ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Explaining Intergenerational Mobility: The Role of Fertility and Family Transfers

Authors	Diego Daruich, and Julian Kozlowski
Working Paper Number	2018-011F
Revision Date	October 2019
Citable Link	https://doi.org/10.20955/wp.2018.011
Suggested Citation	Daruich, D., Kozlowski, J., 2019; Explaining Intergenerational Mobility: The Role of Fertility and Family Transfers, Federal Reserve Bank of St. Louis Working Paper 2018-011. URL https://doi.org/10.20955/wp.2018.011

Published In	Review of Economic Dynamics
Publisher Link	https://doi.org/10.1016/j.red.2019.10.002

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment.

Explaining Intergenerational Mobility: The Role of Fertility and Family Transfers

Diego Daruich
University of Southern California

Julian Kozlowski*
Federal Reserve Bank of St. Louis

October 17, 2019

Abstract

Poor families have more children and transfer less resources to them. This suggests that family decisions about fertility and transfers dampen intergenerational mobility. To evaluate the quantitative importance of this mechanism, we extend the standard heterogeneous-agent life cycle model with earnings risk and credit constraints to allow for endogenous fertility, family transfers, and education. The model, estimated to the US in the 2000s, implies that a counterfactual flat income-fertility profile would—through the equalization of initial conditions—increase intergenerational mobility by 6%. The impact of a counterfactual constant transfer per child is twice as large.

JEL Classifications: J13, J24, J62, D91.

Keywords: Intergenerational mobility, Inequality, Fertility

*ddaruich@nyu.edu; kozjuli@gmail.com. We are grateful to our advisors, Raquel Fernandez, Ricardo Lagos, and Gianluca Violante. For helpful comments, we thank Matthias Doepke, Simon Mongey, Pau Roldan, Edouard Schaal, Laura Veldkamp, and Venky Venkateswaran. We also received valuable feedback from participants at seminars at Brown University, New York University, Di Tella University, ECINEQ 2015, EEA 2015, the 2015 Chicago Summer School on Socioeconomic Inequality, the Fall 2015 Midwest Macro Meeting, SED 2016, and the Families and the Macroeconomy Conference 2018. The views expressed are ours and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

What factors determine intergenerational mobility? We study the sources of intergenerational mobility focusing on the impact of family choices, in particular fertility and parental transfers. Extensive empirical evidence shows that family choices are heterogeneous and correlated with family characteristics: Poor families tend to have more children (e.g., Jones and Tertilt, 2008) and invest fewer resources toward their children than rich families (e.g., Altonji et al., 1997). The heterogeneity in family choices can lead to higher levels of inequality relative to an economy without such heterogeneity because it can increase the differences on education outcomes. Moreover, the correlation between income and family choices can also lead to lower intergenerational mobility, as the children of richer parents have more resources available for education.

To evaluate the quantitative importance of family choices on intergenerational mobility, we build a model in which families choose the number of children and the parental transfers to them. Poor parents have more children because they are time-intensive and the time-opportunity cost is smaller for low-income than for high-income parents. With fewer resources and more children, poor parents also invest less in their kids. Parents influence children's initial conditions regarding skills and economic resources, both of which shape their education choices and later labor income due to capital market imperfections. Specifically, we introduce endogenous fertility, family transfers, and education in the standard heterogeneous agent life cycle model, with idiosyncratic income risk and credit constraints. This extension allows us to study intergenerational dynamics and, in particular, the role of fertility and family transfers in intergenerational mobility, which is our main contribution.

The model incorporates four sources of intergenerational persistence. The following two sources are endogenous and are the main focus of this paper: the choices of (i) number of children and (ii) parent-to-child transfers, both of which affect the resources available for each child's education and, hence, their labor income. Given the objective of quantifying the importance of these two sources, we also incorporate two other standard (and exogenous in our model) sources of persistence to avoid biasing our results: (iii) Initial skills, which affect education and income and are related to parental characteristics to capture, in an stylized manner, pre-birth and early childhood development as sources of persistence[1]; and (iv) adult income shocks, which reduce the importance of initial conditions for labor earnings and, thus, increase intergenerational mobility.

We estimate the model to the US in the 2000s. To estimate the novel elements in our model, we fit moments about the relationship between fertility and income, family transfers, and intergenerational mobility. A set of validation exercises shows that the quantitative model is consistent with both non-targeted moments and new cross-state evidence in the relationship between average income, fertility, and education. In particular, we find that children born in states with smaller fertility differences between income groups are associated with higher high school graduation rates, and this also holds in the model.

The variation in lifetime earnings can be decomposed into differences in initial conditions—defined as the agents' initial state variables—and in labor-income shocks. Huggett et al. (2011) also use a heterogeneous agent life cycle model and find that initial conditions drive most of the lifetime-earnings inequality. Consistent with their results, our model suggests that 56% of the lifetime-earnings inequality in the US can be attributed to initial conditions. A key difference relative to Huggett et al. (2011) is that, in our analysis, initial conditions are endogenously related to parental background through fertility and transfer choices, allowing us to study intergenerational mobility.

We find that both fertility and transfer differentials are important to understand intergenerational mobility. First, we solve an alternative model in which fertility is exogenous and constant across families. This exercise reveals that in the baseline economy, fertility differentials account for 6% of the intergenerational mobility observed in the data. This difference is equivalent to approximately one-third of the standard deviation in intergenerational mobility across commuting zones (Chetty et al., 2014). Second, we simulate an economy in which fertility is endogenous, but transfers from parents to children are exogenous and constant at the

[1]We model initial skills as human capital and school taste at age 16. We incorporate these exogenous sources of persistence to avoid biasing our quantitative results. Without these, our estimation procedure that targets the amount of intergenerational persistence in US, among other things, would have to assign some of the observed persistence to other factors, possibly biasing our results.

average level. Parental transfers play a major role; the impact on intergenerational mobility of this counterfactual constant transfer per child is twice as large as the one from the constant fertility counterfactual. Both of these counterfactuals also reveal that fertility and transfer differentials increase inequality, particularly for young individuals.

Effects, however, are heterogeneous for children with different family backgrounds. Children from low-income and less-educated parents are those with fewest resources and most siblings. They are, therefore, the most affected by the constant fertility or constant transfer counterfactuals. With either counterfactual, children born to bottom-income-quintile parents exhibit a lower probability of remaining in the bottom quintile, while children of top-income-quintile parents have a larger probability of dropping to the bottom quintile. Even though effects are heterogeneous, the aggregate distribution of education shifts toward an economy with fewer high school dropouts and more high school graduates, which decreases income inequality and increases intergenerational mobility.

Both of these counterfactuals operate through the distribution of initial conditions, particularly initial assets (or parental transfers) and human capital. With a counterfactual flat income-fertility profile there are relatively fewer children born from poor households. As fewer children are born with low levels of initial human capital and assets, the initial distribution becomes more homogeneous. This increases access to education and, thus, also intergenerational mobility. With counterfactual constant transfers per child, the initial distribution also becomes more homogeneous, increasing mobility through similar mechanisms. Our findings suggest that to understand social mobility, one should take fertility differentials and family transfers into account.

We evaluate the effect of two simple policies that directly target a reduction on fertility differentials or early inequality in resources, taking into account that fertility and parental transfers will react to these policies. First, we consider a conditional transfer in which each family receives $20,000 only if they have two children. We find that the effect of this policy is similar to the constant-fertility counterfactual, inducing households to have two children and increasing intergenerational mobility. Second, we consider an unconditional transfer in which each household receives $20,000 as part of their initial assets. We find that this policy generates no changes on intergenerational mobility because poor households choose to have more children, which increases both the average fertility and the fertility elasticity. Abbott et al. (Forthcoming) studies parental transfers and detailed government education policies, showing that endogenous parental transfers are important to take into account when evaluating government education policies due to crowding out effects. Even though our model is much more limited than theirs regarding education policies, by introducing endogenous fertility, these two simple policy exercises highlight that fertility, as well as transfers, is not policy invariant and may undo the intended effect of policies that aim to equalize initial assets.

Finally, we also study the role of the exogenous sources of persistence in our model. Adult income shocks and initial human capital are both quantitatively relevant for intergenerational mobility. Among these, however, the stochastic process behind the initial human capital is the most important. Even though we model this as an exogenous—but disciplined by data—process, our results suggest that policies that affect initial human capital may also have sizable effects on fertility and transfer choices. Thus, studies of human capital policies that abstract from family transfers and fertility choices may be biased in their results.

Related Literature This paper relates to two literatures usually studied in isolation: income inequality and intergenerational mobility. However, there is a strong and positive correlation between the two (Corak, 2013). On the one hand, models of inequality typically focus on adult shocks and abstract from endogenous initial conditions (e.g., Keane and Wolpin, 1997; Huggett et al., 2011). On the other hand, models of intergenerational mobility usually focus on initial conditions and abstract from adult income volatility (e.g., Restuccia and Urrutia, 2004; Lee and Seshadri, 2018). Both initial conditions and labor-income volatility generate income inequality. We contribute by providing a model that combines these two sources and assess their relative importance, which also allows for the joint study of inequality and mobility.[2]

There is a large literature that studies intergenerational persistence of human capital and,

[2]The literature on quantitative models combining adult uncertainty and endogenous initial conditions is scarce. See Yum (2018), Daruich (2019), and Lee and Seshadri (2019) for some relevant exceptions.

also, tries to distinguish the role of family background (e.g., genetics or parental investments), environment (e.g., neighborhood, social networks or peers), and government policies (e.g., public education or progressive taxation), among others, in generating this persistence. First, parental investments are well-known to be crucial for human capital accumulation in early childhood and elementary school (e.g., Todd and Wolpin, 2003, 2007; Cunha et al., 2010; Agostinelli and Wiswall, 2016), high school (e.g., Card and Krueger, 1992; Heckman et al., 2018) and college (e.g., Belley and Lochner, 2007; Abbott et al., Forthcoming). This paper only incorporates the role of parents in early stages of development exogenously but endogenously models both high school and college (through endogenous parental transfers and education choices).[3]

Second, environment is usually found to be an important factor behind human capital accumulation (e.g., Brooks-Gunn et al., 1993; Cutler and Glaeser, 1997; Chetty and Hendren, 2018a,b). Given that higher-income parents tend to live in environments more suitable for development, environment is also suggested to be important for intergenerational persistence. Our model is only able to capture this effect in an exogenous reduced-form fashion. Pre-high-school human capital is an exogenous function that depends on parental human capital, which can be interpreted to capture the role of environment as well as parents (through genetics and early education) in early-age human capital persistence. We highlight, however, that our counterfactuals will not be able to capture the indirect effects through this channel. Finally, government policies can help reduce intergenerational persistence (e.g., Herrington, 2015; Holter, 2015; Lee and Seshadri, 2019; Daruich, 2019). Even though we abstract from progressive taxation or targeted education policies, we contribute to this literature by studying the role of fertility and parental transfers (and simple potential policies that may affect these two) for intergenerational persistence.

The closest paper to ours is Huggett et al. (2011), in which the authors use a Bewley model

[3]Given the substantial evidence that suggests that early childhood and, even, in-utero development is crucial (e.g., Heckman and Mosso, 2014; Almond et al., 2018), it would be interesting to study the endogenous interaction between endogenous fertility and early childhood development. However, to properly model early childhood, Yum (2018), Daruich (2019), and Lee and Seshadri (2018) have to introduce various types of skills (i.e., cognitive and non-cognitive), different forms of investment (i.e., time and/or money), and age-dependent skill-production functions. Adding this complexity to a model with endogenous fertility will make the environment even more convoluted and harder to understand the results (not to mention the computational burden). Thus, we left the interaction between fertility and early childhood choices as an open question for future research.

to study the sources of inequality. They find that most income inequality is due to conditions present before entering the labor market. These conditions, however, are exogenous in their analysis, implying that their results are silent about the forces that determine inequality of opportunity. Given the extensive literature on the importance of parental environment and investments for human capital development, modeling family choices is necessary to study the origin of these pre-labor-market conditions. Our model endogenizes these earlier stages of life through choices regarding education, fertility, and family transfers. By incorporating intergenerational linkages, we are able to explore the role of family background forces that may determine inequality and intergenerational mobility.[4]

We study the impact of fertility choices on inequality and social mobility.[5] Our model highlights a quantity-quality trade-off as in Barro and Becker (1989) as a main determinant of initial conditions. There is evidence that poor families have more children than richer ones—i.e., there is a negative elasticity of fertility to income (Jones and Tertilt, 2008).[6] We contribute to this literature by showing that this elasticity is smaller for richer states within the US. We also find that children born in states with larger fertility differentials are associated with lower education outcomes. Our quantitative model is consistent with these empirical findings.[7]

Several papers explore the effect of inequality on human capital accumulation (or growth) through fertility choices (e.g., de la Croix and Doepke, 2003, 2004, 2009; Moav, 2005). In these papers, inequality affects growth negatively through its interaction with fertility choices. We complement this literature by exploiting a similar mechanism between fertility choices and adult

[4]Borrowing constraints increase the importance of parental transfers to afford education in our model (see Appendix C for details). Belley and Lochner (2007) provides evidence that suggests that borrowing constraints have become increasingly important over time and Lochner and Monge-Naranjo (2011) shows that they are helpful to explain the correlation between ability and schooling in the US.

[5]Note that we focus on labor-income inequality and do not look into wealth inequality. Recent literature also finds a decisive role for family background in explaining wealth inequality (Nardi and Yang, 2016; Benhabib et al., Forthcoming). We also abstract from sorting, another force that has been used to generate inequality through families (e.g., Fernandez and Rogerson, 2001; Fernandez et al., 2005).

[6]Jones and Schoonbrodt (2016) show that the fertility rate is positively correlated with the business cycle. We abstract from business cycle dynamics and focus, instead, on the cross-sectional dispersion of fertility choices.

[7]Quantitative models in the fertility literature include Manuelli and Seshadri (2009) and Roys and Seshadri (2017), which are used to explain differences in average fertility rates across countries and long-term economic growth, respectively. Nevertheless, both abstract from uncertainty, and though heterogeneity is allowed in Roys and Seshadri (2017), it is only in the form of constant skill differences across dynasties.

outcomes to study, instead, the effect of fertility differentials on inequality and intergenerational mobility.

The rest of the paper is organized as follows. Section 1 introduces the model, and Section 2 explains its estimation and conducts some validation exercises. Section 3 presents the model's results on inequality and intergenerational mobility. Finally, Section 4 concludes. The Appendices contain additional details.

1 Model

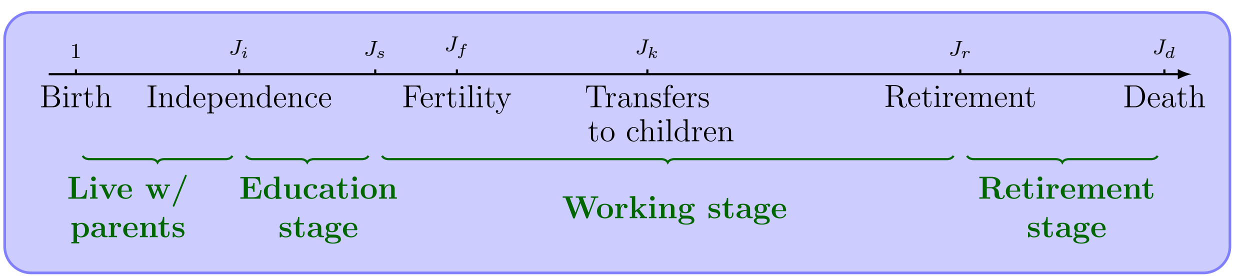
We specify a life cycle economy in a dynastic framework with four main stages. In the first stage, individuals live with their parents. In the second stage, agents decide whether to attend school or start working. Education increases their human capital and modifies their life cycle of income, as well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

There are four main sources of intergenerational persistence in the model. The following two are endogenous: (i) the number of children and (ii) parent-to-child transfers. Both of these sources affect the resources available for education and, hence, their labor income. As we want to quantify the importance of these two sources, we have to also incorporate two other standard (and exogenous in our model) sources of persistence to avoid biasing our results: (iii) initial skills—modeled as human capital and school taste at age 16—which affect education and income and are related to parental characteristics to capture pre-birth and early childhood development as sources of persistence[8]; and (iv) adult income shocks, which reduce the importance of initial conditions for labor earnings and, thus, increase intergenerational mobility. Section 3 studies the importance of each of these forces for education, inequality, and, particularly, intergenerational mobility.

1.1 The individual problem

Figure 1 shows the life cycle of an agent, in which each period in the model refers to two years. Let j denote age at the beginning of the period. From $j = 1$ until $j = J_i$, the child lives with her parents, who choose the child's consumption. At age $j = J_i$, the child becomes independent. Her initial states are assets, human capital, and school taste (or psychic cost). Initial assets are money transfers from her parents. The initial human capital and school taste are stochastic but correlated with the parents' education and human capital. We calibrate the independent age as $J_i = 16$, meaning that the agent has a separate optimization problem from her parents, and there are no more transfers between parents and children after age 16. Nevertheless, parents directly influence children's education choices when they decide their initial assets.[9]

Figure 1: **Life cycle**



[8]Modeling school taste is necessary because pecuniary returns can only account for part of the observed college attendance patterns by human capital (e.g., Abbott et al., Forthcoming; Cunha et al., 2005; Heckman et al., 2006).

[9]Even though assuming that agents are independent at age 16 may seem early in life, we note that even though parents cannot provide state-contingent insurance on some of the initial risk in life, they do provide non-state-contingent insurance through their transfer. Moreover, given that parents in our model are altruistic, they agree with the choices their children make for a given set of initial states (i.e., after that uncertainty is revealed).

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

Education stage: From $j = J_i$ until $j = J_s$, the agent has the option to study. The individual state variables are assets a, human capital h, and school taste ϕ. The first choice the agent makes is the education decision, which is irreversible. All agents become independent as high school dropouts with 10 years of education ($e = 1$). The agent can choose to complete high school ($e = 2$), which takes 1 period (2 years), or be a college graduate ($e = 3$), which takes 1 period to finish high school and 2 periods to complete college (a total of 6 years). Education increases the human capital deterministically as $f^{s,e}(h_0)$.

The cost of education is p_{HS} and p_{Coll} for high school and college, respectively. We also allow for school taste $\phi \in [0, 1]$ to affect the total cost of education, as in the literature (e.g., Heckman et al., 2006; Abbott et al., Forthcoming). Modeling school taste is necessary because resources available to finance schooling and returns to education can only partially account for the observed education patterns. Particularly, we assume that the school taste enters as a separate term in the value function. We scale the school taste ϕ by a different constant in each schooling level $\bar{\psi}_e$. After leaving school, the psychic cost is assumed not to affect any adult outcome. While working, human capital evolves stochastically. We allow for education- and age-dependent idiosyncratic and persistent labor-income shocks. In Section 2, we discuss the estimation of the returns of education and the income process.

Students face borrowing limits $\underline{a}^s_e$ for subsidized loans. High-school students cannot borrow (i.e., $\underline{a}^s_{HS} = 0$). College students have access to subsidized loans at rate $r^s = r + \iota^s$, where $\iota^s < \iota$. To simplify computation, we follow Abbott et al. (Forthcoming) and assume that, after graduation, college student debt is refinanced into a single bond that carries interest rate r^-.[10] Borrowing limit $\underline{a}^s_{Coll}$ and wedge ι^s will be based on federal college loans, to be explained in detail in Section 2.

Formally, let V^s_j and V^w_j be the value of an agent of age j in school and working, respectively. The first choice the agent makes, at age $j = J_i$, is how much education to acquire. The value function at this stage is given by V_{J_i}:

$$V_{J_i}(a, h_0, \phi) = \max\left\{\mathbb{E}\left[V^w_j(a, h_0, 1, z)\right], V^s_j(a, h_0, 2) - \phi\bar{\psi}_2, V^s_j(a, h_0, 3) - \phi\bar{\psi}_3\right\},$$

where V^s_j is defined by

$$\begin{aligned} V^s_j(a, h_0, e) &= \max_{c,a'} u(c) + \beta\tilde{V}^s_{j+1} \qquad (1)\\ & c + a' + p_{\Omega(j)} - h^{sw} w_{\Omega(j)}(1-\tau) = \begin{cases} a(1+r) & \text{if } a \geq 0 \\ a(1+r^s) & \text{if } a < 0 \end{cases} \\ & a' \geq \underline{a}^s_{\Omega(j)}, \\ \Omega(j) &= \begin{cases} HS & \text{if } j = J_i \text{ (i.e., high school age)} \\ Coll & \text{if } j \geq J_i + 1 \text{ (i.e., college age)} \end{cases} \\ \log h^{sw} &= \begin{cases} \log h_0 + \gamma_{j,e} & \text{if } j = J_i \text{ (i.e., high school age)} \\ \log f^{s,HS}(h_0) + \gamma_{j,e} & \text{if } j \geq J_i + 1 \text{ (i.e., college age)} \end{cases} \end{aligned}$$

[10] We assume that fixed payments would have been made for 10 periods (i.e., 20 years) following graduation, so we can transform college loans into regular bonds using the following formula:

$$\tilde{a}^s(a') = a' \times \frac{r^s}{1-(1+r^s)^{-10}} \times \frac{1-(1+r^-)^{-10}}{r^-},$$

where $\tilde{a}^s(a')$ is the function performing this transformation. Stafford college loans, the ones on which our estimation is based, have various repayment plans during which the borrower pays a fixed amount each month. Even though repayment plans typically last 10 years, they can be extended to up to 25 years. As in Abbott et al. (Forthcoming), we choose 20 years for our fixed payment plan in order to fall on the conservative side of our results. The longer the repayment plan, the less important family transfers should be to be able to afford education. Thus, we expect that a shorter repayment plan would increase the role of fertility and family transfers on intergenerational mobility.

$$\tilde{V}^s_{j+1} = \begin{cases} \mathbb{E}\left[V^w_{j+1}\left(a', h_0, e, z\right)\right] & \text{if } e = 2 \text{ and } j = J_i \\ & \text{(i.e., last period of schooling for HS grads)} \\ \mathbb{E}\left[V^w_{j+1}\left(\tilde{a}^s(a'), h_0, e, z\right)\right] & \text{if } e = 3 \text{ and } j = J_i + 2 \\ & \text{(i.e., last period of schooling for college grads)} \\ V^s_{j+1}\left(a', h_0, e\right) & \text{otherwise.} \end{cases}$$

The agent is risk averse and her preferences are represented by an increasing, concave, and positive utility function u.[11] The return on positive savings is $1 + r$. However, if the agent is borrowing, she pays interest rates $r^s > r$ and can borrow up to the limit $\underline{a}^s_{\Omega(j)}$. β is the discount factor. We denote as $w_{\Omega(j)}$ the wage for an agent who is currently in school at level $\Omega(j)$. In particular, we assume that the agent does not work during high school (i.e., $w_{HS} = 0$), and we allow for (part-time or internship) work while in college (i.e., $w_{Coll} \in [0, w]$). Note that the next shock the agent receives is z_0, which is the initial value of the persistent component of the income process. This shock is realized in the first period of work. Hence, we only have to take expectations with respect to that in V^w.

The value of work V^w_j is defined by

$$V^w_j\left(a, h_0, e, z\right) = \max_{c,a'} u\left(c\right) + \beta \mathbb{E}\left[V^w_{j+1}\left(a', h_0, e, z'\right)\right], \tag{2}$$

$$c + a' - hw\left(1 - \tau\right) = \begin{cases} a\left(1 + r\right) & \text{if } a \geq 0 \\ a\left(1 + r^-\right) & \text{if } a < 0 \end{cases}$$

$$\log h = \log f^{s,e}(h_0) + \gamma_{j,e} + z,$$

$$z' = \rho_{z,e} z + \zeta, \quad \zeta \sim N(0, \sigma_{\zeta,e}),$$

$$a' \geq \underline{a}_{e,j}.$$

The income process has a fixed effect as a result of the initial h_0 transformed by the education

[11]The fact that the utility function u is positive is necessary to model altruism in a model with endogenous fertility. As shown by Jones and Schoonbroodt (2010), the implicit assumption that parents enjoy having children requires that the utility function must be always positive or always negative. If we choose the negative case, we need an extra assumption for the value of having zero children. Therefore, we follow the classic approach of u being always positive and assume that having zero children delivers zero utility.

choice, an age-education profile $\gamma_{j,e}$, and an AR(1) idiosyncratic shock z with persistence $\rho_{z,e}$ and innovation variance $\sigma_{\zeta,e}$. The agent can borrow up to the limit $\underline{a}_{e,j}$, and the return on positive savings is $1+r$. However, if the agent is borrowing she pays interest rates $r^- > r$. The return from working is the wage w net of taxes τ. There is no disutility from working, and so the labor supply is inelastic.

Working stage: From $j = J_s$ until $j = J_r$, the agent works and her individual problem is equivalent to (2). In the (exogenously given) fertility period $j = J_f$, however, the agent also chooses the number of children. Once the children become independent (at $j = J_k$), the agent also chooses the transfer to her offspring.

Fertility: We model altruism as in Barro and Becker (1989), in which parents care about the utility of their children. The problem at the age of fertility $j = J_f$ is

$$V_j^w(a, h_0, e, z) = \max_{c, c_k, a', n} u(c) + \beta \mathbb{E}\left[V_{j+1}^w(a', h_0, e, z', n)\right] + b(n)u(c_k) \tag{3}$$

$$c + nc_k + a' + C(h, n) - hw(1-\tau) = \begin{cases} a(1+r) & \text{if } a \geq 0 \\ a(1+r^-) & \text{if } a < 0 \end{cases}$$

$$\log h = \log f^{s,e}(h_0) + \gamma_{j,e} + z,$$

$$z' = \rho_{z,e} z + \zeta, \quad \zeta \sim N(0, \sigma_{\zeta,e}),$$

$$a' \geq \underline{a}_{e,j}, \qquad n \in \left\{0, 1, \ldots, \overline{N}\right\}.$$

In this period, the agent chooses her consumption c, her children's consumption c_k, savings a', and the number of children n, which is a discrete choice. As usual, the agent derives utility from her own consumption and her continuation utility. Furthermore, similar to Roys and Seshadri (2017), the agent is altruistic and derives utility from her children's consumption. The altruistic discount factor $b(n)$ is positive, increasing, and concave.

Raising children is costly, as is reflected in (3). Parents pay the cost $C(h, n)$ in addition to the money spent on children's consumption and transfers. First, the cost is increasing and concave in the number of children n reflecting the increasing cost of having more children as well as the increasing returns to scale. Second, the cost is increasing in h which we interpret, as in the literature stemming from Mincer (1963), as having children being time intensive and sacrificing work time being more costly for high-income parents than for low-income ones.[12] The functional form of C and its estimation are described in Section 2.

Until the agent's children become independent ($j = J_k$), she chooses the children's consumption and pays the cost C. Hence, the problem is equal to (3) but takes the number of children n as given. The transfer to each child φ is assumed to be made in the period before the offspring becomes independent (age $j = J_k$). Moreover, transfers are assumed to be the same for all children and weakly positive. The problem at the age when transfer to children is chosen $j = J_k$ is

$$V_j^w(a, h_0, e, z, n) = \max_{c, c_k, a', \varphi} u(c) + \beta \mathbb{E}\left[V_{j+1}^w(a', h_0, e, z')\right] + b(n)\left\{u(c_k) + \beta \mathbb{E}\left[V_{J_i}(\varphi, h_k, \phi_k)\right]\right\} \tag{4}$$

$$c + nc_k + a' + n\varphi + C(h, n) - hw(1-\tau) = \begin{cases} a(1+r) & \text{if } a \geq 0 \\ a(1+r^-) & \text{if } a < 0 \end{cases}$$

$$\log h = \log f^{s,e}(h_0) + \gamma_{j,e} + z,$$

$$z' = \rho_{z,e} z + \zeta, \quad \zeta \sim N(0, \sigma_{\zeta,e}),$$

$$a' \geq \underline{a}_{e,j}, \quad h_k \sim f^k(h), \quad \phi_k \sim g^k(e), \quad \varphi \geq 0.$$

Notice that unlike (3), the value function at this stage now includes the continuation value of the children V_{J_i}. This is the last period in which parents' choices affect their descendants. As the problem is written recursively, this implies that for every period in which parents' choices affect children's outcomes, the value function of their descendants will be taken into account. This embeds the parental altruism motives. The initial human capital and the school taste of the children are stochastic but correlated with the parents' human capital and level of education,

[12] Appendix D shows that the cost function can be micro-founded in terms of a time cost. Another possible interpretation is that parents with higher human capital spend more resources on their kids' early development and are also more likely to have higher-skilled children. This would be (exogenously) captured in our model by the cost function $C(h, n)$, as well as the correlation between the initial draw h_k and parents' human capital.

respectively.[13] The functional form of the altruism, as well at the stochastic processes of human capital, $f^k(h)$, and psychic costs, $g^k(e)$, are specified in Section 2.

Retirement stage: At $j = J_r$, the agent retires with two sources of income: savings and retirement benefits. These benefits depend on her education level and human capital and are progressive. Formally, the problem at the age of retirement is

$$V_j(a, h_0, e) = \max_{c,a'} u(c) + \beta V_{j+1}(a', h_0, e), \tag{5}$$
$$c + a' = \pi(e, h_0) + a(1+r),$$
$$a' \geq 0,$$

where π are the retirement benefits, which depend on the education and human capital.[14]

2 Estimation

The model is estimated to match household level data so that an agent in the model corresponds to a household with two adults in the data. The number of children n is also in terms of households, so $n = 1$ refers to one household.[15] We use the following three primary data sources: (i) IPUMS US Census; (ii) Panel Study of Income Dynamics (PSID); and (iii) 1979 cohort of the National Longitudinal Survey of Youth (NLSY79). We select a population for which our model can be taken as a reasonable approximation of household behavior and impose two selection criteria on the data. First, as is standard in the literature (e.g., Huggett et al., 2011), we drop household observations with income below a certain threshold. We choose this

[13] Initial skills are exogenously related to parental characteristics to capture, in an stylized manner, pre-birth and early childhood development as sources of persistence. Our model abstracts from the endogenous interaction between endogenous fertility and early childhood development—e.g., time and money investments as in Daruich (2019) or Lee and Seshadri (2018). Adding this complexity to a model with endogenous fertility will make the environment even more convoluted and harder to understand the results (not to mention the computational burden).

[14] We use education, together with the initial level of human capital h_0, as a proxy to approximate average lifetime income with which the retirement benefits are determined. See Section 2 for details.

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

threshold as the one that corresponds to one person working 20 hours a week for the minimum wage (approximately $8,000 total annual household income). Second, there is no decision regarding marriage in our model. Given our focus on fertility, we are interested in two-member households. To avoid differences in income and time availability due to single parenthood, we keep only married households. Additional details on statistics and sample selection are relegated to the Appendices.

Table 1: **Externally estimated parameters**.

Parameter	Value	Description	Source
Demographics			
	2	Time period	
J_i	16	Independent	
J_s	22	Maximum age for education	
J_f	28	Fertility decisions	
J_k	42	Transfers to children	
J_r	68	Retirement	
J_d	80	Death	
Prices			
w_{Coll}	0.56	Wage while in college	Census
τ	12.4%	Payroll tax	Social Security
p_{HS}	0.05	Price of high school	Digest of Education Statistics
p_{Coll}	0.58	Price of college	Delta Cost Project
Financial markets			
r	3%	Interest rate (annual)	Smets and Wouters (2007)
$\underline{a}_1$	-10	Borrowing limit of HS dropout ($1k)	SCF
$\underline{a}_2$	-24	Borrowing limit of HS graduate ($1k)	SCF
$\underline{a}_3$	-34	Borrowing limit of college graduate ($1k)	SCF
ι	10%	Borrowing-saving wedge (annual)	Gross and Souleles (2002)
ι^s	1%	College loan wedge (annual)	NCES
Childcare			
c_3	0.64	Returns to scale	Folbre (2008)
Preferences			
β	0.975	Discount factor (annual)	
γ_c	0.5	Risk aversion	Roys and Seshadri (2014)

Note: SCF refers to the 2001 Survey of Consumer Finances. NCES refers to the report "Student Financing of Undergraduate Education: 1999-2000," from the National Center for Education Statistics. Census refers to the 2000 Census data (IPUMS).

We numerically solve the steady state of this economy. There are several non-convexities due to the discrete choices in education and fertility, so we apply a global solution method.[16] We then compute the ergodic distribution of the economy to match moments from the US in 2000.

We describe below how we parameterize the model economy. Some of the parameters can be estimated "externally," while others must be estimated "internally" from the simulation of the model. Table 1 reports all externally calibrated parameters, while Table 2 summarizes all the internally estimated parameters in the model, as well as the moments used to estimate them. Although the model is highly nonlinear so that all parameters potentially affect all outcomes, the estimation of some parameters relies on some key moments in the data, which we briefly discuss below.

2.1 Simulated Method of Moments

Demographics: A period in the model is two years. Individuals become independent at the age of $J_i = 16$, and they start with the equivalent of 10 years of education. They can go to high school (one period) and then to college (another two periods), and so the maximum age for education is $J_s = 22$. Fertility decisions are made around the average age at first birth, $J_f = 28$.[17] At age $J_k = 42$, one period before the agent's children become independent, she chooses the assets to transfer to her children. Retirement occurs at $J_r = 68$. Death is assumed to occur for all agents at age $J_d = 80$.

Prices: The average annual income at age 42 in the model is one, and prices are normalized such that this is equal to $\$75,630$ in the data. We estimate the wage while in college from IPUMS census data. We focus on individuals between the ages of 18 and 22 years and match the relative earnings of those currently in college to those who are not, leading to $w_{Coll} = 0.56$. We set the annual interest rate to $r = 3\%$ (e.g. Smets and Wouters, 2007). Based on self-reported limits on unsecured credit by family from the Survey of Consumer Finances, we set $\underline{a}_{e,j}$, the borrowing limits for working-age ($J_s < j < J_r$) individuals, to $\{-10,000, -24,000, -34,000\}$ for

[16] We adapt the generalized endogenous grid method proposed by Fella (2014).

[17] The average age of first birth in 2007 was 27.97 for married women according to the National Center for Health Statistics.

high school dropout, high school graduate, and college graduate, respectively. The (annualized) wedge ι for borrowing is set to 10%, which is the average among the values for credit card borrowing interest rates (net of r and average inflation) reported by Gross and Souleles (2002).[18] Appendix C shows that the main results of the paper are robust to setting a lower interest wedge of $\iota = 0.05$.

The yearly price of college is from the Delta Cost Project, where we get $\$6,588$.[19] The yearly price of high school is obtained from the Digest of Education Statistics, using the relative private cost of high school to college. Our estimate of high school cost is about 9% of college cost, which is consistent with the US education system (i.e., relatively low cost of high school when compared to college), leading to a price for high school of \$593. Taking into account that education takes multiple periods and households contain two members, our normalization leads to $p_{HS} = 0.05$ and $p_{Coll} = 0.58$.

College Loans: College students have access to subsidized loans at rate $r^s = r + \iota^s$. According to the National Center for Education Statistics report "Student Financing of Undergraduate Education: 1999-2000," among the undergraduates who borrow, nearly all (97%) took out federal student loans—only 13% took out nonfederal loans. Moreover, the average loan value was similar for both federal and nonfederal cases. Since average values were similar but federal loans were significantly more common, we focus on federal loans for our model estimation. Among federal loans, the Stafford loan program was the most common: 96% of the undergraduates who borrowed took out Stafford loans. The second most common loans were the Perkins loans, but they were much smaller: Only 11% of borrowers used Perkins loans, and average amounts were one quarter of average Stafford amounts. Therefore, we focus particularly on Stafford loans. Stafford offers multiple types of loans, some of which are subsidized and some of which are not,

[18] Gross and Souleles (2002) report an average credit card interest rate of 17% between 1995 and 1998. During this period the average federal funds rate was approximately 7%, so we choose ι such that the annualized wedge is 10%. Our estimated model has 6.2% of households with negative net worth, which is similar to the estimate in Abbott et al. (Forthcoming).

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

so we use the weighted average interest rate to set $\iota^s = 0.009$.[20] The borrowing limit while in college in the model is set to match the cumulative borrowing limit on Stafford loans ($23,000).

School Taste: In this class of models, it is difficult to match the high school dropout rate. Previous studies (e.g., Abbott et al., Forthcoming; Krueger and Ludwig, 2016) introduced nonpecuniary (psychic) costs of education. We assume the agent's school taste (or psychic cost) ϕ is between 0 and 1, which will be scaled by different estimated levels according to the education choice ($\bar{\psi}_2$ and $\bar{\psi}_3$). Its distribution is related to parents' education through the parameter ω. Particularly, we assume that the psychic cost for children of high school-graduate parents is uniformly distributed between 0 and 1. On the other hand, we assume that the probability of high psychic costs for children of high school dropouts is increasing in ω, and decreasing for those of college graduates. Hence, the CDF of school taste is

$$G^k(e, \phi) = \begin{cases} \phi^\omega & \text{if parents are high school dropouts} \\ \phi & \text{if parents are high school graduates} \\ 1 - (1-\phi)^\omega & \text{if parents are college graduates.} \end{cases} \tag{6}$$

The share of high school and college graduates will be particularly informative of $\bar{\psi}_2$ and $\bar{\psi}_3$. For larger values of high school's (college's) taste shocks, we would observe more high school dropouts (less college graduates). In addition, higher correlation between parents' education and a child's school taste implies lower intergenerational mobility of education. We target the trace index of education mobility equivalent to $(3 - \text{trace}(P))/2$, where P is the transition matrix of education. No mobility implies an index value equal to zero, while perfect mobility implies an index equal to one. Thus, the degree of intergenerational mobility of education observed in the data will be informative about ω. Our estimation suggests that psychic costs are higher for children of less-educated parents, which is consistent with estimates in Abbott et al. (Forthcoming) and Krueger and Ludwig (2016).

Education returns: Returns to education vary between high school and college as well as

[20] According to the "Student Financing of Undergraduate Education: 1999-2000," 60% of borrowers took subsidized (interest free) loans, while 40% took unsubsidized ones. Until 2006, the interest rate was 2.3% above the prime rate. Thus, the weighted average interest rate on these loans is 0.9%.

between agents, as suggested by Heckman et al. (2006). Particularly, we specify the human-capital production function to have the nonlinear form

$$f^{s,e}(h) = h + \alpha_e h^{\beta_e}. \tag{7}$$

Income ratios between education groups are informative about the levels of education returns α_e, while the variance of log-income of agents with different education levels are informative about the curvature parameters β_e. In particular, a higher curvature implies a larger variation of income. Table 2 shows that our estimates for high school are $\alpha_2 = 0.16$ and $\beta_2 = 0.56$, while for college they are $\alpha_3 = 0.11$ and $\beta_3 = 0.24$. The returns to education using lifetime earnings and taking into account education costs (for the whole population, not just those that attend school) are 10% per year on average. This return is in line with empirical estimates in the literature summarized in Heckman et al. (2006).

Income process: We assume that labor productivity is $\log h\left(h_0, e, j, z\right) = \log f^{s,e}(h_0) + \gamma_{j,e} + z_j$. We estimate this following the procedure of Abbott et al. (Forthcoming) but using household income and 2-year periods instead; see Appendix B.1 for details. In a nutshell, both the age profile $\gamma_{j,e}$ and the persistent component z_j, which follows an AR(1) process, are estimated independently for each education group. Two comments are appropriate. First, income risk is calibrated to include total earnings variation, encompassing what may be considered both wage shocks and hours-worked (or effort) differences. Second, we match standard statistics of labor earnings. This is necessary to properly evaluate the impact of initial opportunities on income inequality. Otherwise, the comparison could be favorable for initial opportunities.

Opportunity cost of children: The cost function is $C(h, n) = c_1 w h^{c_2} n^{c_3}$.[21] The cost is increasing and concave in the number of children, which we interpret as increasing returns-to-scale in the number of children. We estimate the returns to scale $c_3 = 0.64$, based on Table 6.4 in Folbre (2009). The cost is increasing in h, which we interpret as having children being time intensive and sacrificing work time being more costly for high-income parents. This is the

[21] Appendix D shows that this cost function is the solution of a cost minimization problem of choosing amount of time spent at home and hiring childcare services on the market.

standard time-opportunity cost highlighted in the literature. The estimated values of c_1 and c_2 will be informed by the fertility rates by income deciles.[22] As the opportunity cost of having children increases (i.e., higher c_1), it becomes more costly to raise children, which implies a decreasing mean fertility rate. A lower value of c_2 implies that costs do not increase as much for higher-income parents. Thus, the rate of decay of fertility by income will be particularly informative for c_2. We estimate $c_1 = 0.59$ and $c_2 = 0.66$. We interpret this estimation as higher-income parents using market services (e.g., nannies or extra-curricular programs), which reduce the amount of total (i.e., not necessarily quality) time they need to spend with their children. This functional form is consistent with the empirical evidence (e.g., Sylva et al., 2007) showing that more advantaged families are more likely to use purchased child care. Section 2.2 analyzes the fertility choices and the role of the parameters.

Replacement benefits: The pension replacement rate is based on the Old Age, Survivors, and Disability Insurance federal program. The payroll tax is $\tau = 0.124$, which is the current rate for Social Security. We then use education level and initial human capital to estimate the average lifetime income, on which the replacement benefit is based. See Appendix B.2 for details.

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

$$log(h_0) = \mu_{h_0} + \rho \left[log(h_p) - log(\bar{h_p})\right] + \varepsilon_{h_0}, \tag{8}$$

where $\varepsilon_{h_0} \sim N(0, \sigma_{h_0})$ and $\bar{h_p}$ is the average human capital of parents at age 42.[23] Equation (8) defines f^k, the distribution of the initial draw of human capital in the household problem (4). The variance of log-income at age 28-29 in the model is related to the variance of the initial draw of human capital. As the initial draw becomes more dispersed, this variance increases. The degree of intergenerational mobility of income observed in the data is informative about

[22] Appendix A describes the estimation of fertility rates.

[23] Recall that μ_{h_0} is chosen such that the average annual labor income at age 42 is equal to one.

ρ: A higher persistence in the initial draw leads to a lower intergenerational mobility of income (higher rank-rank coefficient).

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 1, the utility function has to be positive, and therefore $\gamma_c \in [0, 1)$. We follow the literature and assume that $\gamma_c = 0.5$ (e.g., Roys and Seshadri, 2017). Other articles, like Manuelli and Seshadri (2009), that have estimated this parameter also obtain roughly this value. As is standard in the literature (e.g., Barro and Becker, 1989), the altruism function is assumed to be $b(n) = \lambda_n n^{\gamma_n}$. The target moment of average transfers from parents to children (as a share of average income) will be informative about λ_n since parents that value their children more (i.e., higher λ_n) would increase the transfers to them. The curvature of altruism (γ_n) will be informed by the fertility rate of different income deciles. When $\gamma_n = 0$, the marginal value of an additional child is equal to zero, which implies that all parents have (at most) one child. When γ_n is positive, however, the quantity-quality trade-off can generate a negative fertility elasticity.

To summarize, thirteen parameters of the model are estimated using Simulated Method of Moments with 15 target moments. Two parameters, λ_n and γ_n, are related to altruism; σ_{h_0} is the standard deviation of the initial distribution of human capital; ρ relates to the intergenerational persistence of human capital through the initial draw; c_1 and c_2 are the opportunity cost of raising children; α_e and β_e, for $e \in \{2, 3\}$, define the returns to education in high school and college; $\bar{\psi}_e$, for $e \in \{2, 3\}$, defines the distribution (both mean and standard deviation) of the school taste; and ω is related to the correlation of this school taste and parents' education level.

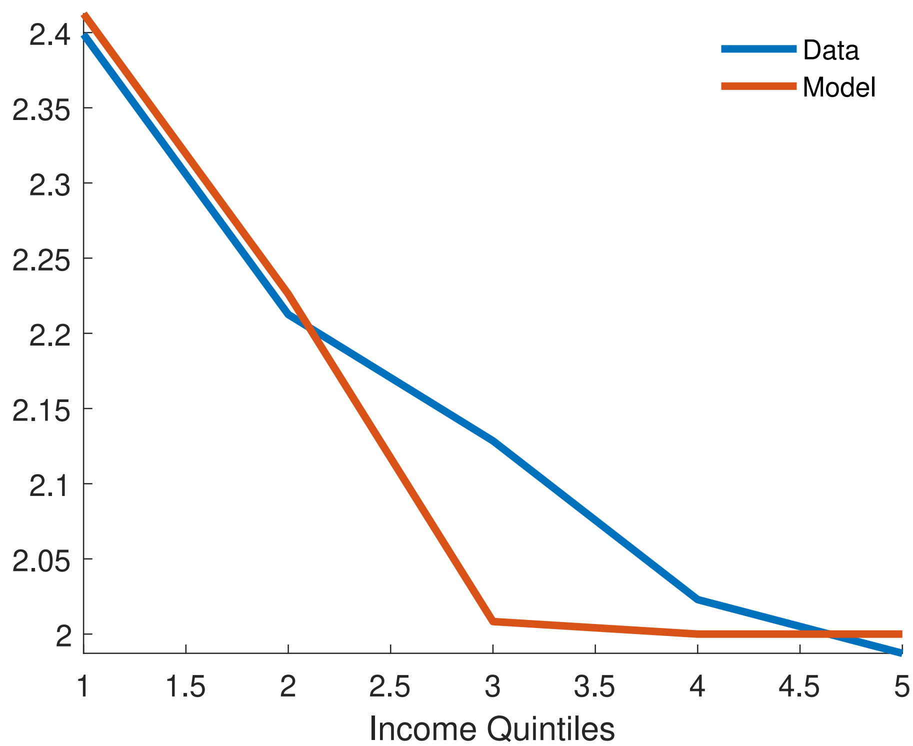
Table 2 shows the estimated parameters and target moments in the simulated economy. Figure 2 shows the moments regarding fertility by income deciles. Parent-to-children transfers (average), as well as fertility rates, are successfully matched, which is necessary given their key roles in our model. As for income inequality, the model displays levels similar to the data. Education shares are well matched in the model. We also obtain levels of intergenerational mobility that are close to the empirical evidence. Appendix C shows that the main results of the paper are robust to the estimated parameters.

Table 2: **Internally estimated parameters**.

Parameter	Value	Description	Moment	Data	Model
Altruism					
λ	0.76	Level	Parent-to-Child Transfers	0.43	0.44
γ	0.25	Curvature	Fertility by income quintiles		
Child cost					
c_1	0.59	Level	Fertility by income quintiles		
c_2	0.66	Curvature	Fertility by income quintiles		
Initial draw of human capital					
ρ	0.19	Intergenerational persistence	Intergenerational mobility of income: Rank-Rank	0.29	0.30
σ_{h_0}	0.25	Standard deviation	Variance of log(income)	0.34	0.32
Education returns					
α_2	0.16	High School: level	Log(Income) Ratio Age 28-29: HS Dropout - HS Grad	-0.43	-0.42
α_3	0.11	College: level	Log(Income) Ratio Age 28-29: College Grad - HS Grad	0.43	0.40
β_2	0.56	High School: curvature	Variance of log(income): HS Grad 28-29	0.27	0.27
β_3	0.24	College: curvature	Variance of log(income): College Grad 28-29	0.26	0.23
School taste					
$\bar{\Phi}_2$	3.98	High School	High School Dropout (%)	9.4	8.2
$\bar{\Phi}_3$	25.71	College	College Graduates (%)	30.5	32.5
ω	1.73	Intergenerational correlation	Intergenerational mobility of education	0.85	0.86

Source: Education shares, income ratios, variances of log(income), and fertility by income deciles are calculated from 2000 Census data (IPUMS). Parent-to-Child transfers refers to the average transfer (as a share of mean income) from Abbott et al. (Forthcoming). Intergenerational mobility of education is measured using the trace index, as reported by Checchi et al. (1999). Intergenerational mobility of income is measured using the rank-rank coefficient reported by Chetty et. al. (2014) for children of married parents.

Figure 2: **Fertility by Income Quintiles: Model and Data**.



Note: The model is estimated to replicate the fertility rate by income quintiles from the 2000 Census.

2.2 Endogenous Fertility

Households choose the number of children based on their state variables: assets (a), education (e), and the permanent and transitory components of the income process (h_0 and z). The heterogeneity on the state-space generates different costs and benefits of having children, implying heterogeneous fertility rates across households. Moreover, fertility will also change due to policy or changes in the economic environment because it is an endogenous variable.[24] First, fertility is decreasing in income, as shown in Figure 2. Second, the top-left panel of Figure 3 shows that, conditional on income, fertility also depends on assets and households with more wealth have more children. Intuitively, children are costly because they are time-intensive and reduce household labor income. Hence, households with low savings are not able or willing to

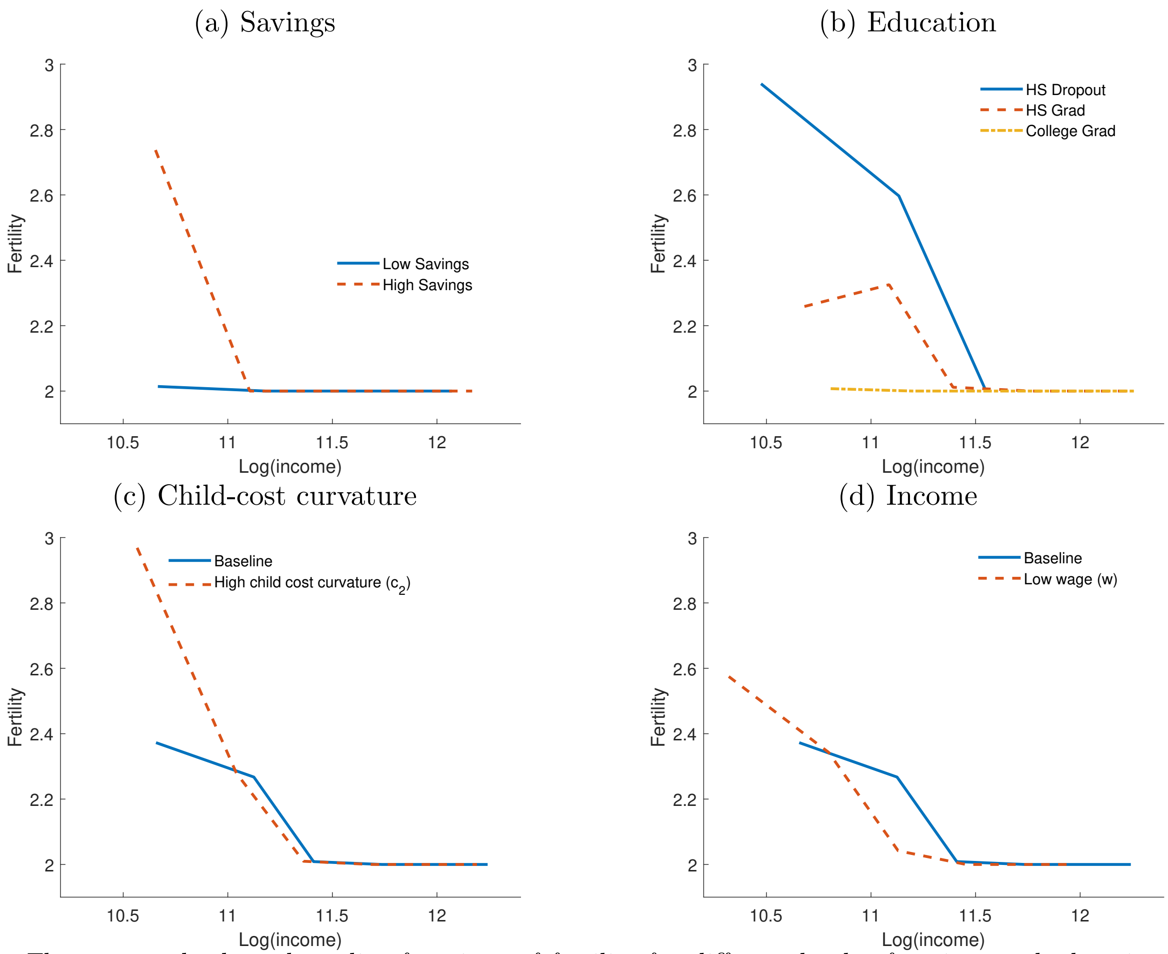
[24] In a simpler model in which fertility is an exogenous function of only human capital, fertility choices would be independent of both other state variables as well as the economic environment. In such a model, counterfactual analyses will not take into account the response of fertility to policy changes that are beyond human capital.

afford as many children.

The top-right panel shows that, conditional on a level of current income, fertility also depends on education and is higher for less-educated households. The main intuition is that the age profile ($\gamma_{j,e}$) is steeper for more-educated households. Hence, the net present value of the cost of having children, conditional on current income, is increasing in the education of parents.

The curvature of the child-cost function (given by c_2) is the key parameter behind the fertility differentials across income groups shown in Figure 2. The bottom-left panel of Figure 3 shows that increasing c_2 leads to larger fertility differentials with high-income households still having two children.

Figure 3: **Fertility Choices**.



Note: The top panels show the policy functions of fertility for different levels of savings and education. The bottom-left panel shows how fertility changes with the curvature (c_2) of the child-cost function. The bottom-right panel shows the fertility rate for different levels of w in the baseline estimation.

2.3 Validation Exercises

We test the validity of the estimation by looking at non-targeted moments. First, we evaluate the model within the steady state. Then, we present new empirical evidence on cross-state variation of fertility, which we use to test the model by moving the economy away from the steady state. Table 3 summarizes the results of the validation exercises.

Inequality: As a measure of inequality, the estimation targets the variance of log-income. Table 3 shows that the model is also in line with another measure of inequality such as the Gini coefficient. Other measures of inequality, such as the coefficient of variation or the top-bottom, are also similar in the estimated model and the data.

Table 3: **Validation exercises**.

Moment	Data	Model
Inequality		
Income Gini	0.38	0.37
Fertility by education		
High school dropouts	2.90	2.64
High school graduates	2.16	2.13
College graduates	2.05	2.00
Fertility elasticity by education		
High school dropouts	-0.22	-0.25
High school graduates	-0.17	-0.12
College graduates	-0.08	0.00
Cross-State Evidence: Regression Coefficients		
Fertility elasticity to avg. income	0.22–0.26	0.07
High-school graduation rate to fertility elasticity	0.03–0.07	0.03

Source: For the income-persistence estimates, we follow Heathcote et al. (2010). Regression coefficients in the data are shown in Tables 4 and 5.

Fertility by education groups: The estimated model also has fertility rates and fertility differentials by education groups that are in line with the data even though these moments were not targets. First, the average fertility rate is decreasing with education both in the model and in the data (second panel of Table 3). Second, to summarize the relation between income and fertility within education groups, we estimate the elasticity of fertility to income for different education groups. Let $\text{inc}_{e,q}$ and $\text{fert}_{e,q}$ be the mean income and fertility rate, respectively, of income quantile q for education group e (high school dropout, high school graduate, or college graduate). We estimate, both in the model and in the data, the following regression:

$$\ln(\text{fert}_{e,q}) = \alpha_e + \beta_e \ln(\text{inc}_{e,q}) + \epsilon_{e,t}, \tag{9}$$

where β_e is the elasticity of fertility to income for education group e.[25] A negative value implies that richer households tend to have fewer children, and values closer to zero imply that fertility rates are less related to income. In line with the data, the model generates elasticities equal to -0.25, -0.12, and 0 for high school dropouts, high school graduates, and college graduates, respectively. Both model and data display a decreasing (in absolute terms) relation with education, though the model displays a faster decreasing relation.

Cross-state variation: We exploit cross-state variation using US Census micro data to test whether the relation between fertility differentials and adult outcomes in the model is in line with the data. For each state and year we estimate the average income and the fertility elasticity to income and show that states with higher average income tend to have a fertility elasticity closer to zero; i.e., fertility is less related to income in richer states. We regress the fertility elasticities on the logarithm of the real average household income, controlling for state and time fixed effects. The regression specification is

$$\text{Fertility Elasticity}_{s,t} = \alpha + \gamma \ln(\text{Avg. Household Income}_{s,t}) + \eta_s + \mu_t + \epsilon_{s,t} \tag{10}$$

$$\ln(fert_{q,s,t}) = \alpha_{s,t} + \beta_{s,t} \ln(inc_{q,s,t}) + \epsilon_{q,s,t}, \tag{11}$$

where $\text{Fertility Elasticity}_{s,t}$ is equal to the estimated $\beta_{s,t}$ from (11). Table 4 shows that the elasticity of fertility is increasing in real average household income.[26] This implies that richer states are associated with smaller fertility differentials. This relationship seems stable and robust to controlling for state fixed effects and time fixed effects.

[25] See Appendix A for details on the definition of the variables as well as additional cross-state figures.

[26] Appendix A shows that these results are robust to using different measures of fertility rates.

Table 4: **How the elasticity of fertility to income changes with average income**.

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

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

We do a simple exercise to show that the model is qualitatively consistent with the cross-state patterns. Recall that in the benchmark calibration, the wage was normalized to one. Consequently, to generate economies with different levels of average household income, as in the data, we move wages such that the real wage (i.e., in consumption terms) is the main change. The size of wage movements is such that average income in the simulations is in the range of the corresponding empirical estimates. Many other things beyond income vary across states and time, so the main focus of this exercise is on the qualitative (rather than quantitative) features of the model.

This exercise involves moving the model away from the steady state to which it was estimated. We also have to adjust the school taste shocks since they enter separately in the utility function and do not scale with income. The school taste maximum values $\bar{\Phi}_2$ and $\bar{\Phi}_3$ are scaled by w^z. We set z such that when wage w is adjusted to match the differences in average income between 1960 and 1970, the average high school dropout rate of children born between those years changes as in the data. Most of our results, however, are robust to alternative adjustments.[27]

The bottom-right panel of Figure 3 shows the relation between fertility and income for economies with different w. In the baseline model, there is a negative relation between fertility and income with high-income households having 2 children. As the wage decreases, the income distribution shifts to the left, with more individuals having more than 2 children. Consequently, fertility

[27] For example, if we do not adjust school taste with wages (i.e., $z = 0$), results are almost unchanged.

differentials become larger as the income distribution shifts to the left. As in the data, the model displays a positive, though smaller, relation between fertility elasticity and average income.

Cross-state variation on fertility and education outcomes: How do education outcomes relate to fertility differentials? According to the forces described in the model, we expect states with larger fertility differentials to have lower average levels of education. Larger fertility differentials imply that poor families have relatively more children, which would lead to a larger share of children being born with scarce resources. Assuming this affects their education, we would then expect to observe lower average education in states with larger fertility differentials. To test this hypothesis, we focus on the education of individuals born in different states in 1960, 1970, and 1980. For clarity, we will look at how we study children born in 1960. First, we use data from 1960 to calculate the fertility elasticity, average household income, and income inequality in each state in their year of birth (i.e., in 1960). Second, we use data from when that generation is 30 years old (i.e., in 1990) to calculate high school graduation rates for each state.[28] Table 5 shows that individuals born in states with larger fertility differentials are associated with smaller high school graduation rates. This result is robust to controlling for mean household income and income inequality present in the state and year in which they were born, as well as for state or year fixed effects.[29] These results suggest that a one-standard-deviation change in the elasticity can explain approximately one-fifth of the standard deviation in high school graduation rates in the data.

The last row of Table 3 shows the relation between education outcomes and fertility differences across income groups. In the data, children born in states with smaller fertility differences between income groups are associated with higher high school graduation rates, with a regression coefficient between 0.03 and 0.07. The model is within this range, with a coefficient of 0.03. We take this as evidence that the model can also capture our main patterns of interest outside of the economy for which the benchmark is estimated.

[28] The choice of the age and year of birth is limited by the timing of the Census data.

[29] This evidence complements the findings in Kremer and Chen (2002), who show income inequality and fertility differentials across education groups are positively correlated across different countries.

Table 5: **How education relates to fertility elasticity**.

	(1)	(2)	(3)	(4)	(5)
Fertility Elasticity	0.0656**	0.0591**	0.0325***	0.0314***	0.0264**
	(0.0265)	(0.0243)	(0.0111)	(0.0107)	(0.0108)
Ln(Avg. Household Income)		0.101***	0.0904***	0.0803***	0.0952**
		(0.0123)	(0.00798)	(0.00848)	(0.0396)
Household Income Gini				-0.357*	-0.319*
				(0.178)	(0.177)
Observations	151	151	151	151	151
R-squared	0.030	0.355	0.902	0.909	0.910
# of States	51	51	51	51	51
State FE	NO	NO	YES	YES	YES
Year FE	NO	NO	NO	NO	YES

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

3 Sources of Intergenerational Mobility

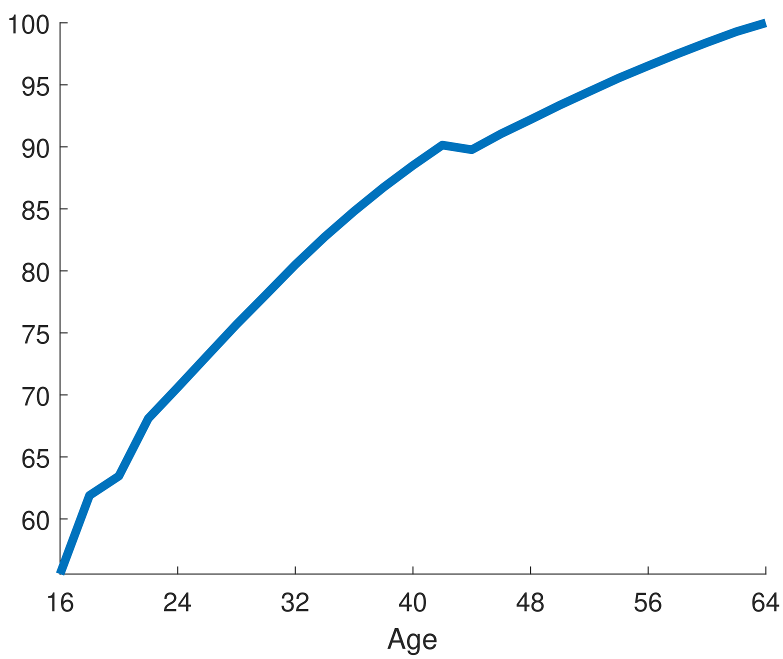
Initial conditions (i.e., initial human capital, school taste, and initial assets) explain 56% of the variation of lifetime earnings in the model. But what are the forces driving the importance of initial conditions? Our estimated model allows us to study the quantitative importance of both endogenous and exogenous sources of income inequality and intergenerational persistence.

We focus on the role of two endogenous family choices: fertility and transfers. In this section we show that an economy without fertility differentials would have 6% higher intergenerational mobility. The impact of a counterfactual constant transfer per child is twice as large. Both of these counterfactuals also lead to lower inequality. We then study the role of three exogenous forces (shocks): initial human capital, school taste, and adult income shocks. Among these, the stochastic process behind initial human capital is the most important determinant of mobility and inequality. Finally, we perform two simple policy experiments to understand policy implications and highlight that endogenous fertility may undo some of the intended policy objectives.

3.1 Variance of Lifetime Earnings

We decompose the variance of lifetime earnings into variation due to initial conditions and variation due to adult income shocks. An agent starts his life with an initial level of human capital (which is imperfectly correlated with the human capital of his parent); initial assets (which are the transfers' choice of the parents); and school taste (which is imperfectly correlated with parents' education). Most of the lifetime earnings inequality is explained by initial conditions in our model. The variance of log lifetime earnings is 0.35, with 56% being explained by the three initial conditions (column 1 of Table 6). Among these variables, the initial human capital is the most important and explains about 38% of the variation, while school taste explains about 16% and initial assets about 5%. After the agent becomes independent, the only sources of variation are the adult income shocks, which explain the remaining 44%. Even though our model choices are different, our results are similar to Huggett et al. (2011). Differently from them, however, we will be able to explore the role of family choices (fertility and transfers) in explaining inequality and intergenerational persistence.

Figure 4: **Variance of future earnings explained by current state**.



Note: For each age, we show the share of the variance of future earnings predicted by the current state variables.

We can decompose the variance of lifetime earnings for the conditions given at different ages (i.e., the state variables at older ages). To understand the decomposition, it is helpful to start at the end of the life cycle and iterate backward. In any period, the state variables determine current labor income. However, future labor income is subject to shocks. Therefore, in the last period of work, all shocks determining labor income have been realized and current state variables explain all (100% of) future earnings. One period before, the agent knows its current labor income, but its next (and last) period's income is subject to an idiosyncratic shock. Iterating backward toward the initial period, the agent faces more uncertainty about future labor income, and as a result, the current state provides less information about future earnings. Figure 4 shows that this decomposition increases with age and converges to 100%. The main takeaway from this figure is that, abstracting from initial conditions, shocks received between ages 20 and 40 seem to be the most important to predict future income.

3.2 Endogenous Forces: The Role of Fertility and Transfers

We start by analyzing the effects of endogenous choices of fertility and transfers on social mobility and income inequality.

The Role of Fertility To study the role of endogenous fertility, we keep the same model and estimated parameters from Section 2 but examine the case of constant fertility (assuming fertility n is exogenous and each household has 2.13 children, equal to the average fertility in the baseline model).

Without fertility differentials, intergenerational mobility would increase by 6% (the rank-rank correlation reduces from 0.298 to 0.281, Table 6). This change amounts to one-third of the standard deviation in intergenerational mobility across commuting zones (Chetty et al., 2014). Effects, however, are heterogeneous for children with different family backgrounds. To understand the heterogeneous effects, define p_{ij}^{s} as the probability of children of parents with income quintile i (at age 40) achieving income quintile j (at age 28) in the economy s, where s can refer

Table 6: **Intergenerational mobility and inequality: Endogenous forces**.

	Benchmark	**Constant fertility**	**Constant transfers**
Fertility and transfers			
Mean fertility	2.13	2.13	2.00
Fertility elasticity	-0.13	0.00	0.00
Mean transfer to children	28,129	24,808	28,129
CV transfers to children	1.02	1.02	0.00
Education			
Dropouts	8.2	6.5	5.9
High-school graduates	59.3	59.4	59.3
College graduates	32.5	34.1	34.7
Mobility			
Intergenerational Mobility: Rank-Rank	0.298	0.281	0.256
Transition: Parent Q1 and Child Q1 (%)	33.8	32.6	31.3
Intergenerational Education Persistence: Trace	0.864	0.874	0.874
Bottom parents (income Q1 & high-school dropouts)			
Children high-school dropouts	21.5	18.7	17.4
Children high-school graduates	68.2	71.0	72.3
Children college graduates	10.3	10.3	10.3
Inequality			
Variance of log Lifetime Earnings	0.347	0.345	0.345
% expl. by all initial conditions	55.6	55.0	54.7
% expl. by human capital	38.4	38.1	38.5
% expl. by transfers	5.4	5.7	0.0
% expl. by school taste	16.0	15.6	15.4
% expl. by adult income shocks	44.4	45.0	45.3
Total	100.0	100.0	100.0

Note: In the case of constant fertility, we keep the same model and estimated parameters from Section 2 but assume fertility n is exogenous and each household has 2.13 children, equal to the average fertility in the baseline model. In the case of constant transfers, we assume, instead, that parents' transfers are constant at the average value in the baseline model.

to the baseline economy or to some counterfactual (e.g., constant fertility).[30] The left panel of Figure 6 shows how the probability of being at the bottom quintile changes in the constant fertility counterfactual: $\Delta p_{i1} = p_{i1}^{\text{Constant Fertility}} - p_{i1}^{\text{Baseline}}$. For clarity, the first bar shows Δp_{11}, which is the change in the probability that a child with a parent in the bottom quintile remains in the bottom quintile. With constant fertility, children born to bottom-income-quintile parents exhibit a one percentage point lower probability of remaining in the bottom quintile (i.e., the probability falls from 33.8% to 32.6%). Children of top-quintile parents, however, have a larger probability of dropping to the bottom quintile (i.e., the probability increases from 6.6% to 7.1%).[31]

Children from low-income and less-educated parents (i.e., those of high school dropout parents in the first income decile) are the ones with most siblings and are, therefore, the most affected by the constant fertility counterfactual. In the baseline economy, 21.5% of these children are high school dropouts. When we force these families to have fewer children, parents choose to increase transfers, and the share of high school dropouts is reduced to 18.7% (second panel of Table 6). Even though effects are heterogeneous, the aggregate distribution of education shifts toward an economy with fewer high school dropouts and more high school graduates, decreasing income inequality and increasing intergenerational mobility. Hence, we find that reducing the negative income-fertility profile can alleviate intergenerational poverty traps.

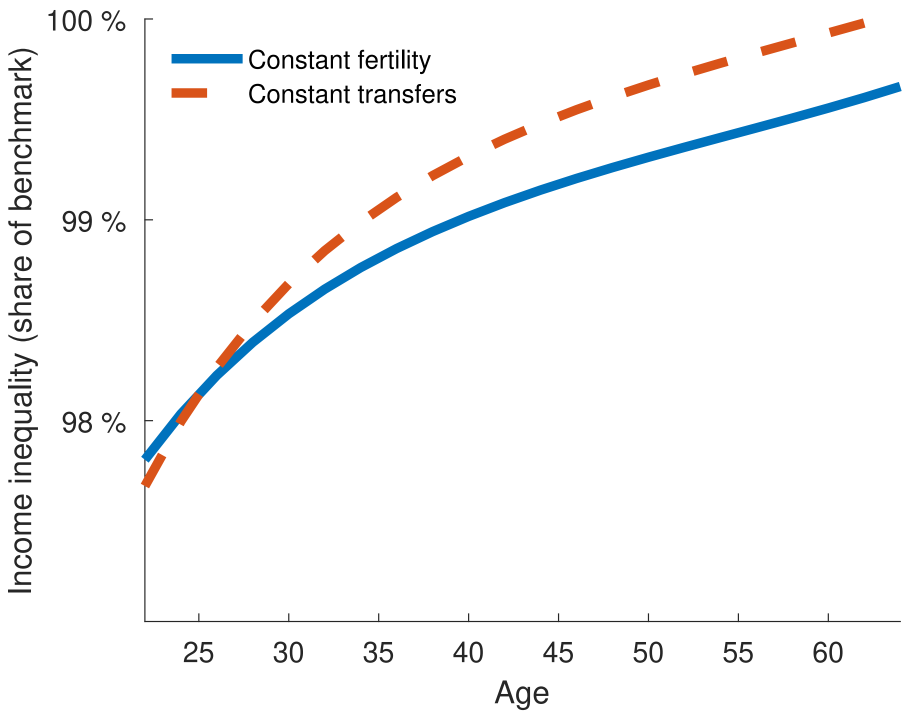
The effect of constant fertility on income inequality varies across the life cycle. The solid blue line of Figure 5 shows the variance of log income by age in an economy with constant fertility relative to the benchmark with endogenous fertility. For younger agents the effects are larger: Income inequality would decrease by 2% at age 25, about 1% at age 40, and about 0.5% at age 60. As agents grow older, the effect of initial conditions (which are more affected by endogenous fertility differentials) is diluted by life cycle income shocks. Fertility choices, thus,

[30] The age choice is the one used for the intergenerational mobility coefficient from the estimation, as calculated by Chetty et al. (2014).

[31] Using income ranks (rather than income levels) to measure mobility implies that when one group has a higher probability of achieving a certain rank, another group must have a lower probability of reaching such rank. In this case, most of the reduction in the probability of reaching the top quintile comes from children of high-income parents because this is the group that has a higher fertility rate in the counterfactual than in the baseline.

have a stronger effect on inequality among younger individuals. Overall, the variance of log lifetime earnings decreases by 0.6% (from 0.347 to 0.345, third panel of Table 6).

Figure 5: **Income Inequality and Family Choices**.



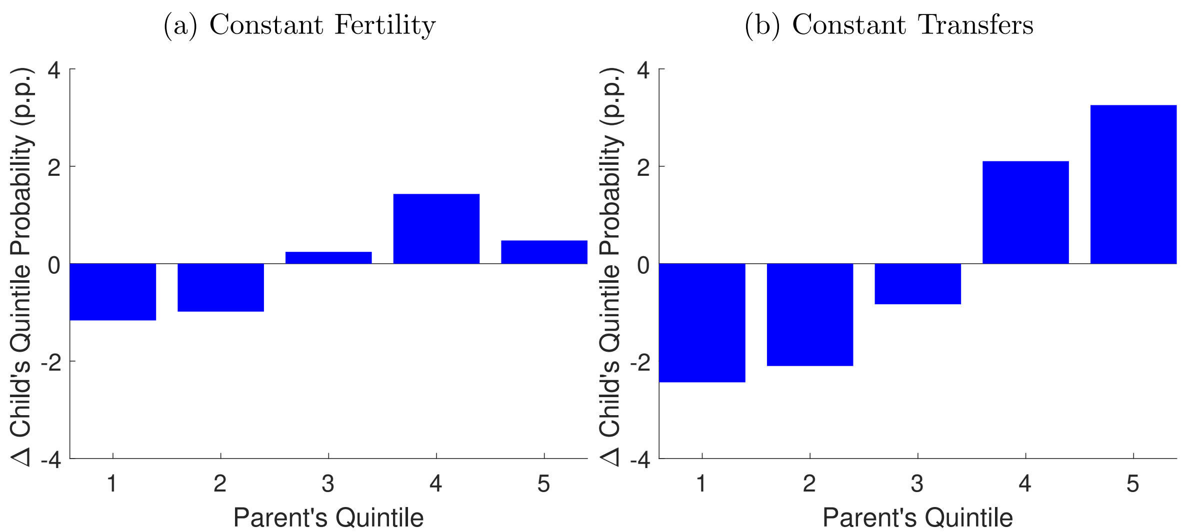
Note: Income inequality is measured using the variance of log income. For each alternative model, we plot the resulting inequality by age as a share of the benchmark economy; i.e., 98% implies a 2% reduction in inequality.

The Role of Family Transfers We evaluate an economy in which parents' transfers are exogenously constant at the average value in the baseline model. With constant transfers, intergenerational mobility increases by 14% (from 0.298 to 0.256). This change amounts to approximately two-thirds of the standard deviation in intergenerational mobility across commuting zones (Chetty et al., 2014). The large role that the heterogeneity of initial assets played in education choices is eliminated, and most lifetime earnings are instead due to characteristics less directly related to parents' income. The increase in mobility is associated with a reduction in the variance of years of education, mostly driven by a decrease in the share of high school dropouts.

Effects are also heterogeneous across children with different family backgrounds. The right panel of Figure 6 shows that children of bottom-income-quintile parents have a 2.5 percentage point lower probability of remaining in the bottom quintile (this probability is reduced from 33.7% to 31.3%). Children of top-income-quintile parents are more likely to fall to the bottom quintile (this probability increases from 6.6% to 9.8%). Table 6 shows that effects are particularly large for children of low-income and less-educated parents (the ones with the most siblings and, hence, the lowest amount of resources per child). The high school dropout rate for this group, for example, falls from 21.5% to 17.4%.

This counterfactual economy also displays lower income inequality. Similar to the case of constant fertility, the effect on income inequality varies across the life cycle and is stronger for younger individuals. The dotted red line of Figure 5 shows that the variance of log income by age in an economy with constant transfers is up to 2% below the benchmark model.

Figure 6: **Changes to Income Intergenerational Mobility**.



Note: Let p^s_{i1} be the probability of children of parents with income quintile i (at age 40) being in income quintile 1 (at age 28) in the economy s, where s can refer to the baseline economy or to some counterfactual (e.g., constant fertility). The left panel shows how these probabilities change in the constant fertility counterfactual: $\Delta p_{i1} = p_{i1}^{Constant\ Fertility} - p_{i1}^{Baseline}$. The right panel repeats the analysis for the case of constant transfers.

3.3 Exogenous forces

We evaluate how social mobility and income inequality are affected by the three exogenous shocks that take place across the agent's life cycle: initial human capital, school taste, and adult income shocks.

Initial human capital Columns 2–4 of Table 7 show economies under alternative scenarios for the initial human capital process described by Equation (8). We focus on three cases: *constant human capital* (all agents start with the same level of human capital regardless of the family background; i.e., $\rho = 0$ and $\sigma_{h_0} = 0$); *i.i.d.* (the initial draw of human capital is not correlated with parents, i.e., $\rho = 0$); and *no uncertainty* (the initial draw is perfectly correlated with parents; i.e., $\sigma_{h_0} = 0$).

When the initial human capital is not correlated with the parents (either constant or i.i.d.), intergenerational mobility increases substantially, the rank-rank coefficient is reduced by about 65% in both cases. The effect on income inequality depends on the specific counterfactual. With constant initial human capital, income inequality is reduced by over 30% because there is no variation in the initial human capital. In the i.i.d. case, however, income inequality is reduced by 7% because the initial human capital does not vary with parental background,s but the shock ε_{h_0} has a positive variance as in the baseline calibration.

When the initial human capital is not correlated with the parents, lower-income parents have more incentives to have more children since these children are more likely to be better off. Thus, mean fertility increases and fertility choices become more correlated with income (the fertility elasticity becomes more negative). This leads to lower transfers per child on average and a larger high school-dropout rate. This result highlights that studies of human capital policies that abstract from family transfers and fertility choices may be biased in their results.

An economy in which the initial level of human capital has no uncertainty but is perfectly correlated with parents' human capital displays lower inequality but higher intergenerational persistence. Income inequality decreases by about 23%, while the rank-rank correlation increases from 0.298 to 0.337. Mean fertility is slightly lower and the income elasticity reduces by one-half.

School taste Columns 5 and 6 of Table 7 evaluate alternative stochastic processes for the school taste shock as given by Equation (6). We focus on two cases: *i.i.d.* (not correlated with parents; i.e., $\omega = 1$) and *constant* (i.e., the same ϕ for all agents, fixed at the mean value). In both cases, intergenerational mobility increases and inequality decreases. The effects, however, are smaller than those given by the initial human capital process.

Adult income shocks We study the role of adult income risk by simulating an economy with no shocks to labor income (fifth column of Table 7). In this environment, inequality is reduced by 54% (the variance of log lifetime earnings decreases from 0.347 to 0.161). The intuition is very simple: Even if agents enter the labor market with similar conditions, income shocks spread them across the income distribution. The absence of labor income shocks, thus, reduces the variance of lifetime earnings.

Labor income shocks also make it possible for agents with disadvantaged initial conditions to earn similar levels of income as agents with better initial conditions. Consequently, we find that an economy without labor income shocks would also have lower intergenerational mobility (the rank-rank correlation increases from 0.298 to 0.354). Even though they are not directly comparable, the magnitude of the effects suggest that labor income shocks may have a large impact on both inequality and intergenerational mobility.

3.4 Policy Experiments

Our results suggest that policies that reduce fertility differentials or early inequality in resources can lead to higher intergenerational mobility. We evaluate two policies that directly target these margins, abstracting from any issues regarding financing. We take into account, however, that fertility and parental transfers will react to these policies. First, we consider a conditional

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

	Benchmark	Initial human capital			School taste		No adult
		constant	iid	no uncertainty	iid	constant	risk
Fertility and transfers							
Mean fertility	2.13	2.17	2.30	2.05	2.24	2.11	2.00
Fertility elasticity	-0.13	-0.20	-0.25	-0.06	-0.19	-0.11	-0.01
Mean transfer to children	28,129	25,355	27,112	24,742	27,243	29,003	19,575
CV transfers to children	1.02	1.13	1.15	0.99	1.11	1.03	0.66
Education							
Dropouts	8.2	23.9	25.7	4.9	21.5	0.0	1.2
High school graduates	59.3	53.1	51.8	59.4	51.4	100.0	64.9
College graduates	32.5	23.0	22.5	35.8	27.1	0.0	33.9
Mobility							
Intergenerational Mobility: Rank-Rank (100)	0.298	0.110	0.102	0.337	0.272	0.273	0.354
Transition: Parent Q1 and Child Q1 (%)	33.8	24.1	24.3	28.2	32.7	32.2	36.0
Intergenerational Education Persistence: Trace	0.864	0.859	0.857	0.865	0.985	1.000	0.897
Bottom parents (income Q1 & high-school dropouts)							
Children high-school dropouts	21.5	39.0	40.8	15.0	24.7	0.0	5.2
Children high-school graduates	68.2	51.9	49.8	75.3	50.8	100.0	84.5
Children college graduates	10.3	9.1	9.4	9.7	24.4	0.0	10.3
Inequality							
Variance of log Lifetime Earnings	0.347	0.236	0.323	0.267	0.345	0.258	0.161
% expl. by all initial conditions	55.6	35.1	54.0	40.3	56.5	39.7	100.0
% expl. by human capital	38.4	0.0	33.1	12.5	36.0	39.1	72.0
% expl. by transfers	5.4	0.3	0.5	8.6	3.7	3.9	7.6
% expl. by school taste	16.0	25.1	17.4	22.2	16.6	0.0	26.3
% expl. by adult income shocks	44.4	64.9	46.0	59.7	43.5	60.3	0.0
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Note: We evaluate how income inequality and social mobility are affected by the three exogenous shocks that take place across the agent's life-cycle: initial human capital, school taste, and adult income shocks. See main text for details.

transfer in which each family receives $20,000 only if they have two children. Second, we consider an unconditional transfer in which each household receives $20,000 as part of their initial assets. Table 8 shows the estimated results.

A conditional transfer program increases the incentives for households to have two children. The second column of Table 8 shows that almost all families have two children and that the elasticity of fertility to income is almost zero. Mean transfer to children increases slightly relative to the baseline given that poor households now have fewer children and more disposable income (due to the policy). The effect of this experiment is similar to the constant-fertility counterfactual of Table 6. Average education levels increase with this policy and intergenerational mobility increases by 4%.

An unconditional transfer, however, generates no changes on intergenerational mobility in our model. When the transfer is unconditional, poor households choose to have more children, which increases both the average fertility and the fertility elasticity. Consequently, average education levels are reduced and mobility is unaffected. This simple exercise highlights that fertility is not policy invariant and may undo the intended effect of equalizing initial assets.

Table 8: **Policy Experiments**.

	Benchmark	**Fertility Transfer**	**Initial Transfer**
Fertility and transfers			
Mean fertility	2.13	2.01	2.59
Fertility elasticity (q10)	-0.13	-0.01	-0.37
Mean transfer to children	28,129	28,727	16,900
Initial assets	28,129	28,727	36,900
CV transfers to children	1.02	0.93	1.79
Education and Mobility			
High-school Dropouts (%)	8.2	7.9	11.6
High-school Graduates (%)	59.3	58.5	58.7
College Graduates (%)	32.5	33.6	29.8
Intergenerational Education Persistence: Trace	0.864	0.868	0.870
Intergenerational Mobility: Rank-Rank	0.298	0.286	0.298

Note: In the case of fertility transfer, there is a conditional transfer of $20,000 for families with two children. In the case of initial transfer, we assume, instead, that the transfer is unconditional to all households when they become independent.

4 Conclusion

Doepke and Tertilt (2016) argue that there is a potentially large role for family economics within macroeconomics. Our results are consistent with this: Those interested in understanding inequality, intergenerational mobility, or inequality of opportunity may need to take fertility differentials and family transfers into account.

This paper analyzes the roots of social immobility and income inequality focusing on the impact of family choices. We use a standard heterogeneous agent life cycle model with idiosyncratic risk and incomplete markets extended to account for the role of families (through endogenous fertility, family transfers, and education) in determining initial opportunities. The model also allows for human capital transmission from parents to children. We propose that fertility differentials between rich and poor households can lead to substantial differences in the resources available for children, which can be important for their adult outcomes.

We find that both fertility and transfer differentials are important to understand intergenerational mobility. Fertility differentials account for 6% of the intergenerational mobility observed in the data, which is equivalent to approximately one-third of the standard deviation in intergenerational mobility across commuting zones. The impact of heterogeneity on parental transfers on intergenerational mobility is twice as large. Both fertility and transfer differentials increase inequality, particularly for young individuals. Even though fertility differentials play a smaller role relative to parental transfers in the US, this may not be the case for other countries with larger fertility differentials. According to our model, this implies that policies that reduce the incentives for poorer households to have children may be successful in reducing inequality and increasing social mobility in such countries.

Consistent with the early-childhood investment literature (Heckman et al., 2010; Gertler et al., 2014), we find that the stochastic process for the initial human capital (i.e., the correlation between parents' human capital and the variance of the process) is an important factor to understand income inequality and social mobility. Policies that are successful in increasing the resources available to all children earlier in life may reduce inequality and improve intergenerational mobility (e.g., Daruich, 2019). The objective of this paper, however, is to study the role of fertility and family transfers, so we model this stochastic process as an exogenous shock disciplined by the data. Our results suggest that policies that affect initial human capital may have sizable effects on fertility and transfer choices so studies of human capital policies that abstract from fertility and family transfer choices may be biased in their results. The effect of such policies in a general equilibrium model with endogenous fertility and early childhood development is left for future research.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (Forthcoming). Education policy and intergenerational transfers in equilibrium. *Journal of Political Economy*.

Agostinelli, F. and M. Wiswall (2016). Estimating the technology of children's skill formation. Technical report, National Bureau of Economic Research.

Almond, D., J. Currie, and V. Duque (2018). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature 56*(4), 1360–1446.

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1997). Parental altruism and inter vivos transfers: Theory and evidence. *Journal of Political Economy 105*(6), 1121–1166.

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

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

Belley, P. and L. Lochner (2007). The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital 1*(1), 37–89.

Benhabib, J., A. Bisin, and M. Luo (Forthcoming). Wealth distribution and social mobility in the us: A quantitative approach. *American Economic Review*.

Brooks-Gunn, J., G. J. Duncan, P. K. Klebanov, and N. Sealand (1993). Do neighborhoods influence child and adolescent development? *American journal of sociology 99*(2), 353–395.

Card, D. and A. B. Krueger (1992). Does school quality matter? returns to education and the characteristics of public schools in the united states. *Journal of political Economy 100*(1), 1–40.

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

Chetty, R. and N. Hendren (2018a). The impacts of neighborhoods on intergenerational mobility i: Childhood exposure effects. *The Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R. and N. Hendren (2018b). The impacts of neighborhoods on intergenerational mobility ii: County-level estimates. *The Quarterly Journal of Economics 133*(3), 1163–1228.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics 129*(4), 1553–1623.

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

Cunha, F., J. Heckman, and S. Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *oxford Economic papers 57*(2), 191–261.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, 05). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica 78*(3), 883–931.

Cutler, D. M. and E. L. Glaeser (1997). Are ghettos good or bad? *The Quarterly Journal of Economics 112*(3), 827–872.

Daruich, D. (2019). The macroeconomic consequences of early childhood development policies. Working paper.

de la Croix, D. and M. Doepke (2003, September). Inequality and Growth: Why Differential Fertility Matters. *American Economic Review 93*(4), 1091–1113.

de la Croix, D. and M. Doepke (2004). Public versus private education when differential fertility matters. *Journal of Development Economics 73*(2), 607–629.

de la Croix, D. and M. Doepke (2009). To segregate or to integrate: Education politics and democracy. *Review of Economic Studies 76*(2), 597–628.

Doepke, M. and M. Tertilt (2016). *Families in Macroeconomics*, Volume 2 of *Handbook of Macroeconomics*. Elsevier.

Fella, G. (2014, April). A Generalized Endogenous Grid Method for Non-smooth and Non-concave Problems. *Review of Economic Dynamics 17*(2), 329–344.

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

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

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

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

Gertler, P., J. J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, and S. Grantham-McGregor (2014). Labor market returns to an early childhood stimulation intervention in jamaica. *Science 344*(6187), 998–1001.

Gross, D. B. and N. S. Souleles (2002). Do liquidity constraints and interest rates matter for consumer behavior? evidence from credit card data*. *The Quarterly Journal of Economics 117*(1), 149–185.

Heathcote, J., F. Perri, and G. L. Violante (2010). Unequal we stand: An empirical analysis of economic inequality in the united states, 1967–2006. *Review of Economic dynamics 13*(1), 15–51.

Heathcote, J., K. Storesletten, and G. L. Violante (2010). The macroeconomic implications of rising wage inequality in the united states. *Journal of Political Economy 118*(4), 681–722.

Heckman, J. J., J. E. Humphries, and G. Veramendi (2018). Returns to education: The causal effects of education on earnings, health, and smoking. *Journal of Political Economy 126*(S1), S197–S246.

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

Heckman, J. J., S. H. Moon, R. Pinto, P. Savelyev, and A. Yavitz (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the highscope perry preschool program. *Quantitative Economics 1*(1), 1–46.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Herrington, C. M. (2015). Public education financing, earnings inequality, and intergenerational mobility. *Review of Economic Dynamics 18*(4), 822–842.

Holter, H. A. (2015). Accounting for cross-country differences in intergenerational earnings persistence: The impact of taxation and public education expenditure. *Quantitative Economics 6*(2), 385–428.

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

Jones, L. and A. Schoonbrodt (2016, October). Baby Busts and Baby Booms: The Fertility Response to Shocks in Dynastic Models. *Review of Economic Dynamics 22*, 157–178.

Jones, L. E. and A. Schoonbroodt (2010). Complements versus substitutes and trends in fertility choice in dynastic models. *International Economic Review 51*(3), 671–699.

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

Keane, M. P. and K. I. Wolpin (1997). The career decisions of young men. *Journal of Political Economy 105*(3), pp. 473–522.

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

Krueger, D. and A. Ludwig (2016). On the optimal provision of social insurance: Progressive taxation versus education subsidies in general equilibrium. *Journal of Monetary Economics 77*, 72 – 98.

Lee, S. Y. and A. Seshadri (2018). Economic policy and equality of opportunity. *The Economic Journal 128*(612), F114–F151.

Lee, S. Y. and A. Seshadri (2019). On the intergenerational transmission of economic status. *Journal of Political Economy 127*(2), 000–000.

Lochner, L. J. and A. Monge-Naranjo (2011, October). The nature of credit constraints and human capital. *American Economic Review 101*(6), 2487–2529.

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

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

Mincer, J. (1963). Opportunity costs and income effects. *Measurement in Economics*.

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

Moav, O. (2005). Cheap children and the persistence of poverty. *The Economic Journal 115*(500), 88–110.

Nardi, M. D. and F. Yang (2016). Wealth inequality, family background, and estate taxation. *Journal of Monetary Economics 77*, 130 – 145.

Restuccia, D. and C. Urrutia (2004). Intergenerational persistence of earnings: The role of early and college education. *The American Economic Review 94*(5), pp. 1354–1378.

Roys, N. and A. Seshadri (2017). On the Origin and Causes of Economic Growth. Working paper.

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

Smets, F. and R. Wouters (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review 97*(3), 586–606.

Sylva, K., A. Stein, P. Leach, J. Barnes, L.-E. Malmberg, et al. (2007). Family and child factors related to the use of non-maternal infant care: An english study. *Early Childhood Research Quarterly 22*(1), 118–136.

Todd, P. E. and K. I. Wolpin (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal 113*(485), F3–F33.

Todd, P. E. and K. I. Wolpin (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human capital 1*(1), 91–136.

Yum, M. (2018). Parental time investment and intergenerational mobility. Working paper.

Appendix

A Empirical Findings: Fertility and Income

In the calibration and validation of the model, we use several moments related to fertility. In this appendix we describe the definition of these statistics and show additional figures.

A.1 Total Fertility Rate and Fertility Elasticity

The model focus on the decisions made by individual households, so we would like a measure of fertility decisions at the household level. Probably the closest measure to this is Children Ever Born (CEB), available from the US Census. This variable asks women 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 40 years ago could be used. Notice also that choosing the upper end of the age that determines the sample can bring issues. For example, if we used women up to any age we might get biased measures of fertility if this is correlated with mortality risk. Last but not least, this variable has even been dropped from the US Census after 1990. Hence, we use an alternative measure of fertility for our main analysis but use CEB to evaluate the robustness of our results.

For the sake of clarity let us introduce the most basic measure of fertility, the Crude Birth Rate (CBR), which is defined as the ratio of births to women alive in one year. A typical issue with the CBR is that it can be too low because a big share of women who have already passed child-bearing age but still bring 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 the child-bearing age, TFR is not affected. Formally, let $f_{a,s,t}$ be the number of children born to women age a in region s and period t divided by the number of women age a in region s and period t. Assume that the child-bearing age extends between ages a_L and a_H.[32] Then the TFR in region s and period t, $\text{TFR}_{s,t}$, is defined as

$$\text{TFR}_{s,t} = \sum_{a=a_L}^{a=a_H} f_{a,s,t}. \tag{12}$$

Typically these age-specific fertility rates are constructed for age bands in increments of 5 years and then summed, with the limits of the sum being $a_L = 15$ and $a_H = 49$.[33] Relative to CEB, the main benefit is that it does not require the data to report how many children each woman has had. Instead, it needs only children under the age of one to be associated with their mothers within the household—a much more standard requirement. Moreover, TFR does not require women to have passed child-bearing age as it focuses on fertility rates, which are not associated with a particular cohort but with 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 used widely in the literature (Kremer and Chen, 2002; Manuelli and Seshadri, 2009).[34]

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

$$\text{TFR}_{q,s,t} = \sum_{a=15}^{a=49} f_{a,q,s,t}. \tag{13}$$

[32]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.

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

[34]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, however, we do our empirical work based on the TFR measure of fertility.

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, because occupations change their relative importance over time (e.g., Autor et al., 2008). Moreover, there is a substantial variation in income across people within a given occupation.[35] Hence, we focus on annual total household income in the year of the sample. To get the appropriate quantile groups, we cannot compare the income level of young and old households, because, 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.[36] This way the TFR for each quantile-region-year can be estimated.

To estimate the fertility elasticity in (9) we define $\text{inc}_{q,s,t}$ and $fert_{q,s,t}$ as the mean income and fertility rate, respectively, of income quantile q in region s and year t.[37] We allow for region identifier s for our cross-state analysis. We estimate

$$\ln\left(fert_{q,s,t}\right) = \alpha_{s,t} + \beta_{s,t}\ln\left(inc_{q,s,t}\right) + \epsilon_{q,s,t}, \tag{14}$$

where $\beta_{s,t}$ will be referred to as the *elasticity of fertility to income* for region s in year t. If this value is negative, richer households tend to have fewer children. Values closer to zero imply that fertility rates are not related to income (at least, according to this specification). Given our main interest on fertility decisions as a function of income, we do not want to mix single-parent households with two-parent households. Hence, we limit our analysis to "marital fertility"; i.e., the fertility of those women who, when they answer the survey, indicate that they are married.

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

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

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

Sample Selection For each year of the US Census, we start with all women belonging to the main family of each household and with non-missing family income. We drop women outside of the "age of fertility"; i.e., 15 to 49 years old. Then, we restrict our attention to those who are either heads of households or spouses of heads of households and report as married. Finally, we drop those who report as in school or whose annual household income (in 2000 US$) is less than $4,000. Each entry of Table A.1 shows the number of women after each selection procedure, in the corresponding year.

Table A.1: **Sample Selection.**

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

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

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

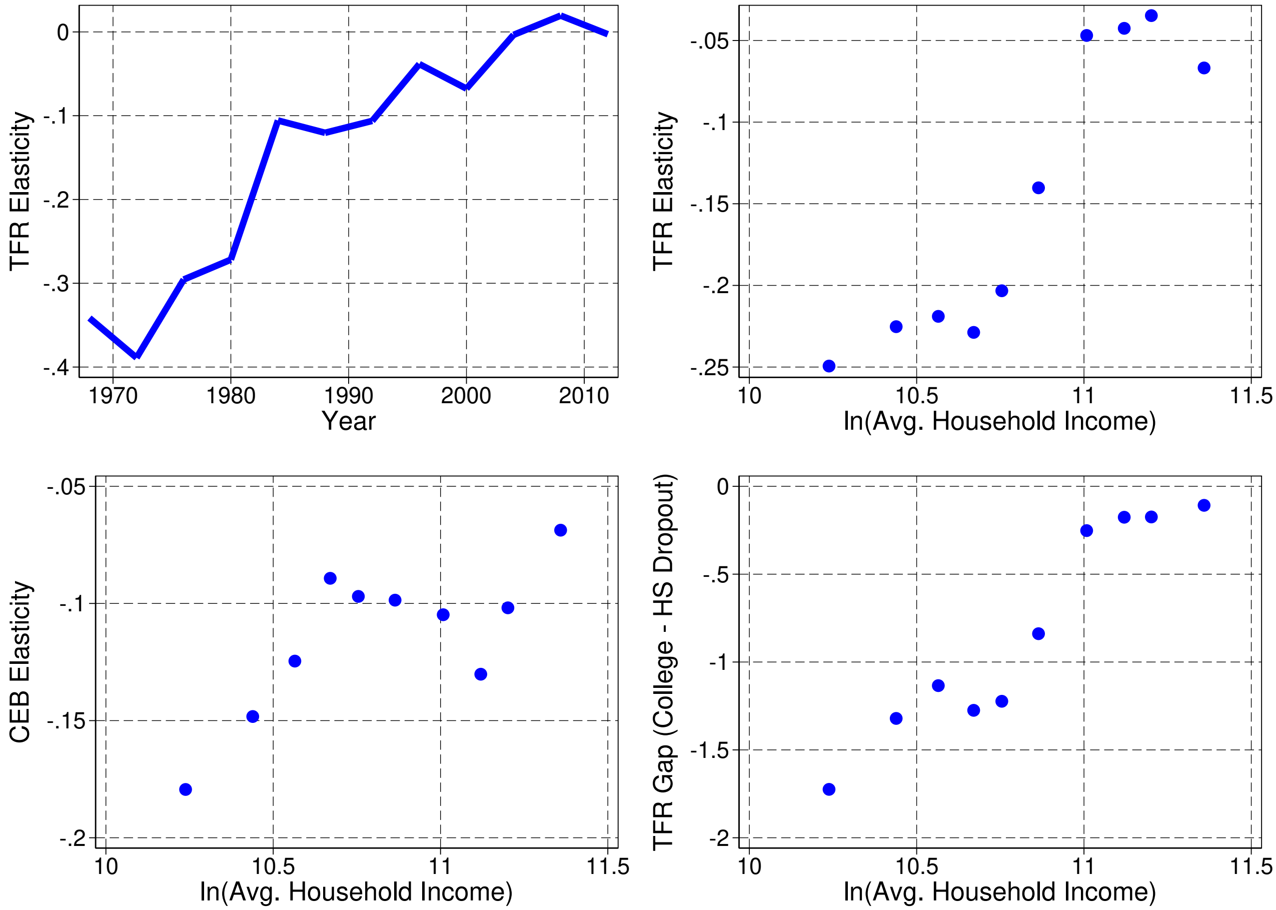
A.2 Fertility and Income

In the validation Table 4, we exploit cross-state variation and report a negative relation between income and fertility. In this appendix we show additional figures to confirm the negative relation between fertility elasticity and average income. First, the top-left panel of Figure A.1 shows the evolution of the elasticity of fertility to income for the US, with its value on the vertical axis. The figure not only confirms that fertility elasticity has been negative since 1968, but also suggests that it has decreased over time, implying that the difference in the number of children between poor and rich households has become smaller.

To better understand what is behind this pattern, we extend our analysis to exploit the cross-state variation using US Census micro data from IPUMS; i.e., for each state s and year t we estimate (14). In the main text, Table 4 performs formal statistical tests. Here, we analyze the data visually. First, we combine all our observations and divide them into deciles according to their levels of real average household income. For each of these groups we calculate the mean household income and fertility elasticity. The top-right panel of Figure A.1 shows that richer states tend to have elasticities closer to zero or, in other words, smaller fertility differentials.

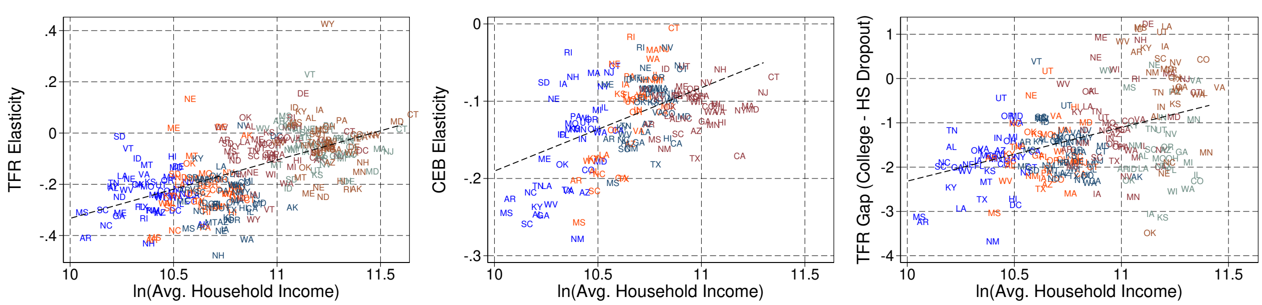
The pattern of higher fertility differentials being associated with lower average household income is robust to alternative measures of fertility. First, instead of using TFR, we use Children Ever Born (CEB) to compute the fertility elasticity as in (14). The bottom-left panel of Figure A.1 reports the results. Second, we use the fertility differences between education groups: We calculate the difference between the TFR of women married to college-graduate men and that of women married to high school dropout men. The bottom-right panel of Figure A.1 reports the results. Qualitative results in both cases are similar to those using the TFR elasticity to income: Richer states are associated with smaller fertility differentials. Figure A.2 reports all the observations used in Figure A.1.

Figure A.1: **Fertility Differentials and Income**.



Note: The top-left panel uses CPS data for years 1968-2013 and observations are grouped in 3-year windows. The rest of the panels use Census data for years 1960, 1970, 1980, 1990, 2000 (TFR elasticity and TFR Gap), and 2010 (TFR elasticity and TFR Gap). We divide observations into deciles according to average household income. For each decile, we calculate the mean level of household income as well as the mean fertility differential measure.

Figure A.2: **Elasticity of fertility to income and GDP: All observations**.



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

B Estimation Details

B.1 Income Process

We follow the estimation procedure of Abbott et al. (Forthcoming) but focus on labor earnings of two-adult households. To overcome the problem that the NLSY provides observations only for young workers (up to age 50 at the time of writing), we use income data from the PSID to estimate age polynomials for different education groups. After removing the age profile, we then estimate the persistent income process using NLSY79 data.

B.1.1 Age Profile

We estimate the age profile by education groups using the PSID data from Heathcote et al. (2010). We start with 6,134 households with heads between 24 and 63 years old. After focusing on households with at least 8 income observations and who do not report extreme changes of income (i.e., annual growth above 400%, or reduction by 66%), we are left with 2,508 households. When we split this sample into 3 education groups, we get a high school dropout sample of 349 households, a high school graduate sample of 1,425 households, and a college graduate sample of 734 households.

Quadratic age polynomials are separately estimated for each education group. Table B.1 reports the deterministic age profile for each education group. In line with Abbott et al. (Forthcoming), we find that the higher the level of education, the steeper the profile.

Table B.1: **Income Age-Profile**

	High-school dropouts	**High-school graduates**	**College graduates**
Age	0.051	0.067	0.122
$\text{Age}^2 \times 1000$	-5.522	-7.312	-13.147

B.1.2 Persistent Shocks

We follow the procedure of Abbott et al. (Forthcoming) but use household income and 2-year periods instead. After we use the age profiles (from PSID to avoid issues regarding the limited observations of older individuals in NLSY) to filter out age effects, we use NLSY79 data to estimate the persistent income process. NLSY79 is particularly useful because it allows us to control for measures (AFQT89) of persistent skills, as those given by h_0 in our model.

The sample selection procedure is as follows. We start with 12,686 individuals, with a total of 317,150 observations. We exclude observations in the army and restrict to those between the ages of 24 and 63. This reduces the number of individuals to 12,685 (272,918 observations). We drop observations with top-coded earnings and drop individuals who change education groups (after age 24) or who have missing information on their AFQT score. This reduces the number of individuals to 11,153 (238,753). We further restrict observations to those with positive hours of labor in the household (but lower than 10,000 annually). We also drop individuals who at least once report hourly wages under half the minimum wage or above \$400. We keep individuals with at least 8 observations of income. This reduces the number of individuals to 7,947 (129,619 observations). After focusing on two-adult households and grouping observations in 2 year periods (like the model), we eliminate observations with wages above \$400, as well as households that report extreme changes of income (i.e., above 400% or reduction by 66%). This leads to a number of 6,580 households with a total 53,968 observations. When we split this sample into 3 education groups, we get a high school dropout sample of 843 households, a high school graduate sample of 4,127 households, and a college graduate sample of 1,610 households.

Income residuals $u_{i,j,e,t}$ are obtained by purging the age component and controlling for AFQT89 from NLSY data. As in Heathcote et al. (2010), we then model the residual $u_{i,j,e,t}$ as the sum of two independent components:

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

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

$$z_{i,j,e,t} = \rho z, e z_{i,j-1,e,t-1} + \zeta_{i,j,e,t}$$
$$\zeta_{i,j,e,t} \sim N(0, \sigma_{\zeta,e}),$$

and $m_{i,j,e,t} \sim N(0, \sigma_{m,e})$ is measurement error (and noise from the point of view of the model). The initial draw is $z_{i,0,e,t} \sim N(0, \sigma_{z0,e})$. We estimate this independently for each education group using a Minimum Distance Estimator, with covariances of wage residuals at various lags for different age groups as moments. Estimates are reported in Table B.2

Table B.2: **Income Process**

	High-school dropouts	**High-school graduates**	**College graduates**
Persistence ($\rho_{z,e}$)	0.88	0.96	0.97
Variance income shocks ($\sigma_{\zeta,e}$)	0.06	0.02	0.03
Initial dispersion ($\sigma_{z_0,e}$)	0.23	0.10	0.07
Variance measurement error ($\sigma_{m,e}$)	0.18	0.17	0.12

B.2 Replacement Benefits: US Social Security System

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

The pension formula is given by

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

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

C Sensitivity of the Results

Table C.1 shows that the main quantitative results of the paper—the increase in intergenerational mobility with constant fertility or transfers—are robust to the estimated parameters. In the baseline calibration, intergenerational mobility increases by 5.7% and 13.8% under constant-fertility and constant-transfer counterfactuals, respectively.

The second panel of Table C.1 shows the sensitivity of the results to the estimated parameters. In each row we increase the corresponding parameter by 5% while keeping the other parameters at the calibrated values. We then calculate the new steady-state and evaluate the effect of the constant-fertility and constant-transfer counterfactuals (as in the baseline). We report the difference in the effects of each alternative calibration relative to the baseline. Values close to zero imply that the effect is similar to the baseline calibration. For example, the row regarding γ says that when γ is 5% higher than in the baseline, the effect of fertility differentials is 1.37% larger than in the baseline (i.e., 5.72% + 1.37% = 7.09%, instead of 5.72%). Similarly, with higher γ, the effect of transfers would be 1.31% smaller. Overall, the effect of constant fertility or transfers on intergenerational mobility are quantitatively robust to changes of these magnitudes in the parameters. More interesting, however, may be to study the direction of the changes. The first row shows that an increase of γ generates a larger effect of constant fertility on intergenerational mobility. With higher γ, the model displays a larger fertility differential, which amplifies the effect of the constant fertility counterfactual.

Table C.1: **Robustness**

	Fertility elasticity	**Effect on Persistence**	
		Constant Fertility	**Constant Transfers**
Baseline	-0.127	-5.72	-13.83
Sensitivity to internally estimated parameters		Relative to Baseline	
Altruism curvature (γ)	-0.205	-1.37	1.31
Altruism level (λ)	-0.244	-1.05	0.13
Child cost level (c_1)	-0.067	2.54	1.98
Child cost curvature (c_2)	-0.147	-1.60	-0.91
Persistence initial draw h_0 (ρ)	-0.125	1.41	1.19
Variance initial draw h_0 (σ_{h0})	-0.132	-0.45	-0.15
Education return high-school level (α_{HS})	-0.124	0.41	-0.62
Education return college level (α_{Coll})	-0.127	-0.73	-0.65
Education return high-school curvature (β_{HS})	-0.132	-0.61	1.00
Education return college curvature (β_{Coll})	-0.130	-0.32	-0.18
School taste correlation (ω)	-0.124	1.18	1.56
School taste high-school (Φ_2)	-0.135	0.83	2.10
School taste college (Φ_3)	-0.131	-0.75	1.75
Sensitivity to external calibration			
Interest rate wedge (ι)	-0.132	1.03	1.75
Discount factor (β)	-0.135	0.14	-2.26

Note: The first panel shows the fertility elasticity and the effects (in percentage points) of constant fertility and transfers on the rank-rank coefficient for the baseline calibration. The second and third panel show, for alternative calibrations, the fertility elasticity and the change in the rank-rank coefficient due to either constant fertility or transfers.

The third panel studies the sensitivity and robustness of the results to some parameters that we calibrated externally. We first studied the role of the interest rate wedge ι since this affects how much agents are willing to borrow (so the importance of parental transfers can change). A lower borrowing wedge should make parental transfers less important for intergenerational mobility. To evaluate this we used our same estimated parameters but reduced ι to 5% (from 10%) and repeated our main experiments from Table 6. Table C.1 confirms that the effect of fertility differentials and transfers on intergenerational persistence is reduced, to 5% (from 6%)

and 12% (14%), respectively. In the low-wedge economy, more households are borrowing (7.7% instead of 6.2%). We believe that, most likely, this explains the reduced important of family background for mobility. The effect on mobility, however, remains close to the baseline result.

We also studied the role of the discount factor β since this would affect how much parents are willing to spend in consumption rather than parental transfers. We focus on an increase in the discount factor from 0.975 to 0.99. This change, however, leads to large effects in the baseline steady state (fertility in particular increases significantly as parents care more about future generations) so we needed to recalibrate the model. After re-estimating the model, the results, shown in the last row of Table C.1, are qualitatively and quantitatively similar to the baseline estimation. There is a slightly larger effect in the constant-transfer counterfactual which is, most likely, driven by the fact that parental transfers are more heterogenous with a larger β: constrained, poor parents still transfer very little (or nothing), while richer parents transfer more than in the baseline (because they care more about the future of their children). The effect, however, of constant transfers on persistence is only 2 percentage point larger than in the baseline.

D Childcare cost function

In this appendix we show that the functional form for the childcare cost is the outcome of a cost-minimization problem in which parents choose how much time spend at home to take care of their children and how much they buy on the market (e.g., by hiring nannies or sending the children to childcare).

Assume a childcare production function $y = x_1^{\alpha} x_2^{1-\alpha}$. The first input is the parental time at home which cost is the opportunity of not going to work $p_1 = (1-\tau)\, wh$. The second input is the amount buy on the market at price $p_2 = w\chi$, where χ is the level of human capital of nannies hired in the market. A family with n children has to produce An^{γ} units of childcare.

The solution of the cost minimization problem to produce y units is

$$C\left(y\right)=y\theta p_{1}^{\alpha}p_{2}^{1-\alpha}$$

where θ is a non-linear function of α. Hence, for a family with n children and human capital h the childcare cost is

$$C\left(h,n\right)=An^{\gamma}\theta\left(\left(1-\tau\right)wh\right)^{\alpha}\left(w\chi\right)^{1-\alpha}$$
$$C\left(h,n\right)=c_{1}wh^{c_{2}}n^{c_{3}}$$

with $c_1 = A\theta\left(1-\tau\right)^{\alpha}\left(\chi\right)^{1-\alpha}$, $c_2=\alpha$, and $c_3=\gamma$, which coincides with the reduced-form functional form assumed on the main text.