 1996, 1998, 2002, 2004, 2008, 2010	LIS
Mexico	1995, 2000	IPUMS
Panama	1980, 1990, 2010	IPUMS
Peru	2004	LIS
Poland	1995, 1999, 2004, 2007, 2010	LIS
Puerto Rico	1970, 1980, 1990, 2000, 2005	IPUMS
Romania	1995, 1997	LIS
Slovak Republic	1992	LIS
South Africa	1996, 2001, 2007	IPUMS
South Korea	2006	LIS
Spain	1980, 1990, 2004, 2007, 2010	LIS
Taiwan	1981, 1986, 1991, 1995, 1997, 2000, 2005, 2007, 2010	LIS
United States	1960, 1970, 1980, 1990, 2000, 2005, 2010	IPUMS (Census)
United States	1968-2013	IPUMS (CPS)
United Kingdom	1994, 1999, 2004, 2007, 2010	LIS
Uruguay	2004	LIS

Table B.1: **Countries included in empirical work.**

Sources: Minnesota Population Center (IPUMS) and Luxembourg Income Study Database (LIS).

$$a' \geq \underline{a}, \quad h' \sim f_{e,j}^v(h).$$

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

$$\frac{\partial u_{J_f}}{\partial c}(C_n + \varphi) = b_n(n) \beta \mathbb{E}_{J_f} [V_{J_i}(h_k, \varphi, \phi) | e, h_{J_f}] + \beta^{J_r - J_f} \frac{\partial \mathbb{E}_{J_f} [V_{J_r}(h, a, \theta) | n, e, h_{J_f}]}{\partial n}, \quad (12)$$

which is (8) in the main text.

D Additional Estimation Details

D.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.

D.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 $\hat{y}(h) = \bar{h}(e) \times h$ 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 $\hat{y}$ is used in (13) to obtain the replacement benefits.

The pension formula is given by

$$\pi(h) = \begin{cases} 0.9\hat{y}(h) & \text{if } \hat{y}(h) \leq 0.3\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(\hat{y}(h) - 0.3\bar{y}) & \text{if } 0.3\bar{y} \leq \hat{y}(h) \leq 2\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(2 - 0.3)\bar{y} + 0.15(\hat{y}(h) - 2\bar{y}) & \text{if } 2\bar{y} \leq \hat{y}(h) \leq 4.1\bar{y} \\ 0.9(0.3\bar{y}) + 0.32(2 - 0.3)\bar{y} + 0.15(4.1 - 2)\bar{y} & \text{if } 4.1\bar{y} \leq \hat{y}(h) \end{cases} \quad (13)$$

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

D.3 Correlation of parents' background and children initial human capital

Mother	Children				
	Q1	Q2	Q3	Q4	Q5
High school dropouts					
Low income	53	21	13	8	5
Medium income	34	31	10	15	9
High income	27	31	14	15	13
High school graduates					
Low income	32	26	14	15	13
Medium income	20	22	19	22	16
High income	14	21	19	24	22
College graduates					
Low income	13	21	18	24	25
Medium income	6	15	20	29	30
High income	5	16	16	30	32

Table D.2: **Correlation of parents' background and children ability (%)**.

Source: NLSY79. Each cell reports the children's test score quintile probability conditional on mother's education and income group.

D.4 Validation of Persistence of Income Process

We estimate an income process similar to [Abbott, Gallipoli, Meghir, and Violante \(2013\)](#), 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.

E Theil-L index: Definition and Decomposition

One such measure is the Theil-L index, defined as

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

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

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

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.