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

Diego Daruich† Julian Kozlowski‡
NYU NYU

July 6, 2015

PRELIMINARY

Abstract

How much of social mobility and income inequality is due to initial opportunities relative to adult income risk? Previous studies have yielded very wide estimates due to data limitations. To provide a more precise answer this article builds on a standard heterogeneous agent life cycle model with idiosyncratic income shocks. 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. Accounting for this is essential for the proper evaluation of initial opportunities, so we extend the model to introduce the role of families through endogenous fertility, family transfers and education. We find that initial conditions as of age 13 account for more of adult income inequality than do labor income shocks. Moreover, fertility differentials and family transfers are found to account for over 50% of the social mobility in the data.

JEL Classifications: D91, J13, J24, J62.

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

*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 and Edouard Schaal. We also received valuable feedback from participants at seminars, especially at Brown and NYU. We alone are responsible for all errors and interpretations.

†ddaruich@nyu.edu.
‡kozjuli@nyu.edu.
1 Introduction

Recent empirical work has increased the interest on social mobility and income inequality. But is inequality mainly due to differences in opportunities determined early in life or to differences in effort and luck experienced over the working lifetime? What factors determine social mobility? And how are inequality and social mobility connected? Answering these questions is of utmost importance. First, it will help evaluate the different policies usually suggested to reduce inequality. Shall we provide insurance against shocks over the working lifetime (e.g. progressive taxation or unemployment insurance) or shall we focus on equalizing initial opportunities (e.g. early childhood investment or improving education access)? Similarly, an answer would help understand if those policies are effective in improving social mobility or if other measures may be necessary. Thirdly, theories of distributive justice typically distinguish ethically acceptable inequalities (e.g. due to differences in effort) from unfair inequalities (e.g. due to endowed characteristics) (Arneson, 1989; Cohen, 1989). Preferences for redistribution are systematically correlated with beliefs on the relative importance of these two in the determination of outcomes (Alesina and Giuliano, 2011). Those with beliefs that inequality is mainly due to differences in endowed characteristics tend to be more willing to accept redistributive policies. Hence any political discussion on redistributive policies is not likely to prosper without answer to these questions. However, providing one has been difficult in empirical applications due to lack of data so this paper uses an extended standard heterogeneous agent life cycle model with idiosyncratic risk to overcome this difficulty. Accounting for the role of families is essential for the proper evaluation of initial opportunities, so we extend the model to allow for endogenous fertility, family transfers and education. 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 also allows for human capital transmission from parents to children, calibrated to match that in the data. We find that initial opportunities, as of age 13, account for more of the inequality observed in the data than do labor income shocks. Moreover, our results suggest that fertility differentials and family transfers generate over 50% of the (lack of) social mobility observed in the data.

Why is inequality a concern? Inequality can be thought of - and has been modeled - as a necessity to provide incentives for people to study or put effort (Krueger and Ludwig, 2015). However, Figure 1, based on Corak (2013) and known as the Great Gatsby Curve, which looks at income inequality and social immobility - the likelihood that someone will inherit their parents' relative position of income level - across (developed) countries has called public attention on one possible concern from income inequality. The Gatsby Curve shows that countries with high levels of inequality are associated with low levels of social mobility. One interpretation has been that (using a popular metaphor) as the rungs

get further apart, it is harder to climb the ladder. As inequality grows, more children may not be capable of improving their income relative to their parents. If children born from poor parents are not able to access education (e.g. due to borrowing limits or detrimental home environments) they may not be able to take advantage of the extra rewards or incentives from studying, suggesting the need for further policy attention.

Figure 1: Gatsby Curve: Intergenerational Mobility and Income Inequality

![Figure 1: Gatsby Curve: Intergenerational Mobility and Income Inequality](image)

**Note:** Income inequality is measured using the Gini index on household income before taxes and transfers. Intergenerational economic mobility is measured as the elasticity between paternal earnings and a son’s adult earnings. **Source:** Corak (2013) and OECD.

However, many interpretations are possible for Figure 1 so a proper framework to study income inequality and social mobility is necessary. Quadrini and Rios-Rull (2015) suggest that there is no conclusive theory to study inequality but that, among the economic literature, the classic Bewley economy is a typical one.² Probably the best example of a Bewley model used to study the sources of inequality is Huggett et al. (2011). However, they find that most of the income inequality is due to conditions given before entering the labor market which are exogenous in their analysis. More importantly, they admit that their results are silent about the forces prior to age 23, which can be considered very important when interested in inequalities of opportunities (or endowed characteristics). Previous empirical studies have tried to estimate the share of inequality that is explained by inequalities of opportunities but have been constrained by lack of data. The problem is that it is not possible to

² Besides a typical Bewley economy, other theories that generate inequality include sorting (Fernández and Rogerson, 2001; Fernández et al., 2005) or models that can generate a Pareto distribution, like, for example, any with exponential growth occurring over an exponentially-distributed amount of time (Jones and Kim, 2014). Models on social mobility are also rare, with an exception being Restuccia and Urrutia (2004). However, they focus on a stylized model without uncertainty and hence are not designed for quantitative analysis, which is our main purpose.
observe all initial conditions that determine opportunities. For example, direct data on the quality of the home environment or resources available for education are not usually available. This has forced these empirical studies to provide very wide estimates due to data limitations (Brunori et al., 2013; Niehues and Peichl, 2014). Their estimates (using the Theil-L index) suggest that between 16 and 75% of adult inequalities are explained by initial inequalities of opportunities. Though interesting, their wide range limits their value for policy analysis. If we believe inequalities of opportunities are close to that lower bound, we might prefer redistributive policies that are focused on late redistribution. On the other hand, if we believe that inequalities of opportunities are closer to the upper bound we might find policies focused on improving the initial distribution of opportunities to be more appropriate. In order to overcome this data limitation, this paper introduces a model in the spirit of Huggett et al. (2011) but that endogenizes these earlier stages of life through choices of education, fertility and family transfers. As we push back the age at which initial conditions are determined, one’s family becomes more important so including fertility and transfers decisions is essential. It would be interesting to go all the way back to age 0 so as to clear the mystery on how actual initial conditions are formed, but our model contributes by making one step in that direction and looking at initial conditions as of age 13. The calibrated model, which is consistent with the empirically observed relation between income inequality and social mobility from Figure 1, will provide a more precise answer to the impact of initial conditions.

Jones and Tertilt (2008) study fertility in the United States and provide substantial evidence that poor families have more children than richer ones. In Section 2, we update their analysis for the US confirming their results and also showing the novel finding that this negative relation between fertility and income gets smaller over time. Moreover, to better understand the origin of this phenomenon we extend the analysis to 28 countries where it seems that as countries get richer, fertility differentials between the (relatively) poor and rich get smaller. To the best of our knowledge, the papers in the literature are not able to explain this finding. In their excellent literature review, Jones et al. (2010) mention that there is no full consensus on the motivations behind fertility choices. However, most fertility models are very simple three-period representative agent ones which abstract from uncertainty issues. This also makes them inappropriate for a quantitative analysis of the impact on income inequality or social mobility. We contribute to this literature by building a full life cycle heterogeneous agents model which allows for uninsurable shocks and is also capable of explaining the reduction in fertility differentials. Since we propose that fertility differentials and family transfers could be generating inequality and social immobility, capturing this finding is important if we want to understand the actual drivers of these family decisions.

3See Morand (1999); de la Croix and Doepke (2003); and Mookherjee et al. (2012). Two exceptions might be Manuelli and Seshadri (2009) or Roys and Seshadri (2014) which move towards a full life cycle model to explain fertility differences across countries. Nevertheless, both of them abstain from uncertainty and, though heterogeneity is allowed in the second one, it is only in the form of constant skill differences across dynasties.
Fertility differentials and family transfers can generate many effects on inequality and social mobility. For the sake of clarity, abstract from family differences and take an economy where fertility differentials and family transfers do not exist. All children are born equal (with the same skills and resources) so they will all choose the same level of education. In such an economy the only source of inequality will be adult risk. Moreover, social mobility would be “perfect”: parents’ and children’s relative position of income would be independent. Now suppose you start from this economy and, all of a sudden, allow for endogenous fertility differentials and family transfers. Assume, as in most of the data, that poor families have more children and transfer less resources than rich ones.\footnote{For a clearer explanation of why this happens in our model see Section 3.} Adults that were lucky and are currently rich will have few children and give them plenty of resources, while those on the opposite end will have many children with few resources. The first group of children will now be able to educate themselves even more than the previous generation, while the second group will not even be able to afford the previous generation’s level of education. As education increases their expected lifetime income, the children of rich households are more likely to remain rich. This will generate a reduction in social mobility. Moreover, income differences would also increase as the sources of inequality now include education access on top of adult risk. This is an extremely simple example but conveys the main mechanism through which fertility differentials and family transfers may affect inequality and social mobility.

The theoretical effect of fertility differentials and family transfers may be interesting in itself, but quantifying their impact is of utmost importance for policy questions regarding redistribution. Hence, we use our model to overcome data limitations and answer our initial questions, focusing on the US in 1960. There are different ways to analyze the impact of initial conditions, but we broadly find that initial opportunities (as of age 13) account for more of the inequality observed in the data than do labor income shocks. More particularly, if we look at inequality of opportunities as computed by Brunori et al. (2013) or Niehues and Peichl (2014), our model suggests that 84% of observed inequality in lifetime earnings is due to inequality of opportunities (very close to their empirically estimated upper bound). On the other hand, if we focus on the variation of lifetime earnings, as does Huggett et al. (2011), our model suggests that 68% of this is due to differences in conditions determined as early as before high school. Similarly to Huggett et al., these results should be understood as applying to the age where our model starts. 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 human capital holds as in the data), it can still shed light on its importance. More applied research (like Gertler et al. (2013)) may be needed to understand how this initial distribution may be improved. Finally, our model also suggests that fertility differentials and family transfers generate over 50% of the social immobility observed in the data. Models interested in understanding social mobility should take these two forces into account.
The article is organized as follows. Section 2 shows the empirical work. Section 3 introduces the model, while Section 4 explains its estimation. Results are detailed in Section 5 and Section 6 concludes. The Appendix contains the list of countries used in the empirical work and some additional figures.

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. If children were considered a normal good, we should observe richer people having more children. However, this is not usually found in the data. On a country time series level, most countries have experienced a decrease in fertility over time (as they become richer). On a cross-country level, richer countries tend to have a smaller number of children. More importantly 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 this relationship with income is robust. We update their analysis for the US using micro data from the Current Population Survey (CPS) between 1968 and 2013.\footnote{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 A.1 in the Appendix details the countries and years used as well as the source for each case.} We show that the negative relationship holds for the US in every year, but it is also the case that it has become more modest over time. Moreover, using international evidence from 28 countries, it seems that as countries get richer, fertility differentials between the (relatively) poor and rich get smaller. Since our model will rely on old-age support motives to explain these observations, we also introduce evidence that old-age support is sizable in the data.

2.1 Fertility and Income

Economic models focus on the decisions made by individual households. Consequently, we would like a measure of fertility decisions at the household level. Probably the closest measure to this is available from the US Census: Children Ever Born (CEB). This variable asks each woman how many children they had had during their lives and allows researchers to compute fertility rates by cohorts. Unfortunately, this variable has some limitations. First, it requires women’s fertility period to be over to be of use for our purposes. Even assuming that child bearing age extends only to forty years old,
using the most current census possible only women born forty years ago could be used. Notice also that choosing the upper end of the age that determines the sample can bring issues. For example, if we used women up to any age we might get biased measures of fertility if this is correlated with mortality risk. Last but not least, this variable is only available for a very 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 \( a_H \). 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 \( a_H = 49 \). 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 theirs mothers within the household - a much more standard requirement. Moreover, TFR does not require for the child bearing age to be complete as it focuses on fertility rates which are not associated with a particular cohort but with the women currently alive. Hence, information on the TFR is more up to date than that of the CEB. For this and other reasons, TFR has been widely used in the literature (Kremer and Chen, 2002; Manuelli and Seshadri, 2009).

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

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

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

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

$$TFR_{q,c,t} = \sum_{a=15}^{a=49} f_{a,q,c,t}.$$  \hspace{1cm} (1)

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 et al., 2008). Moreover, there is a lot of variation in income across people within a given occupation. 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. This way the TFR for each quantile-country-year can be estimated.

Let $inc_{q,c,t}$ be the median income of quantile $q$ in country $c$ and year $t$. Now we are ready combine the information on fertility and income. We estimate

$$\ln (fert_{q,c,t}) = \alpha_{c,t} + \beta_{c,t} \ln (inc_{q,c,t}) + \epsilon_{q,c,t}$$  \hspace{1cm} (2)

where $q$, $c$ and $t$ stand for the same elements as before. $\beta_{c,t}$ 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 lower number of children. Values closer to zero imply that fertility rates are not related to income (at least according to the specification used). Finally, positive values of this elasticity imply that richer household tend to have more children than poorer ones. Figure 2 shows the elasticity of fertility to income for the US, with its value on the vertical axis. It confirms that the fertility elasticity has been negative since 1968 but it also suggests that it has decreased over time, implying that the difference in number of children between poor and rich households has become smaller.


\hspace{1cm} $^{10}$For example, for households within the age group 15-19 years old, income quantiles are defined among other households in the same age group.

\hspace{1cm} $^{11}$Using mean income changes the results slightly, but they qualitatively remain the same.
We would like to better understand what is behind this pattern so we extend our analysis to include 28 countries for a total of 116 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 (using the logarithm of the real GDP per capita (PPP) of each country in each year, as a proxy for development.) Then, for each of these groups we calculate the median logarithm of the real GDP per Capita (PPP) and the median elasticity of fertility to income. Figure 3 shows that richer or more developed countries tend to have smaller elasticities or, in other words, smaller fertility differentials.
Figure 3: Income Elasticity of Fertility by GDP Decile

The regression specification is

\[ \text{Fertility Elasticity}_{c,t} = \alpha + \beta \ln(GDP_{c,t}) + \eta_c + \mu_t + \epsilon_{c,t} \]  

(3)

where Fertility Elasticity\(_{c,t}\) is equal to \(\beta_{c,t}\) from (2). Table 1 shows that the elasticity of fertility is increasing in the level of real GDP per Capita. Once again, this implies that more developed or richer countries are associated with smaller fertility differentials.\(^{13}\) Moreover, this relationship seems stable and robust to controls for country fixed effects as well as decade fixed effects, suggesting that

\(^{12}\)For 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.

\(^{13}\)It is important to remark that for 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.
the pattern from Figure 2 may not be exclusive to the US nor simply due to changes over time. In Appendix Figure A.3 we also show that countries with more inequality are associated with higher fertility differentials than more equal ones. This evidence coincides with that shown by Kremer and Chen (2002) which looked at fertility differentials across education groups (while we focus on differentials across income groups).

Table 1: Income Elasticity of Fertility and GDP per capita PPP

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(GDPpc)</td>
<td>0.207***</td>
<td>0.137***</td>
<td>0.203***</td>
</tr>
<tr>
<td></td>
<td>(0.0423)</td>
<td>(0.0449)</td>
<td>(0.0424)</td>
</tr>
<tr>
<td>Observations</td>
<td>116</td>
<td>116</td>
<td>116</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.160</td>
<td>0.871</td>
<td>0.167</td>
</tr>
<tr>
<td># of countries</td>
<td>28</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>Country FE</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Decade FE</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

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

2.2 Family Transfers

Up to now we have presented evidence on two main facts: (i) Income inequality differs across countries and is associated with lower social mobility; and (ii) Higher levels of GDP per capita are associated with smaller fertility differentials. In Section 3 we introduce a life cycle model that is consistent with these patterns. 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. Moreover, this will be used in the model to capture 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 in the data.

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 transfers annually. The higher end of that range is dominated by developing countries. Moreover, they also present evidence that, among

---

14See Nugent (1985) or Banerjee et al. (2014).
developing countries, more than one-third of the elderly receive transfers from their children. But old-age support is relevant in terms other than money. 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 also use recent Panel Study of Income Dynamics (PSID) information from the 2013 Rosters and Transfers survey to study transfers between parents and children. As suggested by Abbott et al. (2013), data on family transfers is scarce and problematic so more work needs to be done to extract more precise information from this type of sources.15 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.16 The data is limited to children who are over eighteen years old and does not include “long-term” transfers (for example, tuition or buying a house). In its first column 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 big, 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 household income. And transfers from children to parents are almost 60% bigger than those going in the opposite direction.

Table 2: Family transfers over a lifetime

<table>
<thead>
<tr>
<th>Parent → Children</th>
<th>Children → Parent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Money</td>
<td>$38,589</td>
</tr>
<tr>
<td>Hours</td>
<td>1060</td>
</tr>
</tbody>
</table>

Source: PSID Rosters and Transfers, 2013. Children are 18+ years old. It does not include transfers before that age neither those considered “long term” such as those for tuition or buying a house. 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 (high school or college) or start working. Education increases their human capital and modifies their life

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

16 The procedure is similar to the one used to calculate the TFR in (1). 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 up all age groups to obtain the average transfer over a lifetime. Different from Abbott et al. (2013) we also include observations with zero transfers. Note that we do not discount transfers to the present value.
cycle of income (among other things). When education is complete, they enter the second stage, which represents their labor market experience. Idiosyncratic uninsurable income risk makes individual earnings stochastic. Throughout their life they choose saving and consumption expenditures. They can borrow only up to a limit, and save through a non state-contingent asset. During this stage, there is also a period in which they choose how many children to have and how much of their resources to transfer to them. Finally, the last stage consists of retirement where they have three sources of income: savings, retirement benefits and old-age support from their children. We study the partial equilibrium version of this economy where prices and government policies are taken as given. 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 her consumption and at \( j = J_i \) she becomes independent. The initial assets are money transfers from her parent and the initial human capital as well as school taste (or psychic cost) which are stochastic but correlated with the education and human capital of her parent.

From \( j = J_i \) until \( j = J_e \) the agent has the option to study. The individual state variables are savings \( a_j \), human capital \( h_j \) and psychic cost \( \psi \). While on education, she sequentially chooses whether to continue in school or to start working and this 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 she chooses education, her education increases to \( e + 1 \), while human capital increases deterministically by \( f^s_{e,j}(h) \). The cost of education is \( p^s_j \) but, as is common in the literature (Heckman et al., 2006; Abbott et al., 2013), we also allow for school taste or psychic costs \( \psi \) to affect the cost of education.\(^{17}\) Particularly, we assume that these costs are perceived by the agent as increasing the financial cost of education.\(^{18}\) After leaving school, the psychic cost is assumed not to affect any adult outcome. While working, human capital evolves stochastically and is distributed by \( f^w_{e,j}(h) \) where 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.

---

\(^{17}\)We 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.

\(^{18}\)Note that these costs have a marginal value in consumption terms, which can be a better way to interpret them. The benefit of this specification (relative to having the cost in utility terms directly) is that its simplifies the calibration without altering the interpretation.
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,

$$V_j^{sw}(a_j, h_j, e, \psi) = \max \left\{ V_j^s(s_j, h_j, e, \psi), V_j^w(s_j, h_j, e) \right\},$$

where $V_j^s$ is defined by

$$V_j^s(a_j, h_j, e, \psi) = \max_{c_j, a_j+1} u(c_j) + \beta V_{j+1}^{ew}(a_{j+1}, h_{j+1}, e+1, \psi)$$

where $c_j = a_j + p_j^s + \psi = h w_j^s (1 - \tau) + a_j (1 + r)$

and $a_j+1 \geq s$, $h_{j+1} = f_j^s(h_j)$.

The agent is risk averse and her preferences are represented by the increasing, concave and positive utility function $u$.\(^{19}\) She can borrow up to the limit $a$ and the return on savings is $1 + r$. Future is discounted by $\beta$ and note that, in this particular problem, there is no uncertainty. We denote as $w_j^e$ the wage for an agent of age $j$ that is in school. In particular, we assume that the agent does not work during high school ($w_j^s = 0$) and we allow for (part-time or internship) work while in college.

The value of work $V_j^w$ is defined by

$$V_j^w(a_j, h_j, e) = \max_{c_j, a_j+1} u(c_j) + \beta \mathbb{E} \left[ V_{j+1}^{w}(a_{j+1}, h_{j+1}, e) \right],$$

where

$$c_j + a_j+1 = h_j w (1 - \tau) + a_j (1 + r),$$

and $a_{j+1} \geq s$, $h_{j+1} = f_j^w(h_j)$.

The return from working is the wage $w$ net of taxes $\tau$. Note that there is no disutility from working and so the labor supply is inelastic. Also note that the choice of leaving education is irreversible. Once

\(^{19}\)That $u$ is positive is important 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 generates zero utility.
the agent enters the labor force, she cannot return to school.

From \( j = J_e \) until \( j = J_r \) the agent works and her individual problem is equivalent to (5). 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 and, once the children become independent, the money transferred to them. Second, when her parents retire at \( j = J_t \), the agent provides old-age support, transferring a fraction of her current labor income to her parent.

We model altruism à la Barro and Becker (1989) where parents care about the utility of their children. Then, the problem at the age of fertility \( j = J_f \) is

\[
V_{J_f} (a_{J_f}, h_{J_f}, e) = \max_{c_{J_f}, c_{J_f+1}, a_{J_f+1}, n, T} u(c_{J_f}) + \beta \mathbb{E} \left[V_{J_f+1} (a_{J_f+1}, h_{J_f+1}, e, n) \right] + b(n) u(c_{J_f}). 
\]

\( c_{J_f} + nc_{J_f} + a_{J_f+1} + C(h, n, T) = h_{J_f} w (1 - \tau) + a_{J_f} (1 + r) \),

\( a_{J_f+1} \geq a, \ h_{J_f+1} \sim f_{e,J_f}^w (h_{J_f}) \).

In this period the agent chooses her consumption \( c_{J_f} \), her children consumption \( c_{J_f} \), savings \( a_{J_f+1} \) and 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 consumption. The altruistic discount factor \( b(n) \) is increasing and concave.

The transfer to each children \( \varphi \) are assumed to be made in the period before the children independent (age \( j = J_k \)). More, they are assumed to be equal for all children.\(^{20}\)

\[
V_{J_k} (a_{J_k}, h_{J_k}, e, n) = \max_{c_{J_k}, c_{J_k}, a_{J_k+1}, \varphi, T} u(c_{J_k}) + \beta \mathbb{E} \left[V_{J_k+1} (a_{J_k+1}, h_{J_k+1}, e, n, \varphi, h_{J_k}) \right] + b(n) u(c_{J_k}),
\]

\( c_{J_k} + nc_{J_k} + a_{J_k+1} + \frac{n \varphi}{(1 + r)} + C(h, n, T) = h_{J_k} w (1 - \tau) + a_{J_k} (1 + r) \),

\( a_{J_k+1} \geq a, \ h_{J_k+1} \sim f_{e,J_k}^w (h_{J_k}), \ h^c \sim f^c (e, h_{J_k}), \ \psi^c \sim g^c (e) \).

Notice that differently to (6), the value function at this stage now includes the continuation value of

\(^{20}\)The altruism value derived from children depends on their initial assets. Therefore, we assume that (one period in advance) parents set a fund such that their children receive \( \varphi \) when they become independent. Moreover, we note that 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 \( n = 1 \).
the children. This is the last period where parent’s choices affect children’s outcome. As the problem is written recursively this implies that at every period where parent’s choices affect children’s outcomes the value function of their descendants will be taken into account. This refers to the altruistic motives of parents, who care about their children’s later outcomes.

Raising children is costly, which is reflected both in (6) and (7). The parent pays the cost \( C(h_f, n, T) \) 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 parent \( h_j \). The variable \( T \) is an indicator variable, \( T \in \{ \text{hire childcare, stay at home} \} \), to capture the idea that childcare is a good that can be either produced at home or hired in the market. 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 with their own human capital.\(^{21}\)

After the fertility decision the individual problem is similar to (5) with the following caveats. First, for \( J_i \) periods the parent chooses the children’s consumption and pays cost \( C \). Later, at \( j = J_i \) a fraction of the current labor income goes to the agent’s parent 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 eliminate 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 et al., 2005). In our heterogeneous agent model, which includes multiple stages of discrete choices (education and fertility) as well as high dimensionality of the state space, the altruism approach to old-age support is computationally infeasible. Hence, we adopt the rule that children are constrained to transfer an exogenous share \( \xi \) of their income to their parents à la Morand (1999). Note that this transfer should be considered to include both money transferred and time dedicated to parents. Anyway, 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.\(^{22}\) This would make the old-age support motive even more important for poor families.

At \( j = J_r \) the agent retires with three sources of income. He has savings and retirement benefits that depend on the human capital and are progressive as in the US Social Security Administration.

---

\(^{21}\) We 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 3. Finally, note that childcare has no effect on the children’s human capital.

\(^{22}\) The marginal utility of consumption for poor parents would be higher than for richer families, increasing the incentive of children of poor parents to transfer.
Furthermore, 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.\footnote{Given that the data shows families having over four children, this number of states is not small.} 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 are the number of children $n$, the initial assets of their children $\varphi$ and their own education and human capital when the children were at home $h,J$\footnote{The education level and human capital of the parents is essential for the initial draw of human capital and psychic cost of the children. In the solution we use the human capital from the period before the children become independent to reduce the state space in previous periods. The initial assets are very important to determine how long the children will remain in school. Finally, the number of children $n$ is important to determine, among others, the size and uncertainty of the transfer to be received.}. These 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.$^{25}$ Formally, the problem at the age of retirement is

$$ V_j (a_j, h, e, \theta) = \max_{c_j, a_{j+1}} u(c_j) + \beta V_{j+1}^w (a_{j+1}, h, e), $$

$$ c_j + a_{j+1} = \theta + \pi(e, h) + a_j (1 + r), $$

$$ a_{j+1} \geq a, $$

where $\theta$ are the old-age support transfers and $\pi$ are the retirement benefits, which depend on the education and human capital at the age of retirement.$^{26}$

### 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 brings their children more income). In the calibrated version of the model, altruism will be such that it provides incentives to have a rather constant number of children across income groups. Old-age support causes children to be a form of private investment, with different rates of returns across

\footnote{We 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, our assumption does not seem very harmful.}

\footnote{We use education, together with the last level of human capital, as a proxy to approximate average lifetime income with which are the retirement benefits determined.}
families, which can be important as an investment for retirement. However, in general they are a relatively bad one since they are costly in a stage of life where they would actually like to borrow.\footnote{For example, in the US the current average age of first birth is 27 while the income peak is closer to 50 years old.} More importantly, for high income parents children are particularly costly due to the time they require. And their return in future transfers are also not that relevant since they will have plenty other sources of savings. The opposite is true for low income households, who have low time costs and obtain high marginal returns from them, making the old-age support channel much more valuable for poorer households. This way, the altruism channel dominates for richer households while the investment one gives an extra incentive for poorer agents to have children.

To understand these effects consider the first order condition with respect to the quantity of children $n^{28}$

$$u_{Jc}' C_n = b_n (n) \beta J I E Jc \left[ V_0 (h_{0 \theta}, \varphi) | e, h_{Jc} \right] + \beta J R Jc \left[ V_{JR} (h, s, \theta) | n, e, h_{Jc} \right].$$

(9)

The left hand side is the marginal cost of an extra child, that is the marginal childcare cost $C_n$, scaled by the parent’s marginal utility at age of fertility $u_{Jc}'$. The right hand side is composed of two terms. The first is the benefit from altruism, while the second is from old-age support. The altruism channel is standard, with the payoff being the expected value function of the child and the discount factor increasing by $b_n (n)$. The benefit of old-age support is generated by the change in the distribution of the transfers $\theta$. For the sake of clarity, we can decompose the random variable $\theta$ into its conditional mean and a martingale shock $\varepsilon \sim g_\varepsilon$. This implies that

$$\frac{\partial E_{Jc} [V_{JR} | n, e, h_{Jc}]}{\partial n} = \mu_\theta E_{Jc} [u_{JR}'] + \int \left( \int u_{JR}' \frac{\partial g_\varepsilon (\varepsilon | n, e, h_{JC})}{\partial n} d\varepsilon \right) df (h_{JR} | e, h_{JC})$$

where the expected transfer of each child is $\mu_\theta = w (1 - \tau) \xi E \left[ h_{JT} | e, h_{JC} \right]$. 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 for richer agents the old-age support is less important 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 are less important. On the other hand, the altruism benefits do not decrease and therefore dominate the fertility trade-off.\footnote{To understand the trade-off we regard $n$ as a continuous variable in equation (9). Moreover, to simplify the exposition we do not take into account the utility derived by the consumption of the children. This argument is standard but complicates the explanation. A similar argument can be derived from the first order condition with respect to the transfers to children $\varphi$.}
Consequently, this implies that as the economy grows - for instance as $w$ increases - fertility choices will be controlled by altruism instead of old-age support. This provides a framework that 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 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 transfers will 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.\footnote{Once an economy is sufficiently rich, it is possible for children to behave as a normal good. Every household prefers educated children and rich families can afford more of them.}

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.\footnote{Recall that old-age support transfers occurs only at age $J_r$ but social security benefits are received throughout retirement. In Figure 5 we compare the net present value of social security with the stock of savings and the old-age support transfers.} For the poorest quintile old-age support represents around 20\% of resources but for the richest quintile it represents less than a quarter of that. 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 the children are less important. This coincides with the intuition obtained from equation (9): old-age support is more relevant for poor agents in our model.
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 five state variables and several non convexities due to the discrete choices in education and fertility. Therefore we apply a generalized endogenous grid method from Fella (2014). Then we simulate the economy to match moments from United States in 1960. Some of the parameters can be estimated “externally”, while others need to be estimated “internally” from the simulation of the model.

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, four years) and then to college (another period, four years) 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 $J_r = 68$ and therefore parents retire at the child’s age
of \( J_t = 40 \), which is when children transfer money to their parents. 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 (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 borrowing is not possible, \( a = 0 \). The payroll tax is from Krueger and Ludwig (2015), \( \tau = 0.124 \) and is chosen to match the current social security payroll tax. The price of college is from Delta Cost Project, where we get $6588.\(^{31}\) The 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. In this class of models it is difficult to match the high school dropouts rate. Previous studies such as Abbott et al. (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 which 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 are between 0 and \( \bar{\psi} \), allowing for its distribution to be related to her parents education through parameter \( \omega \). 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 \( \omega \), and decreasing for those of college graduates.\(^{32}\) Our estimation suggests that psychic costs are higher for children of less educated parents, which is consistent with previous estimates in the literature. However, we remark that - as we will see in section 5 - they do not play as big a role as in previous models (Keane and Wolpin, 1997; Abbott et al., 2013). We find this very important as these psychic costs are a leftover in the estimation as it is not observable in the data. Hence, we believe that reducing the importance of this element in our models goes in the direction of reducing what we cannot actually explain.

**Education returns:** We base our estimates of the high school returns on sections 3 and 8 from the literature review of Heckman et al. (2006), with a return of 45% (equivalent to an yearly return of

\(^{31}\)We take into account grants and scholarships, such that only private tuition costs are considered. Prices are in 2000 US dollars.

\(^{32}\)Psychic costs are distributed between 0 and \( \bar{\psi} \). The Cumulative Density Function (CDF) for children of high school graduates is uniform, i.e. \( F(x) = \frac{x}{N} \). The CDF for children of high school dropouts is \( F(x) = \left( \frac{x}{N} \right)^\omega \), which assigns higher probability to higher values of psychic cost \( x \) the higher \( \omega \) is. For simplicity, children of college graduates are assumed to have opposite probabilities of children of high school dropouts.
10%). The return of college is allowed to be heterogeneous, as suggested by Heckman et al. (2006). Particularly, we specify the college human capital production function to have the non-linear form $f_{w,s,j}^n(h) = \log(h_{j+1}) = \alpha + \beta \log(h_j)$. Table 3 shows that our estimates are $\alpha = 1.12$ and $\beta = 0.6$. This are associated to average yearly returns (in the whole population, not just those that do study) of 8% with an standard deviation in the population of 6%, which is consistent with estimates from section 8 in Heckman et al. (2006).

**Labor income risk:** We assume that $f_{j}^{e,w}(h) = h(1 + \delta)$ with $\delta \in \{\delta_1, \delta_2, \delta_3\}$ which vary by age $j$ and education $e$. These and their probabilities are estimated 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. Two comments are appropriate. First, this 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 very important in order 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 income age profile is well estimated, including the coefficient of variation of log income as well as two measures of income inequality over the life cycle (the ratio of top 20% to bottom 20% earners and the Gini coefficient). It is seen that they display similar levels and trends both in the model and in the data.

---

**Childcare:** We propose the following functional form for childcare

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

This function allows for non-constant returns to scale in the number of children. We particularly

---

33The return of high school is hard to identify in the data as most individuals currently complete high school in the US and there could be very relevant selection issues on those who do not. Hence, we use 10% as the yearly return which is within the band of estimates, and closer to its lower bound. This means that we are also giving a lower bound to the role of family background with regards to the possibility to afford high school.

34We assume that these are constant before 24 and after 63 as data is problematic for those ages.
assume there increasing returns to scale with $\alpha_n = 0.47$.\textsuperscript{35} 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.\textsuperscript{36} We estimate that if the agent stays at home, her labor income is reduced by $\beta_n = 1/6$ but if she hires childcare her labor income is reduced only by $\beta_c = 1/16$.\textsuperscript{37} The cost of childcare is $(\beta_n - \beta_c) p_c$ with $p_c = $26669.\textsuperscript{38}

**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 equivalent to almost 15%.

**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.\textsuperscript{39} Then, the pension formula is given by

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

(10)

where $\bar{y}$ is approximately $50,000.

**Intergenerational transmission of ability:** We assume that the initial level of human capital is stochastic but correlated with the human capital 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% 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% (even below 5% for those with high income). In our model, each child first draws

\textsuperscript{35}The returns to scale are based on Folbre (2009).

\textsuperscript{36}Based on NLSY79.

\textsuperscript{37}We adapt the estimation from Angrist and Evans (1998) to be consistent with the equilibrium of our model. In particular we assume that in the estimation of Angrist and Evans (1998) high income households hire childcare in the market whereas low income households stay at home.

\textsuperscript{38}The cost of childcare is based on the average wage of nannies from IPUMS Census data.

\textsuperscript{39}With the last level of human capital before retirement $h$ and the education level $\epsilon$, we estimate the average life time income to be $\hat{y}(h) = h(\epsilon) \times h$ with $h$ equal to 0.97, 1.22 and 1.23 for high school dropouts, high school graduates and college graduates, respectively. Then average annual income $\hat{y}$ is used in (10) to obtain the replacement benefits.
her ability quintile from the panel in Figure 7 corresponding to her parents’ education and income tercile (within education group). Then, she draws $h_{Ji}$ uniformly within the given quintile of (positive) truncated normal distribution with parameters $(\mu, \sigma)$ to be estimated. Recall that another source of intergenerational persistence is that we also allow for psychic costs to be correlated with parent’s education level (explained above).

Figure 7: Correlation of parent’s background and children ability (%)

![Figure 7: Correlation of parent’s background and children ability (%)](image)

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 A.2

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 the literature and assume $\gamma_c = 0.5$ (for example, see Roys and Seshadri (2014)).

We are left with seven parameters. Two parameters, $\lambda_n$ and $\gamma_n$, are related to altruism. $\sigma$ is the standard deviation of the initial distribution of human capital. $\alpha$ and $\beta$ define the heterogenous returns to education in college. Finally, $\psi$ defines the distribution (both mean and standard deviation) of the psychic costs, while $\omega$ is related to the correlation with parent’s education level. In order to pin down the value of these parameters we aim to match ten moments from the data, shown in Table 3. First, given our focus on inequality, we target economic measures like the top-bottom income ratio and Gini coefficient as well as more social ones like the distribution of education attendance (dropouts, high school graduates and college graduates). To make sure the (market) value of education is properly identified, we also target the ratio of average incomes between education groups. As we are also interested in social mobility, we target the intergenerational mobility rank-rank coefficient. Finally, to identify the altruism we target the mean fertility and the fertility elasticity.\(^{40}\) Table 3 shows the results of the estimation, while table 4 shows the estimated parameters.

\(^{40}\)Given the very 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 $\frac{1}{\max\{m_j, m_j^2\}}$ where $m_j$ is the value of moment $j$ in the data. Since some moments could potentially
Table 3: Targeted moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dropouts</td>
<td>11.9</td>
<td>13.6</td>
</tr>
<tr>
<td>High school graduates</td>
<td>59.0</td>
<td>56.8</td>
</tr>
<tr>
<td>College graduates</td>
<td>29.1</td>
<td>29.6</td>
</tr>
<tr>
<td>Mean fertility</td>
<td>2.050</td>
<td>2.000</td>
</tr>
<tr>
<td>Fertility elasticity</td>
<td>-0.150</td>
<td>-0.114</td>
</tr>
<tr>
<td>Income top-bottom</td>
<td>7.275</td>
<td>7.302</td>
</tr>
<tr>
<td>Income Gini</td>
<td>0.440</td>
<td>0.439</td>
</tr>
<tr>
<td>Intergenerational Mobility: Rank-Rank</td>
<td>0.341</td>
<td>0.402</td>
</tr>
<tr>
<td>Mean income HS Dropout / HS Grad (Age 40)</td>
<td>0.647</td>
<td>0.781</td>
</tr>
<tr>
<td>Mean income College Grad / HS Grad (Age 40)</td>
<td>1.905</td>
<td>1.656</td>
</tr>
</tbody>
</table>

Source: Fertility elasticity, Gini, income top-bottom 25-29 and income ratios are constructed from IPUMS Census data. Mean fertility is obtained from United Nations Statistics. Education attainment are reported in the Current Population Survey historical time series tables. The intergenerational mobility is from Chetty et al. (2014).

Table 4: Estimated parameters

<table>
<thead>
<tr>
<th>$\gamma_n$</th>
<th>$\lambda_n$</th>
<th>$\sigma$</th>
<th>$\bar{\psi}$</th>
<th>$\omega$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.277</td>
<td>0.103</td>
<td>0.471</td>
<td>0.332</td>
<td>4.930</td>
<td>1.127</td>
<td>0.601</td>
</tr>
</tbody>
</table>

The model performs generally well. Fertility levels and its income elasticity are successfully matched, which is very important given the key role of fertility in our model. Regarding education, the three levels are well matched in the model. With regards to inequality, the model displays similar levels of lifetime inequality as the data. We obtain relatively low levels of social mobility, but we can partially open the transition matrix on which this estimate is based on to get a better sense of why this is the case. Figure 8 shows that the model does a good job in replicating the transition probabilities for all entries except for the probability of remaining at the very bottom or top quantiles.

We have tried other alternatives to solve for this issue, but we have not found significant differences.
Finally, we do a very simple exercise in order to show that the model is consistent with the cross-country patterns described in Figures 1 and 3. 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 (keeping relative prices of education and childcare stable), such that the real wage (i.e. in consumption terms) is the main change. Table 5 shows that the cross-country correlation between GDP per capita and fertility elasticity and between the Gini coefficient and intergenerational mobility, both in the data and the model. This rather simple exercise shows that the model is at least capable of qualitatively capturing these correlations. Moreover, it also does a decent quantitative job, particularly with regards to the correlation between GDP per capita and fertility elasticity. 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.

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>corr(GDP,Fertility Elasticity)</td>
<td>0.827</td>
<td>0.793</td>
</tr>
<tr>
<td>corr(Gini,Intergenerational Mobility)</td>
<td>0.519</td>
<td>0.807</td>
</tr>
</tbody>
</table>

5 Results

We now use the model to answer the questions presented in the introduction, with our first goal being to study the role of initial opportunities on income inequality. There is a large literature - both theoretical
and empirical - on distributive justice, which has recently been summarized by Roemer and Trannoy (2015).\textsuperscript{41} We contribute with a quantitative evaluation (using our estimated model) of some inequality measures proposed there but particularly following the equality of opportunities (EoP) definition used by Brunori et al. (2013). We first define an outcome of interest such as lifetime income or educational attainment. We then divide the population into different types according to their initial conditions (e.g. parent’s income, initial assets or human capital). We evaluate the effect of initial conditions on the outcome of interest by comparing the distribution of outcomes of different types. Equality of opportunities is defined to be achieved when the conditional distributions of outcomes are equalized across types. We first 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 and on variance decompositions.

We start by looking at the income of children and defining the different types according to parents’ income.\textsuperscript{42} The advantage of this definition is that we can do this exercise both in the model and in the data (using NLSY79). 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), where it is seen that the model is able to match the data relatively well.\textsuperscript{43} For example, the distance between the two lines in the model is similar to that in the data. Focusing on the median, the high income parents group has almost twice the income of the poor parents group. More generally, it is seen that 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 an important factor determining children’s income and equality of opportunity does not seem to hold either in the data nor the model.

\textsuperscript{41}The philosophical debate is dominated by the seminal articles Rawls (1982), Arneson (1989) and Cohen (1989). For recent empirical studies see Brunori et al. (2013) or Niehues and Peichl (2014) and the quantitative analysis by Lee and Seshadri (2015).

\textsuperscript{42}In 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 on 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.

\textsuperscript{43}The model fails to capture both tails - particularly the upper one - since it is not designed to fit the extremes of the income distribution.
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. Moreover, life cycle income paths could differ by individuals (for example, due to education groups), making any single year not good enough to proxy lifetime earnings. This suggests that a more relevant outcome measure of income is lifetime earnings which is usually not 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) which are also not available in the data. However, we can use our model to look at our outcome measure of interest (lifetime earnings) and its relation with the true initial conditions in the model (initial transfers, human capital and school taste), which is shown in Figure 10. This exercise suggests that lifetime earnings depend strongly on being above a low threshold of human capital (which provides a major advantage in the labor market and increases income during college). However, once a threshold around the median is passed, differences of initial levels of human capital seem less important. On the other hand, family transfers are still important but on a more continuous way (relevant to afford education). Finally, consistent with other results, psychic costs or school taste do not seem that relevant as differences across groups are relatively small. Importantly, recall that these initial state variables are positively correlated through family background.
This visual inspection suggests that the outcomes’ distributions differ across types. Particularly, there seems to be more inequality of opportunity when types are defined using the actual initial state variables rather than proxies like parents’ income. For a more formal analysis, 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 types and within types. One such
measure is the Theil-L index which is defined as

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

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

\[
T = \sum_{j=1}^{J} \frac{m_j}{m} \ln \left( \frac{E_j[y]}{E[y]} \right) + \sum_{j=1}^{J} \frac{m_j}{m} E_j \left[ \ln \left( \frac{y_i}{E_j[y]} \right) \right] \\
= T_b + \sum_{j=1}^{J} \frac{m_j}{m} T_j
\]

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.

Table 6 summarizes our main results. Focusing for now on the columns Data and Benchmark, the first two panels show that, as discussed in the previous section, the model generates levels of inequality and social mobility consistent with the data. The third panel looks at different measures of inequality of opportunity, with its first part looking at the Theil-L index explained above. The empirical problem with this measure is that it is very difficult to observe all types or outcomes. For instance, Brunori et al. (2013) estimate a lower bound of inequality of opportunity of 16% whereas Niehues and Peichl (2014) estimate an upper bound of 75%. As explained by Brunori et al., this literature focuses on observable initial differences like gender, race or parents’ income. Since the first two do not exist in our model, we replicate their exercise using the only variables in our model that could potentially be observable by Brunori et al.. Identifying parents’ income and education as initial conditions, the fifth row of Table 6 shows that our estimates are very close to the lower bound from these empirical studies. However, the advantage of having a 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 (which are more closely related to lifetime utility and reduce the role of temporary shocks). In this case, the sixth row of Table 6 shows that our actual level of inequality of opportunity (as measured by the relative Theil-L index) is 84%.
Table 6: Social mobility and income inequality

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Benchmark</th>
<th>Exogenous fertility</th>
<th>Constant transfers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$N = 2$</td>
<td>High school</td>
</tr>
<tr>
<td><strong>Income Inequality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income top-bottom 25-64</td>
<td>7.3</td>
<td>7.3</td>
<td>6.3</td>
<td>4.6</td>
</tr>
<tr>
<td>Income gini 25-64</td>
<td>0.44</td>
<td>0.44</td>
<td>0.38</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Social mobility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intergenerational Mobility: Rank-Rank</td>
<td>0.34</td>
<td>0.40</td>
<td>0.24</td>
<td>0.02</td>
</tr>
<tr>
<td>Transition Par Q1-Child Q5 (%)</td>
<td>8</td>
<td>6</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td><strong>Inequality of opportunity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Theil-L relative (%): Income 28-31 (Parent’s income and education)</td>
<td>16 - 75</td>
<td>16.6</td>
<td>4.8</td>
<td>25.0</td>
</tr>
<tr>
<td>Theil-L relative (%): Lifetime Earnings (Initial conditions)</td>
<td>84.1</td>
<td>82.9</td>
<td>88.9</td>
<td></td>
</tr>
<tr>
<td>CV of Lifetime Earnings</td>
<td>0.62</td>
<td>0.60</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>% expl. by initial conds</td>
<td>68.4</td>
<td>68.2</td>
<td>73.6</td>
<td></td>
</tr>
<tr>
<td>% expl. by initial HK</td>
<td>41.6</td>
<td>47.7</td>
<td>44.8</td>
<td></td>
</tr>
<tr>
<td>% expl. by transfers</td>
<td>7.1</td>
<td>4.5</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>% expl. by school taste</td>
<td>6.0</td>
<td>4.1</td>
<td>23.0</td>
<td></td>
</tr>
<tr>
<td>% expl. by interaction</td>
<td>13.8</td>
<td>11.9</td>
<td>5.8</td>
<td></td>
</tr>
<tr>
<td>Variance years of education</td>
<td>6.1</td>
<td>6.5</td>
<td>3.9</td>
<td>7.8</td>
</tr>
<tr>
<td>% expl. by initial conds</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
<tr>
<td>% expl. by initial HK</td>
<td>18.6</td>
<td>26.7</td>
<td>22.3</td>
<td></td>
</tr>
<tr>
<td>% expl. by transfers</td>
<td>52.2</td>
<td>29.1</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>% expl. by school taste</td>
<td>12.4</td>
<td>24.0</td>
<td>75.1</td>
<td></td>
</tr>
<tr>
<td>% expl. by interaction</td>
<td>16.8</td>
<td>20.2</td>
<td>2.5</td>
<td></td>
</tr>
</tbody>
</table>

An alternative way to evaluate the importance of initial conditions is the one followed by Huggett et al. (2011), which is shown in the second part of the third panel of Table 6. We decompose the variance of lifetime earnings into variation due to initial conditions versus variation due to adult shocks. Huggett et al. (2011) found that 61% of variation in lifetime earnings is determined by age 23. Taking into account the role of families, we find that most of this is actually determined even before high school. Table 6 reports that in our model 68% of variation in lifetime earnings is determined by

---

44Such a decomposition makes use of the fact that a random variable can be written as the sum of its conditional mean plus the variation from its conditional mean. As these two components are orthogonal, the total variance equals the sum of the variance in the conditional mean plus the variance the conditional mean.
initial conditions as of age 13.\textsuperscript{45} Regarding policy implications, our model is consistent with the early childhood investment literature which suggests that improving children’s conditions very early in their lives can have a very significant impact on their future outcomes (Heckman et al., 2010; Gertler et al., 2013).

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. The last part of the third panel of Table 6 decomposes the variance in years of education and shows that parents’ transfers play a crucial role, explaining over 50% of the total variance in years of education. This is consistent with the results from Belley and Lochner (2007) who find that family income and wealth are relevant for educational choices. Importantly, school taste or psychic costs do not play a major role in either of these variance decompositions. This is interesting as most of the literature has relied on this unobservable (and hence harder to understand) element to explain education choices.

Finally, we can use our model to ask how important endogenous fertility and family transfers are for social mobility. For this we keep the same model and estimated parameters from Section 4, but move away from the endogenous fertility and family transfers. First, we allow for endogenous family transfers but look at the exogenous fertility case where each parent has three children. Column 3 of Table 6 reports that in such a society without endogenous fertility inequality would be slightly reduced and social mobility would almost double for all our measures. The rank-rank intergenerational mobility measures would be halved while the probability of moving from the bottom to the top quintile would double.\textsuperscript{46} We also 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 very similar conclusion. Alternatively, we can also look at the case where we allow for endogenous fertility but parents’ transfers are exogenously constant.\textsuperscript{47} Similarly to removing endogenous fertility, the fourth column shows that without endogenous transfers inequality would be reduced and social mobility improved, but differences are even bigger now. With constant family transfers, the big role that initial assets used to play in education choices is eliminated and most of lifetime earnings is now due to characteristics less directly

\textsuperscript{45}The later we compute this, the smaller the window for lifetime earnings is, and hence the more initial conditions will be able to explain the outcome of choice. Hence, we would expect to have a lower number the further back in time we go within a same model. However, we have to keep in mind that our model is quite different from Huggett et al. (2011). Particularly, education choices differences due to variation in initial states (human capital, family transfers and school taste) can have a large impact on lifetime earnings which is not directly captured in Huggett et al. (2011).

\textsuperscript{46}We believe that countries or periods with higher fertility differentials than the US in the year 2000 may display a much bigger role for fertility differentials on inequality and social mobility. An older version of this paper was estimated using the US in 1960 and results were in that direction.

\textsuperscript{47}We can fix the amount transferred 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. As this is a partial equilibrium exercise, this ensures that prices and other parameters still remain sensible.
related to parent’s 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 inequality.

To summarize, we find that initial opportunities account for more of income inequality than does adult
income risk over the working life. Our results suggest that taking into account family transfers as well
as the negative income-fertility relationship observed in the data is very important for the analysis of
inequality and social mobility. A model interested in understanding social mobility or capturing the
role of inequality of opportunity, should take fertility differentials and family transfers into account.

Some caveats are worth mentioning regarding our results. First, one concern left for future research is
whether general equilibrium effects could alter the results. It is not evident to us that qualitative results
should change significantly by endogenizing prices in our model, but it could be worth exploring. Even
though this may not be theoretically difficult to solve, we have not dealt with it due to computational
limitations. Second, our results on the role of initial conditions should be considered to apply as of
age 13. It would be interesting to go all the way back to age 0 so as to fully understand how initial
conditions are formed. Our model makes one step in that direction and takes into account how family
transfers are determined. Although our model is calibrated to match the data it is still silent about
the exact way the initial level of human capital is formed. Further research is needed to understand
this.

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 effort and
luck experienced over the working lifetime. In order to overcome data limitations, this paper uses
a standard heterogeneous agent life cycle model with education and idiosyncratic risk extended to
account for the role of families (through endogenous fertility and family transfers) 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. Importantly, the model is able to capture both US and cross-
country evidence on the relation between fertility differentials, income inequality and social mobility.
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. 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 13) account for more of income inequality than does adult income risk over the working life. If we particularly look at inequality of opportunity as computed by Brunori et al. (2013) or Niehues and Peichl (2014), our model indicates that 84% of observed inequality is due to different initial opportunities (very close to their empirically estimated upper bound). On the other hand, if we focus on the variation of lifetime earnings, as does Huggett et al. (2011), our model suggests that 68% of this is due to differences in conditions determined even before high school. Our model also points that fertility differentials and family transfers generate over 50% of the social immobility observed in the data, implying that those interested in understanding social mobility may need to take these two forces into account.

Some comments on policy implications are appropriate now. Our results suggest that improving access to education (through education subsidies for example) may have an important effect in reducing the importance of family transfers, helping reduce income inequality and improve social mobility. Alternatively, old-age support is a motive in our model for fertility differentials which introduces a new potential gain from a generous government support for the elderly. Standard analysis of retirement benefits abstracts from the long run effects these may have on inequality and social mobility through fertility and education. Our model suggests that increasing the retirement benefits (or their progressivity) could reduce the incentives of poor families to have many children. By reducing the share of children born with few resources (and who cannot access education), it may decrease inequality or improve social mobility. It would be interesting to analyze the effect of these alternative policies but doing so requires a general equilibrium model from which we have abstained in order to provide a more detailed analysis of the sources of income inequality and social mobility in the data. Finally, 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 (Cunha et al., 2010) or on ways to improve it (like Gertler et al. (2013) on early childhood investment) is needed to understand how this initial distribution may be modified to reduce later income inequality or promote social mobility.
References


## A Additional Figures and Tables

Table A.1: Countries included in empirical work

<table>
<thead>
<tr>
<th>Country</th>
<th>Years</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>1987</td>
<td>LIS</td>
</tr>
<tr>
<td>China</td>
<td>2002</td>
<td>LIS</td>
</tr>
<tr>
<td>Colombia</td>
<td>2004, 2007, 2010</td>
<td>LIS</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>1992</td>
<td>LIS</td>
</tr>
<tr>
<td>Dominican Republic</td>
<td>1981, 2002</td>
<td>IPUMS</td>
</tr>
<tr>
<td>Guatemala</td>
<td>2006</td>
<td>LIS</td>
</tr>
<tr>
<td>India</td>
<td>2004</td>
<td>LIS</td>
</tr>
<tr>
<td>Mexico</td>
<td>1995, 2000</td>
<td>IPUMS</td>
</tr>
<tr>
<td>Panama</td>
<td>1980, 1990, 2010</td>
<td>IPUMS</td>
</tr>
<tr>
<td>Peru</td>
<td>2004</td>
<td>LIS</td>
</tr>
<tr>
<td>Romania</td>
<td>1995, 1997</td>
<td>LIS</td>
</tr>
<tr>
<td>Slovak Republic</td>
<td>1992</td>
<td>LIS</td>
</tr>
<tr>
<td>South Korea</td>
<td>2006</td>
<td>LIS</td>
</tr>
<tr>
<td>United States</td>
<td>1968-2013</td>
<td>IPUMS (CPS)</td>
</tr>
<tr>
<td>Uruguay</td>
<td>2004</td>
<td>LIS</td>
</tr>
</tbody>
</table>

**Sources**: Minnesota Population Center (IPUMS) and Luxembourg Income Study Database (LIS)
Figure A.1: Income Elasticity of Fertility and GDP: All observations

![Graph showing income elasticity of fertility and GDP]

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

Figure A.2: Income Elasticity of Fertility and Inequality: All observations

![Graph showing income elasticity of fertility and inequality]

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

Figure 1 showed that income inequality differs across countries and that it seems associated with lower social mobility. In the current Section we have shown that poorer countries tend to have bigger fertility
differentials than richer ones. How does income inequality relate to fertility differentials? Intuitively, we would expect that in countries with bigger fertility differentials, stronger inequalities are present 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 very scarce resources while a smaller group is born with a bigger pie to split among fewer hands. Assuming this affects their education or business opportunities, we would expect to observe more inequality in countries with bigger fertility differentials. Figure A.3 looks at these using the same sample as before and calculating income inequality using total household income. Then we divide the observations into deciles according to their levels of income inequality. For each decile we calculate the median level of inequality as well as the median fertility elasticity of income. As expected, Figure A.3 shows that countries with more inequality are associated with higher fertility differentials than more equal ones. This evidence coincides with that shown by Kremer and Chen (2002) which looked at fertility differentials across education groups.

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

Source: IPUMS International and LIS. Methodology is explained in the main text. Figure A.2 in the Appendix includes all the data observations before grouping them in deciles.
Table A.2: Correlation of parent's background and children ability (%)

<table>
<thead>
<tr>
<th>Mother</th>
<th>Children</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school dropouts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low income</td>
<td></td>
<td>51</td>
<td>23</td>
<td>14</td>
<td>05</td>
<td>7</td>
</tr>
<tr>
<td>Medium income</td>
<td></td>
<td>37</td>
<td>24</td>
<td>18</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>High income</td>
<td></td>
<td>28</td>
<td>21</td>
<td>22</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>High school graduates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low income</td>
<td></td>
<td>31</td>
<td>24</td>
<td>17</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>Medium income</td>
<td></td>
<td>17</td>
<td>22</td>
<td>22</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>High income</td>
<td></td>
<td>12</td>
<td>16</td>
<td>23</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>College graduates</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low income</td>
<td></td>
<td>10</td>
<td>19</td>
<td>23</td>
<td>23</td>
<td>25</td>
</tr>
<tr>
<td>Medium income</td>
<td></td>
<td>4</td>
<td>12</td>
<td>19</td>
<td>30</td>
<td>35</td>
</tr>
<tr>
<td>High income</td>
<td></td>
<td>4</td>
<td>12</td>
<td>18</td>
<td>30</td>
<td>37</td>
</tr>
</tbody>
</table>

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