Child Care Subsidies
with One- and Two-Parent Families

Emily Moschini *

JOB MARKET PAPER

November 13, 2018

CLICK HERE FOR LATEST VERSION

Abstract

The implementation of child care subsidies has varied widely across countries and states, as well as over time, ranging from universal to poverty-tested eligibility. I study the implications of eligibility rules for child care subsidies in a general equilibrium, overlapping generations framework with parental investments in child skill. I allow for one- and two-parent families and endogenize family formation. This explicitly incorporates single mothers, who currently parent 20% of children under five in the United States. Using individual-level data from the US Department of Education, I estimate how mother time, father time, and non-parental child care affect child skill for each family structure. These estimates allow me to account for the differential effect of child care subsidies on one- and two-parent families. I find that targeting child care subsidies to one-parent families yields welfare gains that are 1.5 and 2 percentage points higher than universal and poverty-tested subsidies, respectively. Single mother subsidies ultimately support families with low incomes who are sensitive to the price of non-parental child care, but without requiring higher taxes, as in universal subsidies, or introducing a disincentive to invest in skill, as in poverty-tested subsidies.

1 Introduction

Large-scale subsidized child care programs have been implemented around the world.[1] These large-scale programs are built on the success of smaller programs, which provided enriched early childhood environments for children from poor families and have shown potentially major gains from public intervention.[2] Empirical evaluations of large-scale child care subsidies have uncovered heterogeneous treatment effects along dimensions of family income, family structure, and child gender, with conclusions differing by the country or state whose program is being evaluated.[3] Designs of these programs vary, but eligibility is usually means-tested or universal.

In this paper, I pursue a structural macroeconomic approach to analyze large-scale child care subsidies in the United States. The approach allows me to transcend local confounding factors, acknowledge general equilibrium effects on wages, government budgets, and marriage, and to evaluate policy designs that have not been already implemented. Specifically, I examine child care subsidies under three eligibility rules: family-tested (targeted to single mothers), poverty-tested (targeted to the poor), and universal. To do this, I construct a general equilibrium overlapping generations model with families that are heterogeneous in income and marital status. My model takes into account endogenous family formation and altruistic parental investment in child skill (Becker and Tomes (1979)). The three eligibility rules I examine permit a meaningful comparison between a targeting rule a model like mine can evaluate and more commonly implemented policies. In my analysis, I focus on the distribution of welfare gains across gender, skill, marital status, and age.

Comparing these subsidies in general equilibrium allows me to incorporate adjustments in labor income taxes, wages, and marriage decisions in response to policy. This feedback can magnify or mitigate the gains from the subsidy. For example, increases in skill, if accompanied by much higher labor income taxes to fund those increases, may not raise disposable earnings, consumption, or welfare for some groups. This is the case under the universal subsidy. Sim-

---

[1]Examples include: Oklahoma (started in 1998), Quebec (1997), Argentina (1993), Norway (1975), and Denmark (1964).

[2]Two well-known examples are the Abecedarian Program and the Perry Preschool Program. See Anderson (2008), Garcia, Heckman, Leaf, and Prados (2016), and Baker (2011).

[3]Studies of large-scale child care subsidies include Baker, Gruber, and Milligan (2008), Gupta and Simonsen (2010), Berlinski, Galiani, and Gertler (2009), Havnes and Mogstad (2014), Kottelenberg and Lehrer (2017), Blau and Currie (2006).

ilarly, endogenizing wages allows changes in the aggregate stock of skill to affect its rate of return, while endogenizing the family formation decision allows policy to affect the mass and composition of one- and two-parent families raising children in the economy. In fact, I find that the marriage rate decreases under high levels of the child care subsidy under all three eligibility rules. The family formation channel in my model is similar to that of Abbott, Gallipolli, Meghir, and Violante (2018), who examines how college tuition subsidies can affect the composition of marriages, but without allowing for single-parent families to exist in equilibrium. The baseline framework into which I incorporate and endogenize heterogeneous family structures has been used to analyze the interaction of policy and skill investment in many studies, including Restuccia and Urrutia (2004), Lochner and Monge-Naranjo (2011), Guner, Kaygusuz, and Ventura (2016), Gayle, Limor, and Soytas (2017), Daruich (2017), Caucutt and Lochner (2017), Abbott, Gallipolli, Meghir, and Violante (2018), and Lee and Seshadri (2018).

The subsidies I evaluate generate welfare gains by partially addressing the market failures of the environment I construct. This model has three main sources of uninsurable risk: the family a child is born into, the potential spouse an adult meets in the marriage market, and the fiscal externality. The risk of being born into a low-resource family is uninsurable for individuals, because parents and children cannot contract with one another (Cunha and Heckman (2007)). The government, however, can use labor income taxes to recoup the costs of subsidizing investment in children's skill. With regard to marriage market risk, child care subsidies can partly insure individuals against parenting alone or with a low-skill spouse by lowering the cost of being a parent. Even parents who do not receive the child care subsidy benefit from the insurance it provides for the marriage risk of their children. The final missing market is the fiscal externality that arises from an endogenous labor income tax. Taxpayers cannot use individual contracts to encourage others to invest in their children, so that the labor income tax can decrease for everyone. A child care subsidy partially addresses this missing market by lowering the cost of investment in child skill, using funds contributed from everyone in the economy.

In my environment, one- and two-parent families invest in their children using distinct technologies. I estimate these functions using a nationally representative panel dataset from the US Department of Education—the Early Childhood Longitudinal Study, Birth Cohort (ECLS-

B). The ECLS-B is designed to be representative of the entire population of parents in the United States with 9-month-old children in 2001, and follows over 10,000 children from 9 months to kindergarten entry, recording their skill at 9 months, 2 years, 4 years, and 5 years of age. In addition to providing a rich source of information on the evolution of skill in young children, this survey is unique in its emphasis on the role of fathers in child development. It is especially useful because it records parenting activities for both members of a couple over time for a large subset of families in the first three waves of the survey. For almost 1,000 couples over the first three waves of this survey, one can observe the couple's hourly wages, their individual time investments in their children, and their use of non-parental child care. For families with single mothers, the analogous sample contains about 500 families.

The panel nature of the ECLS-B allows me to control for fixed effects within a family when implementing the estimation of skill accumulation parameters. This accounts for unobserved and time-invariant heterogeneity in parenting productivity, child attributes, and local environment. My model specification for two-parent families is related to the skill accumulation technology in Del Boca, Flinn, and Wiswall (2014), who use the Panel Study of Income Dynamics, Child Development Supplement (PSID CDS) to estimate a Cobb-Douglas technology for couples with one or two children with separate time inputs from the mother and the father. Like them, I find that fathers play a non-negligible role in the raising of young children. I provide a second estimation for single mothers, which allows a new comparison of the ways that one- and two-parent families interact with non-parental care. I find that the input composition of investments is not very sensitive to price changes, and that single mothers rely more on non-parental child care for investment than two-parent families do. This makes the cost of investment for single mothers more sensitive to subsidies to the price of non-parental care.

Using my framework to compare family-tested, poverty-tested, and universal subsidies, I find that ex ante welfare for families in the baseline economy increases by 7.7%, with an 85% subsidy targeted to single mothers. Universal subsidies can only reach a level of 6.2% welfare gains ex ante (at an 80% subsidy level), and poverty-tested subsidies provide a 5.7% gain (at an 80% subsidy level). The welfare gains with a family-tested or poverty-tested subsidy are more evenly spread out over age, marital status, gender, and skill groups. Family-tested subsidies, however, have higher levels of gains across most of these groups. Under a universal subsidy,

4

the old are worse off because the costs of the program exceed the gains in the stock of skill (which means that labor income taxes rise). Family-tested subsidies perform best because they do not increase the tax burden, like universal subsidies, nor introduce a disincentive to invest in skill like poverty-tested subsidies.

This paper provides three main contributions. First, I incorporate both one- and two-parent families into an overlapping generations framework to analyze subsidies to investment in children's skill. Second, parallel to this heterogeneity in family structures, I introduce heterogeneity in the technologies that parents use to invest in their children, and provide estimates of these technologies. Finally, I allow the population to endogenously sort into the two structures via a marriage market, so that policy can affect the mass and composition of parents in one- and two-parent families.

The rest of the paper proceeds as follows. In section 2, I lay out the model. Section 3 presents the model parameterization, including the estimation of the skill accumulation technologies. Section 4 reports the subsidy design and level that maximizes ex ante welfare and interprets the distribution of welfare gains under each eligibility rule. Section 5 concludes.

2   The Model

There are four sets of agents in the economy: consumers, a representative firm, the government, and a non-parental care provider. Consumers are grouped into families with either one or two parents. Given prices for labor and capital, the firm chooses labor and capital inputs to maximize profits subject to a free-entry condition. The government chooses labor income taxes to finance lump-sum transfers and non-parental care subsidies. A child support system exists, enforced by the government, where single fathers contribute a lump-sum amount that is redistributed lump-sum and equally to all single mothers. Finally, the non-parental care sector supplies child care at the amount demanded in equilbrium, at a price equal to some fraction of the average hourly wage.

The Life Cycle of Consumers

Each individual lives for four periods of equal length: childhood, parenthood, adult worker, and old worker. During childhood, an individual makes no decisions: she is a passive recipient of consumption and investment chosen by her family. Upon independence, the individual leaves with her skill to start the parenting phase as an independent decision maker.

At the beginning of the parenting phase, before any decisions are made, everyone participates in a marriage market. This market is modelled as a random search with an arrival rate of one: a potential match is drawn from the skill distribution of the other gender in the same generation. Once assigned a potential spouse, the agent compares the expected present discount value of parenting alone or in a couple. The expectation is taken over the initial skill of the child, which is unknown when the agent makes her marriage decision. A marriage requires that both the husband and wife accept the match (the two individuals remain single otherwise). After the marriage market, single mothers and couples draw the initial skill of their two children, which is the same for both children (single mothers and couples raise a son and a daughter). Given their beliefs about how skill affects lifetime utility, parents then choose the level and composition of investment in their children's skills. At the end of the parenting phase, the children leave their parents' state space, and parents enjoy an altruistic return from their lifetime utility.

The problems of the adult worker and the old worker periods differ only in that old workers die at the end of their phase, so in the last period of the old worker phase there is no savings choice. Before old age there are no borrowing constraints. To summarize, the shocks in life are your own initial skill, your gender, the family that raises you, your potential spouse, and the draw of your children's initial skill. Figure 1 illustrates the timing of the phases and the draws of these shocks.

Figure 1: Life Cycle of the Consumer

6
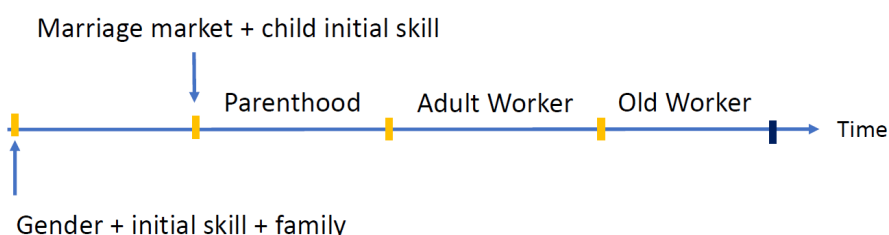
One- and Two-Parent Families

The two family structures are different in that they have different efficiencies of consumption ($\{\phi_s\}$ and $\{\phi_{mc,p}, \phi_{mc,w}\}$ reflected in equivalence scales) and different marginal utilities of leisure (which will be reflected in the parameters of the period utility functions (u_s, u_{mc}), whose functional forms are defined in the model parameterization section). Married couples compromise on their leisure decision in the sense that it has to be the same for both members of the couple. They also use a skill accumulation technology specific to two-parent families to invest in their children (f_{mc}). Similarly, single mothers invest in children using a skill technology specific to them (f_s).

Family Problems

The solution to a family's life-cycle problem is a set of policy functions and value functions. These functions are defined for young adults (before the marriage decision) with skill θ and gender $g \in [1, 2]$ where 1 is father and 2 is mother, and for one and two-parent families at each period $j \in \{1, 2, 3\}$ (with the numbers corresponding to parenthood, adult worker, and old worker, respectively), state z, and the skill of the single parent θ or the skills of the father f and mother m of the couple (θ_f, θ_m). The state space z always contains assets, which are zero at the start of life, and during parenthood also contains the initial skill of the child. The value functions are:

- For young adults: $V_0(\theta, g)$
- For single mothers: $V_j^{SM}(z; \theta)$
- For single fathers: $V_j^{SF}(z; \theta)$
- For married couples: $V_j^{MC}(z; \theta_f, \theta_m)$

The Marriage Decision and Probability of Getting Married

The policy function that maps from the type of the spouse to a yes/no marriage market decision ($d_{mm}(\theta_f, \theta_m, g) \in \{0, 1\}$, given the skill θ_m or θ_f and gender g of the decision maker), is a threshold strategy in the skill of the potential spouse. It solves:

$$d_{mm}\left(\theta_f, \theta_m, 1\right) = \underset{\delta \in \{0,1\}}{\arg\max} \left[\delta \int_{\theta_c} V_1^{MC}\left(z; \theta_f, \theta_m\right) \pi\left(\theta_c\right) d\theta_c + (1-\delta) \int_{\theta_c} V_1^{SF}\left(z; \theta_f\right) \pi\left(\theta_c\right) d\theta_c\right] \quad (1)$$

$$d_{mm}\left(\theta_f, \theta_m, 2\right) = \underset{\delta \in \{0,1\}}{\arg\max} \left[\delta \int_{\theta_c} V_1^{MC}\left(z; \theta_f, \theta_m\right) \pi\left(\theta_c\right) d\theta_c + (1-\delta) \int_{\theta_c} V_1^{SM}\left(z; \theta_m\right) \pi\left(\theta_c\right) d\theta_c\right]$$

where $z = \{0, \theta_c\}$ is the state space in the first period of life which reflects the zero wealth all consumers start with and the initial skill of the child θ_c, and the distribution of initial child skills has a probability density function given by $\pi\left(\theta_c\right)$.

Draws for Marriage Offer, Partner Type, and Child Skill

The expected lifetime utility of a child enters into the parent problem. It is an expectation over outcomes in the marriage market and draws of initial child skill when they become a parent, as a function of their skill and gender g:

$$V_0\left(\theta_f, 1\right) = \int_{\theta_m} \left[\mathbb{I}_d \int_{\theta_c} V_1^{MC}\left(z; \theta_f, \theta_m\right) \pi\left(\theta_c\right) d\theta_c + (1-\mathbb{I}_d) \int_{\theta_c} V_1^{SF}\left(z; \theta_f\right) \pi\left(\theta_c\right) d\theta_c\right] \mu\left(\theta_m, 2\right) d\theta_m \quad (2)$$

$$V_0\left(\theta_m, 2\right) = \int_{\theta_f} \left[\mathbb{I}_d \int_{\theta_c} V_1^{MC}\left(z; \theta_f, \theta_m\right) \pi\left(\theta_c\right) d\theta_c + (1-\mathbb{I}_d) \int_{\theta_c} V_1^{SM}\left(z; \theta_m\right) \pi\left(\theta_c\right) d\theta_c\right] \mu\left(\theta_f, 1\right) d\theta_f$$

where $\mathbb{I}_d \equiv d_{mm}\left(\theta_f, \theta_m, 1\right) \times d_{mm}\left(\theta_f, \theta_m, 2\right)$ indicates a mutual acceptance of the match, $z = \{0, \theta_c\}$ and θ_c is the child's initial skill, drawn independently from an exogenous distribution. In addition, $\mu\left(\theta, g\right)$ is the endogenous distribution over adult skill θ and gender g. In the following subsections, I define the life-cycle problem for each of the three family types: single mother, single father, and married parents. In all the family problems, τ_y is the labor income tax, τ_n is the child care subsidy, T are lump-sum transfers, and T_{cs} are child support payments.

Single Mother Problem

During parenthood, a single parent chooses consumption c, savings a', non-parental care time n, her own time investments in her child q, and leisure ℓ to solve the following problem:

$$V_1^{SM}(a, \theta_c; \theta) = \max_{c,a',n,q,\ell} u_s\left(\frac{c}{\phi_s}, \ell\right) + \beta V_2^{SM}\left(a'; \theta\right) + b\sum_g V_0\left(\theta'_c, g\right) \quad (3)$$

$$c + a' + (1 - \tau_n)\, p_n n \leq (1 + r)\, a + (1 - \tau_y) w\theta\, (1 - \ell - q) + T + T_{cs}$$

$$\ell, n, q \in [0, 1] \qquad \ell + q \leq 1 \quad n + q \leq 1$$

$$\theta'_c = f_s\left(\theta_c, n, \theta q\right)$$

Here, and in what follows, ϕ_s is the consumption-equivalence scale for single-parent families with children, β is the discount factor (patience), b is the altruism parameter, w denotes the wage rate, and p_n denotes the price of non-parental care. During the parenting phase, parental and non-parental time investments in child skill affect the child's skill in the next period according to the single parent production technology. The final child skill enters the objective function of the mother through an altruism term $b\sum_g V_0\left(\theta'_c, g\right)$, which weights the expected lifetime utility of the child $V_0\left(\theta'_c, g\right)$ with the altruism coefficient b. Throughout their childhood, the mother cannot direct investments separately to the son and daughter. Her children are born with the same skill and receive the same investments. For a regression motivating this assumption, see the appendix.

During the adult worker phase, the single mother consumes c, saves a', and enjoys leisure ℓ:

$$V_2^{SM}(a; \theta) = \max_{c,a',\ell} u_s\left(c, \ell\right) + \beta V_3^{SM}\left(a'; \theta\right) \quad (4)$$

$$c + a' \leq (1 + r)\, a + (1 - \tau_y) w\theta\, (1 - \ell) + T$$

$$\ell \in [0, 1]$$

During the old worker phase, the single mother consumes c, and enjoys leisure ℓ. She dies with zero savings:

$$V_3^{SM}(a; \theta) = \max_{c,\ell} u_s\left(c, \ell\right) \quad (5)$$

$$c \leq (1 + r)\, a + (1 - \tau_y) w\theta\, (1 - \ell) + T$$

$$\ell \in [0, 1]$$

9

Single Father Problem

During parenthood, a single father chooses consumption c, savings a', and leisure ℓ to solve the following problem. They also have to pay child support T_{cs} which is received by single mothers:

$$
\begin{aligned}
V_1^{SF}(a;\theta) &= \max_{a',c,\ell} u_s(c,\ell) + \beta V_2^{SF}\left(a';\theta\right) + \zeta(\mu) \qquad (6)\\
\ell &\in [0,1]\\
c + a' + T_{cs} &\leq (1+r)a + (1-\tau_y)w\theta(1-\ell) + T
\end{aligned}
$$

Single fathers do not internalize their child's outcome directly; rather, their payoff includes some function of the aggregate endogenous state $\zeta(\mu)$, which I will specificy in the parameterization section. This specification allows policy to affect single fathers through the distribution of skill μ. This specification also does not allow a single father to use his own time to invest in his children. I start from the assumption that living with your children allows you to invest more time in them. Here, I take that assumption to an extreme and abstract from single father time investments entirely.

During the adult worker phase, single fathers consume c, save a', and enjoy leisure ℓ by solving the following problem:

$$
\begin{aligned}
V_2^{SF}(a;\theta) &= \max_{c,a',\ell} u_s(c,\ell) + \beta V_3^{SF}\left(a';\theta\right) \qquad (7)\\
c + a' &= (1+r)a + (1-\tau_y)w\theta(1-\ell) + T\\
\ell &\in [0,1]
\end{aligned}
$$

During the old worker phase, the single father consumes c and enjoys leisure ℓ. He dies with zero savings:

$$
\begin{aligned}
V_3^{SM}(a;\theta) &= \max_{c,\ell} u_s(c,\ell) \qquad (8)\\
c &\leq (1+r)a + (1-\tau_y)w\theta(1-\ell) + T\\
\ell &\in [0,1]
\end{aligned}
$$

Married Couple Problem

During the parenting phase, married couples choose consumption c, savings a', leisure ℓ, non-parental care time n and parental time inputs $\{q_f, q_m\}$. Like single mothers, married couples cannot direct investment by child gender and enjoy an altruistic return from the lifetime utility of their children as a function of their skill at adulthood.

$$
V_1^{MC}\left(a, \theta_c; \theta_f, \theta_m\right) = \max_{c, a', \ell, n, q_f, q_m} u_{mc}\left(\frac{c}{\phi_{mc,p}}, \ell\right) + \beta V_2^{MC}\left(a', \theta_c'; \theta_f, \theta_m\right) + b \sum_{g_{child}} V_0\left(\theta_c, g\right)
$$

$$
c + a' + (1 - \tau_n)\, p_n n \leq (1 + r)\, a + (1 - \tau_y) w\left[\theta_f\left(1 - \ell - q_f\right) + \theta_m\left(1 - \ell - q_m\right)\right] + T \qquad (9)
$$

$$
\ell, n, q_f, q_m \in [0, 1] \qquad \ell + q_f \leq 1 \quad \ell + q_m \leq 1 \quad n + q_f + q_m \leq 1
$$

$$
\theta_c' = f_{mc}\left(\theta_c, n, \theta_f q_f, \theta_m q_m\right)
$$

Here, $\phi_{mc,p}$ is the consumption-equivalence scale of a family with two adults and two children. During the adult worker phase, married couples consume c, save a', and enjoy leisure ℓ:

$$
V_2^{MC}\left(a; \theta_f, \theta_m\right) = \max_{c, a', \ell} u_{mc}\left(\frac{c}{\phi_{mc,w}}, \ell\right) + \beta V_3^{MC}\left(a'; \theta_f, \theta_m\right) \qquad (10)
$$

$$
c + a' \leq (1 + r)\, a + (1 - \tau_y) w\left[\theta_f\left(1 - \ell - q_f\right) + \theta_m\left(1 - \ell - q_m\right)\right] + T
$$

$$
\ell \in [0, 1]
$$

Here, $\phi_{mc,w}$ is the consumption-equivalence scale of a family with two adults and two children. During the old worker phase married couples consume c and enjoys leisure ℓ. They die with zero savings:

$$
V_3^{MC}\left(a; \theta_f, \theta_m\right) = \max_{c, \ell} u_{mc}\left(\frac{c}{\phi_{mc,w}}, \ell\right) \qquad (11)
$$

$$
c \leq (1 + r)\, a + (1 - \tau_y) w\left[\theta_f\left(1 - \ell - q_f\right) + \theta_m\left(1 - \ell - q_m\right)\right] + T
$$

$$
\ell \in [0, 1]
$$

Government

The government collects revenue from labor income taxes τ_y to finance lump-sum transfers T and non-parental care subsidies τ_n. The variable H is the aggregate supply of labor efficiency units, and N is the aggregate demand for non-parental care:

$$\tau_y w H = T + \tau_n p_n N \tag{12}$$

Representative Firm

The firm chooses capital and labor inputs K_F, H_F to maximize profits, taking prices r, w as given. The variable δ_F is the depreciation rate of capital.

$$\max_{K,H} \quad \left\{ K_F^{\alpha_F} H_F^{1-\alpha_F} - w H_F - (r + \delta_F) K_F \right\} \tag{13}$$

Non-parental Care Sector

The non-parental care sector provides N units of non-parental care at price p_n. The price of non-parental care is set as a constant fraction κ of the average earnings per unit of time:

$$p_n = \kappa \sum_g \int_\theta w \theta \mu (\theta, g) \, d\theta \tag{14}$$

This allows the price of non-parental child care to adjust with the average level of skill in the economy without specifying a production function for non-parental care.

Equilibrium

Given a government policy $\{\tau_n\}$, transfers T, and child support T_{cs}, a stationary equilibrium is defined as:

- Prices $\{r, w\}$ and tax τ_y,

- Individual marriage decision rules for each type and potential spouse,

- Policy functions for each family type for each period of life j and state z,

- Value functions at the beginning of adulthood V_0 and for each family type V^{SM}, V^{SF}, V^{MC},

- A non-parental care price p_n,

- Joint distribution μ over adult skill and gender,

such that:

- Capital and labor markets clear.

- The government balances its budget taking prices as given.

- Family decision rules solve their dynamic problems taking prices and taxes as given.

- The non-parental care price p_n is a constant fraction κ of average hourly earnings.

- The joint distribution μ is stationary.

Market Incompleteness in the Environment

The sources of market incompleteness in this model are:

- The inability of children to insure against the family shock, and the inability of parents to insure against the initial skill of their children.

- The inability of children and their parents to perfectly insure against the marriage market outcome.

- The inability of parents to subsidize other parents' investment decisions.

Because parents can borrow against their own future income, that credit constraint is not a source of market incompleteness. Rather, because it is not possible for anyone (neither the children, nor their parents) to borrow against the future income of the child in order to finance investment in their skill, investment is constrained by the resources available to those currently parenting rather than those who will parent in the future. The government's instruments are labor income taxes, a subsidy for non-parental care, and lump-sum transfers that are equal for everyone. The combination of labor income taxes and non-parental care subsidies offers a way to partially complete the missing markets described above: the government can tax the future earnings of children in order to fund investment in them through subsidized non-parental care (which in turn makes any level of investment in child skill cheaper).

3   Model Parameterization

In this section, I explain how I grouped the model parameters into those chosen externally, those estimated outside the model, and those calibrated inside the model.

3.1 Functional Form Assumptions

Period Utility Functions

Utility functions are defined separately for one- and two-parent families:

$$u_{type}(c, \ell) = \log(c) + \psi_{type} \log(\ell), \quad type \in \{s, mc\} \tag{15}$$

Single Father Altruism Term

I set the single father term $\zeta(\mu)$ equal to the minimum possible outcome in the economy:

$$\zeta(\mu) = \min_{\theta, g} V_0(\theta, g) \tag{16}$$

Reasonable values for child support transfers are not sufficient to reach the targeted marriage rate in the internal calibration without further penalizing the single father path. With this specification, I still make the single father lifetime utility depend directly on the skill distribution in the economy, as with the other parenting problems. By comparison, using the average outcome does not sufficiently penalize the single father.

Distribution of Initial Child Skill

I assume that initial child skill is drawn i.i.d. from $\pi(\theta_c)$, which I set as a uniform distribution. For motivating regressions for the i.i.d. assumption, see the appendix.

Skill Accumulation

I specify the skill accumulation technologies to be nested constant elasticity of substitution (CES) functions with inputs of non-parental time, mother time, and father time (which aggregate to investment I_{type}) and between investment and the current stock of skill (which aggregate to tomorrow's skill). Specifically, the functional forms of the dynamic equation for skill accumulation, and how investment is generated from time inputs for the two family structures, are as follows:

$$f_{type}(\cdot) = \left[\upsilon \left(\lambda I_{type} \right)^\xi + (1 - \upsilon) \theta_c^\xi \right]^{\frac{1}{\xi}}, \quad type \in \{s, mc\} \tag{17}$$

$$I_s = \left[\gamma_s \left(\theta_m q_m \right)^{\eta_s} + (1 - \gamma_s) \left(n \right)^{\eta_s} \right]^{\frac{1}{\eta_s}} \tag{18}$$

$$I_{mc} = \left[(1 - \alpha) \left(\gamma \left(\theta_m q_m \right)^\eta + (1 - \gamma) \left(n \right)^\eta \right)^{\frac{\rho}{\eta}} + \alpha \left(\theta_f q_f \right)^\rho \right]^{\frac{1}{\rho}} \tag{19}$$

The parameters γ, γ_s and α control the relative level of inputs given a price ratio. The values

of η, η_s, and ρ control the percentage change in the ratio of inputs for a percentage change in

the ratio of their prices. With this parameterization, I allow for mother time and non-parental

care time to interact differently across household structures (i.e., I do not impose $\gamma = \gamma_s$ and

$\eta = \eta_s$).

3.2 Externally Chosen Parameters

The length of a lifetime and of each phase are proportional to 20 years of childhood and 60

years of adulthood (death at age 80). The discount (patience) factor is set to a yearly value of

0.96 to match the risk-free interest rate. The capital share of the production technology is set

to a standard value of 0.33, and the depreciation rate of capital is chosen to be 0. Finally, the

consumption equivalence (CE) scales are set using the 1994 scales from the Organisation for

Economic Co-operation and Development (OECD). These scales assign a value of 1 for the first

adult, and 0.5 for the subsequent adults; for each dependent the weight is 0.3. They adjust

money spent on consumption into units of consumption for each member of the household.

Once children leave the family, the equivalence scale for single mothers goes back to 1, and the

scale for couples falls to 1.5. This is summarized in Table 1.

Table 1: Externally Chosen Parameters

Symbol	Name	Value
β	Patience	0.96^{20}
$\{\alpha_F, \delta_F\}$	Production Technology	$\{0.33, 0\}$
$\{\phi_s\}$	OECD CE Scales: 1 adult, 2 children	$\{1.6\}$
$\{\phi_{mc,p}, \phi_{mc,w}\}$	OECD CE Scales: 2 adult, 2 children	$\{2.1, 1.5\}$

3.3 Externally Estimated Skill Accumulation Technologies

In this section, I explain how I estimate the parameters of the skill technology from the ECLS-B dataset. The panel nature of the data allows me to use a fixed-effects estimator to control for heterogeneous productivity shifters on parental time inputs. The equations that I use to estimate the skill accumulation technology parameters are derived from the first order conditions of the parenting problem with respect to the quality time of the father and mother, and non-parental care time. I assume that the aggregator of maternal time and non-parental time is the same functional form for both single mothers and married couples. Married couples differ from single mothers because of the time contribution of the father, which is modelled as an outer aggregator of father time and maternal/non-parental time. After taking first order conditions of the parenting problem with respect to the time investment choices, I take ratios of those equations and then logs. The result is a system of four linear equations. After deriving these equations with model notation, I explain how the objects in my model map into variables in the ECLS-B and motivate my use of the fixed effects estimator. I then provide the estimation equations in terms of observed variables in the ECLS-B. My derivation of the estimation equations is similar to the method of Lee and Seshadri (2018).

3.3.1 Estimation Equations

The equations for the investment aggregators of singles and couples are given by (18) and (19). The partial derivatives of these equations enter the first order conditions of the parenting problem with respect to q, n, for single mothers and q_m, q_f, n for married couples. After taking ratios and then logs of these first order conditions, I get the following linear equations, one for single mothers and two for married couples. It is apparent that the parenting productivities can be separated linearly from the terms of interest.

$$\ln\left(\frac{q}{n}\right) = \left(\frac{1}{\eta_s - 1}\right)\ln\left[\frac{w\theta}{(1 - \tau_n)\,p_n}\right] + \left(\frac{1}{\eta_s - 1}\right)\ln\left[\frac{1 - \gamma_s}{\gamma_s}\right] - \left(\frac{\eta_s}{\eta_s - 1}\right)\ln\left(\theta\right) \quad (20)$$

$$\ln\left(\frac{q_m}{n}\right) = \left(\frac{1}{\eta - 1}\right)\ln\left[\frac{w\theta_m}{(1 - \tau_n)\,p_n}\right] + \left(\frac{1}{\eta - 1}\right)\ln\left[\frac{1 - \gamma}{\gamma}\right] - \left(\frac{\eta}{\eta - 1}\right)\ln\left(\theta_m\right) \quad (21)$$

$$\ln\left(\frac{q_f}{n}\right) = \left(\frac{1}{\rho - 1}\right)\ln\left[\frac{w\theta_f}{(1 - \tau_n)\,p_n}\right] + \left(\frac{\eta - \rho}{\eta\,(\rho - 1)}\right)\ln\left(\frac{q_m}{n}\frac{w\theta_m}{(1 - \tau_n)\,p_n} + 1\right) \quad (22)$$

$$+ \left(\frac{1}{\rho - 1}\right)\ln\left[\frac{(1 - \alpha)}{\alpha\,(1 - \gamma)^{\frac{\rho}{\eta}}}\right] - \left(\frac{\rho}{\rho - 1}\right)\ln\left(\theta_f\right)$$

In the first equation q, n are the quality time and non-parental child care time chosen by single mothers, respectively, $w\theta$ is the hourly wage of single mothers, p_n is the price of non-parental child care, and θ is the parenting productivity of single mothers. In the second and third equations, q_m, q_f, and n are the quality time inputs of mothers and fathers and the non-parental child care time chosen by married couples, $w\theta_m$ is the hourly wage of the mother, $w\theta_f$ is the hourly wage of the father, p_n is the price of non-parental child care time, and $\theta, \theta_m, \theta_f$ are the parenting productivites of single mothers, married mothers, and married fathers, respectively.

To derive estimation equations for the skill aggregator equation (17), I impose intertemporal cost minimization, similar to the approach of Lee and Seshadri (2018). This means that I assume parents set the ratio of the marginal costs of investment (the prices of investment in each period, as derived above) equal to the ratio of marginal productivities of investment in each period, appropriately discounting the costs using interest rates. The result of the following steps will be an equation that expresses the ratio of expenditures over two consecutive periods on the left-hand side, and the ratio of investment prices over two consecutive periods on the right-hand side. First, I explain how I derive the price of skill investment for single and married couples. Second, I construct the equation I will use to estimate parameters υ and ξ in equation (17).

The total expenditures on investment, $X_{type,j}$, is the sum of foregone earnings and non-parental care costs for the time invested in the child. For each family type, the expression for this cost is:

$$X_s = (1 - \tau_n) p_n n + w\theta q_m \tag{23}$$

$$X_{mc} = (1 - \tau_n) p_n n + w\theta_f q_f + w\theta_m q_m \tag{24}$$

Investment expenditures purchase investment input into skill accumulation at a price Λ_{type}:

$$X_{type} = \Lambda_{type} I_{type} \tag{25}$$

An analytical expression for the price of investment Λ_{type} can be derived by taking the ratio of X_{type} and I_{type}. The solution will be a function of parental skill and the prices of inputs, which are values the parent takes as given in this model.

For couples, the price of investment Λ_{mc} is:

$$\Lambda_{mc} = \frac{(1 - \tau_n) p_n + w\theta_f \frac{1}{\theta_f} \Psi_{2,mc} + w\theta_m \frac{1}{\theta_m} \Psi_{1,mc}}{\Psi_{4,mc}} \tag{26}$$

where

$$\Psi_{1,mc} \equiv \left[\frac{w\theta_m}{(1 - \tau_n) p_n} \frac{(1 - \gamma)}{\gamma} \frac{1}{\theta_m} \right]^{\frac{1}{\eta - 1}}$$

$$\Psi_{2,mc} \equiv \left[\frac{1}{\theta_f} \frac{(1 - \alpha)(1 - \gamma)}{\alpha} \frac{w\theta_f}{(1 - \tau_n) p_n} (\Psi_{3,mc})^{\rho - \eta} \right]^{\frac{1}{\rho - 1}}$$

$$\Psi_{3,mc} \equiv (\gamma (\Psi_{1,mc})^{\eta} + 1 - \gamma)^{\frac{1}{\eta}}$$

$$\Psi_{4,mc} \equiv [\alpha (\Psi_{2,mc})^{\rho} + (1 - \alpha)(\Psi_{3,mc})^{\rho}]^{\frac{1}{\rho}}$$

For single mothers, the price of investment Λ_s is:

$$\Lambda_s = \frac{(1 - \tau_n) p_n + w\theta \frac{1}{\theta} \Psi_{1,s}}{\Psi_{2,s}} \tag{27}$$

where

$$\Psi_{1,s} = \left[\frac{1}{\theta} \frac{(1 - \gamma_s) w\theta}{\gamma_s (1 - \tau_n) p_n} \right]^{\frac{1}{\eta_s - 1}}$$

$$\Psi_{2,s} \equiv (\gamma_s (\Psi_{1,s})^{\eta_s} + 1 - \gamma_s)^{\frac{1}{\eta_s}}$$

The assumption of intertemporal cost minimization implies that for two consecutive observations in the data $t - 1$ and t, the ratio of the marginal prices of investment equals the ratio of marginal productivities of that investment. Notice that here I am allowing childhood to have multiple periods as in the data, while in my model I assume it is a single period. This is why

I have shifted from using j to index age in the model to using t to reference the period in the data. Now, for the objects θ_c, X_{type}, I_{type}, and Λ_{type}, I add a time subscript t or $t-1$.

$$\frac{\left(\frac{1}{1+r_t}\right)\Lambda_{type,t}}{\left(\frac{1}{1+r_{t-1}}\right)\Lambda_{type,t-1}} = \frac{\frac{\partial \theta_{c,t+1}}{\partial I_{type,t}}}{\frac{\partial \theta_{j+1}}{\partial I_{type,j-1}}} \qquad (28)$$

where $type \in \{s, mc\}$. Substituting $\frac{X_{type,t}}{\Lambda_{type,t}}$ for $I_{type,t}$ and taking logs of both sides of the above equation gives a linear equation with the ratio of investment expenses on the left-hand side and the ratio of prices on the right-hand side:

$$\ln\left(\frac{X_{type,t}}{X_{type,t-1}}\right) = \frac{1}{\xi-1}\ln[1-\upsilon] + \frac{\xi}{\xi-1}\ln\left(\frac{\Lambda_{type,t}}{\Lambda_{type,t-1}}\right) + \frac{1}{\xi-1}\ln\left[\frac{1+r_{t-1}}{1+r_t}\right] \qquad (29)$$

To control for the last term, which is a ratio of interest rates across periods, I assume a yearly interest rate of 7% (the lending rate to households in 2001) and adjust the values $1 + r_t$ in each period t to be proportional to its distance from 0, the period when the child is born. This means that at age 9 months, the first wave of the survey, the interest rate is $\frac{3}{4} \times (1.07)$, while in the second wave of the survey (when the child is 2 years old) this value is $2 \times (1.07)$. Notice that this last equation does not contain an additive constant that is cancelled out by a fixed effects estimator. Note also that I did not cancel parenting and labor market productivities in the definitions of Λ_{type}, in order to make the mapping from the model to the data more straightforward in the next section. When constructing prices, I recover the parenting productivity from the fixed effects estimations used to find the values of η, γ, η_s, and γ_s. I use those family-level values to build the price of investment at the family level in each period t.

Mapping Model Objects to Objects in the Data   In the data section below, I explain in detail the procedure I use to prepare the ECLS-B data for estimation. After applying the procedure I describe there, I observe the following attributes of family i in wave t: hourly wages $\tilde{w}_{i,f,t}, \tilde{w}_{i,m,t}$ for father f or mother m, hourly price of child care $\tilde{p}_{i,t}$ (which differs across families and across waves), quantity of child care purchased $\tilde{n}_{i,t}$ in units of hours per week, and hours of quality time contributed by the father f and the mother m: $\tilde{q}_{i,f,t}, \tilde{q}_{i,m,t}$. Here I also put the m subscripts on the single mother variables in the data, which is a departure from the model notation. For couples, I estimate η and γ first, and then predict the ratio of mother

time and non-parental child care time at the family level (incorporating the fixed effect levels). This is then substituted into the right-hand side of the estimation equation for ρ, α with father time and non-parental child care time on the left-hand side. For the last estimation equation, I construct investment prices $\tilde{\Lambda}_{type,i,t}$ and expenditures $\tilde{X}_{type,i,t}$ for each family in each wave by using the analogous variables from the data in equations (26) and (27). Individual-level parenting productivities θ, θ_m, and θ_f are not observed in the ECLS-B. I assume that these are time-invariant but can vary across individuals. I denote these unobserved parenting productivities with $\tilde{\theta}_{i,m}$ for the mother and $\tilde{\theta}_{i,f}$ for the father. Rewriting the four equations derived in the previous sections using notation for variables in the ECLS-B yields:

$$\ln\left(\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}\right) = \left(\frac{1}{\eta_s-1}\right)\ln\left[\frac{\tilde{w}_{i,m,t}}{\tilde{p}_{i,t}}\right] + \left(\frac{1}{\eta_s-1}\right)\ln\left[\frac{1-\gamma_s}{\gamma_s}\right] - \left(\frac{\eta_s}{\eta_s-1}\right)\ln\left(\tilde{\theta}_{i,m}\right) \quad (30)$$

$$\ln\left(\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}\right) = \left(\frac{1}{\eta-1}\right)\ln\left[\frac{\tilde{p}_{i,t}}{\tilde{w}_{i,m,t}}\right] + \left(\frac{1}{\eta-1}\right)\ln\left[\frac{1-\gamma}{\gamma}\right] - \left(\frac{\eta}{\eta-1}\right)\ln\left(\tilde{\theta}_{i,m}\right) \quad (31)$$

$$\ln\left(\frac{\tilde{q}_{i,f,t}}{\tilde{n}_{i,t}}\right) = \left(\frac{1}{\rho-1}\right)\ln\left[\frac{\tilde{p}_{i,t}}{\tilde{w}_{i,f,t}}\right] + \left(\frac{\eta-\rho}{\eta(\rho-1)}\right)\ln\left(\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}\frac{\tilde{w}_{i,m,t}}{\tilde{p}_{i,t}}+1\right) \quad (32)$$
$$+ \left(\frac{1}{\rho-1}\right)\ln\left[\frac{(1-\alpha)}{\alpha(1-\gamma)^{\frac{\rho}{\eta}}}\right] - \left(\frac{\rho}{\rho-1}\right)\ln\left(\tilde{\theta}_{i,f}\right)$$

$$\ln\left(\frac{\tilde{X}_{type,i,t}}{\tilde{X}_{type,i,t-1}}\right) = \frac{1}{\xi-1}\ln[1-\upsilon] + \frac{\xi}{\xi-1}\ln\left(\frac{\tilde{\Lambda}_{type,i,t}}{\tilde{\Lambda}_{type,i,t-1}}\right) - \frac{1}{\xi-1}\ln\left[\frac{1+r_t}{1+r_{t-1}}\right] \quad (33)$$

Recovering the Parameters of Interest   For estimation purposes, equations (30) to (33) are rewritten as:

$$\ln\left(\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}\right) = \beta_{1,0} + \beta_{1,1}\ln\left[\frac{\tilde{w}_{i,m,t}}{\tilde{p}_{i,t}}\right] + \beta_{1,2}\ln[\theta_{i,m}] + \epsilon_{i,t} \quad (34)$$

$$\ln\left(\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}\right) = \beta_{2,0} + \beta_{2,1}\ln\left[\frac{\tilde{w}_{i,m,t}}{\tilde{p}_{i,t}}\right] + \beta_{2,2}\ln[\theta_{i,m}] + \epsilon_{i,t} \quad (35)$$

$$\ln\left(\frac{\tilde{q}_{i,f,t}}{\tilde{n}_{i,t}}\right) = \beta_{3,0} + \beta_{3,1}\ln\left[\frac{\tilde{w}_{i,f,t}}{\tilde{p}_{i,t}}\right] + \beta_{3,2}\ln\left(\widehat{\frac{\tilde{q}_{i,m,t}}{\tilde{n}_{i,t}}}\frac{\tilde{w}_{i,m,t}}{\tilde{p}_{i,t}}+1\right) + \beta_{3,3}\ln[\theta_{i,f}] + \epsilon_{i,t} \quad (36)$$

$$\ln\left(\frac{\tilde{X}_{type,i,t}}{\tilde{X}_{type,i,t-1}}\right) = \beta_{4,0} + \beta_{4,1}\ln\left(\frac{\tilde{\Lambda}_{type,i,t}}{\tilde{\Lambda}_{type,i,t-1}}\right) + \beta_{4,2}\ln\left[\frac{1+r_t}{1+r_{t-1}}\right] + \epsilon_{i,t} \quad (37)$$

In an Ordinary Least Squares (OLS) estimation, the assumption that the residual was un-correlated with the regressors would be violated in equations (34) to (36) if parenting produc-tivities and labor market productivity were correlated (unobserved parenting productivities would be absorbed into the residual). In regressions (34) to (36) I use a fixed effects estimator, which controls for the time-invariant parenting productivities by subtracting out the family-level mean and adding back in the population mean. Values for the population mean are re-ported under an assumption about the population average of $\ln(\theta_i)$. Since the productivity of non-parental care is normalized to 1, this level assumption is equivalent to one about the dis-tribution of relative parent/non-parent productivities in the population. That is, I assume this population average is 0. If instead one assumed parents were on average less productive than non-parental care providers, this population average would be negative; if one assumed they were more productive, it would be positive. This would in turn change the level of the share parameters I back out from the intercept, which represents the sum of the common component and the average of individual components. Assuming a higher relative productivity of parents to non-parental child care providers in the production of skill would raise the share parame-ter I infer for non-parental child care; assuming a lower relative productivity would lower the share. Given that I do not offer empirical discipline for this choice, I conduct several robustness checks in the Appendix on the share parameters in the skill accumulation technologies.

The mapping from regression coefficients to the model's parameters is as follows:

$$\gamma_s = \frac{1}{\exp\left(\frac{\beta_{1,0}}{\beta_{1,1}}\right)+1} \quad , \quad \eta_s = \frac{1}{\beta_{1,1}}+1$$

$$\gamma = \frac{1}{\exp\left(\frac{\beta_{2,0}}{\beta_{2,1}}\right)+1} \quad , \quad \eta = \frac{1}{\beta_{2,1}}+1$$

$$\alpha = \frac{1}{\frac{\exp\left(\frac{\beta_{3,0}}{\beta_{3,1}}\right)}{1-\gamma}+1} \quad , \quad \rho = \frac{1}{\beta_{3,1}}+1$$

$$\upsilon = 1 - \exp\left(\frac{\beta_{4,0}}{\beta_{4,1}-1}\right) \quad , \quad \xi = \frac{\beta_{4,1}}{\beta_{4,1}-1}$$

3.3.2 Data

The variables that I need to observe are the ones needed to construct the dependent and independent variables in the regression equations (34) to (37). The subscripts denote family i in wave t, with father f and mother m.

- Parental educational time inputs $(\tilde{q}_{i,f,t}, \tilde{q}_{i,m,t})$
- Non-parental child care time inputs $(\tilde{n}_{i,t})$
- Hourly wages $(\tilde{w}_{i,m,t}, \tilde{w}_{i,f,t})$
- Hourly non-parental child care prices $(\tilde{p}_{i,t})$

There are two data sources that I use to measure these variables: the ECLS-B, which is a panel dataset, and the American Time Use Survey (ATUS), which is a repeated cross-section sampled from the Current Population Survey (CPS).

The Early Childhood Longitudinal Study, Birth Cohort   Data for hourly wages, hourly cost of child care, and time levels of quality time from the parents and non-parental care come from the ECLS-B in hours per week.

The ECLS-B reports incomes, the period of time over which the income was earned (a day, a week, two weeks, etc.), and the hours worked in a week. From these I can construct hourly wages for mothers and fathers, if they report these variables. If they did not report hours worked, I used part-time or full-time status to assign 30 or 40 hours worked per week, respectively. If they did not report earnings because they were out of the labor force (an issue confined mostly to mothers) I imputed their hourly wages using a regression of hourly wages on education, age, and age squared for the sample on which I could construct wages with reported income. I then evaluated this regression for observations where I could see education and age to get imputed hourly wages. Finally, to correct for taxes, I use the slopes (tax rates) from Table 2 of McGrattan and Prescott (2017). It is after-tax hourly income, not the directly reported pre-tax income, that I use in the estimation.

For the price of non-parental care, I use weekly spending on each of the three main kinds of non-parental care providers: relative, nonrelative, and center-based. Of these, I keep the primary source of non-parental care. To calcuate cost per hour of non-parental care, I use the

total cost per time unit for the primary source of non-parental care, and adjust it by the number of weeks that cost represents and the hours per week the child spends in that form of non-parental care. After completing these steps, I have cost per hour of parental and non-parental care providers for one- and two-parent families.

I define quality time as activities with the child that include talking and listening, singing, and reading to them. In the ECLS-B, these are reported as frequency per unit time. To convert these units into hours per week, I impute time per activity using the ATUS. The time investments reports the frequency of different activities that parents do with their children, but it does not contain a time diary. To assign levels of time investments to the frequencies reported in the ECLS-B, I impute time per activity from the ATUS.

The American Time Use Survey Data on levels of time per activity for a parent with a given set of characteristics come from the ATUS. To impute levels of time per activity to the ECLS-B, I use a pooled sample from the ATUS from 2003 to 2016. This provides a time diary along with CPS variables on age, gender, marital status, labor force status, educational attainment, parental status, and child age. I restrict the sample to parents who are between 15 and 55, with a child 4 years or younger. I use information on gender, marital status (married/cohabiting or single), labor force status, and educational attainment, where educational attainment is discretized into those with less than a college degree or a college degree and higher. With this information, I group observations from the pooled sample by their characteristics along the aforementioned dimensions. For each group, I find survey-weighted averages of time spent on an activity for those who report engaging in it. The activities whose average duration I tabulate are time spent reading to the child and time spent talking with and listening to the child.

Imputation After linking parents in the ECLS-B with their appropriate group in the ATUS, I assign the level of time spent reading, and time spent talking and listening, to their respective activities in the ECLS-B; additionally, I assign the level of time spent talking and listening to the singing activity reported by parents in the ECLS-B. I do this because there is no singing activity reported in the ATUS. Once I have imputed time levels per activity associated with the ECLS-B sample, I aggregate levels of quality time per parent in each family in each wave.

This gives me quality time investments, non-parental time investments, after-tax hourly wages, and non-parental child care prices at the child-family pair and wave level. I pool cohabiting and married couples in the data and refer to them collectively as "married couples." Single mothers are defined as mothers in the sample who are a primary caregiver and who do not have a significant other living in the household with them. In addition, I select only families with observations in all three of waves 1, 2, and 3. With this information, I can implement the estimation.

Estimating Sample Moments   Moments from the estimation samples for couples and single mothers are presented in Table 2 and Table 3, respectively. I use wave 3 survey weights provided by the ECLS-B, which are designed to correct for attrition in the sample.

Table 2: Married and Cohabiting Couples in the ECLS-B

Waves 1-3		
Levels		
	mean	sd
Education Time Father: Hours per Week	5.55	2.94
Education Time Mother: Hours per Week	7.99	2.74
Non-parental Care: Hours per Week	30.47	13.94
Ratio of Time: Father/Mother	0.76	0.50
Ratio of Time: Mother/Child Care	0.41	0.50
Hourly Pay Mother, After Tax	13.90	11.82
Hourly Pay Father, After Tax	15.55	11.97
Hourly Cost Child Care	3.93	4.22
Ratio Hourly Cost Child Care to AT Hourly Wage Mother	0.37	0.48
Age Resident Mother	33.05	5.60
Age Resident Father	34.94	6.09
Rates		
Mother: BA or higher	0.55	
Father: BA or higher	0.47	
Marriage Rate	0.92	
Below 100 % Poverty Line	0.01	
Below 185 % Poverty Line	0.08	
Observations	3100	

Note: Sample sizes are rounded to nearest 50, following National Center for Education Statistics (NCES) requirements.

Based on Table 2, the following qualitative points are apparent. First, fathers and mothers spend substantial time per week engaged in education time with their children. Within a couple, fathers invest on average 76% of the mother's time. This will be reflected in the complementarity estimates I find for father and mother time in my skill accumulation technology estimation. Next, I find that reported hours per week in non-parental care are 30 on average,

but there is a high variation in this level. The relative quantities of the two time inputs I measure in this sample, along with hourly prices for each, will translate into large estimated CES shares on non-parental care. The ratio of mother time to non-parental care is on average 0.41, but with a high variation. Married or cohabiting mothers and fathers have hourly wages that are about four times higher than the hourly price of child care they use, although there is large variation in all of these prices within my sample. On average, the price of child care is 37% of the mother's hourly wage. Poverty in the sample of married or cohabiting parents is not common: only 1% are below the poverty line, while only 8% are below 185% of that threshold. Although I pool cohabiting and married couples in my estimation sample, most of these families are married couples: the marriage rate is 92%. Both parents are on average in their 30s, and the fraction in this group with a college degree or more is about half for each gender.

Table 3: Single Mothers in the ECLS-B

Waves 1-3		
Levels		
	mean	sd
Education Time Mother: Hours per Week	7.08	2.97
Non-parental Care: Hours per Week	35.57	12.37
Ratio of Time: Mother/Child Care	0.26	0.28
Hourly Pay Mother, After Tax	9.13	10.78
Hourly Cost Child Care	2.41	2.77
Ratio Hourly Cost Child Care to AT Hourly Wage Mother	0.32	0.40
Age Resident Mother	28.07	6.36
Rates		
Mother: BA or higher	0.14	
Below 100 % Poverty Line	0.22	
Below 185 % Poverty Line	0.42	
Observations	1350	

Sample sizes rounded to nearest 50, following NCES requirements.

Table 3 presents a similar set of statistics for single mothers. Single mothers spend only slightly less time than married mothers in educational activities with their children, but non-parental care time is on average five hours higher than for couples. The ratio of mother's time to non-parental child care time is correspondingly lower than for couples, at 0.26 (compared to 0.41 for the latter). Within a family, the ratio of mother time to non-parental care time is almost half that of married or cohabiting families. The poverty rate is twenty times higher among this group, at 22%; for those below 185% of the poverty line, the rate is about five times higher, at 42%. On average, single mothers make about five dollars per hour less than mothers

parenting in couples. The ratio of this price to the mothers hourly wage is five percent smaller than for married and cohabiting mothers. Finally, the age of single mothers is on average 5 years lower than married or cohabiting mothers, and their educational attainment is one third that of married mothers: only 14% of single mothers have a college degree or more.

The comparisons across these two estimation samples help to establish priors about what the skill accumulation technologies of each should look like. Because single mothers are the only source of parental time for their child, and the amount of non-parental time purchased is so large, one expects to see a larger CES share for non-parental care in the single mother problem, ceteris paribus (I will expand more on that later in this section).

From the estimation equations, one can see that the fixed effect correction adjusts for the logged parenting productivity θ_f or θ_m (the productivity of non-parental care is normalized to 1). When I construct the predicted ratio of mother and non-parental child care time for couples in equation (36), I use fixed effects recovered from equations (35). When constructing investment for the estimation of the outermost aggregator in equation (37), I incorporate parental fixed effects recovered from (34) to (36).

For estimation weights, I used wave 3 weights—for the primary caregiver survey sample to estimate $\{\eta, \gamma\}$ and $\{\eta_s, \gamma_s\}$, and for the father survey sample to estimate ρ, α. Using a fixed effects estimator, a requirement for clustering standard errors is heterogeneity in treatment effects, which is not the case here (see Section 4 of Abadie, Athey, Imbens, and Wooldridge (2017)). However, because I impute hourly wages for parents who are not working using age and education, the wages of those observations may have errors that are correlated at the state level. This imputation is my motivation for clustering the standard errors at the level of state of residence.

3.3.3 Estimation Results

I estimate the mother/non-parental time aggregator separately for married and cohabiting mothers whose husbands also filled out a resident father questionnaire. In Table 4, this corresponds to models 1 and 2. Model 3 estimates the parameters governing how father and mother time combine, and model 4 reports estimates for the aggregator and investment and current skill. Models 1, 2, and 3 use a fixed effects estimator, but model 4 uses an OLS estimator. This

is because I cannot linearly separate the parenting productivities in that equation. These are estimations of equations (34), (35), (36), and (37). Estimation results for alternative nestings are reported in the appendix. The nesting order I have chosen restricts the substitutability of father time with mother time or non-parental care time to be the same. Alternative nestings would require that non-parental care be equally substitutable with mother and father time, or that mother time be equally substitutable with father time and non-parental care time. For couples, I chose the nesting that makes it most comparable with that of single mothers. I use this rule because there is not a clear ranking of alternative nestings by fit, and the parameter estimates for couples do not change in a statistically significant way by changing this restriction with one exception (see the Appendix for estimates with alternative nestings).

Table 4: Estimation Results: Skill Accumulation Technology

	(1) Single Mother + Child Care	(2) Cohabiting Mother + Child Care	(3) Cohabiting Father + Not Father	(4) Outermost Estimation
$log(\frac{\text{after-tax hourly wage mother}}{\text{cost per hour child care}})$	-0.308*** (0.0856)	-0.497*** (0.0522)		
$log(\frac{\text{after-tax hourly wage father}}{\text{cost per hour child care}})$			-0.349*** (0.0509)	
$log(\text{composite term})$			-0.599** (0.199)	
$log(\frac{\Lambda_{type,t}}{\Lambda_{type,t-1}})$				0.545*** (0.0298)
$log(\frac{r_{t-1}}{r_t})$				7.034*** (0.391)
Constant	-1.117*** (0.131)	-0.583*** (0.0659)	-0.836*** (0.0923)	0.770*** (0.0417)
R^2	.1356	.2039	.15	.5053
Observations	1350	3100	3100	2600

Notes: Standard errors are in parentheses, clustered at the level of state of residence.
Sample sizes rounded to nearest 50, following NCES requirements.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

These results yield the parameters in Table 5 for the skill accumulation technologies.

Table 5: Parameters of the Human Capital Accumulation Technology

ρ	α	η	γ	η_s	γ_s	ξ	υ
-1.86	0.07	-1.01	0.24	-2.25	0.03	-1.2	0.82
(0.42)	(0.03)	(0.21)	(0.05)	(0.90)	(0.04)	(0.14)	(0.03)

Note: Standard errors are in parentheses, clustered at the level of state of residence and calculated using the delta method.

Except for the single mother share γ_s, these parameter values are statistically different from zero at the 5% significance level. As for the single mother share, it has a p-value of 0.47. The fact that γ_s is small relative to γ is consistent with my intuition based on sample summary statics.

When I translate the point estimates of η, η_s, and ρ into elasticities of substitution, the estimation yields the values shown in Table 6.

Table 6: Elasticities of Substitution

Father + Mother $\frac{1}{1-\rho}$	Married Mothers + Non-parental $\frac{1}{1-\eta}$	Single Mothers + Non-parental $\frac{1}{1-\eta_s}$
0.35	0.50	0.31
(0.05)	(0.05)	(0.09)

Note: Standard errors are in parentheses, clustered at the level of state of residence.

If the ratio of father over mother wage increases by 1%, the ratio of father time to mother time invested in their child decreases by 0.35%. If the denominator of both ratios changes to non-parental child care, the reaction to a 1% change in the former ratio is the same. For married mothers, a 1% increase in the ratio of the mother's hourly wage to the price of non-parental care decreases the ratio of her time to non-parental time invested in her child by 0.50%. She is more responsive than married men or single mothers to the price of non-parental care. Single mothers are the least responsive: they would adjust the ratio of their time to non-parental care time by only 0.31%. The share values, meanwhile, indicate that the level of her time would already be very low compared to the time inputs from couples raising children in this economy.

For intuition on the share parameters, consider a single mother whose hourly wage is exactly the same as the hourly price of non-parental child care care. Setting the prices equal to

one another provides the clearest intuition for the role of the share parameter, and I chose to use the single mother technology because it has no outermost nesting like the married couple does around the ratio of prices of interest (which includes the price I subsidize in my analysis). The expression for the ratio of her quality time and non-parental child care time is:

$$\frac{q}{n} = \left[\frac{1 - \gamma_s}{\gamma_s} \frac{w\theta}{(1 - \tau_n)\, p_n}\right]^{\frac{1}{\eta_s - 1}}$$

Imposing that the ratio of prices is 1 yields:

$$\frac{q}{n} = \left[\frac{1 - \gamma_s}{\gamma_s}\right]^{\frac{1}{\eta_s - 1}}$$

Solving for γ_s, we have

$$\gamma_s = \frac{1}{\left[\left(\frac{q}{n}\right)^{\eta_s - 1} + 1\right]}$$

Knowing the elasticity of substitution between her time and non-parental child care, the share parameter pins down the ratio of her time and the time in non-parental child care she chooses. If these two levels are equal when the prices are the same, then the share parameters will be $\frac{1}{2}$; if the single mother's quality time inputs exceed the time she purchases in non-parental child care, even when the prices are the same, then the weight on her time will be greater than $\frac{1}{2}$, and vice versa. The degree to which the share parameter needs to adjust in order to explain the ratio of input choices when the price of each input is the same depends on the value of the complementarity parameter η_s. The more complementary the two inputs are, the more the share parameter will need to adjust to explain a large difference in their input levels.

3.3.4 Discussion of Estimation Results

There are three main takeaways from the skill technology estimation. First, the share parameters on father time and single mother time inputs are very low. Even if their hourly wages were the same as the prices of the other inputs in their technology, ratios of investment inputs would still show a relatively low level of inputs coming from their quality time. This is independent of the composition of hourly wages in either group. By contrast, non-parental child

care has a large share in the investment technologies of both one- and two-parent families. If this share were zero, changes in the price of non-parental child care incurred by families (due to a subsidy such as the ones I examine) would have no effect on the price of investment or the skill accumulation of children.

The second takeaway concerns the relative elasticities of the three parental inputs with respect to non-parental care. A 1% increase in the ratio of a single mother's hourly wage to the price of non-parental care causes her to adjust the ratio of her time input to non-parental care by 0.31%. She does not heavily readjust her investment input choices because of the price change. Married mothers, by comparison, adjust their input choices more—by 0.50%. A subsidy to non-parental child care will cause them to shift more heavily towards that input than for single mothers. Fathers parenting in a couple have an intermediate elasticity: they are more responsive to the ratio of input prices than single mothers, but less responsive than their wives. This matters for policy analysis in general equilibrium, such as the one I conduct, because it is informative about what one can expect parents in the economy to do with their time in the presence of a subsidy. All parents will shift inputs away from their own time and toward non-parental care if the latter's price decreases. Ceteris paribus, this shift will not be dramatic; to achieve large changes in time use, the subsidy will have to be large. Time use is relevant because if parents easily substituted away from parenting and (at least partially) into time spent working, this would be a source of expansion in the tax base. My estimates indicate that expansions in the tax base from parents substituting away from parenting quality time and into labor should be expected to be small. Changes in the tax base will have to come instead from changes in the population distribution of skill rather than changes in parental time use.

The third and final takeaway is the role of investment in the outermost aggregator, which takes investment and the current stock of skill to produce tomorrow's skill. The share parameter on investment is high: its role in producing tomorrow's skill dominates the role of the current stock of skill. In addition, the complementarity parameter between investment and skill is $\xi = -1.2$, which means that they are more complementary than allowed in a Cobb-Douglas parameterization. The corresponding elasticity of substitution is 0.45. Consider the multi-period estimation formulation presented above, and the intertemporal cost minimization equation that I use to derive my estimation equation for ξ. If the ratio of investment prices

today and tomorrow changes by 1%, this estimate says that the ratio of investments today and tomorrow will change by 0.45%. The distribution of investments over time is not very sensitive to changes in the price of that investment across periods. In my model, unlike in the estimation data, there is a single period of childhood. In that context, there is no intertemporal cost minimization decision for parents. The interpretation of the complementarity parameter in a one-period technology is that it determines how quickly investment increases with the required adjustment to the current stock of skill. For example, if one wants to increase the current stock of skill by 10 percentage points, a skill accumulation technology that exhibits high complementarity (lower ξ) will require more investment than one with greater substitutability, for a given initial skill θ_c and share parameter υ.

3.3.5 Comparison with Other Findings

There is a large body of work on the estimation of skill technologies (e.g., see Todd and Wolpin (2003) and Cunha and Heckman (2008) for a discussion of specification and estimation issues). Direct comparisons are problematic because specifications vary across studies, but are nevertheless qualitatively informative. The data requirements vary with the estimation equations; mine require information on time inputs and their prices. One widely cited estimation is due to Cunha, Heckman, and Schennach (2010). In that study, the authors use the NLSY79 Children and Young Adults suvey to estimate a CES skill accumulation technology for a two-dimensional skill vector containing both cognitive and non-cognitive skills. Each of these dimensions of skill is allowed to affect the evolution of the other dimension. This skill technology is very general; it nests several specifications examined in other studies. Cunha, Heckman, and Schennach (2010) find substitutability between investment and skill during early childhood. In that study, the authors are not specific about which inputs aggregate into investment in skill.

Lee and Seshadri (2018) estimate a skill technology with a single dimension of skill, but assume that there is no initial draw of skill (or rather, that it is identical for everyone and equal to 0). They thereby impose by assumption a property that Cunha, Heckman, and Schennach (2010) found as a result: that in the first period of life, initial investment and the initial stock of skill are substitutable. For other periods, however, Lee and Seshadri (2018) are unable to reject a Cobb-Douglas specification. This is a higher level of substitutability than what I find,

but lower than that in Cunha, Heckman, and Schennach (2010).

My estimates show complementarity between investments and skill early in life. I do not ignore heterogeneity in the first measure I have of skill at 9 months, as in the specification of Lee and Seshadri (2018), because it doesn't appear to be noise in my data and because I have no observations of the investment that generated that initial stock (motivating regressions for these two claims are in the appendix). If the investment were solely responsible for the skill I observe at 9 months, the skill would already be correlated with family income. That is not what I find. See the appendix of Cunha, Heckman, and Schennach (2010) for a discussion of the implications of substitutability versus complementarity between skill and investment for policy design.

In my framework, children being raised in families with low income are more likely to be children of single mothers, who use a different technology than couples to invest in their children. This technology emphasizes non-parental child care more than the one couples use: the share on parental time is much lower. However, the ability of parents to adjust their own time use in response to a change in child care prices is much higher for couples than for single mothers. Subsidies targeted to single mother families target the families whose costs of investment are most sensitive to the price of non-parental child care. On the other hand, they also target a population which my estimates show will be less willing to substitute away from their own time inputs than couples. This heterogeneity across family structures is uncovered with my specification choice, unlike those used in other studies.

3.4   Other Externally Estimated Parameters

In Table 7, the level of lump-sum transfers T, the level of child support payments T_{cs}, and the price of p_n are set to 7.5% of output, 46% of the average per-family transfer, and 35% of the average mother's wage, respectively. The first target is from the ratio of government transfers to persons for federal benefits from social insurance funds, Supplemental Nutrition Assistance Program (SNAP), supplemental security income, and refundable tax credits to GDP from the National Income and Production Accounts (NIPA) tabulations. The second is the ratio of average child support payments per month per capita to average monthly government transfers per family. The third is the average ratio of hourly price of non-parental care to hourly wages

of mothers in the ECLS-B. See the appendix for further details on the estimation of these parameters.

Table 7: Externally Estimated Parameters

Symbol	Name	Source	Value
T	Transfers	NIPA	7.5% of output
T_{cs}	Child Support	Census and NIPA	46% of T
κ	p_n coefficient	ECLS-B	35% ave. mother's wage

3.5 Internally Calibrated Parameters

The parameters $b, \psi_s, \psi_{mc}, \lambda$ are chosen to bring the model moments in the baseline as close as possible to the moments in the "Data" column of Table 8. The coefficient b controls the degree of altruism; ψ_s, ψ_{mc} are the marginal utilities of leisure for singles and married couples, respectively. The parameter λ is a shifter in the skill technology that scales up investment into efficiency units in the production of skill. The moments I chose these parameters to match in the model compared with the data are the correlation of child skill with family income, the average labor supply, the percent of single mothers raising children under 5, and the average time invested by parents of each family structure type. Note that the correlation of child skill and family income grows to 0.29 by the time the child is 5 years old (ECLS-B). By contrast, when children are 9 months old, the measures of skill available in the data are uncorrelated with family income, although they do have predictive power for later child test scores at age 4. For regressions supporting these points, see the appendix.

Table 8: Internally Calibrated Parameters: Targeted Moments

Moment	Source	Data	Model
Corr(child skill, family income)	ECLS-B	0.32	0.46
Average labor supply	CPS	0.28	0.29
Percentage single mothers	ECLS-B	0.20	0.22
Average time investments	ECLS-B (MM,MF,SM)	(1.75,1.25,1.5)	(2.39,1.87,1.24)

3.6 Untargeted Moments

In Table 9, I compare the implications of this model with five untargeted moments: the correlation of hourly wages in spouse (or cohabiting partners), the correlation of parental time inputs within a couple, the ratio of average hourly wages in married or cohabiting mothers with single mothers, the average ratio of mother and father time, and the average ratio of mother time and non-parental care time.

Table 9: Untargeted Moments (Model Fit)

Moment	Source	Data	Model
Correlation of wages within a couple	ECLS-B	0.28	0.29
Correlation of time inputs within a couple	ECLS-B	0.33	0.43
$\frac{\text{ave. hourly wages, single mothers}}{\text{ave. hourly wages, mc mothers}}$	ECLS-B	0.67	0.85
Average $\frac{\text{father time}}{\text{mother time}}$	ECLS-B	0.77	0.79
Average $\frac{\text{mother time}}{\text{non-parental care time}}$	ECLS-B	0.37	0.19

Notes: Table 9 shows relevant moments for the implications of the model for sorting across family structures, assortative matching in marriage, and patterns of time inputs for investments in child skill. Correlations are Pearson correlation coefficients, weighted with survey weights in the ECLS-B and using the analogous distribution in the model. Ratios use empirical moments from Tables 4 and 5 and analogous moments using the baseline equilibrium of the model.

Table 9 emphasizes that although the random search marriage market I model captures an appropriate degree of assortative matching among spouses: the correlation of wages within couples is a close fit with the data. However, within a couple the model implies a higher correlation of time inputs than seen in the data. This partly reflects the fact that the levels of wages for men and women in the model economy have identical distributions, but in the data there is an unmodelled gender wage gap. To examine how mothers sort into the family structures, the third moment looks at the ratio of their wages across family structures. In the data, single mothers have lower hourly wages than married mothers. I assume that hourly wages are correlated with parenting productivity in my estimation, so the sorting across family structures has implications for the parenting productivities using each of the two technologies in the population. My model endogenously captures the qualitative attribute of the data: mothers

with higher skill tend to parent in a couple. However, the average hourly wages of single and married mothers in the model are closer together than in the data. More higher-wage women parent as single mothers in the model than in the data. The last two moments in Table 9 show the average ratios of time inputs in the model and in the data. The ratio of time inputs within a couple is quite close to the data, but the ratio of mother time with non-parental care time is too low in the model. This reflects the fact that I did not model quality choice, so that only the intensive margin of non-parental care time can adjust.

4   Policy Experiment and Results

I vary τ_n over the interval $[0, 1]$ with three eligibility rules: universal, poverty-tested, and family-tested. Under a universal eligibility rule, every parent receives the subsidy, whereas under a poverty-tested rule only families in the bottom 20% of the baseline income distribution of parents raising children receive the subsidy. In practice, the poor by this definition are entirely single mothers. This is also true if I rank families by CE-adjusted income (dividing income by ϕ_s and $\phi_{mc,p}$ to get a rough per capita measure within each family). Under the family-tested rule, only single mothers are subsidized (20% of parenting families). I chose this approach so that the mass of those eligible for the poverty-tested and single mother subsidy is the same in the baseline, to make the two eligibility criteria more comparable. For each τ_n value I solve for equilibrium and record household decisions for every possible family type, as well as the endogenous population weights assigned to that family type.

A family type is the marriage structure (single mother, single father, or married couple), the productivity type of the parent(s), and the initial skill of the children (every single mother and married couple raises one boy and one girl). An equilibrium output in the model is structured like a survey dataset, with observations at the family type level and the endogenous population weights being analogous to survey weights. I use this information to compute welfare changes under different policies. For each eligibility criteria, I find the subsidy level that implements the largest ex ante welfare gain for children at birth. I compare these maximized welfare gains, looking at the distribution of these gains in the population after birth as well as across initial skill types of children before they know the family they will be raised in.

The rest of this section proceeds as follows. First, I report the formulas I use to calculate welfare gains for different levels and eligibility rules of the non-parental child care subsidy. Second, I report the best attainable outcome for each eligibility rule (universal, subsidies to the poor, and subsidies to single mothers), using ex ante welfare by child initial skill and decompositions of welfare gains by attributes of the population such as age, gender, marital status, and adult skill level. Differences between the effects of the two eligibility rules are apparent with the decompositions I report and analyze in the last subsection of these results. Third, I discuss how some features of my model specificaton are related to the results.

Welfare Measures

I use the following formula to compute ex ante welfare changes in consumption-equivalent units:

$$\Delta W_0\left(\theta, g | \tau_n\right) = \exp\left(\frac{V_0\left(\theta, g | \tau_n\right) - V_0\left(\theta, g | \tau_n = 0\right)}{\sum_{age=1}^{3}\left(\beta^{age-1}\right)}\right) \qquad (38)$$

This is the consumption-equivalent change in ex ante lifetime utility, evaluated at the skill type θ and gender g of the child, at an equilibrium indexed by the subsidy τ_n. Computing lifetime utility $V_0\left(\theta, g | \tau_n\right)$ for ex ante welfare gains requires a choice of distribution over which to take expectations in the subsidized equilibrium. Accordingly, I indicate in each result whether I use the distribution of families in that τ_n equilibrium or in the baseline equilibrium.

The analogous equation to compute ex post welfare gains is:

$$\Delta W_j^{type}\left(z; \overrightarrow{\theta} | \tau_n\right) = \exp\left(\frac{V_j^{type}\left(z; \overrightarrow{\theta} | \tau_n\right) - V_j^{type}\left(z; \overrightarrow{\theta} | \tau_n = 0\right)}{\sum_{age=j}^{3}\left(\beta^{age-j}\right)}\right) \qquad (39)$$

where $type \in \{SM, SF, MC\}$ is the family structure type, z is the state space at that phase of life for that family type, and $\overrightarrow{\theta}$ is the vector of adult skill types, equal to θ in the case of single parents or θ_f, θ_m in the case of married couples. The state space z contains the asset stock; equation (39) uses the policy choice for that household at that age, in the previous period. The object $\Delta W_j^{type}\left(z; \overrightarrow{\theta} | \tau_n\right)$ is the consumption-equivalent change in lifetime utility evaluated at age j, for a given family structure and composition. To aggregate these welfare changes, I use the distribution of families in the subsidized or baseline equilibrium (I am specific in each

case). I construct the distribution of families I use to aggregate welfare using the appropriate joint distribution μ, marriage decision rules, and the exogenous distribution from which the initial skill of the child is drawn.

4.1   Welfare Gains: Overall and Decomposition

I tabulate gains in welfare ex ante (for the unborn), and ex post (by age, skill, marital status, and gender groups). When describing and analyzing the distribution of welfare gains in the economy under different subsidies, I have separated adults into two groups: low and high skill. Low skill is defined as those below the 20th percentile of hourly wages in the baseline equilibrium. This procedure allows for more heterogeneity in the high-skilled group than the low-skilled group, but a finer decomposition does not reveal a qualitatively different result from the one I discuss below.

4.2   Ex Ante Welfare Gains

Table 10 reports ex ante welfare gains for children, overall and by initial skill group (low, medium, and high). In addition, the second column reports the subsidy level and the third column reports the marriage rate at that equilibrium. All welfare gains reported in the last four columns are in consumption-equivalent units, with expectations taken over the distribution of families in the subsidized equilibrium. Overall, single mother subsidies yield welfare gains that are 1.5 percentage points higher than universal subsidies, and 2 percentage points higher than a poverty-tested subsidy. The subsidy levels that achieve maximal gains under the different elibility rules are close to one another, ranging from 80% to 85%.

Table 10: Ex Ante Welfare Gains, Subsidized Family Distribution

Eligibility	τ_n^*	Marriage rate	Welfare Gain Average	Welfare Gain by Skill Low	Medium	High
Universal	80%	73%	6.262	8.109	5.562	5.812
Single Moms	85%	72%	7.722	8.865	7.774	6.477
Poor	80%	73%	5.741	6.186	5.885	5.008

Notes: Table 10 reports, by eligibility rule, the welfare-maximizing subsidy level, the marriage rate, and welfare gains on average and by initial skill of the child. Welfare gains are computed as expected values over outcomes, using the distribution of families in the subsidized equilibrium, and are reported in consumption-equivalent units. On average, single mother subsidies yield the highest gains, followed by universal subsidies and, finally, subsidies directed at the poor.

Gains for all subsidies are decreasing in the child's initial skill. The dispersion of welfare gains across initial skill groups is largest under the universal subsidy and smallest under the single mother subsidy.

Notice that a single mother subsidy does not decrease the marriage rate by much more than one directed at the poor or a universal subsidy (the marriage rate is only 1 percentage point lower under the single mother subsidy). Lowering the cost of non-parental care makes singlehood more attractive for both men and women, and all of the subsidies I study do that.

Table 11: Ex Ante Welfare Gains, Baseline Family Distribution

Eligibility	τ_n^*	Welfare Gain Average	Welfare Gain by Skill Low	Medium	High
Universal	80%	4.345	7.640	3.814	2.113
Single Moms	85%	7.108	8.127	7.204	5.900
Poor	80%	5.912	6.265	6.130	5.124

Notes: Table 11 reports, by eligibility rule, the welfare-maximizing subsidy level and welfare gains on average and by initial skill of the child. Welfare gains are computed as expected values over outcomes, using the distribution of families in the baseline equilibrium, and are reported in consumption equivalent units. The single mother subsidy out-performs universal and poverty-tested subsidies by 2.8 percentage points and 1.2 percentage points, respectively. The baseline distribution of families yields lower gains for medium and high-skilled children under the universal subsidy, but actually slightly increases welfare gains under the poverty-tested subsidy.

In Table 11, I present welfare gains using the distribution of families at the baseline equilibrium. This is a relevant set of statistics because, conceivably, these are the families to whom

policy makers are answerable when the policy is implemented. The welfare gains for this distribution of families show a clear ranking with the single mother subsidy at the top. Under this subsidy, the gains are distributed more evenly across initial skill types than for the other two eligibility rules. For universal subsidies, in particular, the gains that accrue to those born with a high initial skill are much lower compared to those enjoyed by other skill levels. A subsidy to single mothers raises welfare gains on average by 2.8 percentage points relative to a universal subsidy, and by 1.2 percentage points relative to the poverty-tested subsidy. Across skill groups, the gains range from 8.1 to 2.1 percentage points.

By comparing Table 10 with Table 11, one can infer that the welfare gains from a single mother subsidy and one directed at the poor are mostly due to changes in the well-being of families that already exist in the baseline. By comparison, gains from universal subsidies are larger once changes in the distribution of families are incorporated into the welfare analysis. Why is a universal subsidy so much more distortionary to the distribution of families? It is not operating through a greater or lesser effect on the marriage rate, because the marriage rate falls by a similar amount under all three eligibility rules (see Table 10). Rather, different effects on welfare are due to the different ways the subsidies affect the distribution of skill.

Universal subsidies give rise to less heterogeneity in adult skill levels even as inequality across these skills rises. This lack of heterogeneity narrows the set of families that a child can be born into. At the same time, family income matters less than at the baseline for children's outcomes. The result is that a child faces a set of fewer types of families she can be born into, in an environment where the family matters less for her outcome. Under the single mother subsidy, meanwhile, the distribution over families changes by much less: the skill distribution does not collapse into high and low levels, but remains disperse. Effectively, the single mother subsidy and poverty-tested subsidy preserve a middle class while the universal subsidy creates a more polarized society with high- and low-skill levels and few people in between.

4.3   Aggregate Moments: Comparison Across Eligibility Rules

Table 12 contains aggregate moments and equilibrium objects under the three eligibility rules at the welfare-maximizing level of τ_n. These aggregates are informative about the sources of welfare gains reported in Table 9 and Table 10. Under the single mother or poverty-tested subsidy,

the labor income tax falls relative to the baseline by 5.2 and 5 percentage points, respectively. By comparison, a universal subsidy is expensive and expands the tax base less, which is reflected in the large increase in the labor income tax relative to the baseline (an increase of 5.7 percentage points). The price of non-parental care reflects changes in the average hourly wage under each eligibility rule: it rises the most under the single mother subsidy (11.3 percentage points) and the least under the universal subsidy (7.4 percentage points).

Table 12: Macroeconomic Moments and Equilibrium Objects Across Eligibility Rules

	r	τ_y	p_n	Y	C	H	K	L	N	Q_f	Q_m
Universal	0.998	1.057	1.074	1.048	1.390	1.046	1.051	1.336	1.672	1.144	1.168
Single Mothers	0.999	0.948	1.113	1.072	1.429	1.071	1.075	1.332	1.545	1.198	1.263
Poverty Tested	0.999	0.950	1.081	1.049	1.399	1.047	1.051	1.331	1.479	1.250	1.299

Notes: Table 12 reports macroeconomic moments and equilibrium objects across eligibility rules at the welfare-maximizing level of τ_n. Moments are reported as fractions of the baseline level, to facilitate interpretation in units of percentage change. Because the price of non-parental care is not an aggregate endogenous state but simply a function of it in this model, I categorize it as a macroeconomic moment. Moving from left to right, the rest of the columns report aggregate output, consumption, labor efficiency units, capital stock, leisure, non-parental care time, quality time from fathers, and quality time from mothers. All aggregates use the distribution over families in the subsidized equilibrium.

Output, consumption, hours, and the capital stock (columns 5 to 12 of Table 12) increase the most under the single mother subsidy. By contrast, leisure and the use of non-parental care increase the most under the universal subsidy, while parental quality time inputs increase the most under the poverty-tested subsidy. All of these subsidies generate welfare gains by increasing skill in the economy, which allows for higher consumption and higher leisure, even as quality time inputs from parents increase. For the universal subsidy, the expense of the program exceeds the revenue it generates due to expansions in the tax base. Under the other subsidies, meanwhile, even parents who do not receive the subsidy have higher incomes due to lower tax levels. During the parenting phase, these higher incomes can be used to invest in child skill.

Note that complementarity between parental time and non-parental time in the production of child skill means that, for a fixed level of parenting productivity, a change in the ratio of prices does not get completely reflected as a change in the ratio of inputs: that is, parenting

time will not be completely crowded out by non-parental care. However, as the distribution of skill in the population increases, the opportunity cost of parenting time also goes up, while the productivity of non-parental time remains fixed. The ratio of time inputs into child skill development will reflect the increased value of parental time, as well as the lowered price of non-parental child care. Note that even in the presence of a lower marriage rate, time investments from fathers have increased, relative to the baseline, under all three subsidies. This is due to increased levels of investment and increases in family income compared to the baseline.

Table 13: Income and Earnings Levels and Gini Coefficient Across Eligibility Rules

	Income		Earnings		Pre-tax Income		Pre-tax Earnings	
	mean	Gini	mean	Gini	mean	Gini	mean	Gini
Universal	1.040	1.154	1.044	1.043	1.044	1.182	1.051	1.043
Single Mothers	1.071	1.150	1.083	1.038	1.067	1.175	1.074	1.038
Poverty Tested	1.049	1.107	1.059	1.023	1.045	1.127	1.051	1.023

Notes: Table 13 reports average income and earnings, both before and after tax, as well as the Gini coefficient for each measure. Averages are calculated using the distribution over families in the subsidized equilibrium. Income and earnings are reported after labor income taxes are applied. Pre-tax income contains lump-sum transfers. Income is the sum of labor earnings and income from wealth. Because everyone starts the parenthood period with no wealth, earnings and income are the same in that phase of life. Relative to the baseline, both income and earnings increase the most under the single mother subsidy.

Table 13 describes the behavior of the earnings distribution relative to the baseline under the three eligibility rules. A single mother subsidy increases the mean of income and earnings, and it also increases the dispersion of income and earnings in the population (as measured by the Gini coefficient). Its advantage over the other subsidies is that it increases the level of income and earnings the most, but it does not increase dispersion the most. The latter point is an advantage because, behind the veil of ignorance, dispersion in income is analogous to risk over family outcomes (even as families become less important for the outcomes of kids). Inequality as represented by the Gini coefficient, meanwhile, increases the least under a poverty-tested subsidy compared with the two other eligibility rules. Universal subsidies yield the smallest increase in income and earnings and the largest increase in inequality for all income measures. Notice that for the poverty-tested and single mother subsidy, after-tax earnings and income are higher relative to the baseline averages than pre-tax income and pre-tax earnings. This ranking

reflects the decrease in the labor income tax reported in Table 12. For universal subsidies, where the labor income tax increases relative to the baseline, this is not the case. Note also that, by comparing the Gini coefficients of before- and after-tax income, one can see that the labor income tax acts to decrease dispersion. The dispersion of earnings is unaffected because the labor income tax is linear.

4.4 Decomposition of Welfare Gains

Table 14 decomposes welfare gains by marital status and age across eligibility rules at the welfare-maximizing level of τ_n. Note that a universal subsidy yields welfare losses for the old and working single fathers, while a single mother or poverty-tested subsidy improves welfare for all but older single fathers. This is due to the lower labor income tax, and the free time single mothers and married couples are able to use to accrue savings that they can live off of later in life.

Table 14: Welfare Gains by Age and Marital Status

| | Universal Subsidy | | | |
	Average	Single Fathers	Single Mothers	Married Couples
Parenthood	2.299	1.594	7.053	1.550
Worker	0.0312	-1.218	0.075	0.254
Old	-0.483	-1.842	-0.128	-0.297
All	0.616	-0.488	2.333	0.502

| | Single Mother Subsidy | | | |
	Average	Single Fathers	Single Mothers	Married Couples
Parenthood	3.225	2.918	9.330	2.103
Worker	0.966	0.306	2.825	0.734
Old	0.638	-0.020	2.804	0.347
All	1.610	1.068	4.986	1.061

| | Poverty Tested Subsidy | | | |
	Average	Single Fathers	Single Mothers	Married Couples
Parenthood	2.474	2.320	6.646	1.750
Worker	0.732	0.301	1.249	0.716
Old	0.444	0.00	1.062	0.413
All	1.217	0.874	2.986	0.960

Notes: Table 14 shows the welfare gains for each eligibility rule at the optimal subsidy level for that rule. Gains are reported at each age, overall and by family structures. Welfare gains are in consumption equivalent units, and are aggregated using the distribution of families at the subsidized equilibrium. Qualitatively, the patterns noted in Table 14 across age and marital status also play out in Table 15. Here, one can see that the difference in welfare gains across skill levels is lowest under the single mother subsidy.

The tax burdens of funding the universal subsidies are high enough that the subsidy to the price of child care does not allow a similar increase in savings to cushion old age. Older single fathers are not directly affected by the price of non-parental care, but gain from decreased labor income taxes during the working phase regardless of any effect on the saving rate. In old age, the slight decrease in the interest rate leads to a decrease in their income, which was compositionally more reliant on savings in the baseline during that period.

Table 15 provies the finest decomposition of welfare gains that I analyze. The universal subsidy implements welfare gains for parents, but the old and working single fathers experience welfare losses due to increased labor income taxes required to fund the program. Single mother subsidies improve the welfare of single mothers throughout their life, and also of ineligible married couples through lower taxes on labor income and insurance against poor mar-

riage market outcomes for their children. The pattern of gains from poverty-tested subsidies is quite similar to ones directed at single mothers, with a few differences: losses for single fathers during old age are slightly less under the poverty-tested subsidy.

All of the maximized welfare gains are implemented with a high level of the subsidy: the values range from 80% for universal and poverty tested, to 85% in the case of the single mother subsidy. The low elasticity of the skill accumulation technology means that it takes large changes in the price of non-parental care to make large changes in the relative time contribution from parents. The maximum for gains from these policies occurs when the tax burden for funding them begins to offset gains from insurance. Everyone in this economy becomes a parent and participates in the marriage market; they all face the same risks, and skill investment is the only way to insure against those risks. Welfare gains continue to increase with the subsidy until it is very high because it provides insurance against poor outcomes, which in this environment can happen to anyone.

Table 15: Welfare Gains by Marital Status and Skill

| | Universal Subsidy | | | | | | | |
| | Married Fathers | | Single Fathers | | Married Mothers | | Single Mothers | |
	Low	High	Low	High	Low	High	Low	High
Parenthood	2.217	0.722	1.626	1.594	4.440	0.573	7.135	6.684
Worker	0.591	-0.164	-0.381	-1.222	1.530	-0.177	-0.287	1.722
Old	0.157	-0.860	-0.295	-1.850	1.867	-1.028	-0.404	1.125

| | Single Mother Subsidy | | | | | | | |
| | Married Fathers | | Single Fathers | | Married Mothers | | Single Mothers | |
	Low	High	Low	High	Low	High	Low	High
Parenthood	2.148	2.094	2.885	2.919	2.184	2.100	9.525	9.229
Worker	0.832	0.714	0.274	0.307	0.757	0.733	2.306	3.093
Old	0.538	0.309	-0.0507	-0.0197	0.470	0.342	2.198	3.117

| | Poverty Tested Subsidy | | | | | | | |
| | Married Fathers | | Single Fathers | | Married Mothers | | Single Mothers | |
	Low	High	Low	High	Low	High	Low	High
Parenthood	1.775	1.743	2.294	2.320	1.798	1.746	8.424	5.330
Worker	0.781	0.699	0.298	0.301	0.852	0.707	1.633	0.966
Old	0.495	0.390	-0.013	0.005	0.602	0.400	1.473	0.757

Notes: Table 15 shows the welfare gains for each eligibility rule at the optimal subsidy level for that rule. Gains are reported by age, marital status, gender, and skill (high or low). Welfare gains are in consumption equivalent units, and are aggregated using the distribution of families at the subsidized equilibrium. Under a universal subsidy, gains accrue to all parents. Past parenthood, however, gains turn into losses at older ages for single fathers, high-skill married people, and low-skill single mothers. For the single mother subsidy, most workers still experience gains, although single fathers are worse off during old age. Finally, for poverty-tested subsidies, the pattern of gains is similar to that of single mother subsidies: gains are negative only for single fathers during old age, slightly less negative for the poverty-tested subsidy.

4.5 Model Specification: Effects on Welfare Gains

Welfare gains from the subsidy schemes that I study are rather large, on the order of 5% to 9%.

This is not an unusual range in the literature on the welfare gains from subsidizing skill accumulation in early childhood, or even later in life (Daruich (2017), Abbott, Gallipolli, Meghir, and Violante (2018)). There are fundamental incomplete markets in the environment I specify, a feature that I share to some extent with these papers, and which child care subsidies address: the inability of a child to choose her parents, the inability of parents to control their child's initial skill, and the inability of parents and children to insure against risk in the family formation

stage of the child. The magnitude of the uninsurable risk from each of these sources determines the welfare gains from a policy that helps to provide insurance. The way I have specified this model allows for each to be a significant source of risk.

Because I have found that my data do not allow me to predict the initial skill of the child using family attributes, in the model I have not made the distribution from which the initial child skill is drawn depend on parental attributes. This choice magnifies the degree of risk a child faces once they know their initial skill but are still uncertain about their family: if skill were very persistent across generations, knowing your initial skill would be very informative about the sort of parents you would have.

In addition, because I have found that I cannot predict initial skill with the measures of investment that I have in the ECLS-B, I have not modelled choices of parents that can affect the initial skill of the child (see the appendix for regressions motivating this and other modelling choices).[4] In addition to the fact that initial skill is not persistent across generations, this means that all parents face the same risk over the initial skill of their children: they differ in their ability to compensate for a low initial skill by investing more. A model with persistence in initial skill would reduce this source of risk and open up other policy discussions (see Abbott, Gallipolli, Meghir, and Violante (2018), and Rustichini, Iacono, and McGue (2016)).

Risk from the marriage market arises from several sources. The marriage market that I model is a simple one: potential spouses are matched randomly, and positive assortative matching by skill occurs because adults have the power to reject the spouse they match with. A marriage market where search was more directed would reduce the risk faced by a new adult over the potential spouse they meet, by giving them some control over it. Here, I only allow parents to affect family formation by ensuring that their child has enough skill to not be rejected by the spouse they meet. The individual can only appear on the marriage market with the skill she has as an adult and make her decision. I interpret the risk added by the presence of the marriage market as an upper bound: with a directed search framework, this risk would be mitigated.

Another source of risk in the marriage market stems from the way I set up the married

---

[4]The assumption of no prenatal investments and the assumption of no genetic persistence in skill are distinct. For an excellent discussion of studies on the role of prenatal care, see Corman, Dave, and Reichman (2018). For a discussion of studies on the relationship between genes and cognitive skills, see National Scientific Council on the Developing Child (2010) and the citations therein.

couple problem. Spouses are not allowed to bargain over the share of household utility they receive. Instead I emphasize economies of scale in consumption (through consumption equivalence scales), and gains from teamwork in investment in the couple's children through the skill technology. By contrast, a framework with bargaining would have allowed the spouse who gains the most from the match relative to being single to convince their potential spouse to marry them, with the promise of a larger fraction of household utility in return (Choo and Siow (2006), Reynoso (2017)). These endogenous shares would act like prices. The missing market for transferring utility between spouses is absorbed into the risk faced on the marriage market, and increases the gains from insuring against a poor outcome there.

5 Conclusion

In this paper, I build and estimate a framework which incorporates a relevant dimension along which to target and analyze early childhood education policy in general equilibrium: family structure. I allow prices, tax rates, the distribution of skill, and family formation to adjust endogenously in response to the subsidy, which makes explicit how these aggregate endogeneous states evolve across equilibria. After estimating the technologies that single mothers and couples use to invest in their children, I apply the model to compare the welfare effects of universal subsidies to child care with subsidies targeted at the poor and at single mothers. I find that targeting child care subsidies to one-parent families yields welfare gains that are 1.5 and 2 percentage points higher than universal and poverty-tested subsidies, respectively. Single mother subsidies target families with low incomes who are sensitive to the price of non-parental child care, but without requiring higher taxes or introducing a disincentive to invest in skill.

In future research, I plan on imposing more discipline on the single father problem specification. Most of the work done on this front has focused on divorced spouses, rather than on those who never marry, as in my model (Del Boca and Flinn (1995), Tartari (2015)). In addition, the non-parental care sector can be further elaborated, so that the elasticity of supply for non-parental care can be disciplined and its implications for large-scale subsidies examined (Blau (1993)). As it stands, this paper provides a useful framework for analyzing child care

policies in an environment with heterogeneous family structures, with implications for how such policies will affect the marriage rate, output, and welfare. This paper also provides new estimates on how parental time inputs interact with non-parental child care time to affect child skill accumulation during early childhood.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017). When should you adjust standard errors for clustering? *Working Paper*.

Abbott, B., G. Gallipolli, C. Meghir, and G. Violante (2018). Education policy and intergenerational transfers in equilibrium. *Cowles Foundation Discussion Paper No. 1887R2*.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluaton of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

Baker, M. (2011). Innis lecture: Universal early childhood interventions: what is the evidence base? *Canadian Journal of Economics 44*(4), 1069–1105.

Baker, M., J. Gruber, and K. Milligan (2008). Universal childcare, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *The Journal of Political Economy*, 1153–1189.

Berlinski, S., S. Galiani, and P. Gertler (2009). The effect of pre-primary education on primary school performance. *Journal of Public Economics 93*(1-2), 219–234.

Blau, D. and J. Currie (2006). Pre-school, day care, and after-school care: Who's minding the kids? In E. A. Hanushek and F. Welch (Eds.), *Handbook of the Economics of Education*, Chapter 20, pp. 1163–1276. Elsevier.

Blau, D. M. (1993). The supply of child care labor. *Journal of Labor Economics 11*(2), 324–347.

Caucutt, E. and L. Lochner (2017). Early and late human capital investments, borrowing constraints, and the family. *HCEO Working Paper Series No. 2017-040*.

Choo, E. and A. Siow (2006). Who marries whom and why. *Journal of Political Economy 114*(1), 175–201.

Corman, H., D. M. Dave, and N. Reichman (2018). Effects of prenatal care on birth outcomes: Reconcicling a messy literature. *NBER Working Paper No. 24885*.

Cunha, F. and J. Heckman (2007). The technology of skill formation. *AEA Papers & Proceedings 97*(2), 32–47.

Cunha, F. and J. Heckman (2008). Formulating, identifying and esitmating the technology of cognitive and noncognitive skill formation. *The Journal of Human Resources 43*(4), 738–782.

Cunha, F., J. Heckman, and S. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Daruich, D. (2017). The macroeconomic consequences of early childhood development policies. *Job Market Paper*.

Del Boca, D., C. Flinn, and M. Wiswall (2014). Household choice and child development. *Review of Economic Studies 81*, 137–185.

Del Boca, D. and C. J. Flinn (1995). Rationalizing child-support decisions. *The American Economic Review 85*(5), 1241–1262.

Garcia, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2016). The life-cycle benefits of an influential early childhood program. *NBER Working Paper No. 22993*.

Garcia, J. L., J. J. Heckman, and A. L. Ziff (2018). Gender differences in the benefits of an influtential early childhood program. *European Economic Review 109*, 9–22.

Gayle, G.-L., G. Limor, and M. Soytas (2017). What is the source of the intergenerational correlation in earnings? *Working Paper*.

Guner, N., R. Kaygusuz, and G. Ventura (2016). Child-related transfers, household labor supply and welfare. *Working Paper*.

Gupta, N. D. and M. Simonsen (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics 94*(1-2), 30–43.

Havnes, T. and M. Mogstad (2014). Is universal child care leveling the playing field? *Journal of Public Economics 127*, 100–114.

Kottelenberg, M. J. and S. F. Lehrer (2014). The gender effects of universal child care in canada: Much ado about boys? *Working Paper*.

Kottelenberg, M. J. and S. F. Lehrer (2017). Targeted or universal coverage? assessing heterogeneity in the effects of universal childcare. *Journal of Labor Economics*.

Lee, S. Y. and A. Seshadri (2018). On the intergenerational transmission of economic status. *Journal of Public Economics (forthcoming)*.

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

McGrattan, E. R. and E. C. Prescott (2017). On financing retirement with an aging population. *Quantitative Economics 8*(1), 75–115.

National Scientific Council on the Developing Child (2010). Early experiences can alter gene expression and affect long-term development: Working paper no. 10.

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

Reynoso, A. (2017). The impact of divorce laws on equilibrium in the marriage market. *Job Market Paper*.

Rustichini, A., W. G. Iacono, and M. McGue (2016). The contribution of skills and family background to educational mobility. *The Scandinavian Journal of Economics 00*(0), 1–30.

Tartari, M. (2015). Divorce and the cognitive achievement of children. *International Economic Review 56*(2), 597–645.

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.

A Appendix

A.1 Motivating Regressions for Model Specification

To motivate my model's specification, this section provides several regressions using ECLS-B
data.

Table 16: Predicting Initial Skill with Family Attributes + Prenatal Care

	(1) Married Couples	(2) Single Mothers
Child is Female	0.0139**	0.0160
	(0.00476)	(0.00806)
B.A.: Mother	-0.00668	-0.0334*
	(0.00620)	(0.0129)
B.A.: Father	0.00302	
	(0.00735)	
Hourly Wage: Mother	-0.0000434	0.0000304
	(0.0000820)	(0.0000346)
Hourly Wage: Father	-0.0000552	
	(0.0000629)	
Month Began Prenatal Visits	0.00104	0.0000998
	(0.00245)	(0.00222)
No. Prenatal Visits	0.000982	-0.000219
	(0.000780)	(0.000892)
Constant	0.434***	0.450***
	(0.0109)	(0.0146)
R^2	0.0076	0.0118
Observations	2450	900

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Sample sizes are rounded to nearest 50, following NCES requirements.

In Table 16, I regress a child's test scores at age 4 on family attributes, child initial skill,
and child gender. I do this separately for children of married couples and single mothers,
because family attributes include those of the father for the former but not the latter (different
explanatory variables).

Table 17: Time Investments by Child Gender

	Married Couples		Single Mothers	
	(1)	(2)	(3)	(4)
	Tot. Parental Time	N Time	N Time	Total Time
Child is Female	-0.0845	-0.216	-1.156	0.0434
	(0.212)	(0.782)	(0.850)	(0.499)
Child Test Score [0,1]	8.528***	8.281***	1.023	9.945***
	(0.598)	(2.219)	(2.477)	(1.389)
B.A.: Father	2.104***	-0.895		
	(0.250)	(0.904)		
B.A.: Mother	1.850***	2.745**	1.578	3.484***
	(0.247)	(0.881)	(1.473)	(0.556)
Hourly Wage: Father	0.0105	-0.0248*		
	(0.00588)	(0.0118)		
Hourly Wage: Mother	0.00922*	-0.0146	-0.0638**	-0.0122
	(0.00369)	(0.0131)	(0.0244)	(0.0107)
Constant	-2.921***	10.59***	29.68***	18.85***
	(0.824)	(3.134)	(3.443)	(1.905)
R^2	.13	1.1e-02	5.0e-03	1.7e-02

Standard errors in parentheses. N stands for non-parental child care.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In Tables 17 and 18, I report regression analyses I use to motivate two modelling assumptions: parents do not target investments by child gender, and initial skill endowments (at 9 months) affect skill outcomes later in life (at 4 years of age). Table 17 reports four models, each with a time input choice as the dependent variable. The first two are for married couples, the second two for single mothers. Time investments are predicted by attributes of the parents (hourly wages and educational attainment) and attributes of the child (current skill). Child gender is not a statistically significant predictor of parental time inputs, according to Table 17. There is some evidence in other studies that parenting behavior and treatment effects of the program vary by the gender of the child (see Garcia, Heckman, and Ziff (2018), Kottelenberg and Lehrer (2014)), but I do not see this my analysis.

Table 18: Predicting Final Skill with Initial Skill + Time Investments

	(1) Married Couples	(2) Single Mothers
Initial Test Score (9 Mo.): Stdzd	0.174***	0.179**
	(0.0396)	(0.0582)
Hourly Wage: Mother	-0.0000204	0.00164**
	(0.0000955)	(0.000598)
Hourly Wage: Father	0.00121***	
	(0.000299)	
Child is Female	0.0173*	0.0199
	(0.00779)	(0.0112)
B.A.: Mother	0.0573***	0.105***
	(0.00981)	(0.0162)
B.A.: Father	0.0626***	
	(0.0109)	
Constant	0.143***	0.122***
	(0.0200)	(0.0264)
R^2	0.1690	0.1104
Observations	2450	1000

Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Sample sizes rounded to nearest 50, following NCES requirements.

Table 18 reports two models, one for married couples and one for single mothers. The dependent variable of both in both models is the final skill of the child at age 4. Explanatory variables include the initial skill of the child, gender of the child, indicators for parental educational attainment (BA or higher), and parental hourly wages. Initial test scores are statistically significant predictors for final test scores, and so are parental attributes related to their skill. This motivates including heterogeneity in initial skill endowments in my model.

A.1.1 Correlation of Child Skill and Family Income

Table 19: Correlations of Skill and Family Income

	(1) Predicting Initial Skill	(2) Predicting Final Skill
Initial Income	7.23e-08	0.000000695***
	(7.09e-08)	(0.000000148)
Flag: Present in model 2 sample	-0.00543	
	(0.00532)	
Final Income		0.000000628***
		(0.000000116)
Initial Score		0.146***
		(0.0358)
Constant	1.456***	0.997***
	(0.00522)	(0.0526)
R^2	.0012	.1207
Observations	1500	1700
Correlation	.0165	.3216
Correlation p-value	.5104	0

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Sample sizes rounded to nearest 50, following NCES requirements.

Table 19 reports a slightly different version of the same qualitative points made with Tables 16, 17, and 18. The dependent variables in the two models are initial skill and final skill, with both family structures pooled. Initial income at 9 months has no predictive power for the initial skill score. At age 4, however, final skill can be predicted with income (both at age 9 months and 4 years) and the initial test score. In this table I also report the correlations of the dependent variable for the model (initial skill for model 1, final skill for model 2) with family income in the same period (initial and final, respectively). These correlations jump from zero to 0.32. The latter number is my target for the correlation of child's skill at the end of childhood with family income. Note that the measure of family income I use here is income before labor income taxes but including transfers. I use the analogous object in the model in the internal calibration step.

A.2 Regression Results with Alternative Nestings

As discussed in the body of this paper, alternative nestings of the CES function imply different restrictions on the elasticity of substitution between inputs. In this section of the Appendix, I report skill accumulation technology estimation results analogous to the one reported in the text, but for different nestings of the technology (Tables 20 and 21). For each nesting, I first give the functional form it assumes. In the titles of the different specifications, M stands for mother, F stands for father, and N stands for non-parental care. I give the innermost nesting of the couple problem (F,M) or (F,N), followed by the outermost one (N, not N), (M, not M). The main qualitative points survive in the alternative nestings. The first alternative nesting presented below, however, does illustrate that restricting the substitutability of non-parental child care with mother and father time to be the same implies unrealistically large complementarity values for mother and father time.

Table 20: Nesting: (F,M) + (N, not N)

	(1) Single Mother + Childcare	(2) Cohabiting Mother + Cohabiting Father	(3) Childcare + Not Childcare	(4) Outermost Estimation
$log(\frac{\text{after-tax hourly wage mother}}{\text{cost per hour child care}})$	-0.308*** (0.0856)		-0.559*** (0.0507)	
$log(\frac{\text{after-tax hourly wage father}}{\text{cost per hour child care}})$		-0.0547 (0.0349)		
$log(\text{composite term})$			-0.514*** (0.120)	
$log(\frac{\Lambda_{type,t}}{\Lambda_{type,t-1}})$				0.469*** (0.0292)
$log(\frac{r_{t-1}}{r_t})$				6.642*** (0.379)
Constant	-1.117*** (0.131)	-0.427*** (0.00543)	-0.179 (0.107)	0.729*** (0.0405)
R^2	.1356	.0022	.2294	.4566
Observations	1450	3100	3100	2600
complementarity	-2.25	-17.28	-.7895	-.8825
complementarity (se)	.9046	11.66	.1625	.1035
complementarity (p-value)	.0129	.1384	1.2e-06	1.5e-17
share	.0258	4.0e-04	.4207	.2534
share (se)	.0361	.0021	.052	.028
share (p-value)	.4749	.844	6.3e-16	1.6e-19

Standard errors in parentheses, clustered at the level of state of residence.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Sample sizes are rounded to nearest 50, following NCES requirements.

Table 21: Nesting: (F,N) + (M, not M)

	(1) Single Mother + Childcare	(2) Cohabiting Father + Childcare	(3) Cohabiting Mother + Not Cohabiting Mother	(4) Outermost Estimation
$log(\frac{\text{after-tax hourly wage mother}}{\text{cost per hour child care}})$	-0.298**	-0.363***		
	(0.0882)	(0.0657)		
$log(\text{composite term})$		-0.815**		
		(0.232)		
$log(\frac{\text{after-tax hourly wage father}}{\text{cost per hour child care}})$			-0.445***	
			(0.0458)	
$log(\frac{\Lambda_{type,t}}{\Lambda_{type,t-1}})$				0.493***
				(0.0251)
$log(\frac{r_{t-1}}{r_t})$				6.841***
				(0.384)
Constant	-1.132***	-0.313**	-1.015***	0.746***
	(0.135)	(0.0940)	(0.0650)	(0.0430)
R^2	.1262	.227	.1406	.4925
Observations	1450	3100	3100	
complementarity	-2.351	-1.757	-1.248	-.9714
complementarity (se)	.9908	.4994	.2314	.0976
complementarity (p-value)	.0176	4.3e-04	7.0e-08	2.4e-23
share	.022	.2768	.0927	.2296
share (se)	.0339	.0607	.032	.0236
share (p-value)	.5157	5.2e-06	.0038	2.6e-22

Standard errors in parentheses, clustered at the level of state of residence.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Sample sizes are rounded to nearest 50, following NCES requirements.

A.3 Description of Data Sources

A.3.1 The Early Childhood Education Longitudinal Survey, Birth Cohort

The ECLS-B follows a nationally representative sample of families with a child who was 9 months old in 2001. It was designed and collected by the United States Department of Education. Using birth-certificate data from the National Center for Health Statistics, over 14,000 births were selected within Primary Sampling Units. Children of mothers younger than 15 were excluded from the sampling frame. There are 5 waves: wave 1 is the 9-month old data collection round, wave 2 occurs at 2 years, wave 3 at 4 years, and waves 4 and 5 at kindergarten entry. If the focal child was not in kindergarten when wave 4 was collected, the surveyors went back and collected data the next year when they were enrolled. In addition, if a child repeated

kindergarten, their scores were also collected in wave 5 in addition to wave 4. Each wave contains several instruments; these are different self-administered questionnaires (SAQs) for different people in the child's life, in addition to the child-level data. Table 22 summarizes these instruments in each wave of the survey.

Table 22: The Structure of the ECLS-B

Instrument	Wave 1	Wave 2	Wave 3	Wave 4+5
1.	Parent Interview	Parent Interview + SAQ	Parent SAQ	Parent SAQ
2.	Resident Father	Resident Father SAQ	Resident Father SAQ	ECEP Interview[1]
3.	Nonresident Father	Nonresident Father SAQ	Preschool Center Director SAQ	Teacher
4.		Child Care Provider	Preschool ECEP SAQ[1]	WECEP Interview [2]
5.		Center Director		

[1] Early Care and Education Provider
[2] Wrap-around Care Early Care and Education Provider

In each wave of the survey, the primary care provider (usually the mother) and the resident father fill out detailed questionnaires on the activities they do with their kids and at what frequency (once a week, twice a week, once a month, etc.). In addition, they report age, educational attainment, income, hours worked, the number of hours the child spent in non-parental care, what type of care that was (relative, non-relative, center-based), and the cost of that care.

I define quality time as the total amount of time spent (1) reading to the child (2) talking with or listening to the child (3) singing to the child. To map from frequencies of activities to levels of quality time supplied by parents, I impute amount of time per activity using data from the ATUS. The imputation uses common characteristics observed across both samples: gender, marital status (married/cohabiting or single), labor force status, and educational attainment. Here educational attainment is less than a college degree, or a college degree or more. For hourly wages, I use time spent working and income to compute the pre-tax levels, and then Table 2 of McGrattan and Prescott (2017) to correct for labor income taxes. For hourly prices of non-parental care, I use total cost of child care and total hours in child care for the primary source of non-parental care reported by the primary caregiver of the survey child subject.

A.3.2 American Time Use Survey

I use a pooled sample from the ATUS from 2003 to 2016, which is provided with CPS variables on age, gender, marital status, labor force status, educational attainment and family structure

along with a time diary. I restrict the sample to parents who are between 15 and 55, with a child 4 years or younger. I use information on gender, marital status (married/cohabiting or single), labor force status, and educational attainment, where educational attainment is recoded as less than a college degree or a college degree and higher.

A.3.3 Imputation

In the baseline estimation, I focus on parents of children under 4 years old who are aged 55 and younger. In the ECLS-B this is respondents of waves 1 to 3. I additionally restrict observations to those who report age, educational attainment, and how often they participated in the activities of interest explained above. I further restrict the sample by excluding outliers: I only use those observations for whom after-tax hourly wages are greater than 0.1 cents and no greater than 200 dollars an hour. Finally, I only consider families observed in each and all of the first three waves of the survey, so that a family-specific mean can be computed for the fixed effects estimator.

Given these characteristics, I find survey-weighted averages of time spent on an activity for those who report engaging in it. The activities whose average duration I tabulate are (1) time spend reading to your child (2) time spent talking and listening. I assign the level of time spent reading and talking and listening to their respective activities in the ECLS-B; additionally, I assign the level of time spent talking and listening to the singing activity reported by parents in the ECLS-B. I do this because there is no singing activity reported in the ATUS.

Once I have imputed time levels per activity, I aggregate levels of quality time per parent in each family in each wave. This gives me quality time investments, non-parental time investments, after-tax hourly wages, and non-parental child care prices at the child-family pair and wave level. With this information, I can implement my estimation.

A.4 Summary Statistics for Raw and Estimation Sample in ECLS-B and Imputation Sample in ATUS

In the following two subsections I report sample summary statistics for the raw data and the estimation sample. For the raw data I report by wave of the sample, and for the estimation

sample I pool all the viable observations in the three waves and report statistics on those. I also report summary statistics for the ATUS sample I used to impute time levels to observed frequencies of activities in the ECLS-B sample.

A.4.1   Summary Statistics ECLS-B

The following summary statistics describe the ECLS-B data, pooled across family structures after the imputation from the ATUS, by wave and before data cleaning. The population moments I use in the internal calibration for the fraction of parents who are single mothers comes from this sample. The fraction in the sample that also reports variables necessary for estimation is larger than these population moments. Notice that the fraction below 185 % of the poverty line in the pooled sample is quite high, at 50%. This drop to 40% by the time the child is age 4. Averages here include observations for which the response is 0. This explains why the average age of the father is now lower than the mother's.

Table 23: Summary Statistics Raw Sample ECLS-B: Wave 1

	Wave 1				
	Obs.	mean	sd	min	max
Education Time Father: Hours per Week	10700	2.88	3.37	0.00	12.67
Education Time Mother: Hours per Week	10700	7.43	3.15	0.00	12.75
Below Poverty Line	10700	0.26	0.44	0.00	1.00
Below 185% of Poverty Line	10700	0.50	0.50	0.00	1.00
Marriage Rate	10700	0.64	0.48	0.00	1.00
Cohabitation Rate	10700	0.78	0.42	0.00	1.00
Single Mother	10700	0.21	0.40	0.00	1.00
Nonparental Care: Hours per Week	10700	14.32	18.53	0.00	70.00
Ratio of Time: Father/Mother	10400	0.44	0.64	0.00	9.93
Ratio of Time: Mother/Childcare	5200	0.54	0.96	0.00	12.75
Hourly Pay Mother, After Tax	5000	11.13	21.46	0.00	788.00
Hourly Pay Father, After Tax	6600	13.43	17.12	0.00	827.40
$\frac{\text{Cost per hour childcare}}{\text{After-tax hourly wage mother}}$	3500	3.41	177.25	0.00	10479.04
Hourly Cost Childcare	8250	1.25	3.37	0.00	99.50
Age Res. Mother	10700	28.46	6.64	-1.00	68.00
Age Res. Father	10700	25.16	14.83	-9.00	75.00
Mother: BA or higher	10700	0.26	0.44	0.00	1.00
Father: BA or higher	10700	0.25	0.43	0.00	1.00

Sample sizes rounded to nearest 50, following NCES requirements.

Table 24: Summary Statistics Raw Sample ECLS-B: Wave 2

			Wave 2		
	Obs.	mean	sd	min	max
Education Time Father: Hours per Week	10700	3.41	4.12	0.00	12.67
Education Time Mother: Hours per Week	10700	7.94	4.04	0.00	12.75
Below Poverty Line	10700	0.00	0.00	0.00	0.00
Below 185% of Poverty Line	10700	0.00	0.00	0.00	0.00
Marriage Rate	10700	0.59	0.49	0.00	1.00
Cohabitation Rate	10700	0.69	0.46	0.00	1.00
Single Mother	10700	0.19	0.39	0.00	1.00
Nonparental Care: Hours per Week	10700	14.04	18.51	0.00	80.00
Ratio of Time: Father/Mother	9350	0.46	0.57	0.00	9.83
Ratio of Time: Mother/Childcare	4800	0.48	0.76	0.00	12.67
Hourly Pay Mother, After Tax	5100	11.64	26.28	0.00	1004.70
Hourly Pay Father, After Tax	6200	15.14	52.93	0.00	3546.00
$\frac{\text{Cost per hour childcare}}{\text{After-tax hourly wage mother}}$	3900	348.68	21649.58	0.00	1345064.75
Hourly Cost Childcare	9000	1.05	2.31	0.00	64.00
Age Res. Mother	9850	29.79	7.02	-1.00	70.00
Age Res. Father	9850	26.20	15.38	-9.00	76.00
Mother: BA or higher	10700	0.33	0.47	0.00	1.00
Father: BA or higher	10700	0.32	0.47	0.00	1.00

Sample sizes rounded to nearest 50, following NCES requirements.

Table 25: Summary Statistics Raw Sample ECLS-B: Wave 3

			Wave 3		
	Obs.	mean	sd	min	max
Education Time Father: Hours per Week	10700	2.22	2.80	0.00	12.67
Education Time Mother: Hours per Week	10700	5.15	3.78	0.00	12.75
Below Poverty Line	10700	0.20	0.40	0.00	1.00
Below 185% of Poverty Line	10700	0.39	0.49	0.00	1.00
Marriage Rate	10700	0.52	0.50	0.00	1.00
Cohabitation Rate	10700	0.59	0.49	0.00	1.00
Single Mother	10700	0.17	0.37	0.00	1.00
Nonparental Care: Hours per Week	10700	16.53	16.52	0.00	96.00
Ratio of Time: Father/Mother	8100	0.48	0.54	0.00	6.57
Ratio of Time: Mother/Childcare	7100	0.44	0.60	0.00	12.67
Hourly Pay Mother, After Tax	5000	12.58	23.18	0.00	1063.80
Hourly Pay Father, After Tax	5400	15.95	16.97	0.00	506.57
$\frac{\text{Cost per hour childcare}}{\text{After-tax hourly wage mother}}$	3200	253.07	10139.75	0.00	419839.25
Hourly Cost Childcare	7500	2.19	3.83	0.00	64.71
Age Res. Mother	9000	32.28	7.56	-1.00	82.00
Age Res. Father	9000	28.16	16.36	-9.00	83.00
Mother: BA or higher	10700	0.41	0.49	0.00	1.00
Father: BA or higher	10700	0.39	0.49	0.00	1.00

Sample sizes rounded to nearest 50, following NCES requirements.

A.4.2 ATUS Imputation Data

For the imputation of time levels per activity, I used a pooled sample from the ATUS 2003-2016. I divide parents of children 3 and under into bins by labor force status, educational

attainment, marital status and gender. For each bin, I find the average time spent reading, doing educational activities, and playing with each bin's children conditional on having done each activity during the observation period. This is the level I assign to the ECLS-B sample for time spent talking and listening or time spent singing (educational hours) or time spent reading (reading hours). Table 23 displays group averages.

The population of these groups in the ATUS, even over such a large sample period, can be quite small. For example, only 85 married men not in the labor force with less than a BA reported reading to their children under 3 between 2003-2016. There were 35 single women with greater than a college degree (BA), not in the labor force, who reported spending educational time with their children aged 3 and under.

Table 26: ATUS Imputation: Group Means (hours per week)

Group	Reading Hours	Obs	Educational Hours	Obs
Married Male in LF lt BA	0.512	612	0.459	371
Married Male in LF gte BA	0.468	1318	0.528	337
Married Male not in LF lt BA	0.694	85	0.359	43
Married Female in LF lt BA	0.457	778	0.603	472
Married Female in LF gte BA	0.458	1799	0.565	601
Married Female not in LF lt BA	0.556	860	0.640	580
Married Female not in LF gte BA	0.565	1062	0.521	413
Single Female in LF lt BA	0.458	273	0.541	198
Single Female in LF gte BA	0.361	134	0.326	44
Single Female not in LF lt BA	0.501	195	0.670	194
Single Female not in LF gte BA	0.621	34	0.409	35

A.4.3 Measures of Skill in the ECLS-B

There are three skill measures from the ECLS-B that I use in the regressions presented in Tables 16 to 18, one for waves 1,2 and 3. In wave 1, the ECLS-B reports test scores from the Bayley Short Form - Research Edition, which is a shortened version of the Bayley Scaled of Infant Development, Second Edition (BSF-R and BSID-II, respectively). The latter exam is the standard one for measuring development in children under 42 months of age. The BSF-R is a shortened version of the BSID-II, asking only some of the questions. Its scores are then re-scaled to make them comparable with scores from children who receive the BSID-II. For the initial test score variable used in Tables 16 to 19, I take the scale scores of the BSF-R at 9 months (in the first wave of the ECLS-B), which are reported both for mental and motor development. I then take

the average of the two, and next I standardize them to lie between 0 and 1. For the test score value in the second wave, used in Table 17, I do the same procedure with the BSF-R scores recorded at 2 years of age (in wave 2 of the ECLS-B). By age 5, when the child is 48 months, the BSID-II and its subset exam the BSF-R are no longer an age-appropriate measures of development for children. Instead, the ECLS-B reports a new assessment battery that covers cognitive development in the domains of language, literacy, color knowledge, and mathematics. This is reported as the ECLS-B Direct Cognitive Assessment in several formats. I use the overall scale score of the Direct Cognitive Assessment and standardize it to lie between 0 and 1. This is the final test score used in the regressions of Table 18.

A.5   Child Support Payments from Single Fathers and Transfers as a Fraction of GDP

To discipline child support payments from single fathers, I use Table 3.12 from NIPA, Census tabulations on family counts in 2001, and the US Census Bureau report "Custodial Mothers and Fathers and Their Child Support: 2001". From US Census counts, I set the number of families in the US in 2001 to 71,787,347. Table 1.1.5 from NIPA gives the total amount of transfers in 2001 by components: I sum federal benefits from social insurance funds, Supplemental Nutrition Assistance Program (SNAP), supplemental security income, and refundable tax credits billion dollars. This totals 790 billion dollars. Combined with the number of families, this gives about 1,100 dollars per family in transfers in 2001. From the Census Bureau report, the average level of child support due to custodial mothers was 5,101 dollars in 2000. This makes the average child support payment due as a fraction of the average transfer to families about 0.46. Transfers as a fraction of GDP use Table 1.1.5 from NIPA in addition to Table 3.12. Total GDP in 2001 was 10,581 in billions of dollars. This makes transfers as a fraction of GDP 0.075.

A.5.1   Labor Supply in the CPS

To calculate the average labor supply for the data moment in the internal calibration, I use Annual Social and Economic Supplement of the CPS from 2000. I define hours worked per week as the self-reported hours worked last week. I assign a 0 for this value for those who

report not being in the labor force. I then find the average hours worked per week, weighted by the supplement household weight, for those between the ages of 20 and 80. This is a sample of 85,399 individuals. The resulting moment value is 28 hours of average labor supply, or 0.28 when expressed as a fraction of 100 disposable hours per week.

A.6 Robustness Checks

Work in progress.