DISCUSSION PAPER SERIES

Actors in the Child Development Process

Daniela Del Boca
Christopher Flinn
Ewout Verriest
Matthew Wiswall

DISCUSSION PAPER SERIES

Actors in the Child Development Process

Daniela Del Boca
University of Torino, Collegio Carlo Alberto and IZA

Ewout Verriest
NYU

Christopher Flinn
NYU and Collegio Carlo Alberto

Matthew Wiswall
University of Wisconsin-Madison and NBER

ABSTRACT

Actors in the Child Development Process[*]

We construct and estimate a model of child development in which both the parents and children make investments in the child's skill development. In each period of the development process, partially altruistic parents act as the Stackelberg leader and the child the follower when setting her own study time. We then extend this non-cooperative form of interaction by allowing parents to offer incentives to the child to increase her study time, at some monitoring cost. We show that this incentive scheme, a kind of internal conditional cash transfer, produces efficient outcomes and, in general, increases the child's cognitive ability. In addition to heterogeneity in resources (wage offers and non-labor income), the model allows for heterogeneity in preferences both for parents and children, and in monitoring costs. Like their parents, children are forward looking, but we allow children and parents to have different preferences and for children to have age-varying discount rates, becoming more "patient" as they age. Using detailed time diary information on the allocation of parent and child time linked to measures of child cognitive ability, we estimate several versions of the model. Using model estimates, we explore the impact of various government income transfer policies on child development. As in Del Boca et al. (2016), we find that the most effective set of policies are (external) conditional cash transfers, in which the household receives an income transfer given that the child's cognitive ability exceeds a prespecified threshold. We find that the possibility of households using internal cash transfers greatly increases the cost effectiveness of external cash transfer policies.

JEL Classification:	J13, D1
Keywords:	time allocation, child development, household labor supply

Corresponding author:
Daniela Del Boca
Department of Economics and Statistics
University of Turin
Via Lungodora 100 T
Torino 10153
E-mail: daniela.delboca@carloalberto.org

1   Introduction

In most models of the human capital investment, the subject of the investments is either a young and passive agent or an adult making investments to increase their own productivity. For example, Becker's model of fertility and investment in children views parents as choosing a quantity of offspring and an average quality (see, e.g., Becker and Tomes (1976)). However, the canonical model of educational choice and on-the-job investment views the person in whom investments are being made as the decision-maker (e.g., Becker (1964), Ben-Porath (1967)).[1]

Within the child development literature, the vast majority of theoretical models and empirical studies have considered children to be passive agents, especially when they are very young, and view parents as the only active decision-makers.[2] Empirical evidence suggests that the effect of parental investments on child development declines during adolescence (Carneiro et al. (2003), Del Boca et al. (2014)), whereas the effect of time spent in the child's self-investment (e.g., homework) increases (Cooper et al. (2006)). It is during adolescence, in fact, when most teenagers begin taking responsibility for their own actions, and when they begin to increase their investments in themselves (Dauphin et al. (2011), Lundberg et al. (2009), Kooreman (2007), Del Boca et al. (2017), Fiorini and Keane (2014)). A natural and potentially important question to address is the evolution of the individual from a passive receiver of investment inputs into one whose actions partially determine their own rate of cognitive development.

In this paper, we attempt to better understand this process and the role that incentives from parents and agents external to the household might play in it. We develop a model of child and parent interactions, in which both are active agents in the child's development process. We estimate this model using detailed time diary information on the time both parents and children spend investing in the child's human capital, and use the estimates of this model to study how policies could affect intra-household interactions and the path of skill developoment.

The starting point for our analysis is the observation that throughout childhood and adolescence the amount of time children spend "investing" alone (without their parents) increases markedly. Using data from 1997-2007 Child Development Supplement (CDS) of the Panel Study of Income Dynamics (PSID), we find that the average number of hours children spend studying alone increases from about 1.4 hours per week (ages 6-8) to about about 6 hours per week for ages 12-15, and in a quarter of households, teenagers are spending 10 or more hours per week studying on their own.[3]

We find that this child self-investment time also varies markedly with parental in-

---

[1]In some cases models of on-the-job investment in which there exist search frictions, it is assumed that workers and firms jointly invest in the general human capital of the worker and/or match-specific capital (Wasmer (2006), Bagger et al. (2014), Lentz and Roys (2015), Flinn et al. (2017)). In this paper we limit attention to the development of cognitive human capital prior to labor market entry, so that the joint investments are limited to parents and children.

[2]This literature is surveyed in Currie and Almond (2011).

[3]See Table 3 and Appendix tables. The CDS-PSID data and our time measures are described more fully in later sections.

come. Figure 1 shows the average hours of self-investment time per week by quartiles of parents' income. The average study time for the children of the highest income quartile households is nearly twice as large as that for the lowest quartile households (8.5 vs. 4.5 hours per week). This variation in study time is potentially an important contributor to differences in cognitive development if it represents a productive investment. The evidence on the productivity of child self-investment, much of it likely due to homework assignments in school, generally concludes that this time does in fact have a positive impact on development.[4] Our model contains mechanisms that allow for behaviors that are consistent with these observed relationships. The model emphasizes the advantages that better-educated and wealthier parents have in producing child quality directly and indirectly through the design of incentives offered to the child to increase her study time.

Our model of parent and child interaction considers a case in which the parents operate as a single decision-making unit, and we assume that they are able to choose their period t actions (labor supply, time with children, goods expenditure on children) prior to the child choosing hers (self-investment time). Thus, household decision-making has a Stackelberg structure within periods, not unlike the one used in (mostly) static models of inter-firm competition. The main differences here are that (1) the model is dynamic with fully forward-looking agents, (2) the parents act altruistically with respect to the child, and (3) parents and children both value the "public" good of the child's cognitive ability. The parent's altruism and shared public goods moderates the differences between the (private) objectives of the parents and the objectives of the child.

By explicitly incorporating the child's actions into her development process, a number of previously unexamined factors explaining the dispersion in cognitive outcomes at the end of adolescence can be examined. When modeling the human capital investment decisions of adults (in themselves or in their children), it is typically assumed that a common discount factor is applied to an additively separable lifetime welfare function. Studies by developmental psychologists (e.g., Steinberg et al (2009)) have convincingly demonstrated that the capacity to delay gratification changes markedly over adolescence, particularly around puberty. Because the motivation to invest depends critically on how forward-looking an agent is, it is important to allow for changes in this characteristic over the development period, and we use existing evidence to inform our parameterization of the child's age-varying discount factor sequence.

---

[4]There is an active debate in the education community about the role of homework in the learning process, and parents and children have their own views on homework that partially determine its measured effectiveness. In an often-cited overview and meta-analysis of the research on the relationship between homework and cognitive outcomes, Cooper et al. (2006) find that virtually all studies, whether they be based on cross-sectional survey data or small-scale experimental designs, find a positive (and significant) association between homework time and measures of cognitive performance for older children and adolescents. The fact that even the (small-scale) experimental studies find such a relationship means that this relationship is most probably not solely due to selection or omitted covariates. In the PSID-CDS data, child self-investment time is strongly positively correlated with cognitive test score measures.

We also allow the skill development production function parameters to change across development periods, reflecting the possibility that time with parents may be declining in productivity as children age, while the productivity of investment by children themselves may be increasing. In addition, our model of the skill development process allows for persistence in skill development ("skill begets skill"), for both time and goods (monetary) investments, and for the productivity of parental time to vary by the parent's level of human capital, as measured by their schooling attainment.

In our baseline specification, parents make their investment and consumption choices, as well as the level of expenditure on the private consumption good of the child. Given these choices, the child then chooses the amount of time to spend in study (self-investment), with the remainder of time outside of school consumed as leisure. In addition to allowing heterogeneity in household resources (wage offers for the mother and father, and non-labor income), the model allows for heterogeneity in preferences in leisure, consumption, and child "quality" across households, and also within households, since parents and children can "disagree" on the relative value of the child's human capital.

We generalize the model by allowing parents to incentivize their children to provide more self-investment time, which requires a monitoring cost borne by the parents. This structure allows even greater scope for parent and child interactions, and we show that by using this incentive scheme an efficient intra-household outcome is attained. However, due to the monitoring cost, not all households choose to use this incentivize scheme. We think of this, loosely speaking, as a "parenting style."

More broadly, our analysis centers on the the interaction between parents, children, and actors external to the household (which we will often refer to as a social planner) during the child development process. The interaction between parents and children has been frequently studied and is deemed to have important consequences for the welfare and development of children.[5] Relatively recent contributions by economists to this topic include Weinberg (2001), Hotz and Pantano (2015), Hao et al. (2008), Akabayashi (1996, 2006) and Lizzeri and Siniscalchi (2008). Weinberg (2001) develops a model in which altruistic parents provide for their child's consumption and attempt to influence behaviors of the child that lead to long-term beneficial outcomes for her. Weinberg shows that poor parents, i.e., those for whom the floor on their child's consumption is binding, have less scope for influencing the actions of their children. As a result, Weinberg's analysis predicts that less well-off parents are more likely to resort to non-pecuniary incentive mechanisms, such as corporal punishment, to influence their children. Akabayashi (1996, 2006) develops an incentive model of the parent-child re-

---

[5]Other related studies focus on adolescents' risky behaviors. Hao et al. (2008) show that all parents, including the very altruistic, have incentives to penalize older children for their risky behaviors in order to dissuade their younger ones from engaging in such behaviors. Hotz and Pantano (2015) consider a model in which parents impose more stringent disciplinary environments in response to their earlier-born children's poor performance in school in order to deter such outcomes for their later-born offspring. They provide evidence that the school performance of children in the NLSY-C declines with birth order as does the stringency of their parents' disciplinary restrictions.

lationship showing the potential for child-abuse. Although our paper does not include consumption floors that inhibit some parents' ability to positively influence their child's behavior, we do allow for a mechanism that restricts the ability of some households to use (positive) incentive schemes for their children, and households that do not use them will generally experience lower growth in the child's ability. As we shall show, wealthier households are more likely to use this incentivize device, which tends to further increase disparities in child outcomes at the end of the development period.

Given that at least some households actively employ incentive systems, it is of interest to investigate how conditional or unconditional transfers offered to the household by agents external to it modify household behavior, including the "parenting style" chosen. In terms of designing effective conditional cash transfers systems (CCTs), households enjoy considerable advantages with respect to external agents with an objective of increasing cognitive ability within a population of households. First of all, the parents can be assumed to know their own child's characteristics, in terms of both ability and preferences. This is a substantial advantage with respect to an external agent who may only know the distributions of ability and preferences of children and their parents in the population. A second advantage is in terms of the potential observability of the effort of the child, which is a direct input to the production process. While no parent can say with certainty how much "effective time" their child spends in study or other activities designed to increase her cognitive ability, they clearly have more information than does an agent external to the household. External agents will largely be constrained to offer outcome-based incentives, at least in more developed societies, and outcomes are likely to contain substantial amounts of noise which further complicates the design of effective mechanisms (Del Boca, et al. (2016)).

It is helpful to think of the parents as the "agents" within a standard principal-agent setting, with the social planner playing the principal. The principal's objective is to increase the level of the cognitive ability of children in the population by incentivizing the household members, parents and child, to provide more productive inputs than they would in baseline. All three parties value the child's cognitive ability to different degrees, and the problem the social planner faces is to design simple policies to achieve its goal given knowledge of the distributions of preferences and resources within the target population. We only consider very simple transfer policies of the type that have been implemented in various forms in the past. The most sophisticated policy we consider is a CCT policy in which the planner pays any household a fixed amount if their child's ability increases by a prespecified amount between ages t and $t + 1$. Such threshold-based policies produce strategic behavior in multi-period settings (see, e.g., Weitzman (1980), Macartney (2016)), and for this reason we only consider policies of this type that are unanticipated by the target population and last only for one period. Even within this limited class of policies, we find that relatively inexpensive CCT policies offered by the planner may be very cost effective when the parameters characterizing the policy are set optimally. Most importantly, we find that one of the reasons for the cost effectiveness of these policies is in promoting the use of more effective parenting styles, as characterized by the percentage of parents using an "Internal" CCT, henceforth referred

4

to as an ICCT, and the elasticity of the reward offered by parents with respect to the self-investment time of the child. Thus, we discover an important complementarity between an output-based CCT offered by the social planner, henceforth referred to as an "External" CCT (ECCT), and the use of ICCTs by parents.

The plan of the paper is as follows. In Section 2, we describe the model and characterize the solution in the situations in which the parents can or cannot offer incentives to their children for self-investment. Section 3 contains a description of the data used to estimate the model and presents descriptive statistics. In Section 4 we discuss the estimation method that we employ and the identification of the primitive parameters describing the model. Section 5 presents the model estimates and discusses within-sample fit. In Section 6, we discuss the design of efficient mechanisms to increase the cognitive outcomes of children, given our model estimates. We also explore the manner in which transfers and ECCTs impact household decisions and welfare. We conclude in Section 7.

2 Model

2.1 Model Primitives

The model consists of two (period) utility functions, one for the parents (jointly) and one for the child. The production technology for the cognitive ability of the child has a Cobb-Douglas form, as in Del Boca et al. (2014), with intertemporal linkages captured by the dependence of period $t + 1$ cognitive ability on period t cognitive ability, in addition to inputs chosen by the parents and the child in period t. There is no lending or borrowing in the model, so that expenditures on investment goods for the child and household consumption in any period t are equal to the sum of all income sources for the household in period t, including the labor income of each parent and the non-labor income of the household. We now describe the components of the model more formally.

2.1.1 Environment

There are two agents, the parents and the one child. Although we consider two parent households, and allow for different time allocation decisions for each parent, we assume the parents act as a unitary decision-maker. Both the parents and the child make decisions during each of the M (annual) child development periods, starting with the birth of the child at period $t = 1$ and ending at period $t = M$, when this phase of the child development process concludes. We can think of period $M + 1$ as the beginning of the next phase of the young adult's cognitive development process, which starts with an initial cognitive ability level of the child of k_{M+1}. We might think of the next phase as being one in which the child is the sole decision-maker in her investment decisions.

2.1.2 Preferences

We begin by describing the preferences of the parents, considered as one player, and the child. Parents are allowed to be altruistic, although we preclude the child from acting altruistically toward the parents. All utility functions are assumed to be Cobb-Douglas, which greatly simplifies the solution of the model in the presence of so many choice variables.

The child's period t private utility is given by

$$
\begin{aligned}
u_{c,t} &= u_c(l_{c,t}, x_t, k_t) \\
&= \lambda_1 \ln l_{c,t} + \lambda_2 \ln x_t + \lambda_3 \ln k_t,
\end{aligned}
$$

where $l_{c,t}$ is the child's leisure (or play) time in period t, x_t is her consumption of a private good purchased in the market, and k_t is the child's cognitive ability at the beginning of period t. The preference weights are all strictly positive and $\sum_{i=1}^{3} \lambda_i = 1$, an inconsequential normalization. As will be true with respect to the parents' utility function, we allow for the fact that the vector λ may be heterogeneous in the population of households with a child.

Parents have both a "private" utility function and are altruistic toward their child. The parents' "private" utility is given by

$$
\begin{aligned}
u_{p,t} &= u_p(l_{1,t}, l_{2,t}, c_t, k_t) \\
&= \alpha_1 \ln l_{1,t} + \alpha_2 \ln l_{2,t} + \alpha_3 \ln c_t + \alpha_4 \ln k_t,
\end{aligned}
$$

where $l_{i,t}$ is the leisure of parent $i \in \{1, 2\}$ in period t, c_t is the level of consumption of a private good valued only by the parents, each α_j is strictly positive, and $\sum_{j=1}^{4} \alpha_i = 1$. We define this as the parents' private utility function because it does not include the altruism component. The "total" period t utility of the parents is given by

$$
\tilde{u}_{p,t} = (1 - \varphi) u_{p,t} + \varphi u_{c,t},
$$

where $\varphi \in [0, 1]$ indicates the extent of the parents' altruism, with $\varphi = 1$ indicating "pure" altruism on the part of the parents, and with $\varphi = 0$ indicating that they exhibit no altruistic behavior toward the child. We note that this specification resembles that of a Benthamite social welfare function for the household, with $(1 - \varphi)$ being the weight given the parents and φ being the weight given to the child. In this sense, we can think of the parents as having preferences consistent with those of a social planner.

Substituting the expression for the child's utility into the parents' utility yields the parents' period t utility function

$$
\tilde{u}_{p,t} = \tilde{\alpha}_1 \ln l_{1,t} + \tilde{\alpha}_2 \ln l_{2,t} + \tilde{\alpha}_3 \ln c_t + \tilde{\alpha}_4 \ln k_t + \tilde{\alpha}_5 \ln l_{c,t} + \tilde{\alpha}_6 \ln x_t,
$$

where

$$\begin{aligned}
\tilde{\alpha}_1 &= (1-\varphi)\alpha_1 \\
\tilde{\alpha}_2 &= (1-\varphi)\alpha_2 \\
\tilde{\alpha}_3 &= (1-\varphi)\alpha_3 \\
\tilde{\alpha}_4 &= (1-\varphi)\alpha_4 + \varphi\lambda_3 \\
\tilde{\alpha}_5 &= \varphi\lambda_1 \\
\tilde{\alpha}_6 &= \varphi\lambda_2.
\end{aligned}$$

This expression makes clear that, due to parental altruism, the child's consumption, leisure, and human capital are "public" goods, that both agents, the parents and the child, enjoy. We note that for $\varphi = 0$, the parent's total utility is simply equal to their private utility and the parents place no value on the leisure or private consumption of the child. For $\varphi = 1$, the parents' utility function is the same as the child's, which implies that all household choices would be made so as to maximize the child's welfare, narrowly defined. In such a case, both sets of players have the same objective, and the problem reduces to a single agent problem. However, this case produces several counterfactual implications, such as that the parents would consume no leisure and would set $c_t = 0$ in every period.

An additional crucial aspect of preferences, in particular with regard to dynamic decision-making, are the agents' discount factors. We allow the child's discount factor to vary with their stage in the development process. This is indicated by the fact that the child's discount factor is indexed by t, $\beta_{c,t}$, with $\beta_{c,1} \leq \beta_{c,2} \leq \ldots \leq \beta_{c,M}$. This non-standard generality follows studies by developmental psychologists (e.g., Steinberg et al. (2009)) which have demonstrated that the capacity to delay gratification changes markedly over adolescence. The parent's discount factor, β_p is not indexed by time and is assumed to be constant over the course of the development process. For children, the age-varying discount factor implies that they value their investments in human capital (versus current leisure and consumption) differently as they age, and in particular younger children value it less than older children.

Our assumption regarding the child's discount factor does not imply any type of time-inconsistent behavior, such as is generated in models with hyperbolic discounting (e.g., Fang and Silverman (2009)). Given our assumption that the child's discount factor sequence is nondecreasing in the child's age combined with our assumption of rational decision-making on the part of both parents and children, a young child of age t makes decisions using the age t discount factor $\beta_{c,t}$, however their decisions reflect their knowledge that the future values of discount factors they use are no less than the current value. Although this assumption imposes a large degree of rationality on young children, its practical significance may not be great if the current discount factor is low and larger discount factors only emerge gradually. By making this assumption we avoid problems of time inconsistency, which may be particularly troublesome when there are strategic interactions between agents.

We assume the parents and the child also value the level of human capital produced at the end of the development period, k_{M+1}. With two sets of agents, we allow there to be two different terminal value weights, $\psi_{c,M+1}$ and $\psi_{p,M+1}$, for the child and the parents, respectively. These terminal weights in period $M+1$ are given by

$$
\begin{aligned}
\psi_{c,M+1} &= \xi_c \lambda_3 \\
\psi_{p,M+1} &= \xi_p \alpha_4,
\end{aligned}
$$

where ξ_c and ξ_p are positive constants to be estimated. The terminal period payoffs for agent i are given by $\psi_{i,M+1} \ln k_{M+1}$, $i = c, p$.

2.1.3 Child Development Technology

The other primitive of the model is the production technology. The law of motion for child quality is given by

$$
\ln k_{t+1} = \ln R_t + \delta_{1,t} \ln \tau_{1,t} + \delta_{2,t} \ln \tau_{2,t} + \delta_{3,t} \ln \tau_{12,t} + \delta_{4,t} \ln e_t + \delta_{5,t} \ln \tau_{c,t} + \delta_{6,t} \ln k_t,
$$

where $\tau_{i,t}$ is the amount of time parent $i \in \{1,2\}$ spends in child investment when the other parent is not present, $\tau_{12,t}$ is the time the parents spend in child investment when the parents are together, e_t is the amount of child investment goods purchased in period t, and $\tau_{c,t}$ is the time the child spends in self-investment. Note that the production function parameters are indexed by the child's age, t, allowing them to vary with the child's stage of development. Descriptive information from the CDS leads us to believe that the value of the child's time in self-investment, $\tau_{c,t}$, is increasing in the age of the child.

In this specification of the production technology, we interpret R_t as total factor productivity (TFP), and the Cobb-Douglas parameters are allowed to be time-varying functions of time-invariant, observable characteristics of the parents (e.g. parental education). We provide more details on the specification when we discuss econometric issues below.

2.1.4 Constraints

Each parent has an amount of time T to allocate to leisure, investment in the child, and market work. The time spent in the labor market by parent $i = 1, 2$ in period t is given by $h_{i,t}$, where we allow either or both parents to not work at all. The wage offer available to parent i in period t is given by $w_{i,t}$, and the household's non-labor income in period t is given by I_t. Then the total income of the household in period t is

$$
Y_t = w_{1,t} h_{1,t} + w_{2,t} h_{2,t} + I_t.
$$

There are no capital markets available for transferring consumption between periods, so that the household spends all of its period t income on investment goods for the child,

8

the private consumption of the child, and the private consumption of the parents.[6] Then

$$Y_t = e_t + x_t + c_t,$$

where we have assumed that the prices of all investment and consumption goods are equal to 1.

In terms of the time constraints, in each period $t = 1, ..., M$, for the parents we have

$$T = l_{i,t} + h_{i,t} + \tau_{i,t} + \tau_{12,t}, \ \ i = 1, 2,$$

and for the child we have

$$T = l_{c,t} + \tau_{p,t} + \tau_{c,t} + s_t,$$

where $\tau_{p,t} = \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$, the total time spent with the parents in investment activities, and where s_t is time spent in school at age t. We assume that s_t is exogenously determined, with any child quality gains to schooling reflected in the TFP factors $R_1, ..., R_M$. In what follows we will define $\tilde{T}_t \equiv T - s_t$, which is the child's discretionary time outside of formal schooling.

2.2   Coordination

We begin by assuming the following (noncooperative) decision-making structure. In each period, the parents are the first movers, and they choose all actions except for the time spent in self-investment by the child, which is chosen by the child. Because we assume that all household income is received by the parents, it is natural to have them make all money expenditure decisions. We also assume that parents can choose the amount of time they spend with the child, and that the child is obligated to spend that time with the parents.[7]

In any period t, the child is the Stackelberg follower and the parents are the Stackelberg leaders. Given the parent's choices $a_{p,t}$, the child allocates her available time between leisure in period t and self-investment, $\tau_{c,t}$. The decision rule of the child (reaction function) will therefore have the form $\tau_{c,t}^*(a_{p,t})$. The parents choose their actions $a_{p,t}$ given the child's time allocation rule (reaction function), including their three time investments with the child, their labor supplies in the period, $h_{1,t}$ and $h_{2,t}$, and their allocation of household income across their own private consumption, c_t, the private

---

[6]The private consumption goods of the parents may include goods that are best thought of as public, in the sense that the child also profits from them. This includes housing, heat, transportation services, etc. To simplify the exposition of the model, we have assumed that the child does not perceive these expenditures as contributing to their welfare, so that, in this sense, they are private consumption expenditures of the parents.

[7]Any parent of adolescents could easily take issue with this assumption, but practically speaking, the assumption will not be that objectionable. This is due to the fact that in the data, time with parents decreases dramatically over the development period, beginning when the child enters formal schooling.

consumption of the child, x_t, and expenditures on child investment goods, e_t.[8] Parents and the child are forward-looking, and fully account for the impacts of current actions on future state valuations. As discussed above, children make decisions fully understanding that their valuation of future events will be increasing as they age.

In each period t of the development process, the household's state variables are the vector $\Gamma_t = (w_{1,t} \ w_{2,t} \ I_t \ k_t)$, which are the wage draws of the parents in the period, the non-labor income received by the household, I_t, and the beginning of period value of child "quality," or cognitive ability in our application.[9] In order to simplify the problem, the wage and non-labor income processes are assumed to be exogenous with respect to household actions, although the wage draw of parent i in period t will only be observed if parent i supplies a positive amount of time to the labor market in period t.

The value function for the child in periods $t = 1, \ldots, M$ is given by

$$V_{c,t}(\Gamma_t|a_{p,t}) = \max_{\tau_{c,t}|a_{p,t}, \mathbb{C}_{c,t}} u_c(l_{c,t}, x_t, k_t) + \beta_{c,t} \mathbb{E} V_{c,t+1}(\Gamma_{t+1}|\tau_{c,t}, a_{p,t}),$$

where $a_{p,t}$ denotes the actions of the parents in period t, which occur prior to the child's selection of the value $\tau_{c,t}$, and $\mathbb{C}_{c,t}$ denotes the choice set of the child in period t, which was defined above. The parents' problem is similarly structured. Being the leaders in terms of action choices in period $t = 1, \ldots, M$, the parents' problem is

$$V_{p,t}(\Gamma_t) = \max_{a_{p,t}|\tau_{c,t}^*(a_{p,t}), \mathbb{C}_{p,t}} \tilde{u}_p(l_{1,t}, l_{2,t}, c_t.k_t, l_{c,t}, x_t) + \beta_p \mathbb{E} V_{p,t+1}(\Gamma_{t+1}|a_{p,t}),$$

where the parents' actions are chosen given the child's period t reaction function $\tau_{c,t}^*(a_{p,t})$ and their choice set $\mathbb{C}_{p,t}$.

The functional forms we have specified produce solutions that have attractive properties from the point of view of solving for the decision rules of the agents and for the estimation of model parameters. However, because of the strategic aspects of the game played between parents and the child, the solutions for the parents' problem no longer have a closed form as they did in Del Boca et al. (2014). We leave these solution details to Appendix A, and we focus here on some key aspects of the interaction between the parents and the child.

The child's reaction function determining their optimal self-investment time $\tau_{c,t}$ for any period t is given by

$$\tau_{c,t}^*(\tau_{p,t}) = \gamma_t(\tilde{T}_t - \tau_{p,t}) \tag{1}$$

---

[8]If the parents are not altruistic, i.e., $\varphi = 0$, the parents' contribution to the child's private consumption is not well-defined. That is one technical motivation for at least allowing a minimal level of altruism on the part of the parents, although we believe most parents exhibit considerable altruism with respect to their children.

[9]Although the child's discount factors are changing with age, as is the cognitive ability production technology, these functions are common to all households with the same observed characteristics, and are not explicitly included in the list of state variables.

where $\tilde{T}_t = T - s_t$ is the time available to the child outside of formal schooling time in period t, s_t, and $\tau_{p,t}$ is total time spent with the parents in period t, with $\tau_{p,t} \equiv \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$. The proportion of uncommitted time in period t that is devoted to self-investment is given by

$$\gamma_t = \frac{\Delta_{c,t}}{\lambda_1 + \Delta_{c,t}} \in (0,1),$$

where

$$\Delta_{c,t} = \beta_{c,t} \psi_{c,t+1} \delta_{5,t},$$

which is the present value of the marginal increase in the child's time investment in period t. It captures the "return" to self-investment from the perspective of the child. This return is the product of the child's discount rate $\beta_{c,t}$, the future value of the child's human capital $\psi_{c,t+1}$, and the productivity of self-investment time in producing human capital. The scalar $\psi_{c,t+1}$ is defined as

$$\psi_{c,t+1} = \frac{\partial \mathbb{E}V_{c,t+1}}{\partial \ln k_{t+1}},$$

so that $\psi_{c,t+1}$ represents future utility flows from the child's human capital as perceived by the child (including the preference for own human capital, λ_3, and future discount factors and technology parameters). Appendix A provides the explicit solution for $\psi_{c,t+1}$.

The child's reaction function (1) can be used to determine the rate at which parental time "crowds-out" child self-investment time, with

$$\frac{\partial \tau_{c,t}^*(\tau_{p,t})}{\partial \tau_{p,t}} = -\gamma_t.$$

This expression indicates that every hour of parental time reduces the child's self-investment time by $\gamma_t \in (0,1)$. γ_t then indicates the degree of parental "crowd-out" of child self-investment time. From this expression, some intuition for the child's behavior in response to her parents' time spent with her is clear. Impatient children ($\beta_{c,t} \approx 0$) have a $\gamma_t \approx 0$ and invest little of their own time in skill development. In this case there is a low degree of crowd-out. As children become more patient ($\beta_{c,t}$ increases) and as they become more productive in increasing their own cognitive ability (i.e., as $\delta_{5,t}$ increases), children spend more time in self-investment for any given amount of parental time and the degree of crowd-out by parents' time investments increases.

Although the parents' choices of actions no longer have closed form solutions, it remains true that all choices in the model are unique. The reaction function of the child in terms of her self-investment time is a linear function of the amount of time that the parents choose to spend with her. Substituting this reaction function into the parents' choice problem makes this a single-agent optimization problem that produces unique solutions for the parents' choices in the period, $a_{p,t}$. The choices made by the parents and child in each period are unique and well-defined under our functional form assumptions.

2.3   Expanding the Policy Space: Internal CCTs

We now reconsider the model after allowing the parents to offer the children contracts that have the possibility of increasing parental welfare. These policies, to a limited extent, mimic the kind of policies that an external social actor may offer the household. They will differ in important ways, however, which may give the parents an advantage over a social planner in implementing effective conditional cash transfer (CCT) schemes.

There are essentially two major advantages enjoyed by the parents. First, a conditional cash transfer offered by an external agent, who is assumed to be unable to observe the investments of the household members, must be based on an output measure.[10] This means that there is a delay between the investment costs of the household and the reward. While some parents may have a fairly high discount factor, so that such a delay in receiving payment is not a major disincentive to effort choices, children may have low discount factors, and the prospect of a delayed reward may reduce their incentive to increase current period investments. If the child could be paid immediately for increases in their current period investments, this will make it less costly to incentivize children to reach any given targeted "quality" level.

Second, the parents have more information concerning the child's preferences (and their own) than does an external agent. This allows them potentially to design more effective CCTs than can an external agent. As we will see, the parents will be able to design an internal CCT (ICCT) that allows them to obtain their preferred outcome subject to the child being no worse off than she is in the Stackelberg equilibrium. The parents have significant informational advantages with respect to the planner. However, the choices of the planner and the parents in incentivizing child development cannot be directly compared, since, in general, they have different objectives. In the case of the external agent, we will assume that their only concern is with some function of the distribution of child quality in the targeted population, which clearly differs from the objective of the parents, which is to maximize their discounted stream of payoffs. These payoffs are only partially a function of their own child's quality.

We now demonstrate how the use of an ICCT allows the household to achieve a cooperative, meaning efficient, outcome. In this two-agent setting, the set of efficient outcomes can be found by allowing each agent to make all of the decisions, subject to a constraint that the other agent receive a payoff no less than some minimal level. More formally, let $V_i(a_1, a_2), i = 1, 2$, be the payoff to agent i given the actions of herself and the other agent. We have some baseline behavior, which in our case is given by the Stackelberg equilibrium outcomes. In the Stackelberg equilibrium, the parents (agent 1) choose all actions with the exception of the child's time investment, τ_c, which is

---

[10]It is true that a number of well-known CCT experiments, such as PROGRESA in Mexico, provide payments to households for what may be considered as inputs into the production technology of cognitive development. In that experiment, most payments to parents were given for keeping their children in school. Even in more developed settings, inputs may be incentivized. One prominent example is Opportunity NYC-Family Rewards, which incentivized parents to take children to doctors and dentists. However, the largest incentives were targeted towards outputs, such as a child's grades and examination performance in school.

identified as a_2. Then the welfare outcomes in the Stackelberg equilibrium are given by

$$V_1^S = V_1(a_1^S, a_2^S); \ V_2^S = V_2(a_1^S, a_2^S),$$

where a_j^S are the (unique) Stackelberg actions of player j. In order to trace out the Pareto frontier associated with efficient outcomes, we allow one agent, agent 1 for example, to choose both a_1 and a_2, subject to the constraint that agent 2 has no lower payoff than $\underline{V}_2$. Then one point on the Pareto frontier is given by $(V_1^P(\underline{V}_2), \underline{V}_2)$, where

$$V_1^P(\underline{V}_2) \ = \ \max_{a_1, a_2} V_1(a_1, a_2)$$
$$\text{s.t. } V_2(a_1, a_2) \ \geq \ \underline{V}_2.$$

As we will show below, in our application the constraint $V_2(a_1, a_2) \geq \underline{V}_2$ will always be strictly binding. Then the Pareto frontier is defined by the set of points

$$(V_1^P(\underline{V}_2), \underline{V}_2), \ \underline{V}_2 \in \aleph,$$

where $\aleph$ is a set of feasible potential welfare levels of agent 2.

Now we discuss how the ICCT fits into this setting. The mechanism we envisage is a mapping from the observed effort choice of the child, which is their study time $\tau_{c,t}$ in period t, into the private consumption of the child in the period, x_t. We first observe that the parents can formulate a simple mapping that allows them to manipulate the child to choose any $\tau_{c,t}$ that they desire, for any given (feasible) value of child consumption x_t, *conditional on the child participating in the mechanism.* This incentive scheme is not unique. In what follows, we utilize the incentive function,

$$x_t = \exp(b_t + r_t \ln \tau_{c,t}), \ t = 1, ..., M.$$

which contains the two parameters, b_t and r_t, and has some appealing features from the point of view of characterization of the household equilibrium and computation. For any choice of b_t and r_t, and for any strictly positive $\tau_{c,t}$, this function restricts the consumption of the child to be positive in all periods, which is required under our functional form assumptions regarding parental and child utility.

In any development period t, given the child's participation in the incentive scheme, her best response to this policy function is given by

$$\tau_{c,t}^*(\tau_{p,t}, r_t) \ = \ \frac{\lambda_2 r_t + \Delta_{c,t}}{\lambda_1 + \lambda_2 r_t + \Delta_{c,t}}(\tilde{T}_t - \tau_{p,t})$$
$$= \ \gamma_t(r_t)(\tilde{T}_t - \tau_{p,t}),$$

where

$$\gamma_t(r_t) = \frac{\lambda_2 r_t + \Delta_{c,t}}{\lambda_1 + \lambda_2 r_t + \Delta_{c,t}}.$$

We note that when $r_t = 0$, this reaction function is that of the baseline Stackelberg equilibrium in which the parents make a fixed transfer of x_t to the child that is not

tied to the child's investment time. We see that the function $\tau_{c,t}^*$ is continuous and differentiable in r_t, and that

$$\frac{\partial \tau_{c,t}^*}{\partial r_t} = \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2 r_t + \Delta_{c,t})^2} > 0,$$

so that the function is monotone increasing and invertible. Thus, for any value of $\hat{\tau}_{c,t}$ that the parents desire, there exists a unique value of r_t that generates it, subject to the child participating in the incentive scheme.

Since the parents' choice of $\hat{\tau}_{c,t}$ is attained using only one parameter of the incentive function, r_t, the other parameter, b_t, can be used to select the value $\hat{x}_t$ desired by the parents. Then given $\hat{r}_t(\hat{\tau}_{c,t})$, we have

$$x_t = \exp(b_t + \hat{r}_t(\hat{\tau}_{c,t}) \ln \hat{\tau}_{c,t}),$$

so that for any feasible $\hat{x}_t$, we have

$$\hat{b}_t = \ln \hat{x}_t - \hat{r}_t \ln \hat{\tau}_{c,t}.$$

For any feasible choice of the parents, $(\hat{x}_t, \hat{\tau}_{c,t})$, there is a unique pair $(\hat{b}_t, \hat{r}_t)$ that implements it.

Proposition 1 *The child's participation constraint is always binding under the ICCT.*

Proof. Whether or not an ICCT is employed, the parents make all decisions in the household in each period given the reaction function of the child that they face. In either case, the parents' problem is

$$V_{p,t}(\Gamma_t) = \max_{a_{p,t} | \tau_{c,t}^*(\tau_{p,t}, r_t), \mathbb{C}_{p,t}} \tilde{u}_p(l_{1,t}, l_{2,t}, c_t, k_t, l_{c,t}, x_t) + \beta_p V_{p,t+1}(\Gamma_{t+1} | a_{p,t}),$$

where $a_{p,t}$ are the actions chosen by the parents at time t facing the choice set $\mathbb{C}_{p,t}$ and the reaction function of the child $\tau_{c,t}^*(\tau_{p,t}, r_t)$. In the case in which an ICCT is used, $r_t \neq 0$. However, in either case, the objective function of the parents is linear in the log of the child's time investment, which is $\ln(\tau_{c,t}) = \ln(\gamma_t(r_t)(\tilde{T}_t - \tau_{p,t})) = \ln \gamma_t(r_t) + \ln(\tilde{T}_t - \tau_{p,t})$ and in the log of the child's leisure, $\ln(l_{c,t}) = \ln(\tilde{T}_t - \tau_{p,t} - \gamma_t(r_t)(\tilde{T}_t - \tau_{p,t})) = \ln(1 - \gamma_t(r_t)) + \ln(\tilde{T}_t - \tau_{p,t})$. These are the only terms of the parents' objective function in period t that involve the reaction function of the child. Then all of the parents' first order conditions used to determine their actions in period t are chosen independently of the value of r_t. This implies that the parents will make all of the same choices in their (unconstrained) first-best equilibrium with the exception of r_t, which is equal to 0 in the Stackelberg equilibrium. In particular, this means that the parents choose the same level of child consumption in both cases, x_t^*. Then given this value, after r_t^* is chosen, we have

$$b_t^* = \ln x_t^* - r_t^* \ln(\gamma_t(r_t^*)(\tilde{T}_t - \tau_{p,t}^*)),$$

14

where the asterisks denote the optimal values under the ICCT. Thus all choices with the exception of r_t^* are the same in the Stackelberg and the ICCT cases.

Since the child receives the same amount of consumption, x_t^* in either equilibrium, whenever her leisure consumption differs from the level that she "freely" chooses in the Stackelberg equilibrium, her welfare level is lower in her parents' first-best equilibrium than it is in the Stackelberg equilibrium. This means that the parents' first best equilibrium choices are not achievable when the child can exercise the credible option of choosing the leisure level from the Stackelberg equilibrium. ∎

We now know that when using an ICCT, the parents will be constrained to alter their first-best choices so as to satisfy the participation constraint of the child. It will immediately follow that the constrained first-best choice of the parents will have the attractive property of efficiency.

Proposition 2 *The parents' constrained ICCT choice is efficient.*

Proof. When the parents are able to make all choices for both parties, the solution by definition is efficient since any other feasible choice must reduce the parents' welfare. In this case, the parents make all choices for the household under the ICCT, subject to the participation constraint of the child, which enforces the child's welfare to be no less than it is in the Stackelberg case. Thus the parents' constrained efficient choice lies on the Pareto frontier and corresponds to the point at which the parents receive all of the welfare surplus generated by the household moving from an inefficient outcome to an efficient one. ∎

We have shown that the parents weakly prefer the use of an ICCT to the noncooperative Stackelberg equilibrium. Given the continuity of the distribution of household preferences, the probability of indifference is zero, so that with probability one parents achieve a strictly higher welfare outcome under an ICCT than without one.

2.4   Subgame Perfection with an Internal CCT

We now consider whether the Internal CCT is subgame perfect, in the sense that the mechanism can be supported in this finite horizon setting where the parents may have the opportunity to renege on the promised level of child consumption given the child's choice of $\tau_{c,t}$ under the contract characterized by (b_t, r_t).

We begin by considering the case in which the parents utilize an ICCT in the last period of child development, M. We have already established that, with probability one, the parents cannot achieve their first best choice in any period in which the ICCT is utilized due to the fact that the child's participation constraint is binding at this point. This means that the parents have an incentive not to honor the terms of the contract in period M. This is a common issue in any finite horizon setting.

Our response to the question of commitment in the last period (or any other) is to rely on the special relationship between children and parents. We have assumed that parents are altruistic, and in fact if they were "perfectly" altruistic in our setting, there

would be no question of commitment since their preferences would be identical to the child's. Even when altruism is "imperfect," it may be reasonable to posit that parents would not renege on a promise made to their children since it would cause them a great deal of disutility. While such an argument could not possibly be made in the case of two competitive firms, it may be more reasonable here.

Assume that reneging on the promised consumption made in the ICCT in period M would yield the parents a disutility of $\zeta > 0$, so that if they renege on their ICCT promise in terms of period M child consumption, their period M value is given by

$$V_{p,M}^R(\Gamma_M) = \max_{x_M|\tilde{a}_{p,M}} u_{p,M}^R(\Gamma_M) - \zeta + \beta_p V_{p,M+1}(\Gamma_{M+1}),$$

where $\tilde{a}_{p,M}$ are all of the other actions chosen by the parents in period M, including the contracted value of $\tau_{c,M}$ from the incentives offered to the child, excluding the amount of consumption actually given to the child in period M. By reducing the consumption of the child below the promised amount, the parents can increase their own private consumption. Notice that the future value of the parents, $V_{p,M+1}(\Gamma_{M+1})$, is assumed not to change. This is due to the fact that all other choices affecting child cognitive ability production are assumed to remain the same whether or not the parents renege on the child's promised consumption level under the ICCT. Thus the only argument that changes is the child's period M consumption level. Let $\hat{u}_{p,M}^R$ denote the maximized value of parental utility in period M if the parents renege and let $\hat{u}_{p,M}$ denote their value of utility in period M if they do not renege, with $\hat{u}_{p,M}^R > \hat{u}_{p,M}$ with probability 1. Then the difference in the parents' value at time M from reneging and not reneging is given by

$$V_{p,M}^R(\Gamma_M) - V_{p,M}(\Gamma_M) = \hat{u}_{p,M}^R - \zeta - \hat{u}_{p,M}.$$

Then the parents will not renege, meaning their compliance with the ICCT is guaranteed, when

$$\zeta > \hat{u}_{p,M}^R - \hat{u}_{p,M}.$$

In other words, the disutility cost from reneging on the ICCT in period M must be greater than the short-term utility gain from the parents' increase in consumption.

We can easily extend this argument to early periods, $t = 1, ..., M - 1$. We will assume that the child uses a grim-trigger strategy so that the parents non-compliance with an ICCT in any period t results in the child not being willing to engage in an ICCT in any future period. Since the parents' welfare is unambiguously greater under an ICCT in any period in which they choose to use one, this means that this represents an additional non-negative cost to reneging on an ICCT in any period $t < M$. Denote this cost by $\mathbb{C}_t$. Then if the parents would never optimally use an ICCT in periods $t + 1, ..., M$, then $\mathbb{C}_t = 0$, while if they would employ an ICCT in any of the periods $t + 1, ..., M$ if the choice was available to them, then $\mathbb{C}_t > 0$. Thus, the difference in the value of reneging on a period t ICCT and not reneging is given by

$$V_{p,t}^R(\Gamma_t) - V_{p,t}(\Gamma_M) = \hat{u}_{p,t}^R - \zeta - \mathbb{C}_t - \hat{u}_{p,t}.$$

Then a sufficient condition for the parents to not renege on the child's promised consumption under an ICCT in any period t, $r = 1, ..., M$, is

$$\zeta > \hat{u}_{p,t}^{R} - \hat{u}_{p,t}, \ t = 1, ..., M.$$

We think of the disutility of disillusioning the child by not paying the promised reward, ζ, as a characteristic of a parent somewhat akin to altruism. For purposes of estimation, we have assumed that all parents share the same altruism parameter, φ, and for purposes of this discussion we have assumed that all parents share the same disutility of disillusioning their children, ζ. It may be that parents differ in their value of ζ, in which case some parents may not be able to offer ICCTs because their direct utility cost of reneging, ζ, is too low to make such a contract offer credible.

2.5 Choice of Using an Internal CCT

From the parents' point of view, an internal CCT delivers a higher level of welfare than does the case in which they make consumption tranfers to the child not explicitly tied to her choice of investment time. Then shouldn't all parents use such devices? Papers by Akabayashi (1996, 2006), Weinberg (2001), Lizzeri and Siniscalchi (2008), Cosconati (2009), and Hotz and Pantano (2015) specifically focus on the choices parents make among various forms of interaction with their children. When modeling parent-child relationships, it is highly desirable to account for population heterogeneity in these forms of behavior, and to make this choice endogenous. We describe how we can accomplish this in a very simple and tractable manner within our framework.

We denote the value of the parents' period t problem when using an ICCT by $\bar{V}_{p,t}(\Gamma_t)$, whereas the value of the parents' problem without an ICCT is denoted by $V_{p,t}(\Gamma_t)$. From our discussion above, we know that

$$\bar{V}_{p,t}(\Gamma_t) \geq V_{p,t}(\Gamma_t)$$

for all t and all Γ_t. In the case of an ICCT, the parents must monitor and measure the investment time of the child, which we believe is not a costless activity.[11] In any period t in which an ICCT is implemented, the parents' must pay a utility cost ω_t, with the population distribution of ω_t given by $R_t(\omega_t|Z_t)$, where the distribution R_t can depend on some observable or unobservable household characteristic vector Z_t and where the support of R_t is $\mathbb{R}_+$, and we assume that R_t is continuously differentiable on its support,

---

[11]This is reminiscent of the often debated assumption of efficiency in cooperative household models. In the work of Chiappori and coauthors, the assumption is made that all household choices are consistent with efficient behavior. These efficient choices are typically not best responses, and it is natural to question whether sufficient monitoring exists within the household to implement such arrangements. Lundberg and Pollak (1993) introduce the notion of cooperation costs in their analysis of household behavior, in which some households may not cooperate if the utility costs of doing so are sufficiently high. In a very stylized dynamic setting, Del Boca and Flinn (2012) formulate and estimate a model in which in equilibrium some households behave cooperatively and others do not.

yielding the density $r_t(\omega_t|Z_t)$. In terms of the parents' period t utility function, it is modified to be of the form

$$\hat{u}_{p,t} = u_{p,t} - \chi[\iota_t = 1]\omega_t,$$

where ι_t is an indicator variable that takes the value 1 when the parents use an ICCT in period t and that takes the value 0 if not. Because the utility cost is independent of all choices made in the period, the choices made in period t if an ICCT is used will be independent of the value of ω_t. The solution to the parents' period t problem in this case is exactly as was described above. By the same token, if an internal CCT is not used, the choices made are those made in the Stackelberg equilibrium.

We have added another choice to the parents' choice set in period t, which is ι_t. By the structure of the problem, it is straightforward to characterize this choice. Since decisions in all future periods are independent of the value of child quality, and since the only dynamic implication of the choice of using an ICCT or not is in terms of the child quality sequence, the current period choice of the parents to use an ICCT has no effect on future decisions. Because of this, the use of an ICCT in period t does not depend on whether an ICCT will be used in periods $t+1, ..., M$. The use of an ICCT in period t, will affect the distribution of next period's state vector, Γ_{t+1}. Then the value of using an ICCT in period t is

$$u_{p,t}(a_{p,t}^I) - \omega_t + \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^I, \Gamma_t)),$$

where we have dropped the child's actions a_2 from this expression since the parent effectively makes all decisions when an ICCT is utilized. The value of not using an ICCT in period t is

$$u_{p,t}(a_{p,t}^S, a_{c,t}^S) + \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^S, a_{c,t}^S, \Gamma_t)).$$

Then the parents will use an ICCT if

$$u_{p,t}(a_{p,t}^I) - \omega_t + \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^I, \Gamma_t)) > u_{p,t}(a_{p,t}^S, a_{c,t}^S) + \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^S, a_{c,t}^S, \Gamma_t))$$
$$\Rightarrow u_{p,t}(a_{p,t}^I) + \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^I, \Gamma_t)) - u_{p,t}(a_{p,t}^S, a_{c,t}^S) - \beta_p V_{p,t+1}(\Gamma_{t+1}(a_{p,t}^S, a_{c,t}^S, \Gamma_t)) > \omega_t.$$

Since the left-hand side of the second line is always positive, every household has a positive probability of implementing an ICCT in any period t.

To recap, the parents collectively have two choices regarding behavior. By paying the utility cost ω_t in period t, they are able to implement an efficient household allocation of resources in which they receive all of the surplus. If they do not pay this utility cost, they obtain a welfare level generated in the Stackelberg (inefficient) equilibrium. Their welfare under the first arrangement in the efficient equilibrium always exceeds their Stackelberg welfare level. Since child quality is a public good, the amount produced in the cooperative equilibrium will typically exceed the amount produced in the Stackelberg equilibrium.[12] A social planner solely concerned with the distribution of

---

[12]As noted above, there could be preference draws for which the efficient solution would involve a reduction in the child's human capital. Given our estimates, the likelihood of such draws is essentially zero.

child quality in the population would generally prefer that all households utilize ICCTs. We shall return to this point in Section 6, below.

3 Data and Descriptive Statistics

We utilize data from the Panel Study of Income Dynamics (PSID) and the first two waves of the Child Development Supplements (CDS-I and CDS-II). The PSID is a longitudinal study that began in 1968 with a nationally representative sample of about 5,000 American families, with an oversample of black and low-income families. In 1997, the PSID began collecting data on a random sample of the PSID families that had children under the age of 13 (CDS-I). Data were collected for up to two children in this age range per family. The CDS collects information on child development and family dynamics, including parent-child relationships, home environment, indicators of children's health, cognitive achievements, socio-emotional development, and time use. The entire CDS sample size in 1997 is approximately 3,500 children residing in 2,400 households. A follow-up study with these children and families was conducted in 2002-03 (CDS-II), and a third wave was released in 2007 (CDS-III). These children were between the ages of 8-18 in 2003. No new children were added to the study after the initial wave (Hofferth et al. (1998)).

3.1 Sample Selection

All of our results are based on a selected sample of households that satisfy several criteria. First, included households are intact over the observed period (i.e. only stable two-parent households). Second, households have either one or two children. We select only one child from each household (see below). Third, all children are biological; there are no adopted children and no step-parents. Fourth, all selected children are at least three years old during the first wave of the CDS (CDS-I in 1997), because we need access to an initial Letter Word (LW) score observation. Fifth, all selected children have an observed LW score in 1997 and in 2002. Some of these also have an observed LW score in 2007 (CDS-III) as well, although it is not required. Sixth, if a household has two eligible siblings satisfying the previous two requirements, we select the youngest. This has two potential advantages: parental labor supply is probably more responsive to the age of the youngest sibling than the age of the oldest sibling, and we also have a higher chance of observing the youngest sibling in CDS-III, which enriches the total sample. We keep only data observations for selected children whose age is between 0 and 16 at the interview date for waves 1 through 3.

Overall, these selection criteria yield a sample of $N = 247$ children (or households). Appendix B contains more detailed information about the sample selection procedure. Tables 1, 2 and 3 contain some descriptive statistics of our sample of households. Next, we discuss the various data components in more detail.

3.2 Child's Time Allocation

Starting in 1997, children's time diaries were collected along with detailed assessments of children's cognitive development. For two days within a week (one weekday and either Saturday or Sunday), children (with the assistance of the primary caregiver when the children were very young) filled out a detailed 24 hour time diary in which they recorded all activities during the day and who else (if anyone) participated with the child in these activities. At any point in time, the children recorded the intensity of participation of their parents: mothers and fathers could be actively participating or engaging with the child or simply around the child but not actively involved. In this paper, we partition productive child time into four categories of inputs: (1) active time with the mother alone, τ_1, (2) active time with the father alone, τ_2, (3) active time with both parents, τ_{12} and (4) the child's productive self-investment time, τ_c. Active time with the mother is defined as the total time where the mother is actively participating with the child, while the father is either not around or only passively participating. Active time with the father is defined analogously. Joint parental time is defined as the total time when both parents are simultaneously participating actively with the child. As such, the total parental time in investment activities ($\tau_p = \tau_1 + \tau_2 + \tau_{12}$) comprises all activities where at least one parent is actively involved. Finally, the child's self-investment time is defined as any "productive" activity where neither parent is actively involved, such as doing homework, studying, reading, solving puzzles or playing educational games. One or both of the parents could be present in the home during this time, but neither were actively interacting with the child. For each type of child investment time, we construct a weekly measure by multiplying the daily hours by 5 for the weekday report and 2 for the weekend day report (using a Saturday and Sunday report adjustment) and summing the total hours for each category of time.

It is likely that much of the child self-investment time we observe is due to homework assignments. Although the time spent in homework assignments is partially under the control of teachers or others outside of the household, the survey data created by and analyzed in Cooper et al. (1998) make clear that there is considerable variability in the amount of homework assigned by teachers in the same grade level, and that all homework assigned is not completed. Their survey instrument included questions only related to the completion of homework, and therefore does not measure any other type of "self-investment.", such as additional reading a child might do outside of homework assignments. For these reasons we believe that our measure of self-investment time is considerably broader than the homework time measures used in previous studies, such as those surveyed in Cooper et al. (2006).[13] Although teachers' homework assignments are an important factor, we argue that the level of self-investment observed in a household is largely determined by the actions of the parents and the child.

---

[13] The data produced by the National Educational Longitudinal Study (NELS) and other secondary databases that are most frequently used in assessing the relationship between study time and outcomes specifically ask for time spent in completing homework. Thus broader measures of child self-investment time cannot be constructed using these data sources.

Table 3 shows how these four types of time investment evolve as children age. We see a clear decrease in both mothers' and fathers' active time with their children as they age, although this decreasing trend is not so obvious for joint parental time. As children age, we see a clear increase in their average self-investment time. Taken together, the evolution of these four time investment choices highlights the intuitive notion that children become more independent as they age.

In the bottom row of Table 3, we show how the child's school time (s_t) evolves with child age. Although we observe school time in the CDS data, we believe these data to be relatively noisy. Table B-1 shows the detailed distribution of reported school time for every child age. Given the implausibly wide data range of these reported school times, we only use the median of these reported values (conditional on child age t), and use that as a measure to define the child's effective time endowment at age t as $\tilde{T}_t = 112 - med(s_t)$. Appendix B contains more details on how we constructed school time.

3.3   Parental Labor Supply, Wages and Income

By linking the time survey data from the CDS to the labor supply and income data from the PSID, we can complete the parents' weekly time allocation. We define weekly labor hours of mothers and fathers (h_1 and h_2, respectively) as the total yearly reported number of labor hours for each spouse, divided by 52. Parental "leisure" is the residual time not spent working or with the selected child. This "leisure" time can include time spent exclusively with a non-sample child, if the household has more than 1 child. Note that we do include any time spent with the sample child, even if that time is also spent with siblings.[14] From the perspective of the sample child, this definition of parental "leisure" is inconsequential: time spent exclusively with other children does not directly influence their own development.

Using the main PSID data, we define hourly wages ($w_{1,t}$ and $w_{2,t}$) as the total yearly reported income from labor for each spouse, divided by their total yearly hours. Non-labor income (I_t) is defined as the total yearly household income minus the yearly labor incomes reported for each spouse.

The monetary values have been deflated and are expressed in terms of 2007 dollars. All wage and income information is used in estimating the model. Note that since the PSID was only administered every two years after 1996, we do not have yearly data for labor supply and income. For the time use surveys and the children's test scores which we only observe every 5 years, these gaps in the data become even more salient. Our econometric implementation (which relies on the Method of Simulated Moments, discussed in Section 4) accomodates this data structure and makes use of all the available information we have.

Table 1 contains the unconditional averages of the parents' hourly wages and weekly non-labor income. Table 2 summarizes the parents' labor supply behavior for various

---

[14]See Del Boca et al. (2014) for an analysis of the allocation of time across siblings.

child ages, both on the extensive and intensive margins. Whereas fathers' labor supply does not seem to vary much with the child's age, mothers work fewer hours when the child is very young (age 3). Note that the composition of our sample includes households with two children for which, depending on the observable time survey and test score data for the children, we do not always select the youngest sibling. This may partially explain why we do not see a very clear upward trend in maternal labor force participation as the child ages, as has been documented in other studies. Finally, the first two rows of Table 3 contain the unconditional average work hours of either spouse, i.e. averaging over both working and non-working parents.

3.4   Skill Measures

Given the wide range of ages to which the Woodcock Johnson Letter-Word (LW) test was administered, we use this test as our measure of child development. We use the raw scores on this exam rather than the age-standardized scores, and allow our model describe the development process. The test contains 57 items (so that in terms of our discussion in Section 4, $NQ = 57$), and the range of possible raw scores is from 0 to 57. Our econometric framework accommodates the discreteness and the particular floor and ceiling of this test score measure. Table 1 shows the unconditional average test scores taken from each CDS wave. Figure 2 shows how the average test scores increase smoothly with child age.

4   Econometric Implementation and Identification

In this section, we discuss the functional form assumptions we make use of in estimation, how we use the combined PSID and CDS data to identify the parameters characterizing the model, and describe the estimation method.

4.1   Econometric Specification

4.1.1   Preferences

Household preferences are assumed to be fixed over time. However, we do allow for a rich distribution of heterogeneity in the utility functions of the parents and the child. Each household's utility parameters (α for parents, λ for child) are assumed to be an i.i.d. draw from the distribution $G(\alpha, \lambda; \theta)$, where G is a parametric distribution function characterized by the finite-dimensional parameter vector θ. For the parents' preferences, the four-dimensional vector $\alpha = (\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4)'$ is defined such that $\sum_j \alpha_j = 1$, $\alpha_j > 0$, $j = 1, \ldots, 4$. Similarly, the child's preferences are given by a three-dimensional vector $\lambda = (\lambda_1 \ \lambda_2 \ \lambda_3)'$, where $\sum_j \lambda_j = 1$, $\lambda_j > 0$, $j = 1, 2, 3$. These restrictions are standard and ensure that utility is increasing in each argument and the summations impose an inconsequential normalization on the utility functions of the parents and the child.

We allow a large degree of flexibility in the distribution G. The distribution is generated by the latent vector of random variables $v = (v_1\ v_2\ ...\ v_5)$, where v has a multivariate normal distribution with mean vector μ_v and covariance matrix Σ_v. The normal random variables v map into the preference parameters as follows:

$$\alpha_1 = \frac{\exp(v_1)}{1 + \exp(v_1 + v_2 + v_3)}, \quad \alpha_2 = \frac{\exp(v_2)}{1 + \exp(v_1 + v_2 + v_3)},$$

$$\alpha_3 = \frac{\exp(v_3)}{1 + \exp(v_1 + v_2 + v_3)}, \quad \alpha_4 = \frac{1}{1 + \exp(v_1 + v_2 + v_3)},$$

$$\lambda_1 = \frac{\exp(v_4)}{1 + \exp(v_4 + v_5)}, \quad \lambda_2 = \frac{\exp(v_5)}{1 + \exp(v_4 + v_5)}, \quad \lambda_3 = \frac{1}{1 + \exp(v_4 + v_5)}.$$

This mapping ensures that all preference parameters are positive and that $\sum_{i=1}^{4} \alpha_i = \sum_{i=1}^{3} \lambda_i = 1$. The vector μ_v and the matrix Σ_v contain a total of 20 free parameters, so that the distribution of G is quite flexibly parameterized. In practice, we found it difficult to obtain precise estimates of these 20 parameters, and as a result we have imposed a few restrictions on the parameters appearing in the covariance matrix of v. In particular, we have fixed $\sigma_{1,5} = \sigma_{2,5} = \sigma_{3,4} = 0$. Although this constrains certain elements of the v vector to be independently distributed under our normality assumption, the nonlinear mapping of v into the parameter space describing preferences generates dependence between all preference parameters, but with dependence between the parameters α_1 and λ_2, α_2 and λ_2, and α_3 and λ_1 more restricted. These parameter pairs consist of valuations of leisure by one agent (either the parents or the child) and the value of consumption of the other, which we do not expect to be strongly related.

We also estimate three additional preference parameters (ξ_p, ξ_c, κ). The first two parameters enter into the terminal (heterogeneous) valuations of child human capital, k_{M+1}, by the parents $(\psi_{p,M+1} = \xi_p \alpha_4)$ and the child $(\psi_{c,M+1} = \xi_c \lambda_3)$, respectively. The third parameter, κ, governs the distribution of the instantaneous utility costs associated with implementing an internal CCT. The utility cost, ω, is assumed to stay fixed over time but can vary across households.[15] For each household, the value of ω is drawn from an exponential distribution with mean and standard deviation κ. These three parameters $(\psi_{p,M+1}, \psi_{c,M+1}, \kappa)$ are assumed to be the same for all households. Taken together, this implies that the distribution of household preferences is characterized by 20 free parameters.

Finally, we allow for differences in the discount factors between parents and children, allowing for different degrees of "patience." The child's discount rate is allowed to be age-varying, with the child's level of patience increasing as she ages. The discount factors are determined outside of the model using experimental data generously supplied by Laurence Steinberg of Temple University. These data are described in Steinberg et al. (2009).

---

[15]It is straightforward computationally to allow there to be time varying cost ω , but we believe that this characteristic of households is likely to be highly persistent.

4.1.2 Production Technology

As noted above, we allow the production function parameters to vary with the age of the child and with respect to the educational levels of the parents in household h. We economize on parameters by assuming that the input-specific productivity parameters are given by

$$
\begin{aligned}
\delta_{h,1,t} &= \frac{\exp(\gamma_{1,0} + \gamma_{1,1}(t-1) + \gamma_{1,2}s_{h,1})}{1 + \exp(\gamma_{1,0} + \gamma_{1,1}(t-1) + \gamma_{1,2}s_{h,1})}, \\
\delta_{h,2,t} &= \frac{\exp(\gamma_{2,0} + \gamma_{2,1}(t-1) + \gamma_{2,2}s_{h,2})}{1 + \exp(\gamma_{2,0} + \gamma_{2,1}(t-1) + \gamma_{2,2}s_{h,2})}, \\
\delta_{h,3,t} &= \frac{\exp(\gamma_{3,0} + \gamma_{3,1}(t-1) + \gamma_{3,2}s_{h,1} + \gamma_{3,3}s_{h,2})}{1 + \exp(\gamma_{3,0} + \gamma_{3,1}(t-1) + \gamma_{3,2}s_{h,1} + \gamma_{3,3}s_{h,2})}, \\
\delta_{4,t} &= \frac{\exp(\gamma_{4,0} + \gamma_{4,1}(t-1))}{1 + \exp(\gamma_{4,0} + \gamma_{4,1}(t-1))}, \\
\delta_{5,t} &= \frac{\exp(\gamma_{5,0} + \gamma_{5,1}(t-1))}{1 + \exp(\gamma_{5,0} + \gamma_{5,1}(t-1))}, \\
\delta_{6,t} &= \frac{\exp(\gamma_{6,0} + \gamma_{6,1}(t-1))}{1 + \exp(\gamma_{6,0} + \gamma_{6,1}(t-1))}, \\
R_t &= \gamma_{7,0} + \frac{\gamma_{7,1} - \gamma_{7,0}}{1 + \exp(-\gamma_{7,2}(t - \gamma_{7,3}))}
\end{aligned}
$$

where h indexes the household and $s_{h,1}$ and $s_{h,2}$ are the mother and father's years of schooling, respectively. The productivity of the input is restricted to be monotonic in the age of the child, with the characteristics of the parents potentially entering as shifters of the profile. The parental schooling levels, meant to roughly capture the human capital levels of the parents, only appear in the δ productivity parameters connected with the time inputs of the parents (with mother's schooling appearing in the mother's time productivity parameter, father's schooling in the father's time productivity parameter, and both parents' schooling in the joint time productivity parameter). Our specification is intended to capture the possibility that parents with higher human capital can provide higher quality time inputs to their children. For the Total Factor Productivity process (R_t), we have adopted a flexible generalized logistic parametrization which imposes monotonicity in the child's age.

4.1.3 Wage Offers

The parental hourly wage offer processes are assumed to have the following structure for each parent $i = 1, 2$, in household h

$$
\ln w_{h,i,t} = \eta_{0,i} + \eta_{1,i} age_{h,i,t} + \eta_{2,i} s_{h,i} + \varepsilon_{h,i,t}
$$

where $age_{h,i,t}$ is the age of parent i when the child is age t. $\eta_{1,i}$ is a the age coefficient for parent i, and $\eta_{2,i}$ is the labor market "return" to schooling for parent i. The wage shocks

are assumed to be serially uncorrelated draws from the following joint distribution:

$$(\varepsilon_{h,1,t}, \varepsilon_{h,2,t}) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{w_1}^2 & \sigma_{w_1,w_2} \\ \sigma_{w_1,w_2} & \sigma_{w_2}^2 \end{bmatrix} \right)$$

where σ_{w_1,w_2} $(= \rho_{12}\sigma_{w_1}\sigma_{w_2})$ is the contemporaneous covariance between the parental wage shocks. Any correlation between these disturbances $(\rho_{12} \neq 0)$ could arise through assortative mating on unobservable determinants of wages and from the parents inhabiting the same local labor market. Although we assume the shocks are independent over time, this is not implied by the model structure and is not a necessary assumption for identification of the key model parameters. We can allow for temporal dependence at some computational cost.

4.1.4 Non-Labor Income

We restrict non-labor income $I_{h,t}$ to be non-negative for all households and periods. We assume non-labor income is a function of the parent's ages and their schooling. In particular, we first define the probability of household h having a strictly positive non-labor income level in period t as

$$Pr(I_{h,t} > 0) = \frac{\exp(\zeta_{h,t})}{1 + \exp(\zeta_{h,t})}$$

where

$$\zeta_{h,t} = \mu_1 + \mu_2 age_{h,1,t} + \mu_3 age_{h,1,t}^2 + \mu_4 age_{h,2,t} + \mu_5 age_{h,2,t}^2 + \mu_6 s_{h,1} + \mu_7 s_{h,2}.$$

If non-labor income is positive, the level of non-labor income is determined as follows:

$$\ln I_{h,t} = \mu_8 + \mu_9 age_{h,1,t} + \mu_{10} age_{h,1,t}^2 + \mu_{11} age_{h,2,t} + \mu_{12} age_{h,2,t}^2 + \mu_{13} s_{h,1} + \mu_{14} s_{h,2} + \varepsilon_{h,t}^I,$$

where the shock $\varepsilon_{h,t}^I$ is assumed to be i.i.d. $N(0, \sigma_I^2)$ for all h, t.

The non-labor income parameters can be identified outside of the remaining model structure because the non-labor income process is strictly exogenous and $I_{h,t}$ is observed for all households. We estimate the 15 parameters governing this non-labor process in a first step, and bootstrap all parameters of the full model repeating this first step for each bootstrap data draw in order to account for sampling variation in this initial step. To simulate the non-labor income process we use simulation, first drawing from a Bernoulli distribution with probability $Pr(I_{h,t} > 0)$ to determine if income is positive or 0, and, if positive, then drawing again from a Normal distribution to simulate the level of non-labor income.

4.1.5 Measuring Child Quality

Rather than assume the stock of child quality (or skills) are perfectly measured in our data, we allow for only imperfect measures. To derive the mapping between unobserved

(latent) child quality, k_t, and measured child quality, k_t^* in our data, we build on the approach utilized by psychometricians (see, e.g., chapter 17 in Lord and Novick, 1968). Consistent with prior research on this subject, we consider child quality to be inherently unobservable to the analyst, although we assume that it is observable by household members, since it is a determinant of the household utility level and is a (potential) input into the decision-making process.

Most cognitive test scores, such as the one used in our empirical work, are simple sums of the number of questions answered correctly by the test-taker. If a child of age t has a quality level of k_t, the probability that they correctly answer a question of difficulty d is $p(k_t, d)$. It is natural to assume that p is non-decreasing in its first argument for all d and is non-increasing in its second argument for all k_t. Taking the model to data, we assume that the test used in the empirical analysis consists of equally "difficult" questions, and we drop the argument d for simplicity.

Given a cognitive ability test consisting of NQ items of equal difficulty, the number of correct answers, $k_t^* \in \{0, 1, \ldots, NQ\}$ is distributed as a Binomial random variable with parameters $(NQ, p(k_t))$. Note that the randomness inherent in the test-taking process implies that the mapping between latent k and observed k^* is stochastic, and that a child of latent skill k_t has a positive probability of answering each question correctly. For a child of latent skill k_t, the expected number of questions she answers correctly is given by $p(k_t)NQ$. Our measurement model then achieves two goals: (i) we map a continuous latent child quality defined on $(0, \infty)$ into a discrete test score measure imposing the measurement floor at 0 and ceiling at $NQ < \infty$, and (ii) we allow for the possibility of measurement error so that a child's score may not perfectly reflect her latent quality. Previous research has often used linear (or log-linear) continuous measurement equations, e.g., Cunha and Heckman (2008), Cunha et al. (2010), Agostinelli and Wiswall (2016). Our approach differs from this in using a measurement process that explicitly recognizes the discrete and finite nature of the test score measure.

In order to identify the model, we do have to take a position on the form of the function $p(k_t)$. In addition to it being nondecreasing in k_t, we would like it to possess the properties: $\lim_{k_t \to 0} p(k_t) = 0$ and $\lim_{k_t \to \infty} p(k_t) = 1$. We choose the following function that satisfies these restrictions:

$$p(k_t) = \frac{\exp(\lambda_{0,t} + \lambda_{1,t} \ln k_t)}{1 + \exp(\lambda_{0,t} + \lambda_{1,t} \ln k_t)} \tag{2}$$

where $\lambda_{0,t}$ and $\lambda_{1,t} > 0$ are measurement parameters.[16]

---

[16]An alternative, but equivalent, formulation that makes clear the role of measurement error is to write that a child of latent quality k answers a test item correctly if $\lambda_0 + \lambda_1 \ln k + \epsilon > 0$, and incorrectly otherwise. ϵ in this formulation is the stochastic measurement error. If ϵ takes on an i.i.d. extreme value distribution, then the probability of answering a test item correctly takes the familiar form given by (2).

4.2   Identification and Estimation

We next discuss how the primitive parameters of the model can be recovered using the observed data. Given the complexity of the model, we focus on some key identification issues and provide some basic intuition, rather than attempting to provide a rigorous proof of identification. Since we do not utilize a likelihood-based estimator and since the model is nonlinear in parameters, a rigorous demonstration of identification is problematic in any case. In general, our identification strategy relies on the availability of rich data from the PSID-CDS on parental and child time allocation, wages, non-labor income, and measures of child cognitive ability. However, even with these data, we face four major identification challenges: (1) the classic non-random selection problem associated with the observation of wage offers only for those who are employed, (2) the fact that children's skills/quality are not directly observed but only imperfectly measured in our data, (3) the household's choice of whether to use an ICCT is not observed directly, and (4) the existence of gaps in the time series coverage of the panel data that we utilize.

Wage Offers and Non-Labor Income   The wage offer processes of the parents, although exogenous, cannot be estimated directly outside of the model due to the well-known problem of endogenous selection. The log wage offer process is given by

$$\ln w_{h,i,t} = Z^w_{h,i,t}\eta_i + \varepsilon_{h,i,t}, \; h = 1, ..., N; \; i = 1, 2;$$

where $Z^w_{h,i,t}$ is the vector of observed covariates of parent i in household h at time t, η_i is a conformable parameter vector, and $\varepsilon_{h,i,t}$ is normally distributed with mean 0 and variance $\sigma^2_{w_i}$, and where $E(\varepsilon_{h,1,t}\varepsilon_{h,2,t}) = \sigma_{w_1,w_2}$. The disturbances are otherwise independently distributed over time and across households. Although we have access to wage observations for multiple periods, wage observations are non-randomly missing due to the significant number of corner solutions associated with labor supply choices. This type of systematic selection is particularly troublesome when preferences are treated as random in the population. In this case, observing a parent not supplying time to the labor market is consistent with that parent's wage offer being low, the household utility function weight on that parent's leisure being high, that parent's time with the child being highly productive, or for a number of other reasons. In order to "extrapolate" preferences and wages when a large number of households have at least one parent out of the labor force requires parametric assumptions on both parents' wage offer functions. Under our model specification, we can "correct" our estimator of model parameters for the non-randomly missing data using the DGP structure from the model. In this case, both the wage processes and the parameters characterizing preferences and production technologies must be simultaneously estimated.

This is not the case with respect to non-labor income. Since we have assumed that this process is strictly exogenous, and that $I_{h,t}$ is observed for all time periods in which non-labor income data is available for household h, we can estimate this process outside

of the model. Identification is achieved both through time-series and cross-sectional variation in the PSID data.

Production Technology   It is typically straightforward to consistently estimate production function parameters if the output level and all inputs are observed, and if the choice of inputs is not a function of any stochastic component of the production technology. For our cognitive ability production technology, we allow for mis-measured child skills (the output of the skill technology) and an additively separable disturbance term that is independently distributed with respect to the inputs. Identification first requires solving the generic problem of indeterminacy due to the fact that latent child quality/skill k does not have any natural units. This can be accomplished by normalizing for one age $\tilde{t}$ the $\lambda_{0,\tilde{t}}, \lambda_{1,\tilde{t}}$ measurement parameters. We choose the normalization $\lambda_{0,\tilde{t}} = 0$, $\lambda_{1,\tilde{t}} = 1$, thus fixing the location and scale of latent skills at all ages k_t to the age $\tilde{t}$ measure. Second, we cannot separately identify an age-varying measurement system (parametrized by the $\lambda_{0,t}, \lambda_{1,t}$ parameters) from age-varying production function primitives $\delta_{1,t}, \ldots, \delta_{6,t}$ and R_t. We solve this under-identification problem by assuming that the measurement system is age-invariant as in Agostinelli and Wiswall (2016), and assume $\lambda_{0,t} = 0$ and $\lambda_{1,t} = 1$ for all t.

For purposes of discussion, we will assume that we have access to the normalized measures of child quality at two ages, t and $t+1$. Then we can write the log production function for household h as

$$\ln k_{h,t+1} = \ln R_t + \ln Z_{h,t}\delta_t + \phi_t \ln k_{h,t} + \varepsilon_{h,t},$$

where $\varepsilon_{h,t}$ is the disturbance term for household h at age t, $\ln R_t$ is the age t log of TFP, $\ln Z_{h,t}$ is the row vector of log time and money inputs selected by household h in period t, δ_t is a conformable column vector of technology parameters, and ϕ_t is a scalar parameter measuring persistence of the process at age t. Under our modeling assumptions, households are heterogeneous in preferences but all share the same production technology, with the exception that the productivities of certain time inputs are allowed to depend on two parental characteristics, the schooling levels of the mother and father, in a parametric manner. For now, we will ignore this dependency, but will conclude this section by arguing that this generalized production function is still identified given our parametric assumptions.

Under the assumptions of Cobb-Douglas preferences and a Cobb-Douglas growth technology, we have seen that the input choices of a household at age t are not a function of the current level of child cognitive ability, so that conditional on all other household characteristics, $Z_{h,t}$ and $k_{h,t}$ are independent. Of course, unconditionally they are not, since both depend on state variables that exhibit temporal dependence.[17] We

---

[17]There exist random preferences in the population that follow a parametric distribution, and the preference vector drawn by household i is assumed to be time-invariant. These preference draws will have a large impact on the investment choices of household h at each child age. Another important generator of temporal dependence in Z and k is the schooling levels of parents, which affect both wage offers in every period and non-labor income levels.

assume that the disturbance term $\varepsilon_{h,t}$ is mean independent of the inputs and current period ability level, so that $E(\varepsilon_{h,t}|Z_{h,t}, k_{h,t}) = 0$, $\forall(h,t)$, and we assume that $\varepsilon_{h,t}$ is independently distributed over time for each household. These are not particularly strong assumptions, since we know that the investment decisions of the household are independent of TFP, and the term $\ln R_t + \varepsilon_{h,t}$ can be viewed as household-specific TFP.

Since all of the arguments in $Z_{h,t}$ are available in the PSID and CDS data, given measures of $k_{h,t+1}$ and $k_{h,t}$, OLS estimates of the parameter vector $(\ln R_t \; \delta_t \; \phi_t)$ are both unbiased and consistent. Even though for each household h we only have two or three points in time at which k and Z are measured, consistently estimating the age-specific parameter vector is possible given sufficiently large numbers of households at each age t, viewing the entire sample as constituting a synthetic cohort of children and parents progressing through the development process.

We have described the best case scenario for estimating the age-specific production function parameters, and now we describe some of the complications we face in the actual estimation of these parameters.

Systematically Missing Data   There are two types of problems that we face with respect to missing data. One is the gaps in the CDS data that make it impossible to use successive observations on child quality along with inputs to identify the production parameters directly. We observe an imperfect measure of child quality in 1997 (a score on a cognitive test), along with the factor utilization levels in that year, but we only observe the next measure of child quality five years later in 2002; for some households we also have a third observation in 2007. In between these dates, input decisions have been made and levels of child quality have been determined; these input decisions depend on wage and non-labor income draws in the intervening years.

One approach to accommodating the gaps in the time series of our panel data is to reduce the number of decision-making periods, collapsing the dynamic model to just a few periods (a single period for early and another for late childhood, say). Given the rapid changes in child development, even during a single year, we instead prefer to assume that the decision-period frequency is annual. In this case, we measure the initial child quality and input decisions at age t, and only observe subsequent child cognitive ability at age $t + 5$. After repeated substitution, we have

$$\ln k_{h,t+5} = a_1(\{\ln R_s\}_{s=t}^{t+4}) + a_2(\{Z_{h,s}\}_{s=1}^{t+4}) + a_3 \ln k_{h,t} + a_1(\{\varepsilon_{h,s}\}_{s=t}^{t+4}), \qquad (3)$$

where

$$
\begin{aligned}
a_1(\{X_s\}_{s=t}^{t+4}) &= X_{t+4} + \phi_{t+4}X_{t+3} + \phi_{t+4}\phi_{t+3}X_{t+2} + \phi_{t+4}\phi_{t+3}\phi_{t+2}X_{t+1} \\
&\quad + \phi_{t+4}\cdots\phi_{t+1}X_t, \\
a_2(\{Z_{h,s}\}_{s=t}^{t+4}) &= Z_{h,t+4}\delta_{t+4} + \phi_{t+4}Z_{h,t+3}\delta_{t+3} + \phi_{t+4}\phi_{t+3}Z_{h,t+2}\delta_{t+2} \\
&\quad + \phi_{t+4}\phi_{t+3}\phi_{t+2}Z_{h,t+1}\delta_{t+1} + \phi_{t+4}\cdots\phi_{t+1}Z_{h,t}\delta_t, \\
a_3 &= \phi_{t+4}\cdots\phi_t.
\end{aligned}
$$

The main problem confronting us is the lack of information on the sequence of inputs $Z_{h,t+1}, ..., Z_{h,t+4}$. The input choice of household h at any time t depends on a subset of the state variables characterizing the household, which we denote by $\tilde{\Gamma}_{h,t} = (w_{h,1,t} \ w_{h,2,t} \ I_{h,t} \ \Upsilon_h)$, where Υ_h denotes the values of the preference parameters for the household, which is a draw from the preference parameter distribution G. As we know, the input decision in period t does not depend on the value of $k_{h,t}$ or of previous values $k_{h,t-1}, k_{h,t-2}, ...$, conditional on $\tilde{\Gamma}_{h,t}$. The state variables $w_{h,1,t}, w_{h,2,t}, I_{h,t}$ are functions of exogenous shocks and observable household characteristics that evolve deterministically over time. Let the value of these covariates at any time t be given by $H_{h,t}$. As we have seen above, we have written the wage offer processes of the parents and the non-labor income process as parametric functions of $H_{h,t}$. For the moment, assume that the parameters of these three exogenous processes are known, and further assume further that the preference parameter distribution G is also known. Then we can form the expected values

$$E(Z_{h,s}|H_{h,s}), \ s = t+1, ..., t+4,$$

where the expectation is taken with respect to the shocks to the wage and income processes at time s and the preference parameter distribution G. Then we can define the expectation of the function a_2 as follows:

$$Ea_2(\{Z_{h,s}\}_{s=t}^{t+4}|\{H_{h,s}\}_{s=t+1}^{t+4}) = E(Z_{h,t+4}|H_{h,t+4})\delta_{t+4} + \phi_{t+4}E(Z_{h,t+3}|H_{h,t+3})\delta_{t+3}$$
$$+ \ \phi_{t+4}\phi_{t+3}E(Z_{h,t+2}|H_{h,t+2})\delta_{t+2} + \phi_{t+4}\phi_{t+3}\phi_{t+2}E(Z_{h,t+1}|H_{h,t+1})\delta_{t+1} + \phi_{t+4}\cdots\phi_{t+1}Z_{h,t}\delta_t.$$

Then (3) becomes

$$
\begin{aligned}
\ln k_{h,t+5} = & \ a_1(\{\ln R_s\}_{s=t}^{t+4}) + Ea_2(\{Z_{h,s}\}_{s=t}^{t+4}|\{H_{h,s}\}_{s=t+1}^{t+4}) + a_3 \ln k_{h,t} \\
& +a_1(\{\varepsilon_{h,s}\}_{s=t}^{t+4}) + [a_2(\{Z_{h,s}\}_{s=t}^{t+4}) - Ea_2(\{Z_{h,s}\}_{s=t}^{t+4}|\{H_{h,s}\}_{s=t+1}^{t+4})],
\end{aligned}
\tag{4}
$$

where the expression on the last line, which is the composite disturbance term, is mean-independent of $Z_{h,t}$ and $\{H_{h,s}\}_{s=t+1}^{t+4}$. In this case, OLS estimation of (4) yields consistent estimates of the parameters characterizing the function a_1 and the combination of parameters

$$\delta_{t+4}, \ \phi_{t+4}\delta_{t+3}, \ \phi_{t+4}\phi_{t+3}\delta_{t+2}, \ \phi_{t+4}\phi_{t+3}\phi_{t+2}\delta_{t+1}, \ \phi_{t+4}\cdots\phi_{t+1}\delta_t, \ t = 3, ..., 12.$$

We have a reasonably large number of children in each age group $t = 3, ..., 12$ in 1997. This means that as we vary the initial t, we have consistent estimates of $\delta_7, \delta_8, ..., \delta_{16}$, among other combinations of parameters. The $\delta_s, \ s = 7, ..., 16$, are parametric functions of s. Ignoring the dependence on parental education for the moment, they are functions of only two parameters for each input, since the production parameter associated with an input j $(j = 1, ..., 5)$ at age t is

$$\delta_{j,t} = \frac{\exp(\gamma_{j,0} + \gamma_{j,1}(t-1))}{1 + \exp(\gamma_{j,0} + \gamma_{j,1}(t-1))}.$$

Having consistent estimates for $\delta_{j,t}$, $t = 7, ..., 16$, allows us to consistently estimate $\gamma_{j,0}$ and $\gamma_{j,1}$ for each input j. The same argument applies with respect to the identification of the scalar parameters, ϕ_s, since these are specified as a two-parameter function of s. The $\ln R_s$ process is specified as a four-parameter function of s. Given the consistent estimates of the ϕ sequence, these four parameters are identified as well. In fact, all of the parameters in the production technology are over-identified under our functional form assumptions.

Preference Parameters   In this subsection, we make the argument that under certain ideal conditions regarding data availability, the distribution of preferences G is nonparametrically identified. This is due to the simple structure of the choice problem and the fact that, given all of the state variables of the problem, including household preferences, the actions of the household are uniquely determined. Our argument is slightly complicated by the possibility that the household may be behaving "non-cooperatively" or cooperatively, and there is no sample separation information in the data that indicates which behavioral rule the household is utilizing. We will begin by assuming that all households are operating in the non-cooperative Stackelberg equilibrium.

The choices of household h in period t are summarized by the actions $a_{h,t}$. The investment process that determines the sequence of $\{k_{h,s}\}_{s=1}^{M+1}$ is the only endogenous dynamic process in the model, so that the decision rules determining time allocations in the household in any period s, $s = 1, ..., M$, are very simple to characterize given knowledge of the state variables in each period s, and no decisions actually depend on the value of $k_{h,s}$ in any period. Due to the fact that we introduced the joint time of the parents, $\tau_{h,12,t}$ as an argument in the production technology, there is not a closed form solution to the household choice problem in period t, as there was when this argument was not present in the more restrictive model contained in Del Boca et al. (2014). Nonetheless, there is a unique mapping from the state variables of the problem, including preferences, in period t, $\Gamma_{h,t} = (w_{h,1,t} \; w_{h,2,t} \; I_{h,t} \; k_{h,t} \; \alpha_h)$, where a_h denotes the vector of preference parameters for household h, and $a_{h,t}$, so that $a_{h,t} = a^*(\Gamma_{h,t})$. If all actions $a_{h,t}$ were observable in period t, as well as $w_{h,1,t}$, $w_{h,2,t}$, and $I_{h,t}$ ($k_{h,t}$ need not be observed since period t decisions are not a function of $k_{h,t}$), then a^* is invertible in terms of the preference parameter vector, with

$$\alpha_h = (a^*)^{-1}(a_{h,t}|w_{h,1,t} \; w_{h,2,t} \; I_{h,t}).$$

The argument for the invertibility of this function follows Del Boca and Flinn (2012). Given this invertibility in terms of α_h, observations from only one period for each sample household can be used to back out the vectors α_h, $h = 1, ..., N$. Given that households are randomly sampled from the population of all intact households with children, then the empirical distribution of the α_h is a nonparametric maximum likelihood estimator of G, or $\hat{G}_N$. Asymptotically, we have $\lim_{N \to \infty} \hat{G}_N = G$ at all points of continuity of G. Since we are assuming that G is absolutely continuous everywhere on its domain, then $\hat{G}_N$ is a consistent estimator of G.

Missing Data and the Estimation of G In point of fact, the PSID and the CDS do not include information on all of the household's choices in any period t. In particular, we do not observe all of the actions chosen by a household h in any period t. Given the CDS data, we do have an observation on all of the time allocations, including hours worked of the parents, time parents spend with the child, and the child's time "investing" by herself. We do not have direct information on the private consumption of the parents and the child, $c_{h,t}$ and $x_{h,t}$, respectively. In order to simplify the computation of the model solution and to enhance identification of the model, we have assumed that there are no capital markets available for borrowing or saving. This means that total expenditures in the period are equal to total income, so that

$$x_{h,t} + c_{h,t} + e_{h,t} = w_{h,1,t}h_{h,1,t} + w_{h,2,t}h_{h,2,t} + I_{h,t}.$$

Since we observe all of the arguments on the right-hand side, and since a measure of $e_{h,t}$ is available in the CDS data for the years in which it is administered, we only observe total household consumption, $x_{h,t} + c_{h,t}$, and not the individual amounts consumed by the parents and the child within each household h.[18] By our assumptions regarding the time endowment of the parents and childs and the types of activities to which time can be devoted, we can infer the quantity of leisure consumed by parents and the child as the time remaining after all other activities have been accounted for.

If $x_{h,t}$ and $c_{h,t}$ were observed for all households, and if all parents in the population of intact households with children supplied positive amounts of time to the labor market in some period t (so that $h_{h,1,t} > 0$ and $h_{h,2,t} > 0$), then G is nonparametrically identified. Since we only observe, $x_{h,t} + c_{h,t}$ and since we do observe corner solutions in the data, G is not nonparametrically identified. We have parameterized the distribution G in a flexible manner, but with enough explicit restrictions on the relationships between preference parameters draws to enable a separate valuation of x and c even when we do not observe the separate consumption expenditures. It is also important to note that we have not attempted to estimate the altruism parameter, but instead are fixing at the value of $\varphi = \frac{1}{3}$. Attempting to estimate this parameter would make identification of the parameters characterizing G even more challenging.

We do not use the information on $e_{h,t}$ provided in the CDS in estimating the model, which increases the need for parametric restrictions on G. We did not use this information because we thought that the reported values were far too low to be believable. In our opinion, this is due to respondents under-reporting indirect expenses associated with the child's presence in the household, such as expenditures on housing and food.

---

[18]We are not explicitly considering the additional problem that the information on the wages, work hours, and non-labor income of the parents is not collected in the same year as is the information taken from the CDS, which includes the time allocations in investment and the expenditures on the child for investment, $e_{h,t}$. The first CDS interview is conducted in 1997, whereas the other information taken from the core survey of the PSID is available in 1996 and 1998. Since wages, work hours, and non-labor income are reasonably stable over such a brief time period, the fact that the information is not collected at exactly the same point in time should not cause major problems in inferring the stable preference weights associated with the arguments of the parents' altruistic utility function.

In our accounting framework, such expenses would be considered investments in the child.

Unobserved "Parenting Style" Even under ideal conditions, that is, when all decisions of the parents and the child are observed, the fact that households may either be "non-cooperative" or "constrained-efficient" in their behavior is problematic in terms of the non-parametric identification of G. This problem is discussed at length in Del Boca and Flinn (2012), where households choose between these two different modes of behavior and where the choice of the household was unobservable. In their nonparametric setting, they showed that there existed two mappings from the observed actions to the parameters characterizing the household, one corresponding to each mode of behavior. In this case, each mapping must be used to determine the parameter values that would have generated it under the assumed form of behavior, and then each is evaluated to determine whether the assumed mode of behavior would have been the one chosen under those parameter values. In some cases, each of the two generated parameter vectors is consistent with the assumed behavioral rule being preferred by the household, which is essentially a case of multiple equilibria. This problem does not occur if G is being estimated parametrically. Thus, even if information on all household choices $a_{h,t}$ was available to us, parametric assumptions on G are required to resolve the indeterminacy associated with the unobserved behavioral regime of the household.

Monitoring Cost Distribution We do not observe whether a particular household is using an ICCT in any given period. If we did have knowledge of the "parenting" style used, the estimation of a single parameter cost distribution would be fairly straightforward. For each household, first compute the difference in period t value functions of the parents when they use an ICCT and when they do not. Denote these differences by $\Delta V_{h,t}$, where the subscript h denotes the household, $h = 1, ..., N$. From our discussion in Section 2, we know that these values are all nonnegative. Now assume that the distribution of the cost of using an ICCT, ω, is a one-parameter distribution defined on the non-negative real line, and for simplicity assume that it is an exponential distribution, so that $R(\omega) = 1 - \exp(-\sigma\omega)$, with $\sigma > 0$. Then the probability that household h employs an ICCT in period t is $p(\Delta V_{h,t} > \omega) = 1 - \exp(-\sigma\Delta V_{h,t})$. If the sample proportion of households using an ICCT in period t is defined as $p_N(\iota_t = 1)$, this is given by

$$N^{-1} \sum_{h=1}^{N} p(\Delta V_{h,t} > \omega) = N^{-1} \sum_{h=1}^{N} (1 - \exp(-\sigma\Delta V_{h,t})).$$

The expression on the right hand side of the equation is absolutely continuous and monotone increasing in the parameter σ, with $\lim_{\sigma\to 0} N^{-1} \sum_{h=1}^{N} (1 - \exp(-\sigma\Delta V_{h,t})) = 0$ and $\lim_{\sigma\to\infty} N^{-1} \sum_{h=1}^{N} (1 - \exp(-\sigma\Delta V_{h,t})) = 1$. Thus for any $p_N(\iota_t = 1) \in (0,1)$, there

exists a unique value of σ, $\hat{\sigma}(p_N(\iota_t = 1))$, such that

$$p_N(\iota_t = 1) = N^{-1} \sum_{h=1}^{N} (1 - \exp(-\hat{\sigma}(p_N(\iota_t = 1))\Delta V_{h,t})).$$

Since $\text{plim}_{N \to \infty} \, p_N(\iota_t = 1) = p(\iota_t = 1)$, the proportion using an ICCT in the population, then $\text{plim}_{N \to \infty} \, \hat{\sigma}(p_N(\iota_t = 1)) = \hat{\sigma}(p(\iota_t = 1)) = \sigma$. Thus $\hat{\sigma}(p_N(\iota_t = 1))$ is a consistent estimator of σ.

This argument has been largely heuristic in the sense that we have implicitly assumed that the $\Delta V_{h,t}$ are known. Of course, these quantities are all estimated, and are complex functions of most of the parameters that characterize the model. They are not, however functions of σ. If consistent estimates of the $\Delta V_{h,t}$ are available, the identification argument we have made is applicable. Given knowledge of the use of an ICCT by each household in the sample, the parameter σ can be consistently estimated in a second stage, after the $\Delta V_{h,t}$ have been estimated.[19]

This identification argument is not directly applicable in the present setting, since the use of an ICCT by an household is unknown. Instead, the marginal distribution of choices and outcomes represents a mixture of the two situations, one in which an ICCT is being used and the other in which it is not. In addition to all of the other parameters of the model, the parameter σ is an important component of the "endogenous" mixing distribution.[20] When $\sigma \to \infty$, then the cost distribution becomes degenerate at the point $\omega = 0$, and all households would utilize an ICCT in every period. When $\sigma \to 0$, all households would not use an ICCT in any period, and the Stackelberg equilibrium would be the one characterizing all observed choices in the data. To the extent that allowing a type distribution allows more flexibility in describing the data and improves model fit, we can expect to recover an estimate of σ that is strictly positive and bounded from above.

When the use of an ICCT in period t by any household is unobserved, it is no longer possible to estimate σ in a second stage after all other primitive parameters have been estimated. Instead, all parameters must be estimated simultaneously. An example of this type of estimation can be found in Del Boca and Flinn (2012).

Simulation Estimation using the DGP   As we have made clear in the preceding discussion, due to a variety of missing data problems, it is necessary to estimate all model parameters simultaneously, with the exception of those characterizing the non-labor income process of the household. These parameters are estimated in a first-stage, and are used throughout the simulation process that we now describe.

---

[19]This is the route followed in Flinn and Mullins (2015), for example, when estimating a schooling cost distribution.

[20]The term endogenous in this case refers to the fact that the primitive parameters of the model determine the distribution of types, which in this case refers to whether an ICCT is being used in period t. Exogenous mixing, which is commonly utilized, would assert that there is a distribution of types that is not a function of the other primitive parameters of the model.

In discussing the identification of the cognitive ability production process, we assumed that the preference parameter distribution G was known, along with the wage and income process parameters. This information was needed in order to compute the expected values of the a_2 functions. We do not actually use the regression function (4) to estimate the parameters of the model. Instead, we simulate the entire data generating process (DGP) of the model for given draws of all of the random variables characterizing the DGP. Using R simulation paths for each household, we then construct a number of sample moments from the overall sample of $N \times R$ histories that are compared with the same moments computed from the actual data consisting of N histories, one for each sample household.

We require the wage offer functions for the two parents, with the rth replication for household h at time t given by

$$w_{h,i,t}(r) = \exp(Z_{h,i,t}^w \eta_k + \varepsilon_{h,i,t}(r)), \ h = 1, ..., N; \ i = 1, 2; \ t = 1996, ..., 2010; \ r = 1, ..., R,$$

where $Z_{h,i,t}^w$ are the covariates for parent i in household h in year t, η_i is a conformable vector of parameters for parent i, $\varepsilon_{h,i,t}(r)$ is the rth draw of the disturbance for parent i in household h in year t, and where R is the number of replications for each household in the sample. The parameters to be estimated include the vectors η_1 and η_2, as well as the three parameters that characterize the contemporaneous covariance matrix of $(\varepsilon_1 \ \varepsilon_2)$, Σ_ε. Similarly, in each period R non-labor income draws are generated for each household, with these given by

$$I_{h,t}(r) = Q(Z_{h,t}^I, \hat{\mu}, \varepsilon_{h,t}^I(r)), \ h = 1, ..., N; \ t = 1996, ..., 2010; r = 1, ..., R,$$

where $Z_{h,t}^I$ are the covariates in the non-labor income function for household h in period t, $\hat{\mu}$ is a conformable parameter vector of estimates of μ (obtained from previous estimation of the non-labor income process), and $\varepsilon_{h,t}^I(r)$ is the rth replication draw for (h, t) from a Normal distribution with mean 0 and variance σ_I^2. The R preference distribution draws for each household h are taken from the distributions $G(\alpha \ \lambda | \Gamma_G)$, where Γ_G is the vector of parameters characterizing the preference distribution. We have assumed that household preferences are time-invariant and are identically and independently distributed over households.

We conclude by describing the simulation of the cognitive ability process of the child in household h. In 1997 when the child is age t, we have access to their test score $k_{h,t}^*$, which is the number of correct answers out of 57, and $k_{h,t}^*$ follows a binomial distribution with parameters $NQ = 57$ and $p(k_{h,t})$, the probability of a correct answer. We represent our initial prior for $p_{h,t}$ as the (non-informative) Beta distribution with parameters $(1, 1)$, which is Uniform on the interval $[0, 1]$. We then observe the test score $k_{h,t}^*$, allowing us to update our prior and produce a posterior distribution on $p_{h,t}$, which is also Beta (a conjugate distribution for the Binomial). The posterior distribution for $p_{h,t}$ is then Beta with parameters $(1 + k_{h,t}^*, (NQ - k_{h,t}^*) + 1)$.

From the posterior distribution of $p_{h,t}$, we draw R pseudo-random draws of $p_{h,t}$. Let $\tilde{p}_{h,t}^r$ denote the rth draw of p from this distribution. Given our assumed mapping of k to p, can invert $\tilde{p}_{h,t}^r$ to obtain

35

$$k_{h,t}^r = \exp(\lambda_1^{-1}\{\ln(\frac{\tilde{p}_{h,t}^r}{1 - \tilde{p}_{h,t}^r}) - \lambda_0\}).$$

As discussed above, the values of λ_1 and λ_0 are not identified, and we have fixed them at $\lambda_0 = 0$ and $\lambda_1 = 1$.

From this initial value of $k_{h,t}^r$, we then begin the construction of this particular sample path. When we get to the period of the second measurement (in 2002) when the child is age $t' > t$, the test score is drawn from a binomial distribution with parameters $(NQ, p(k_{h,t'}))$, as described above. We can invert this test score to determine the value of $k_{h,t'}$.

4.3 Computation of the Model Solution

Unlike DFW (2014), our model does not have a simple closed form solution for all endogenous choices. Instead we use a mixed numerical and analytic solution, in which we compute some endogenous choices on a fine grid and compute other solutions analytically given these choices. We then substitute these choices into the utility function to determine the relative utility of each grid point, and use the maximum utility value choice as the optimal choice. We leave the details of the multiple-step model solution to the Appendix. We also report experiments showing that our mixed analytical/numerical solution approximates the "near" exact solution (using a very fine grid) quite well.

4.4 Estimator

The family data that are available consist of a sample of households with observed characteristics X, which includes time-invariant and time-varying characteristics, as well as information pertaining to children interviewed at various ages (where child age is indexed by t). The observed household variables include parental characteristics, such as the education and the ages of the parents when the child was born. For each mother and father in the household we observe: hours worked, hours spent with children (both alone and with the other parent), and repeated measures of child quality. We also have measures of the child's own self-investment time. In addition, we observe: (accepted) wages for both parents and total household income, as described below. Although data on some child specific expenditures are available, we do not utilize them in forming the estimator.[21]

---

[21]We made this decision because the distribution of reports of child-specific expenditures had what we considered to be too much mass in the left tail of the distribution. Our interpretation is that respondents were not properly attributing some of the household expenditures on public goods to children. In any case, we think that it is difficult for any person, even an economist, to properly impute these values, and hence did not utilize them when forming the estimator. The implied estimates of money expenditures on children are larger than those reported, but we think that they are more representative of total expenditures on children when public good expenditures are "properly" assigned.

The estimator utilizes simulation methods. The Appendix details the simulation and estimation procedure. In brief, given a trial vector of primitive model parameters and the initial observed sample characteristics of each household, we simulate S sequences of agent actions (time allocation and expenditures) and child quality given S draws from the parametric random variable distributions. Using the simulated data set, we then compute the analogous simulated sample characteristics to those determined from the actual sample. The characteristics of any simulated sample are determined by Ω, the vector of all primitive parameters that characterize the model, and the actual vector of pseudo-random number draws made in generating the sample paths. Denote the simulated sample characteristics generated under the parameter vector Ω by $\tilde{M}_S(\Omega)$. The Method of Simulated Moments (MSM) estimator of Ω is then given by

$$\hat{\Omega}_{S,N,W} = \arg\min_{\Omega}(M_N - \tilde{M}_S(\Omega))'W_N(M_N - \tilde{M}_S(\Omega)),$$

where M_N is the vector of sample moments, $\tilde{M}_S$ are the analogous model simulated moments determined by the model primitives Ω and the length S vector of pseudo-random draws, and W_N is a symmetric, positive-definite weighting matrix.

The moments we use include the average and standard deviation of test scores at each child age, the average and standard deviation of hours of work for mothers and fathers at each child age, the average and standard deviation of child investment hours for mothers and fathers at each child age, and the average and standard deviation of time investment by the child at each age. In addition, we use the average and standard deviation of accepted wages and the correlation in wages across parents. We also compute a number of contemporaneous and lagged correlations between the observed labor supply, time with children, child quality, wages, and income. It is important to note that while we do not observe child inputs, labor supply, wages, and income in the same periods, our simulation method allows us to combine moments from various points in the child development process into a single estimator. A full list of the moments we utilize is included in the Appendix.

5 Model Estimates

5.1 Household Preference Parameters

Given the identification issues discussed in the previous section, we assume that the altruism parameter is fixed to $\varphi = \frac{1}{3}$ for all households. This value is relatively consistent with the broad range of parental altruism parameter estimates that are mainly found in the household macroeconomics literature.[22] This implies that parents value their child's utility about half as much as their private utility. The annualized discount factor is

---

[22]Altruism parameters are estimated, or calibrated, most often in macroeconomic studies of household behavior. Moreover, it is usually the case that the altruism parameter is defined as $u_P + \varphi u_c$, whereas we have defined the parents "final" utility as a weighted average of their private utility and the childs utility, or $(1-\varphi)u_p + \varphi u_c$. In a model of the child's decision to leave the household, Kaplan

fixed to $\beta_p = 0.9$ for parents. For children, the discount factor increases with child age (see e.g. Steinberg et al. (2009)). More specifically, we fix $\beta_{c,t} = 0.705$ if $t \le 13$, $\beta_{c,t} = 0.822$ if $t \in \{14, 15\}$, and $\beta_{c,t} = 0.9$ at the final age $t = 16$. Any investments made by the parents and the child after the final age $M = 16$ are not explicitly modeled, but can be captured by the utility terms specified for this final period, $\psi_{i,M+1}$, for $i = p, c$. Note that while periods are in years, the model is specified in terms of weekly decisions, which are considered to be invariant within each yearly planning period.

Table 4 presents the estimates of the preference parameters for parents and children. Although preferences of the parents and the child are assumed to be time-invariant within each household, we allow for population heterogeneity in preferences across households. Since the primitive parameters of the distributions of parental and child preferences are difficult to interpret, we instead report the means and standard deviations of the population distributions of these preferences. The estimates suggest that the average household puts the most weight on the father's leisure and private consumption (about 0.30), followed by the mother's leisure (about 0.24) and finally child quality (about 0.16). However, this relative order is altered when taking into account the fact that all parents are altruistic, which causes them to care more about child quality than anything else (on average). Indeed, the average *total* parental weight on child quality, $\tilde{\alpha}_4 = (1 - \varphi)\alpha_4 + \varphi\lambda_3$ is approximately 0.24, whereas the average total weight on e.g. parental consumption, $\tilde{\alpha}_3 = (1 - \varphi)\alpha_3$ drops to about 0.20. Conversely, the average child attaches a relatively high weight to her own leisure and quality level ((about 0.42 and 0.41, respectively), and values her private consumption the least (about 0.17). When looking at the dispersion of the estimated preference distributions, we find that the standard deviations of the parental preferences are relatively low (in the range $0.01 - 0.05$) compared to the standard deviations of the child's preferences (ranging from $0.05 - 0.22$). This stands in contrast to DFW (2014), who found relatively high amounts of dispersion in some preference weights.

The estimates related to the household preference correlation matrix suggest that there is substantial positive correlation of preferences (1) among spouses, (2) among children and (3) between parents and their children. We find a large correlation between mother's leisure and father's leisure of about 0.30, which is consistent with a theory of positive assortative matching in the marriage market. The parents' preference for consumption is negatively correlated with the preference for the mother's leisure (about -0.69), and weakly positively correlated with the preference for the father's leisure (about 0.09). Second, the child's preferences for leisure and private consumption are positively correlated (about 0.69). Finally, we find some evidence of "intergenerational"

(2012) finds an estimate of φ of 0.039. In Boar's (2018) study of parents' precautionary savings for the child's risky income, the calibrated value of φ is 0.201. In his study of bequests and inter-vivos transfers, Nishiyama (2002) finds a value of φ of 0.512. Finally, Barczyk and Kredler (2018), in a study of long-term care decisions for elderly parents, find that the calibrated value of parental altruism is 0.692. As can be seen, the range of the estimates is large, and depends on the application and specifics of the model. Our value of $\varphi = \frac{1}{3}$ in our weighted-average formulation implies a value of $\varphi = 0.5$ in the $u_p + \varphi u_c$ specification, which puts it in the middle of the range of values found in these studies.

preference correlation as well. Although we have restricted the possible dependence in these parameters for the purposes of estimation, we find that the child's preference for leisure is positively correlated with both the mother's and the father's tastes for leisure (about 0.47 and 0.52, repectively). Conversely, the child's taste for private consumption does not seem to be correlated with the parents' taste for private consumption (about 0.005). The reader should bear in mind that we use no direct observations on the parents' or the child's consumption in estimation, so that this relationship is primarily identified through functional form assumptions.

Although difficult to interpret, we estimate that the parents' terminal payoff to child quality (ξ_p) is around 5.20, which is significantly larger than the estimate of the child's terminal payoff to child quality, $\xi_c = 0.81$. Both of these parameters are relatively imprecisely estimated, which is unsurprising since they are identified through functional form assumptions.

Finally, we estimate that the parameter of the (exponential) ICCT cost distribution is approximately 110, implying that the mean and standard deviation of the ICCT cost is around $\kappa = 0.009$ utils. What this implies in terms of simulated household behavior can be seen in Figure D-9c, which shows how the average fraction of households choosing to implement an ICCT varies with the child's age t. Consistent with some of the literature on parental rewarding behavior (e.g., Bonke (2013)), we find that the proportion of parents who actively incentivize their children (for example, to study more or to receive higher grades in school) by offering monetary rewards increases with the child's age. Since the utility cost associated with implementing an ICCT is assumed to remain constant within a given household, this increasing age trend can be explained purely in terms of terms of increasing utility benefits to such incentive schemes.

5.2 Child Quality Technology Parameters

In Section 4, we discussed how the Cobb-Douglas productivity parameters for household h associated with each productive input j $(\delta_{h,j,t})$ are allowed to change monotonically with the age of the child t, by specifying an intercept $(\gamma_{j,0})$, an age slope parameter $(\gamma_{j,1})$ for each input, and parental education slope parameters $(\gamma_{1,2}, \gamma_{2,2}, \gamma_{3,2}, \gamma_{3,3})$ for the three parental time inputs. The total factor productivity parameters (R_t) are estimated using a more flexible generalized logistic specification. Table D-2 and Figure 3 show the estimates of these production technology parameters. Due to the exponential transformation, the raw estimates of the intercepts and slopes relating to each input are difficult to interpret. A larger (i.e. less negative) intercept indicates a higher initial productivity at the first relevant child age, $t = 3$. A positive (negative) age slope estimate implies that the productivity increases (decreases) monotonically with child age. In Panel 3a, we plot the estimated sequences of the four time inputs as a function of the child's age, averaging across all households (i.e. given average parental education levels). We find that active parental time inputs are highly productive at young ages, but become less productive as the child ages. Consistent with the literature, we estimate that maternal time is the most productive of all inputs for very young

39

children. Joint parental time is initially slightly less productive than maternal time, but becomes relatively more productive after age 5. Paternal time is less productive than the other two parental time inputs at all child ages. The absolute levels and decreasing time trends in the productivity of maternal and paternal time are in accordance with the findings of DFW (2014). Unsurprisingly, the productivity of the child's self-investment time is initially very low. When the child reaches adolescence, however, self-investment time gradually becomes more productive and matters about as much as the individual parental time inputs, but remains less productive than joint parental time. Given our specification of the model, these relative trends make intuitive sense. Once children start formal schooling and become more influenced by teachers and fellow students, their time allocation starts to shift away from spending time with parents, and other inputs start to become more relevant.

Panels 3b and 3c show the degree to which parental education levels shift the productivity parameters at various child ages. Although years of schooling enters continuously into the parametric specification, the graphs focus on high school graduates versus college graduates (i.e. 12 years versus 16 years of schooling) for simplicity. We notice little or no effect of maternal schooling on the mother's individual time productivity. Conversely, the father's education level positively affects the father's time productivity. For joint parental time, we estimate a small but positive and roughly equal effect of both the father's as well as the mother's education level. This suggests that overall, higher educated parents spend more quality time with their children, ceteris paribus.

Panels 3d and 3e show the estimated productivity parameters of the two remaining inputs as a function of child age: material child investments ($\delta_{4,t}$) and lagged child human capital ($\delta_{6,t}$). There is a slight negative trend in the productivity of child expenditures (although this is hard to compare to the time inputs due to the different units of measurement). Conversely, we estimate a strong positive time trend for the productivity of lagged capital. The strong and increasing level of persistence in the estimated production technology suggests that the self-productivity of skills - or the principle that "skills beget skills" - becomes more salient as children accumulate more human capital. This pattern is consistent with DFW (2014).

Panel 3f shows the estimated total factor productivity process (R_t) as a function of child age. We find that there is an overall positive trend in the TFP, starting out at approximately 0.47 for the youngest ages, then reaching an inflection point around age 8, and finally stabilizing at around 1.01 at age 12.

5.3 Wage and Non-labor Income Process Parameters

Table 6 provides the estimates of the wage and non-labor income processes. As specified in Section 4, the (log) wage offer distribution for each spouse depends linearly on their observable characteristics, i.e. age and completed level of education.[23] We estimate that, for both spouses, each year of education increases the wage by almost 10 percent.

---

[23]More complex polynomials, including e.g. a squared age term, a birth cohort effect, or an interaction between age and education, turned out to have small and imprecisely estimated coefficients.

Given the fairly small age range of the parents in our sample, we estimate a flat earnings profile that is similar for both spouses, where each additional year amounts to a mean wage increase of about 1.3 percent. Since most parents work full-time in all observed periods (especially fathers), we could interpret this as the average return on experience. When compounded over the relevant parental age range of $30 - 50$, this would amount to a total mean wage increase of about 30 percent for each spouse, which is not unreasonable. The remaining three wage parameter estimates $(\sigma_{w_1}, \sigma_{w_2}, \rho_{12})$ determine the variability of the periodic wage shocks of each spouse, which are allowed to be correlated with the wage shocks of the other spouse. We estimate that this correlation is strongly positive at 0.54. This could again be interpreted as evidence of positive assortative matching in the marriage market. Our estimates also suggest that the fathers' wage shocks are more volatile than the mothers' shocks.

Since there is no endogenous selection on non-labor income as there is on wages, the parameters of the non-labor income process can be estimated separately from the rest of the model. We refer to Section 4 for more details on the two-step procedure used to estimate the non-labor income process, which is tailored to (1) fit the large observed fraction of households in the data who report a zero non-labor income in any given year, and (2) capture the observable heterogeneity in non-labor income draws. The first seven parameter estimates in part (c) of Table 6 indicate that the probability of having a strictly positive non-labor income in any period increases strongly with the father's education level, and is slightly concave in the father's age in that period. The remaining parameter estimates show that, conditional on having a strictly positive non-labor income draw, the mean non-labor income increases by about 4.5 percent each year, and does not seem to be significantly impacted by the education levels of the parents. The large standard deviation of the disturbances is not surprising given the small number of covariates used in the empirical specification.

5.4 Within-sample Fit

5.4.1 Test Scores

Figure D-4 plots the average Letter Word test score as a function of the child's age, where the solid line represents actual CDS data and the dotted line represents the simulated test scores under the model. The estimated model captures the S-shaped trend in the measured test scores almost exactly. [24]

---

[24]In general, there are two sources of "noise" in the average level of measured child quality at each age. First, the average child quality level at each age is an aggregate of the child quality levels for our sample of heterogeneous households. Second, we are displaying the simulated measure of child quality using our stochastic measurement model, and the simulated measure is noisy due to the fact that the number of simulation draws is finite.

5.4.2 Time Allocations

Regarding the endogenous time allocation decisions of parents and children, the model is able to fit most of the basic patterns in the data. Table 7 displays the sample fit of the parental labor supply choices, the three types of parental time investments and the child's self-investment time, broken down into four non-overlapping child age bins.

The model correctly predicts that mothers of older children are more likely to supply some positive amount of labor to the market,. For fathers, the model slightly overpredicts the labor force participation rates, which could be due to the relatively low variation in paternal labor supply behavior in the data. When looking at the subset of working parents, the model fits the average weekly hours of supplied labor well for both spouses. For working mothers, the model again captures the positive trend in labor supply as children age.

Overall, the model fits the average levels of the three types of parental time investments observed in the data very closely for each of the child age bins. Importantly, the strong negative trends in the observed maternal and paternal active time investments are reflected in the simulated data. The increasing trend in the child's self-investment time is also captured by the model, although the levels are slightly off at some ages, possibly due to some outliers in the data.

Figures D-5 and D-6 present a more detailed fit of the labor supply choices and the various time investments, broken down for each child age. Overall, the relatively close fit between the data moments and the simulated moments confirms that the results in Table 7 would be robust to different choices of the child age bins.

5.4.3 Wages and Non-labor Income

Table 8 shows how the simulated model is able to fit the basic moments of the wage and non-labor income data fairly precisely. Any differences between the data moments and the simulated moments in this table are due to the fact that we are not only targeting these unconditional moments, but also several other moments of the wage and non-labor income distributions where we condition on parental age and education. Figure D-7 shows a more detailed breakdown of the parents' hourly wages in the data and under the model, as a function of either the respective parent's age or their completed education level. Although these conditional moments are relatively noisy due to small sample issues in the data, we find that the overall fit is reasonably good. For mothers, panels D-7a and D-7b indicate that the model captures the increasing wage trend in both the mother's age as well as her education level. For fathers, panels D-7c and D-7d show that the model also captures the overall increasing trends, although the data become very noisy for older fathers due to the usual small sample issues.[25]

As shown in Table 8, the basic unconditional moments of the exogenously estimated non-labor income process are matched very precisely, including the fraction of strictly

---

[25]We do not simulate wages for very young parents, since the model is only specified for children of age 3 and above.

positive non-labor income observations. Figure D-8 shows how various conditional moments are also replicated relatively well. In the first and second rows, we condition on the father's age and mother's age, respectively. In the third and fourth rows, we condition on the father's and mother's finished education levels. In each of the panels, we find that the simulated non-labor income moments fit the data well, whether we consider the entire sample (left-hand side panels), or focus on the subsample of observations where the non-labor income is strictly positive (middle panels). The right-hand side panels show how the simulated data capture the increasing trends in the probability of reporting a strictly positive non-labor income.

5.4.4   Expenditures, Leisure and Internal CCT Use

In Figure D-9, we present some simulated moments of endogenous variables which are either not explicitly used in the estimation or not observed in our data set. Therefore, none of these moments are targeted in our estimation procedure. Each of the four panels shows average simulated values as a function of the child's age. Panel D-9a shows how the average weekly household income (Y_t) increases steadily with child age, and how this trend is reflected in each of the three expenditure components, although child expenditures (e_t) seem to stabilize at older ages, since this input gradually becomes less productive for older children. Overall, child consumption amounts to about 21 percent of the total household budget, which is in line with the literature.

Panel D-9b plots the weekly (residual) leisure time enjoyed by each household member. Although mothers tend to spend more time with children than fathers, the lower female labor supply causes the average mother's leisure time to be higher than the father's, by approximately 12 hours per week. Due to the increase in the child's exogenously defined school time at age 6, the child's residual leisure time takes a sharp drop there. Although irrelevant for the estimation procedure or any of the other results, this sudden drop in child leisure could be partially smoothed out by adding the observation that children's average sleep time (which is exogenously defined as 56 hours per week for all ages) decreases as they age. Panel D-9c shows the average fraction of households choosing to implement an ICCT at each t. Given the estimated preference, technology and ICCT cost parameters, this fraction increases from about 12 percent at age 3 to around 72 percent at age 15. Panel D-9d shows how the average reward function elasticity (r_t), chosen by the subset of parents who implement an ICCT, increases from around 0.04 at age 3 to around 0.24 at age 15. Therefore, our estimates support the notion that, as children age, (1) more parents choose to incentivize their children by offering a contingent reward system, and (2) the average reward system offers higher, and not lower, incentives to study. Finally, Panels D-9e and D-9f break down the use of ICCTs by household (permanent) income tercile. We notice that although the average reward elasticity (conditional on it being positive) does not depend significantly on household income, the fraction of households using an ICCT at any given child age does. In particular, households in the highest income tercile are up to 12 percentage points more likely to use an ICCT than households in the lowest income tercile.

6 The Impact of External Policies on Household Behavior and Welfare Outcomes

We now study the effect of external resource transfers made to the parents on the behavior of household members and household outcomes, with particular reference to child development. There are two important problems that a social planner faces in attempting to improve aspects of the child cognitive ability distribution by making monetary transfers to households. The planner does not know the preferences of the household members, but rather is only assumed to know the distribution of preferences in the population. Since we assume that the social planner's preferences are only defined with respect to the distribution of child cognitive ability, that gives this problem a principal-agent structure, in that there is asymmetric information regarding household preferences and actions and because the preferences of household members and the planner are different. By making a transfer to the parents (the agents), the planner attempts to influence the distribution of cognitive outcomes in the most cost-efficient way while taking the actions of the household into account.

The planner is limited in terms of the types of goods that can be transferred, which are restricted to money or in-kind transfers of child investment goods, e_t. We restrict attention to non-coercive policies, which we define as any policy under which the welfare of the parents (assumed to be the recipients of all transfers) is no less than it would be in the absence of the policy. In the case of policies that take the form of income transfers to the household contingent on its behavior, these triggering events must be observable to the social planner. Since we assume that the internal actions of the household are not directly observable by the planner, this means that any external CCT, or ECCT in what follows, must be based on observed child quality in the period(s) before and after the one-period policy intervention.

In the case of ECCTs, we will consider only threshold-based policies, which specify a threshold $\underline{k}_{t+1}(k_t)$, that is the lowest level of child quality in period $t+1$ for which the parents receive the reward ϕ_C. We allow this threshold to be a function of the child's cognitive ability when the policy is announced, so that we can consider growth-based policies, such as $\underline{k}_{t+1}(k_t) = (1 + \rho_C)k_t$, where ρ_C is the growth rate required of a household with current child quality k_t to receive the transfer ϕ_C.

Before proceeding, we mention a few restrictive features of our analysis. The potential policy space for the planner is enormous, and we only consider a small set of policies within this space. We do not consider explicit targeting of sub-populations based on potentially observable characteristics, such as family income or demographic characteristics. The reader can think of this targeting as having already been done, with the population we consider being the restricted sub-population of interest to the planner. Within this group of households with children, we only consider policies that transfer an identical amount of resources to each household. This may clearly be inefficient if there is substantial variation in resources across households within the target population. We can think of the initial targeting in terms of observable characteristics,

such as household income or wage rates, as having substantially reduced the variability in these characteristics within the targeted population.

We will also only consider policies that are targeted to households containing a child of age t. We limit our attention to "one-shot" policies of this form primarily because of the complications that arise when examining household behavior when there are multi-period ECCT policies. It is well known (e.g., Weitzman (1980)) that when facing a sequence of incentive contracts of this form, the household will engage in strategic behavior by altering its investments so as to enhance its ability to receive the reward in future periods. A nice example of this phenomenon in the case of incentive contracts offered to educators can be found in Macartney (2016). Although there are no complications in extending unrestricted and restricted transfer policies to a multi-period setting, since our primary interest is in comparing the performance of these three types of policies, we limit our attention to a single period for all policy types. The age upon which we primarily focus is age 15, so that $t = 15$. We will be examing the impact of these various policies on child cognitive ability at age 16.

6.1   Unrestricted Income Transfers

The most basic type of transfer that we consider consists of a money transfer of ϕ_U to parents of a child of age t. There are no restrictions on how the parents can spend the money, so that in the period within which it is received, the transfer increases household non-labor income to $I_t + \phi_U$. Since we do not allow for savings and borrowing in the model, this additional income must be spent in period t. Since child cognitive ability is a normal good to the parents, this will result in an increase in child ability in the following period with respect to its level if the transfer had not be received. From the viewpoint of maximizing household welfare (from the point of view of the parents), this is the (weakly) preferred type of transfer policy by the household of the ones that we will consider.[26]

Using the point estimates of the model parameters, we focus on the growth in child quality between ages 15 and 16 in the baseline and under the unrestricted transfer policy, while varying the level of the transfers.[27] Let g_t denote the average growth rate in child latent human capital between age t and $t + 1$, which is given by

$$g_t = \frac{1}{N \times R} \sum_{r=1}^{R} \sum_{h=1}^{N} \frac{k_{h,t+1}^r - k_{h,t}^r}{k_{h,t}^r}$$

where $k_{h,t+1}^r$ is the optimal level of child quality at age $t+1$ in household h in simulation round r (which is after the transfer has been received in period t), N is the number of

---

[26]This comparison is with a restricted transfer policy that offers an in-kind transfer of ϕ_U of child-investment goods and with a conditional cash transfer policy that pays ϕ_U when the household satisfies the requirements for receiving the transfer.

[27]Recall that we are assuming that all households receive the same amount of the transfer, ϕ_U. A more effective policy given the planner's objective may be to vary the transfer amount by the observable characteristics of the household.

households, and R is the number of simulation rounds. Given our estimates, the growth rate in child quality between the ages of 15 and 16 in the absence of policy intervention has a mean of approximately 12.3 percent (see part (a) in Table 9).

In Table 9 we examine the results of an unrestricted transfer of $170 dollars to week per sample household on a variety of household outcomes, most importantly, the cognitive ability of the child. Given our emphasis on the use of ICCTs, we also analyze the impact of this transfer in our baseline model with costly ICCTs and in the case when ICCTs are not available to any household, which can be thought of as the case in which the cost of their use is indefinitely large. In the baseline case, the unrestricted transfer has a relatively small positive effect on the level of child quality at age 16, 0.32 percent. This is in spite of the fact that the investment time of both parents individual and jointly increases noticeably, as do child expenditures. However, at this age, the impacts of parental time investments on the growth in the child's cognitive ability are modest at best. Expenditures on child investment goods have increased by 3.13 percent in response to the transfer, but child time investments actually decrease by a small amount.

The most striking impact of the transfer is on the labor supply of the parents, which falls by almost 8 percent for mothers and by 5.3 percent for fathers. Most of this time is spent on leisure consumption by the parents, which increases by 2.8 percent for mothers and 3.8 percent for fathers. Household consumption also increases markedly, with the parents' consumption increasing by 3.15 percent and the child's by 3.1 percent. The current period utility of both sets of agents increases by a small amount at ages 15 and 16. We see that the proportion of households using an ICCT is unchanged, and the reward elasticity r_{15} set when an ICCT is used is also essentially unchanged.

The impact of an unconditional transfer of $\phi_U = 170$ in the case in which no households utilize an ICCT is largely the same. The increase in the growth rate in child quality is slightly greater, but this is largely explained by the fact that the average level of k_{15} is much lower if households are precluded from using ICCTs. The proportional drop in child investment time is greater than in the case with costly ICCTs, but smaller in absolute terms since the average level of study time is so much lower in the no-ICCT case than in the costly ICCT case (2.71 hours versus 6.59 hours).

6.2   Restricted Income Transfers

A restricted income transfer can be thought of as an in-kind transfer in which the planner transfers child investment goods (e) directly to the household. This is a restricted transfer since the method in which it can be used by the household is (potentially) limited, unlike in the case of an unrestricted transfer. In this case, the parents can still re-optimize their entire set of choices, but under the additional constraint that the total expenditures on child goods cannot be less than the amount of the targeted transfer of ϕ_R. Such directed transfers are "restrictive" in the sense that they limit the parents' ability to increase their private consumption or leisure. The only impact on the household's period t problem is that (1) the period t non-labor income is now equal to

$I_t + \phi_R$, and (2) the constraint that $e_t \geq \phi_R$. At the household's period t solution, this constraint may be binding for some households but not for others, depending on the parents' preferences, other primitive parameters of the model, and the period t values of the state variables, Γ_t. If the parents are not using an Internal CCT, the optimal child expenditures in the presence of a restricted transfer policy can be defined as follows:

$$e_t^* = \max\{\frac{\beta_p \psi_{p,t+1} \delta_{4,t}}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t, \ \phi_R\}$$

where $Y_t = w_{1,t} h_{1,t} + w_{2,t} h_{2,t} + I_t + \phi_R$. If the parents are using an Internal CCT, we use a numerical solver to find the optimal child expenditures. Computationally, we treat the restrictive transfer ϕ_R as an unrestrictive transfer (as described in the previous subsection), and then verify whether the expenditure constraint is satisfied.

No household can be worse off in terms of parental welfare by accepting such a transfer. When the constraint that $e_t^* \geq \phi_R$ is not binding, then the unrestricted transfer is equivalent to an unrestricted transfer. When it is binding, the household must still be better off than in the absence of the transfer, since even at the pre-transfer choices, if all of the transfer is spent on additional investment in the child, child quality will increase relative to its pre-transfer level, and this will increase parental welfare since child quality is a normal good. Although these transfers will never lead to a decrease in the level of child expenditures, there can still be significant crowding-out by the parents in the sense that a large portion of the transfer may be diverted to other goods valued by the household. In comparison to an unrestricted transfer policy, this type of policy would be expected to be more costly to administer due to the monitoring required to ensure that the money transfer is spent on child investment, or making in-kind transfers directly.

In Table 10 we examine the impacts of a restricted transfer $\phi_R = 170$, the same amount of transfer income upon which we focused in Table 9 when the transfer was unrestricted (i.e., $\phi_U = 170$). There are extremely large differences between the two tables. In the case of costly ICCTs, child quality at age 16 increases by 1.13 percent with respect to the case of no transfer. The growth rate in child quality increases by over 12 percent. The time investments of the parents increase in a similar way as in the case of the unrestricted transfers, albeit a bit more markedly except for the joint time investment of the parents. The most noticeable change with respect to the earlier table is in terms of the money expenditures on the child, which increases by over 50 percent. If baseline expenditure on e_{15} had been greater than or equal to 170 for all households, then there is no difference between a restricted transfer of 170 or an unrestricted transfer of 170. We see instead, that average expenditures on e_{15} were 122.51 dollars before the policy intervention, and this is what accounts for the extreme large increase in e_{15}. Child self-investment time also rises markedly, under the policy, by 4.21 percent relative to the baseline case.

The impact of the transfer ϕ_R on household consumption is not as large as the similar money transfer of ϕ_U due to the fact that the parents are constrained in terms of how the transfer can be spent. Although the labor supply of both parents declines

relative to baseline, the decrease is not as large as in the unrestricted transfer case due to the restriction on the use of the transfer. As in the case of an unrestricted cash transfer, the restricted transfer has (almost) no impact on the proportion of households using an ICCT. However, it has a very large impact on the size of the reward elasticity, r_{15}. In the case of a transfer of ϕ_R, the reward elasticity increases by 11.68 percent relative to baseline, in comparison to no change in the unrestricted transfer of ϕ_U. This is what accounts for the large increase in τ_c under the ϕ_R policy.

When ICCTs are not available to the household, the impacts on child quality at age 16 are roughly the same as in the costly ICCT case. There are smaller increases in parental time investments than in the ϕ_U case, and a slightly smaller decrease in child investment time. In comparison with the costly ICCT environment, the impact on parental labor supply and the leisure of household members is slightly more muted. This is also true regarding consumption and utility levels, although the child has a slightly higher utility level at age 15 when the parents are unable to offer an ICCT.

6.3 External Conditional Cash Transfers

An External Conditional Cash Transfer (ECCT) specifies a requirement for a household to receive a monetary reward of ϕ_C. In the case of an internal incentive scheme (or ICCT), the parents specify a reward function $x_t^*(\tau_{c,t}; \Gamma_t)$, which is a function of the child's time input into the development process, $\tau_{c,t}$, and the set of characteristics describing the household environment when the child is of age t, Γ_t. As in DFW (2016), we will only consider a few simple types of these reward systems, although the ones we do consider could be (and have been) implemented in real world settings. We assume that the requirement that a household must meet in order to receive the cash award of ϕ_C in period $t+1$ can be written in terms of a function of the form $\underline{k}_{t+1}(k_t)$, and that the household will receive the reward if and only if $k_{t+1} \geq \underline{k}_{t+1}(k_t)$. The two leading examples of the function $\underline{k}_{t+1}(k_t)$ are $\underline{k}_{t+1}(k_t) = \underline{k}_{t+1}$ and $\bar{k}_{t+1}(k_t) = (1 + \rho_C)k_t$. In the first case, the reward is assumed to be a constant quality level specific to age $t+1$ children, one which is independent of the history of the child's abilities, $\{k_s\}_{s \leq t}$. In the second case, all households are required to achieve a growth rate of at least ρ_C in child ability between ages t and $t+1$ in order to receive the reward.

Since two sets of agents are involved in producing child quality, it is natural to consider whether the payoff ϕ_C should be paid to a single agent (parents or the child) or whether it might be optimally divided between the two. We restrict our attention to the case in which parents receive the entire award. This is somewhat of a natural choice under our informational assumptions. Since the external agent can only observe child cognitive ability at each age $t = 1, ..., M$, it cannot utilize input-based CCTs, as the parents potentially can. Therefore it cannot design a CCT that would remunerate both parents and the child individually for the quantities of inputs they supply to the development process.

We will be particularly interested in the manner in which the existence of external output-based CCTs paid to the parents affects the parental decision to use an ICCT.

Given our model estimates, virtually all parents in our sample use ICCTs to incentivize their children to self-invest more (rather than less), so that ICCTs produce higher levels of child ability at age $t + 1$. Since the parents are responsible for the ICCT decision, it seems natural for the external agent to make the payments ϕ_C to the parents, in order to increase the likelihood that the household uses an ICCT. Given our model estimates, we are able to evaluate the extent to which external and internal CCTs are complements in the production of child cognitive ability. If we interpret the use of an ICCT as a "parenting style," our results illustrate the manner in which the policies of external agents can influence the pattern of interaction between parents and children inside the household. Therefore, the possibility of using ICCTs has the potential to greatly amplify the impact of government policy on child development.

6.3.1 No Internal CCTs

We begin by considering the impact of an ECCT on household decisions in the absence of ICCTs, so that each household behaves in an inefficient manner. We have already seen that the scope for the parents to influence the choices of the child in this setting is limited. In particular, we know that they only influence the child's choice of their time input through their crowd-out effect on her available time, which is given by $\tilde{T}_t - \tau_{p,t}$. Thus, the only way that the parents can increase the child's time input is by reducing their own, that is, any increase in their own total time input in child production leads to a decrease in the child's study time. Their only other avenue to increase child quality is through increased expenditures on child investment goods, e_t.

In order to examine the impact of an external CCT in this case, we utilize the framework developed in DFW (2016). If the reward of ϕ_C is received in period $t + 1$, the parents' non-labor income in that period is increased from I_{t+1} to $I_{t+1} + \phi_C$. The first case to consider is when the parents' choices in period t, and the child's response in terms of her choice of $\tau_{c,t}$ in the absence of the external CCT would result in a level of child quality in period $t + 1$, k_{t+1}^*, that is no less than the threshold required to receive ϕ_C, so that $k_{t+1}^* \geq \underline{k}_{t+1}$. In this case, the household essentially receives an unrestricted income transfer in period $t+1$ without changing any of their original period t decisions.

The more interesting case is when $k_{t+1}^* < \underline{k}_{t+1}$. In this situation, the parents face a binary choice. First, they can make their original choices (and the child hers) so as to produce the quantity k_{t+1}^*, which means that they will not receive the reward. Second, assuming no uncertainty in the production process, they can alter their decisions (and the child's response) so as to receive the reward ϕ_C in the following period. This requires them to suffer a utility loss in period t for an increase in the continuation value of the parents in period $t + 1$, since in that period they will have a higher level of non-labor income and a higher level of child quality than in the absence of the policy. We define the maximum value of period t utility subject to the constraint that the household

49

produce k_{t+1} units of child quality in period $t+1$ (without using an ICCT) as

$$J_{p,t}^0(\Gamma_t, k_{t+1}) = \max_{a_{p,t}} \tilde{\alpha}_1 \ln l_{1,t} + \tilde{\alpha}_2 \ln l_{2,t} + \tilde{\alpha}_3 \ln c_t + \tilde{\alpha}_4 \ln k_t + \tilde{\alpha}_5 \ln l_{c,t} + \tilde{\alpha}_6 \ln x_t$$

$$\text{subject to: } k_{t+1} = f_t(k_t, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, \tau_{c,t}^*(\tau_{p,t}))$$

$$T = l_{i,t} + h_{i,t} + \tau_{i,t} + \tau_{12,t}, \ i = 1,2$$

$$c_t + e_t + x_t = w_{1,t} h_{1,t} + w_{2,t} h_{2,t} + I_t,$$

$$\tau_{c,t}^*(\tau_{p,t}) = \gamma_t(\tilde{T}_t - \tau_{1,t} - \tau_{2,t} - \tau_{12,t}).$$

This formulation of the problem can be used to define the household's problem when there is no External CCT, since

$$V_{p,t}^0(\Gamma_t) = \max_{k_{t+1}} J_{p,t}^0(\Gamma_t, k_{t+1}) + \beta_p V_{p,t+1}^0(\Gamma_{t+1})$$

and

$$k_{t+1}^*(\Gamma_t) = \arg\max_{k_t} J_{p,t}^0(\Gamma_t, k_{t+1}) + \beta_p V_{p,t+1}^0(\Gamma_{t+1}),$$

where the period $t+1$ vector of state variables, Γ_{t+1}, includes k_{t+1}. The value $k_{t+1}^*(\Gamma_t)$ is the baseline level of period $t+1$ child quality in the absence of the ECCT program.

When $k_{t+1}^* < \underline{k}_{t+1}$, the value of increasing child quality to $\underline{k}_{t+1}$ is given by

$$V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) = J_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) + \beta_{p,t} V_{p,t+1}^0(\Gamma'_{t+1}),$$

where the state variable vector Γ'_{t+1} includes the value of child quality $\underline{k}_{t+1}$ and non-labor income level of $I_{t+1} + \phi_C$. Then, in the absence of internal CCTs, the household will increase child quality to $\underline{k}_{t+1}$ if and only if

$$V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) > V_{p,t}^0(\Gamma_t)$$

$$\Rightarrow J_{p,t}(\Gamma_t, \underline{k}_{t+1}) + \beta_p V_{p,t+1}^0(\Gamma'_{t+1}) > J_{p,t}^0(\Gamma_t, k_{t+1}^*) + \beta_p V_{p,t+1}^0(\Gamma_{t+1})$$

$$\Rightarrow V_{p,t+1}^0(\Gamma'_{t+1}) - V_{p,t+1}^0(\Gamma_{t+1}) > \beta_p^{-1}[J_{p,t}^0(\Gamma_t, k_{t+1}^*) - J_{p,t}^0(\Gamma_t, \underline{k}_{t+1})].$$

Both sides of the last line are strictly positive when $\underline{k}_{t+1} > k_{t+1}^*$.

6.3.2 Relationship between External CCTs and the Use of Internal CCTs

In order for a household facing the requirement to produce more child quality in period $t+1$ than it would in the absence of the external CCT, it is obviously the case that its use of one or more of the inputs in the period t production process must be increased relative to its baseline level. In the single agent case considered in DFW (2016), the parents would increase all inputs in an efficient manner (i.e., with minimum loss in period t utility). In the two-agent case considered here, and in the absence of an internal CCT, an increase in total parental time investments in the child will mechanically lead to a reduction in the time the child spends investing in herself. Of course, if the value of the child's own time investments are low, this may not be a major impediment to reaching

the target $\underline{k}_{t+1}$ with minimal loss of current period utility. However, when this is not the case, increasing parental time investments and decreasing child time investments may require a significant increase in purchased inputs in period t, e_t.

Under an ICCT, the parents are able to influence the input chosen by the child, $\tau_{c,t}$ by choosing a reward function, subject to an incentive compatibility constraint. In this case, the parents are not subject to the same mechanical crowd-out effect that is in effect when there is no ICCT. This allows them to increase the time investments of the child as they increase their own in any situation where it is optimal to do so. The period utility cost of implementing an internal CCT is given by $\omega_t > 0$ in period t.[28] In order to clarify the decisions in this case, it will be useful to distinguish between the payoff values to the household (1) whether they produce the $\underline{k}_{t+1}$ required to receive ϕ_C in period $t+1$ and (2) whether they use an ICCT in period t. Let the value of the parents' problem when using an ICCT be given by $V_{p,t}^1(\Gamma_t)$, where this value does not include the fixed (to the parents) implementation cost ω_t (in utils). As we discussed previously, we know that $V_{p,t}^1(\Gamma_t) > V_{p,t}^0(\Gamma_t)$ for all Γ_t. Then in the absence of an external CCT, the household implements an ICCT if and only if

$$V_{p,t}^1(\Gamma_t) - V_{p,t}^0(\Gamma_t) > \omega_t.$$

In the case for which no internal CCT is used, we defined $V_{p,t}^0(\Gamma_t, \underline{k}_{t+1})$ as the value of the households' problem when it produced $\underline{k}_{t+1}$ over the course of period t. Analogously, we define $V_{p,t}^1(\Gamma_t, \underline{k}_{t+1})$ as the value of the households' problem when it produces $\underline{k}_{t+1}$ during period t and utilizes an ICCT in period t; as above, this does not include the utility cost, ω_t. We know that $V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) > V_{p,t}^0(\Gamma_t, \underline{k}_{t+1})$ for all target levels $\underline{k}_{t+1}$. In period t, the household makes choices along two dimensions: (1) whether to use an internal CCT and (2) whether to produce a level of child quality exactly equal to the required amount, $\underline{k}_{t+1}$. We will now discuss the conditions required for the various outcomes to be observed. We will not explicitly condition on the fact that $k_{t+1}^* < \underline{k}_{t+1}$, where k_{t+1}^* is the unconstrained choice in the absence of an ECCT, when parents endogenously choose whether to use an ICCT. Whenever this inequality is not satisfied, the household can never be better off producing the quantity $\underline{k}_{t+1}$, that is, the targeted value of $\underline{k}_{t+1}$ will never be chosen by those households who will receive ϕ_C next period with no change in their period t decisions. If a household receives ϕ_C in period $t+1$, either $k_{t+1} = \underline{k}_{t+1}$ or $k_{t+1} > \underline{k}_{t+1}$, and in the latter case, we say that ϕ_C received in period $t+1$ is an unrestricted transfer.[29]

Each household must fall into one of the following four distinct cases.

---

[28] In the estimation of the model, we will assume that the cost of running an ICCT is fixed over time within a household, so that $\omega_1 = ... = \omega_M$. There is a distribution of these time invariant costs across households.

[29] It is, in principle, possible for a household to freely choose exactly $\underline{k}_{t+1}$ with no ECCT present, but since preferences are continuous random variables, this is a zero-probability event.

1. (No ICCT, $k_{t+1} \neq \underline{k}_{t+1}$) In this case, we have

$$
\begin{aligned}
V_{p,t}^1(\Gamma_t) - V_{p,t}^0(\Gamma_t) &\leq \omega_t \\
V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t) &< 0 \\
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t) &\leq \omega_t
\end{aligned}
$$

The first line is the condition that, in the absence of the ECCT, the parents would not implement an ICCT. The second line is the condition that the household, in the absence of an ICCT, would choose not to produce the level $\underline{k}_{t+1}$ required to receive ϕ^{CCT}. The last line is the condition that value of the household producing $\underline{k}_{t+1}$ using an ICCT relative to not using an ICCT and not receiving the reward is less than the cost of the ICCT.

2. (ICCT, $k_{t+1} \neq \underline{k}_{t+1}$) In this case, we have

$$
\begin{aligned}
V_{p,t}^1(\Gamma_t) - V_{p,t}^0(\Gamma_t) &> \omega_t \\
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^1(\Gamma_t) &< 0
\end{aligned}
$$

The first line is the condition that the unconstrained choice of k_{t+1} yields a higher value using an ICCT than when not using one. The second line is the condition that producing exactly $\underline{k}_{t+1}$ using an ICCT yields a lower value than not producing $\underline{k}_{t+1}$. Recall that this condition covers also those individuals for whom $k_{t+1}^* > \underline{k}_{t+1}$, for which the reward of ϕ_C is the same as an unrestricted transfer of the same amount, $\phi_U = \phi_C$.

3. (No ICCT, $k_{t+1} = \underline{k}_{t+1}$) Here we have the conditions

$$
\begin{aligned}
V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t) &\geq 0 \\
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) &\leq \omega_t.
\end{aligned}
$$

The first condition is that, in the no ICCT state, the value of producing $\underline{k}_{t+1}$ exceeds the value of producing k_{t+1}^*. The second condition is that the value of producing $\underline{k}_{t+1}$ using the ICCT is less than the value of producing $\underline{k}_{t+1}$ without using the ICCT.

4. (ICCT, $k_{t+1} = \underline{k}_{t+1}$) The conditions are

$$
\begin{aligned}
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^1(\Gamma_t) &\geq 0 \\
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t, \underline{k}_{t+1}) &> \omega_t \\
V_{p,t}^1(\Gamma_t, \underline{k}_{t+1}) - V_{p,t}^0(\Gamma_t) &> \omega_t
\end{aligned}
$$

The first condition is that, when using an ICCT, the value of producing $\underline{k}_{t+1}$ is no less than the value of not producing it. The second condition is that the value of producing $\underline{k}_{t+1}$ is greater when using an ICCT than when not. The third condition is that the value of producing $\underline{k}_{t+1}$ using the ICCT exceeds the value of not using an ICCT and not producing $\underline{k}_{t+1}$ units of quality.

6.3.3 Quantitative Results: Growth-based ECCTs

In what follows, we let k_{t+1}^* and k_{t+1}^{**} denote the equilibrium levels of latent child capital at age $t+1$ in the baseline and in the presence of an ECCT, respectively. In a growth-based ECCT, the planner specifies a triple of policy parameters (t, ρ_C, ϕ_C). The first parameter denotes the age at which the policy is announced and implemented. As mentioned above, to avoid having to consider issues of intertemporal strategic behavior, we restrict the ECCT policies to be limited to one period and to be unanticipated by the population of households. Throughout the empirical exercise, we set $t = 15$. At this age the child's time investment is more valuable in producing cognitive ability, and from our estimates, a large proportion of parents are employing ICCTs. The second parameter denotes the target growth rate of the child's (latent) human capital, with the target for a family with age t child quality of k_t given by

$$\underline{k}_{t+1} = (1 + \rho_C)k_t$$

The last parameter, ϕ_C, denotes the financial reward received by the household in period $t+1$ if $k_{t+1} \geq \underline{k}_{t+1}$.

We will be especially interested in comparing the efficiency of the three transfer programs: unrestricted, restricted, and conditional, where efficiency will be defined as the increase in average child ability relative to baseline for a given per capita transfer payment, with all of our comparisons involving a one-time policy involving families of 15 year olds. Thus, we will be comparing the values of k_{16} in the baseline with its value under the policy. The payoff to the unrestricted and restricted policies is relatively straightforward to determine. Since child quality is a normal good in all households, the average child quality in the targeted households is a continuous and increasing function of ϕ_U or ϕ_R.

In the case of an ECCT, it is somewhat more complicated to determine the cost of a program because of the variability in the take-up rate. In this case, we begin by positing a given desired improvement in average growth with respect to baseline. If baseline average growth in cognitive ability between year t and year $t+1$ among our target population of size N is defined as $g_t = N^{-1} \sum_{h=1}^{N} \frac{k_{h,t+1}^* - k_{h,t}}{k_{h,t}}$, then we specify a desired growth rate of $g_t^{**} = N^{-1} \sum_{h=1}^{N} \frac{k_{h,t+1}^{**} - k_{h,t}}{k_{h,t}}$, so that we look for the program parameters that minimize average per capita cost and that produce an average child quality growth of g_t^{**}, where the planner restricts himself to choosing $g_t^{**} \geq g_t$. It is important to note that the program parameters that are chosen are those that satisfy

$$(\rho_C, \phi_C)^*(g) = \arg \min_{\substack{\rho_C, \phi_C \\ \text{s.t. } g_{t+1}^{**}(\rho_C, \phi_C) = g}} \frac{N_C(\rho_C, \phi_C)}{N} \phi_C, \tag{5}$$

where N_C is the number of households that receive ϕ_C, N is the size of the target population, and $g_{t+1}^{**}(\rho_C, \phi_C)$ is the average growth in child cognitive ability in the target population under the ECCT. The minimand is average per capita cost. The take-up rate is $N_C(\rho_C, \phi_C)/N$.

Because it is difficult to determine the per capita cost of an optimally-designed ECCT chosen to achieve a growth rate of g, we have solved (5) for a variety of values of g, and then interpolated the relationship between per capita cost and g in the figures described below, which examine the relationship between per capita cost and the growth in child quality in period $t + 1$ relative to baseline.

In Table 11 we examine the impact of an ECCT offering $\phi_C = 250$ dollars in the event that the growth in child quality between ages 15 and 16 is at least $\rho_C = 20$ percent, for the benchmark model of a costly ICCT. In the comparison with the situation of no ICCT, the growth requirement is reduced to $\rho_C = 17.5$ percent because of the lower baseline growth rate when ICCTs are excluded. In comparison with the $\phi_U = \phi_R = 170$ cases in Tables 9 and 10, we see dramatic increases in the average level of k_{16} under the ECCT policy. With respect to the baseline, the proportional increases in the time investments of the parents and the child are approximately an order of magnitude larger under the ECCT than under restricted transfers. Because of the restriction on the use of the transfer, the expenditures on child investment are higher under the restricted transfer case.

The parents' time spent in leisure and at work decline markedly, unlike the situation under restricted or unrestricted transfers to the household. The child's time investment increases by 30.6 percent, as parents are increasingly likely to use ICCTs and to increase the reward elasticity. The proportion of households using an ICCT increases by 21.7 percent, and the average reward elasticity of all households using an ICCT increases by 103.59 percent. The consumption of parents and the child fall signficantly at age 15, but recall that meeting the requirement of the ECCT increases household income when the child is 16 and increases the terminal value of both parents and the child by increasing the terminal value of child quality. We see that the utility of both parents and the child at age 16 are increased under the policy.

When there is no possibility of using an ICCT, the ECCT policy still has a substantial impact on average child quality at age 16, increasing it by 2.92 percent. This is less than what is achieved when it is possible to utilize an ICCT, largely because it expands the parents' ability to influence the investment decision of the child. With no ICCT, increases in the parents' time investments in the child only serve to crowd out the child's time investments. With an ICCT, the parents can counteract this by increasing the incentive offered to their child to study. In the baseline case of a costly ICCT, child study time increased by over 30 percent even in the face of dramatic increases in the parents' own investment time. Instead, in the situation with ICCTs, the child's time investment falls by 4.84 percent in response to the large increase in the parents' time investments. Other responses of the parents and the child in the no ICCT case are not as dramatically different from the baseline situation of costly ICCTs.

In Table 12 we compare the impacts of various ECCTs that vary in their per capita cost and have been designed to meet a particular target gain in average k_{t+1} with respect to the baseline of no ECCT. Column (0) contains the baseline situation for the costly ICCT case. The average growth in latent child ability between ages 15 and 16 is 12.3 percent, with a standard deviation of 4.1 percent. We find that 72 percent of

households are using an ICCT when their child is 15 years of age. Among households using an ICCT at age 15, the average reward elasticity is 0.239.

Column (1) contains the results from imposing an ICCT that produces a gain in the average child quality growth of of 2.1 percentage points relative to baseline, or $g_{t+1}^{**} = 14.4$ percent. This gain is substantially greater than what was realized under either the unrestricted or restricted transfers of $\phi_U = \phi_R = 170$, and yet the per capita cost of this gain is only 36, or 21 percent of the amount given to all families under the unconditional transfers. The optimal combination of ρ_C and ϕ_C in this case is $\rho_C = 0.185$ and $\phi_C = 75$, so that even those families who qualify for the transfer receive 95 less than under the unconditional transfers that we examined in Tables 9 and 10. We see that the standard deviation of child quality at age 16 slightly increases under this particular ECCT.

The threshold-based CCT pays all parents whose child's latent ability at age 16 satisfies $k_{16} \geq \underline{k}_{16}$. In some cases, the child would have met the requirement even in the absence of the program, that is, whenever $k_{16}^* \geq \underline{k}_{16}$. In this case, the transfer is the same as an unrestricted transfer $\phi_U = 75$. In other cases, when $k_{16}^* < \underline{k}_{16}$, the parents may decide to devote additional resources to child investment so as to meet the threshold, and in this case we have $k_{16}^{**} = \underline{k}_{16}$. In order to judge the effectiveness of the incentive device, it is important to separate these two cases when we examine the overall take-up rate. In panel (c) of Table 12, we see that the take-up rate for this program is 48 percent. In order to receive the transfer, 42.1 percent of the target population have to increase their pre-policy levels of investment so as to qualify for the transfer. Under this metric, the incentives built into the program are quite effective.

Our main interest is in the relationship between ECCTs and parents' decisions to employ ICCTs. In panel (d), we see that there is an increase in the proportion of households using an ICCT of 6.1 percentage points with respect to baseline, even with the rather low transfer amount of $\phi_C = 75$. Prior to the policy, among the subset of households who would adjust their baseline level of k_{16}^* in order to qualify for the transfer, 81.6 percent were using an ICCT. However, in the presence of the ECCT, the proportion of these households that use an ICCT in order to receive the reward, in which case $k_{16}^{**} = \underline{k}_{16}$, is 96.7 percent. We see that the ICCT is a crucial element of reducing the current period utility loss to the parents of qualifying for the reward through the incentivization of their child to spend more time studying. This is also reflected in changes in the average reward elasticity among those households using an ICCT. In the baseline, this average was 0.239, while under the policy the average reward elasticity increases to 0.309.

In columns (2) through (4) of Table 12, we report results for more expensive ECCTs. In column (2), the gain in average growth relative to baseline is 3.2 percentage points, or $g_{t+1}^{**} = 15.5$ percent, and the per capita cost of this program is 72.2. The optimal target growth rate for this targeted increase in average quality is $\rho_C = 0.195$, and the reward is $\phi_C = 135$. In comparison with the baseline of no program, we see that besides increasing the average level of k_{16} by 2.39 percent, the standard deviation of k_{16} decreases slightly from 5.493 to 5.480. The take-up rate increases to 53.5 percent,

with 93.5 percent of the households receiving the transfer increasing k_{16}^* to $\underline{k}_{16}$. Of the individuals who would increase k_{16}^* to $\underline{k}_{16}$ in order to receive the transfer, 78.2 percent were using an ICCT before the policy. After the policy, the use of ICCTs among this subset of incentivized households increases to 97.8 percent. The reward elasticity also increased substantially among households using an ICCT, to 0.347 from 0.309 under the less ambitious and costly ECCT in column (1).

These patterns continue to be observed as we increase the size of the targeted change in average child quality. In columns (3) and (4), the increase in the average growth rate g_t with respect to baseline is 4.2 and 4.8 percentage points, respectively. The per capita costs of these programs are 115.4 and 146.3, still far less than the unconditional policies that we considered. We see that under both of these policies, the substantial increase in average child quality also occurs with a small reduction in inequality, as measured by the standard deviation of k_{16}. As was noted when discussing the less ambitious ECCTs, the increased utilization of ICCTs is crucial in producing a relatively high take-up rate for modest program costs. As the reward threshold increases, we also see substantial increases in the reward elasticities, as noted earlier.

6.4 Comparison of the Effectiveness of the Policies

We conclude our policy analysis by examining the cost effectiveness of the unrestricted, restricted, and conditional cash transfer programs in terms of the gain in the average growth rate in child quality with respect to baseline (i.e., $g_t^{**} - g_t$) by per capita cost. The results are summarized in Figure 4.

In the Figure, we plot six curves, which correspond to the three types of policies with and without costly ICCTs. As we have seen previously, the existence of an unrestricted transfer of $\phi_U = 170$ had no impact on the proportion of households using an ICCT. In the case of a restricted cash transfer of $\phi_R = 170$, we saw that there was a negligible impact on the proportion of households using an ICCT. In contrast, conditional cash transfer programs of various types, in terms of ρ_C and ϕ_C, increased the usage of ICCTs markedly. Therefore we expect that the cost-effectiveness advantage of the conditional cash transfer policy to decrease when ICCTs are not available to households.

We see large differences between the programs in terms of cost effectiveness, whether or not ICCTs are available to be utilized by households. For example, for a per capita expenditure of 100, an unconditional cash transfer produces well less than an 0.3 percentage point average gain in child quality at age 16, whereas a restricted transfer of 100 produces an average gain of approximately 0.5 percentage points. Instead, the average gain for an optimally-designed ECCT is approximately 3.8 percentage points . This advantage decreases if households do not have access to ICCTs, but is still substantial even in that case.

We have considered very expensive programs as well, up to a per capita cost of 400. In this case, unrestricted transfers yield an average gain in child cognitive ability at age 16 of less than 1 percentage point, while restricted transfers of that amount are able to produce slightly more than a 3 percentage point gain. An ECCT of that amount

is able to produce more than a 6 percentage point gain in our baseline model (with ICCTs available). However, in all cases, we should note that the gains with respect to increased per capita expenditures are decreasing rapidly at high levels of transfers. This indicates the limits of these types of programs in a setting in which the social planner must face intermediaries (parents) with very different objectives.

We have been comparing policies almost exclusively in terms of their impacts on average child cognitive ability, but inequality is an important consideration, as well. Even if the superiority of ECCT programs was clearly established in terms of average quality improvements, we might feel differently about such programs if they greatly increased inequality in the child cognitive ability distribution. In Figure 5, we plot the standard deviation of k_{16} under the various programs with respect to the standard deviation without any policy. We find that the ECCT program actually acquits itself well under this criterion, as well. The ECCT decreases variability at any reasonably costly program level with respect to baseline. Restricted cash transfers actually tend to increase inequality in k_{t+1}, whereas unrestricted cash transfers have little effect on inequality. Therefore, the case in favor of ECCTs (especially when ICCTs are available) is strengthened. In addition to be the most cost-effective way to increase average child quality (among the small set of policies that we have considered), they also seem to reduce inequality in the distribution of cognitive ability.

7 Conclusion

We observe that investments of children in themselves, primarily by studying independently, increase as the child matures. We develop a model of the child's cognitive growth in which both parents and children take an active role. Within our modeling framework, there exist three primary mechanisms that produce such a result. The first is associated with changes in the productive value of the child's study time as she ages, which is consistent with our model estimates. We estimate that the child's time inputs become more valuable than the parents' at later stages of the development period. A second mechanism is the increase in the child's valuation of future events, reflected in a non-decreasing discount factor sequence over the development period. This phenomenon has been well-documented in research conducted by developmental psychologists. As the child becomes more forward-looking, the incentive to invest in their own human capital increases. The third mechanism is related to the choice of parenting styles. By paying a cost, parents can incentivize their children to invest their time as the parents prefer through linking the child's consumption to the amount of time they spend studying. Since the value of this incentive mechanism increases with the child's age as their study time increases in value, and since parents tend to place a higher value on child ability than does the child, the use of these contracts increases over the development period and produces, generally speaking, higher levels of cognitive ability than would be produced in the absence of this mechanism.

Our estimates of the model parameters are reasonable, by and large, and fit the

data acceptably well given the limited sample size available to us ($N = 247$) and the relative complexity of the model. We estimate that an increasing share of households use incentive mechanisms to motivate their children to study more as their children age. At the end of the development period (age 16 in our case), over 80 percent of households utilize this incentive mechanism. Using our baseline parameter estimates, we show that the costliness of using this mechanism has sizable effects on the population distribution of cognitive ability in the population at the end of the development process. For households in which the child is 15 years old, we estimate that 72 percent of households use ICCTs, and that the growth rate of child quality is 12.3 percent. When we preclude the possibility of using an ICCT, the growth rate drops to 10.5 percent.

We show that the choice of parenting styles is important when considering the design of transfers to households that are designed to increase investments in child cognitive ability. In threshold-based reward systems, in which the parents are given a monetary transfer if their child exhibits an increase in cognitive ability greater than some predetermined amount, the use of an ICCT can alleviate the utility cost of reaching the threshold and receiving the reward. In our policy experiments, we show that the increase in the use of ICCTs in the presence of these types of external reward systems can be substantial. Without the potential to use such rewards, these programs would be less cost effective from the perspective of the social planner.

References

[1] Agostinelli, F. and M. Wiswall, 2016. "Estimating the technology of children's skills formation", NBER Working Paper 2244.

[2] Akabayashi, H., 1996. "On the role of incentives in the formation of human capital in the family", Dissertation, University of Chicago.

[3] Akabayashi, H., 2006. "An Equilibrium Model of Child Maltreatment", *Journal of Economic Dynamics and Control*, Elsevier, vol. 30(6), pages 993-1025, June.

[4] Bagger, J. F. Fontaine, F. Postel-Vinay, J.M. Robin 2014 "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics", *American Economic Review*, 104, 6:1551-96.

[5] Barczyk, B. and M. Kredler, 2018. "Evaluating Long-Term Care Policy Options, Taking the Family Seriously", *Review of Economic Studies*, 85(2), 766-809.

[6] Becker, G. S., 1962. "Investment in Human Capital: A Theoretical Analysis", *Journal of Political Economy*, University of Chicago Press, vol. 70:9-29.

[7] Becker, G. S. and N. Tomes, 1976. "Child Endowments and the Quantity and Quality of Children", *Journal of Political Economy*, University of Chicago Press, vol. 84(4): S143-62, August.

[8] Ben-Porath, Y., 1967. "The Production of Human Capital and the Life Cycle of Earnings", *Journal of Political Economy*, 75(4): 352-365.

[9] Boar, C., 2018. "Dynastic Precautionary Savings", Working paper.

[10] Bonke, J., 2013. "Do Danish children and young people receive pocket money?", Study Papers 57, Rockwool Foundation Research Unit.

[11] Carneiro, P. and J.J. Heckman, 2003. "Human Capital Policy", in: J. Heckman and A. Krueger, Inequality in America: What Role for Human Capital Policy?, MIT Press.

[12] Castillo M., Ferraro P., Jordan J. and R. Petrie, 2011. "The today and tomorrow of kids: Time preferences and educational outcomes of children", *Journal of Public Economics*, 2011, vol. 95, issue 11, pages 1377-1385.

[13] Cooper, H., Lindsay, J. J., Nye, B., and S. Greathouse, 1998. "Relationships among attitudes about homework, amount of homework assigned and completed, and student achievement", *Journal of Educational Psychology*, 90(1), 70-83.

[14] Cooper, H., Robinson J. C. and E.A. Patall, 2006. "Does homework improve academic achievement? A synthesis of research, 1987–2003", *Review of Educational Research*, 76(1): 1–62.

[15] Cosconati M., 2009. "Parenting style and the development of human capital in children", Dissertation available from ProQuest, AAI3363272.

[16] Cunha, F., Heckman, J.J., and L. Lochner, 2006. "Interpreting the Evidence on Life Cycle Skill Formation", in E. Hanushek, S. Machin, and L. Woessman,(eds.), *Handbook of the Economics of Education*, Vol. 1, 697-812, Amsterdam: Elsevier.

[17] Cunha F., and J.J. Heckman, 2008. "Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation", *Journal of Human Resources*, 43(4): 738-782.

[18] Cunha, F., Heckman, J.J., and S.M. Schennach, 2010. "Estimating the technology of cognitive and noncognitive skill formation", *Econometrica*, 78(3), 883-931.

[19] Currie, J., and D. Almond, 2011. "Human capital development before age five", in David Card and Orley Ashenfelter (eds.), *Handbook of Labor Economics*, vol. 4B, 1315-1486, Amsterdam: Elsevier.

[20] Dauphin, A., El Lahga, A.R., Fortin, B., and G. Lacroix 2011. "Are Children Decision-Makers within the Household?", *Economic Journal*, 121:871-903.

[21] Del Boca, D. and C. Flinn, 2012. "Endogenous household interaction", *Journal of Econometrics*, 166, issue 1, p. 49-65.

[22] Del Boca, D., Flinn, C., and M. Wiswall, 2014. "Household choice and child development", *Review of Economic Studies*, 81:137-185.

[23] Del Boca, D., Flinn, C., and M. Wiswall, 2016. "Transfers to families with children and child outcomes", *Economic Journal*, 526: F138-F183, October 2016.

[24] Del Boca, D., Monfardini, C. and C. Nicoletti, 2017. "Parental and child time and cognitive development of adolescents", *Journal of Labor Economics,* 35(2):565-608.

[25] Fang, H. and D. Silverman (2009). "Time-inconsistency and Welfare Program Participation: Evidence from the NLSY", *International Economic Review*, 50: 1043-1077.

[26] Fiorini, M., and M. Keane, 2014. "How the Allocation of Children's Time Affects Cognitive and Noncognitive Development", *Journal of Labor Economics*, 32(4): 787 - 836.

[27] Flinn, C.J., Gemici A., S. Laufer, 2017. "Search, Matching, and Training", *Review of Economic Dynamics*, 25:260-297.

[28] Flinn, C. and J. Mullins, 2015. "Labor Market Search and Schooling Investment", *International Economic Review*, 56 (2): 359-398.

[29] Hao L., J. Hotz and Ginger Z. Jin, 2008. "Games Parents and Adolescents Play: Risky Behaviour, Parental Reputation and Strategic Transfers", *Economic Journal*, vol. 118(528): 515-555.

[30] Hofferth, S., Davis-Kean, P., Davis, J., and Finkelstein, J., 1998. "User guide for the Child Development Supplement of the Panel Study of Income Dynamics", http://psidonline.isr.umich.edu/CDS/usergd.html.

[31] Hotz, V. and J. Pantano, 2015. "Strategic parenting, birth order, and school performance",*Journal of Population Economics*, Springer;European Society for Population Economics, vol. 28(4), pages 911-936, October.

[32] Kaplan, G., 2012. "Moving Back Home: Insurance Against Labor Market Risk", *Journal of Political Economy*, vol. 120, no. 3.

[33] Kooreman, P., 2007. "Time, money, peers, and parents; some data and theories on teenage behavior", *Journal of Population Economics*, 20(1): 9-33.

[34] Lentz, R. and N. Roys, 2015. "Training and Search On the Job", NBER Working Papers 21702, National Bureau of Economic Research.

[35] Lizzeri, A. and M. Siniscalchi, 2008. "Parental Guidance and Supervised Learning", *Quarterly Journal of Economics*, 123(3): 1161-1195.

[36] Lord, F. M. and M. Novick, 1968. "Statistical Theories of Mental Test Scores", Reading MA: Addison-Welsley Publishing Company.

[37] Lundberg, S. and R. A. Pollak, 1993. "Separate Spheres Bargaining and the Marriage Market", *Journal of Political Economy*,101, no. 6, Dec. 1993, 988-1010.

[38] Lundberg, S., J. Romich, and K. Tsang, 2009. "Decision-making by children", *Review of Economics of the Household*, 7: 1-30.

[39] Macartney, H., 2016. "Incentives in Education and Teacher Effort", mimeo.

[40] Nishiyama, S., 2002. "Bequests, Inter Vivos Transfers, and Wealth Distribution", *Review of Economic Dynamics*, 5(4), 892-931.

[41] Steinberg, L., Brian, A., Cauffman, E., Graham, S., Wooland, J. and A. Banich, 2009. "Age Differences in Future Orientation and Delay Discounting", *Child Development*, 80:28-44.

[42] Wasmer, E. 2006. "General versus Specific Skills in Labor Markets with Search Frictions and Firing Costs", *American Economic Review*, American Economic Association, vol. 96(3), pages 811-831.

[43] Weinberg, B., 2001. "An Incentive Model of the Effect of Parental Income on Children", *Journal of Political Economy*, 109(2) (April), pp. 266-280.

[44] Weitzman, M., 1980. "The Ratchet Principle and Performance Incentives", Bell Journal of Economics, 11, issue 1, p. 302-308.

[45] Woodcock, R., and M. Johnson, 1989. "Tests of Achievement, Standard Battery" [Form B], Chicago, IL: Riverside Publishing (Houghton Mifflin).

Figure 1: Child Self-investment Time by Income Quartile and Child Age
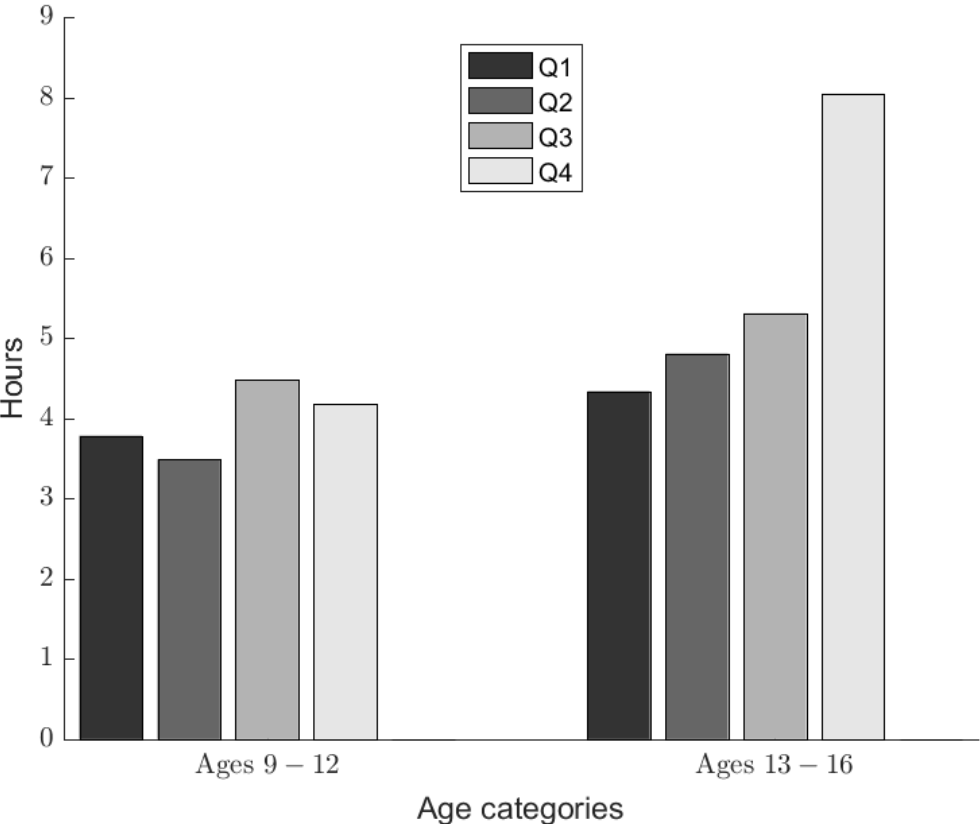
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Notes: Income quartiles are defined based on weekly household income, after aggregating all observations for a given household into a single income measure. The bars represent the average weekly child self-investment time for households within each of the four income quartiles, when the child is either between $9-12$ years old or between $13-16$ years old.

Figure 2: Average Child's Letter Word Score
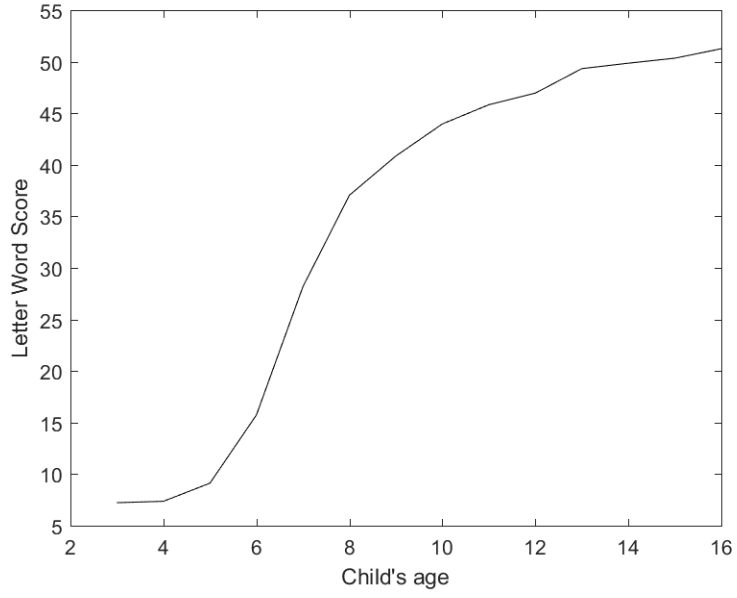
Source: CDS combined sample from 1997, 2002 and 2007 interviews.

Figure 3: Estimated Technology Parameters by Child Age

(a) All Time Inputs (avg. across all households)
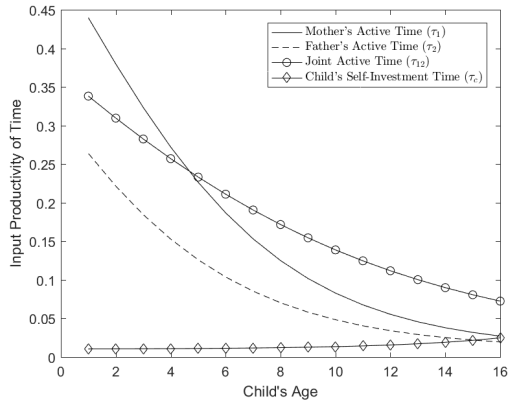
(b) Indiv. Parental Time, High School vs. College
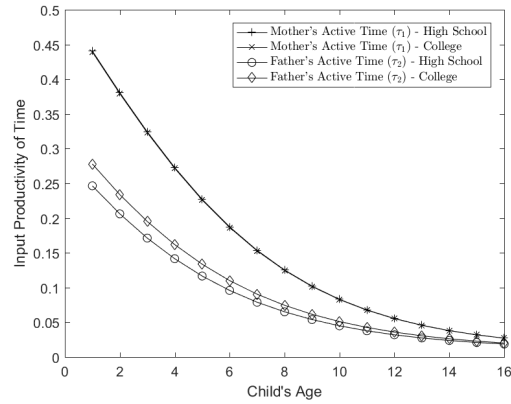
(c) Joint Parental Time, High School vs. College
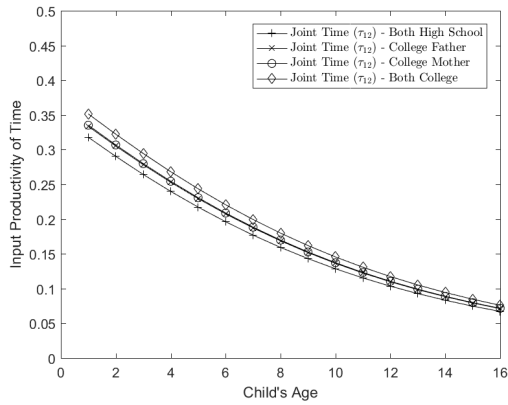
(d) Child Goods
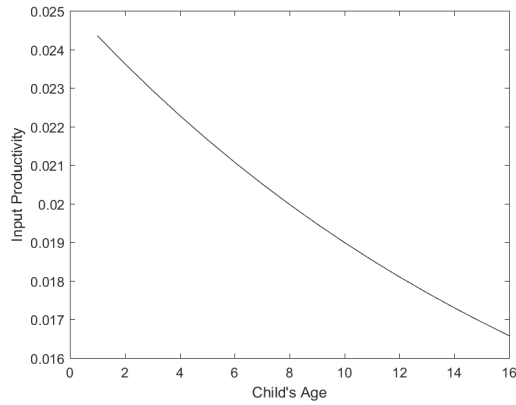
(e) Lagged Child Quality
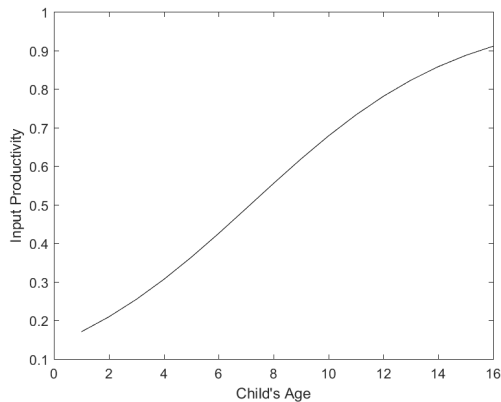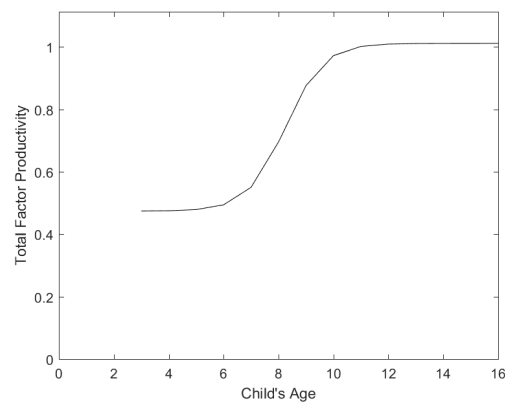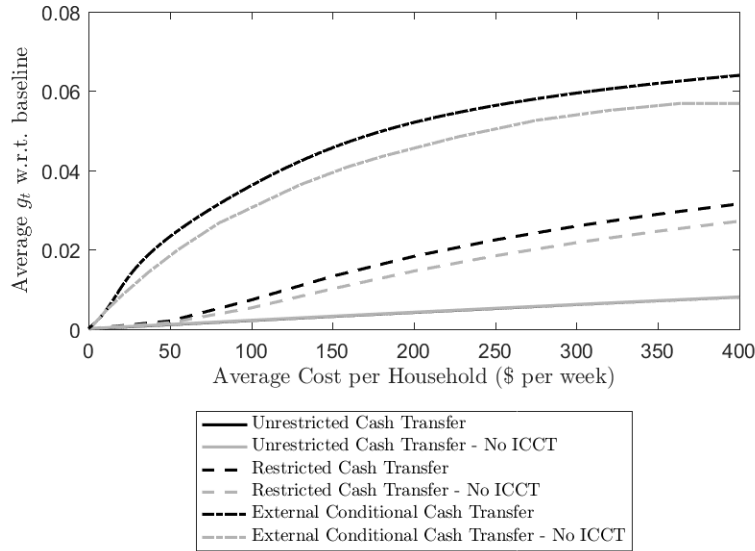

(f) Total Factor Productivity

64

Figure 4: Counterfactual Simulations: Cost-effectiveness of each policy

Notes: Each line plots the increase in the average growth rate in the child's latent value of child quality between the ages of 15 and 16, under a variety of counterfactual policy experiments where each household receives a one-time transfer at the start of either period $t = 15$ (for the UCT and RCT) or $t = 16$ (for the ECCT). The horizontal axis shows the average cost per household, in dollars per week. For the ECCT, this is equal to the product of the transfer amount and the take-up rate. The solid lines correspond to pure income transfers, where the subsidy simply accrues to the household's non-labor income (I_t), without any restrictions on how the money should be spent. The dashed lines correspond to targeted transfers, where the subsidy can only be spent directly on child goods during the period in which the transfer occurs (e_t). The dash-dotted lines correspond to External CCTs, where we only plot the convex hull created by the "optimal" policies, i.e. those achieving the largest gains in average growth of child quality at every average cost level. In order to see how the presence of Internal CCTs affects the effectiveness of each type of policy, we compare policy results for the benchmark model with endogenous (costly) Internal CCTs (i.e. the black lines) to those implied by the model where Internal CCTs are not available (i.e. the grey lines). Note that the baseline growth in child quality is lower for the model without Internal CCTs. For comparison, the vertical axis therefore plots the increase in the average growth rate relative to the respective baseline growth rate for each of the two models. All experiments were done using $R = 10$ simulated data sets.

Figure 5: Counterfactual Simulations: Distributional effects of each policy ($\sigma(k_{t+1})$)
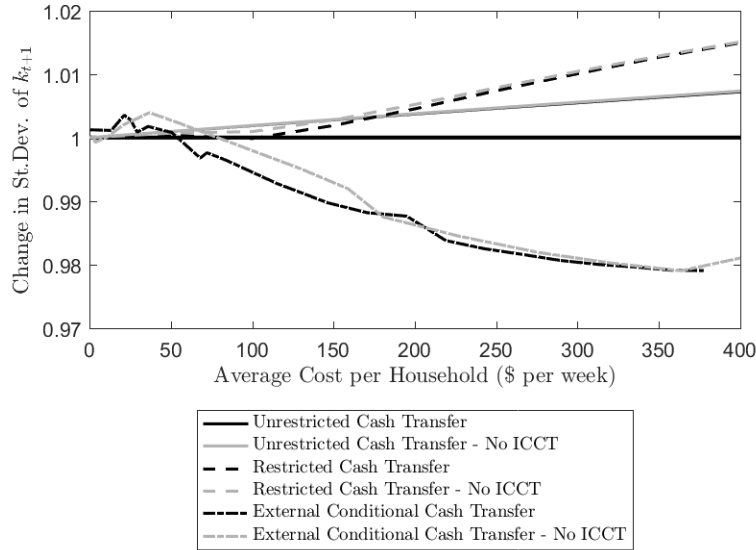
Notes: Each line plots how the standard deviation of the child's latent value of child quality at the age of 16 (i.e. $\sigma(k_{16})$) changes under a variety of counterfactual policy experiments, where each household receives a one-time transfer at the start of either period $t = 15$ (for the UCT and RCT) or $t = 16$ (for the ECCT). The horizontal axis shows the average cost per household, in dollars per week. For the ECCT, this is equal to the product of the transfer amount and the take-up rate. The solid lines correspond to pure income transfers, where the subsidy simply accrues to the household's non-labor income (I_t), without any restrictions on how the money should be spent. The dashed lines correspond to targeted transfers, where the subsidy can only be spent directly on child goods during the period in which the transfer occurs (e_t). The dash-dotted lines correspond to External CCTs, where we only plot the results for the "optimal" policies, i.e. those achieving the largest gains in average growth of child quality at every average cost level. In order to see how the presence of Internal CCTs affects the effectiveness of each type of policy, we compare policy results for the benchmark model with endogenous (costly) Internal CCTs (i.e. the black lines) to those implied by the model where Internal CCTs are not available (i.e. the grey lines). Note that the baseline standard deviation of child quality is different for the model without Internal CCTs. For comparison, the vertical axis therefore plots the proportional change in the standard deviation relative to the respective baseline standard deviation for each of the two models. All experiments were done using $R = 10$ simulated data sets.

Table 1: Descriptive Statistics

	Mean	Std.	N
PSID-CDS: 1997, 2002, 2007			
Mother's age in 1997	36.91	5.38	247
Father's age in 1997	39.08	6.44	247
Mother's education	14.63	2.09	247
Father's education	14.19	2.31	247
Child's age in 1997	6.76	2.33	247
Letter Word raw score in 1997	23.37	15.49	247
Letter Word raw score in 2002	46.60	5.40	247
Letter Word raw score in 2007	50.44	3.42	109
PSID-Core: 1986 - 2010			
Mother's work hours per week	23.29	16.45	2466
Father's work hours per week	42.64	11.52	2407
Mother's hourly wage	20.13	13.32	1852
Father's hourly wage	28.79	18.38	2306
Non-labor income per week	88.77	160.29	2356

Notes: Parental work hours, wages and non-labor income statistics are averaged over all years where the child is between 0 and 16 years old, ranging from 1986 to 2010.
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Table 2: Parental Labor Supply by Child Age

Child Age	Fraction Working > 0 Hours				
	3	4-5	6-8	9-11	12-15
Mothers working	0.796	0.766	0.787	0.816	0.846
Fathers working	0.990	0.985	0.982	0.985	0.961
Both working	0.781	0.746	0.758	0.794	0.804
Child Age	Avg. Hours Working (> 0 Hours)				
	3	4-5	6-8	9-11	12-15
Mothers	27.77	28.87	28.97	29.99	31.77
Fathers	43.57	43.14	44.56	43.33	43.76

Notes: Upper half of the table shows labor force participation rates. Bottom half shows average labor hours conditional on working positive hours.
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Table 3: Time Allocation by Child Age (Average Hours per Week)

Child Age	3	4-5	6-8	9-11	12-15
Mother's Work Hours	22.10	22.12	22.81	24.47	26.89
Father's Work Hours	43.13	42.49	43.78	42.70	42.03
Mother's Active Time	26.12	20.41	14.32	11.18	6.88
Father's Active Time	13.68	7.92	5.28	4.95	4.18
Joint Parental Time	7.17	10.59	9.59	10.10	7.88
Child's Self-Investment Time	0.00	0.67	1.42	3.37	6.01
School Time	11.25	11.89	27.77	31.43	35.17

Notes: Parental work hours, wages and non-labor income statistics are averaged over all years where the child is between 0 and 16 years old, ranging from 1986 to 2010.
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Table 4: Preference Parameter Estimates

(a) Parental preferences	Mean		Std.	
	Estimate	SE	Estimate	SE
Mother's Leisure (α_1)	0.240	(0.00195)	0.034	(0.00152)
Father's Leisure (α_2)	0.301	(0.00143)	0.022	(0.00191)
Consumption (α_3)	0.298	(0.00261)	0.007	(0.00147)
Child Quality (α_4)	0.161	(0.00066)	0.052	(0.00133)

(b) Child preferences	Mean		Std.	
	Estimate	SE	Estimate	SE
Leisure (λ_1)	0.425	(0.00428)	0.221	(0.00320)
Consumption (λ_2)	0.165	(0.00757)	0.050	(0.00194)
Child Quality (λ_3)	0.410	(0.00338)	0.184	(0.00133)

(c) Correlations	Estimate	SE
Correlation of α_1 and α_2	0.295	(0.00383)
Correlation of α_1 and α_3	-0.688	(0.01255)
Correlation of α_1 and λ_1	0.467	(0.00174)
Correlation of α_2 and α_3	0.086	(0.02009)
Correlation of α_2 and λ_1	0.516	(0.00723)
Correlation of α_3 and λ_2	0.005	(0.01029)
Correlation of λ_1 and λ_2	0.688	(0.02187)
Correlation of α_1 and λ_2	0	-
Correlation of α_2 and λ_2	0	-
Correlation of α_3 and λ_1	0	-

(d) Other preferences	Estimate	SE
Altruism (φ)	0.333	-
Parents' Terminal Valuation (ξ_p)	5.198	(0.08400)
Child's Terminal Valuation (ξ_c)	0.810	(0.00983)
Mean/Std. of ICCT Costs (κ)	0.009	(0.00135)

Notes: SEs are standard errors computed using a cluster bootstrap sampling each household with replacement. Parameters without SE are assumed (not estimated) values.

Table 5: Technology Parameter Estimates

		Estimate	SE
Mother's Active Time (δ_1)	Intercept $\gamma_{1,0}$	-0.244	(0.03502)
	Slope $\gamma_{1,1}$	-0.252	(0.00683)
	Mother's Educ. $\gamma_{1,2}$	-0.001	(0.00007)
Father's Active Time (δ_2)	Intercept $\gamma_{2,0}$	-1.662	(0.04973)
	Slope $\gamma_{2,1}$	-0.239	(0.00846)
	Father's Educ. $\gamma_{2,2}$	0.042	(0.00276)
Joint Parental Time (δ_3)	Intercept $\gamma_{3,0}$	-1.259	(0.06380)
	Slope $\gamma_{3,1}$	-0.133	(0.00127)
	Mother's Educ. $\gamma_{3,2}$	0.020	(0.00219)
	Father's Educ. $\gamma_{3,3}$	0.018	(0.00100)
Child Expenditures (δ_4)	Intercept $\gamma_{4,0}$	-4.219	(0.17291)
	Slope $\gamma_{4,1}$	-0.053	(0.00118)
Child's Self-Investment Time (δ_5)	Intercept $\gamma_{5,0}$	-7.930	(0.13529)
	Slope $\gamma_{5,1}$	0.249	(0.00942)
Last Period's Child Quality (δ_6)	Intercept $\gamma_{6,0}$	-1.644	(0.01502)
	Slope $\gamma_{6,1}$	0.264	(0.00170)
Total Factor Productivity (R_t)	$\gamma_{7,0}$	0.47365	(0.00677)
	$\gamma_{7,1}$	1.01128	(0.00414)
	$\gamma_{7,2}$	1.44493	(0.13486)
	$\gamma_{7,3}$	8.24483	(0.10487)

Notes: Productivity parameters take the form $\delta_{i,t} = 0.01 + 0.99 \frac{\exp(\gamma_{i,0} + \gamma_{i,1}(t-1))}{\exp(\gamma_{i,0} + \gamma_{i,1}(t-1))}$, for all $i = 1, ..., 6$ and $t = 1, ..., 16$. Total Factor Productivity parameters take the form $R_t = \gamma_{7,0} + \frac{\gamma_{7,1} - \gamma_{7,0}}{1 + \exp(-\gamma_{7,2}(t - \gamma_{7,3}))}$. SEs are standard errors computed using a cluster bootstrap sampling each household with replacement.

Table 6: Wage and Income Parameter Estimates

(a) Mother's Log Wage Offer	Estimate	SE
Intercept ($\eta_{0,1}$)	0.838	(0.00714)
Mother's Age ($\eta_{1,1}$)	0.013	(0.00022)
Mother's Education ($\eta_{2,1}$)	0.098	(0.00073)
Standard Deviation of Shock (σ_{w_1})	0.445	(0.01300)
Correlation with Father's Wage Shock (ρ_{12})	0.602	(0.01705)
(b) Father's Log Wage Offer	**Estimate**	**SE**
Intercept ($\eta_{0,2}$)	1.364	(0.00446)
Father's Age ($\eta_{1,2}$)	0.013	(0.00029)
Father's Education ($\eta_{2,2}$)	0.097	(0.00030)
Standard Deviation of Shock (σ_{w_2})	0.542	(0.05838)
(c) Latent Non-Labor Income	**Estimate**	**SE**
Logit - Intercept (μ_1)	-4.905	(0.11379)
Logit - Mother's Age (μ_2)	0.000	(0.00009)
Logit - Mother's Age Squared (μ_3)	0.000	(0.00007)
Logit - Mother's Education (μ_4)	0.000	(0.00002)
Logit - Father's Age (μ_5)	0.115	(0.00337)
Logit - Father's Age Squared (μ_6)	-0.001	(0.00010)
Logit - Father's Education (μ_7)	0.243	(0.00673)
Conditional - Intercept (μ_8)	2.207	(0.15222)
Conditional - Mother's Age (μ_9)	0.000	(0.00004)
Conditional - Mother's Age Squared (μ_{10})	-0.000	(0.00004)
Conditional - Mother's Education (μ_{11})	0.003	(0.00849)
Conditional - Father's Age (μ_{12})	0.045	(0.00413)
Conditional - Father's Age Squared (μ_{13})	-0.000	(0.00005)
Conditional - Father's Education (μ_{14})	-0.001	(0.00073)
Standard Deviation of Shock (σ_I)	1.278	(0.03250)

Notes: SEs are standard errors computed using a cluster bootstrap sampling each household with replacement.

Table 7: Sample Fit of Time Allocations by Child Age

(a) Probability Work > 0 Hours

Child Age	Mother		Father	
	Data	Simulated	Data	Simulated
3-5	0.777	0.751	0.987	0.988
6-8	0.787	0.798	0.982	0.989
9-11	0.816	0.812	0.985	0.991
12-15	0.846	0.844	0.961	0.993

(b) Hours Worked if Work > 0 Hours (Avg.)

Child Age	Mother		Father	
	Data	Simulated	Data	Simulated
3-5	28.45	27.04	43.30	42.31
6-8	28.97	27.63	44.56	43.91
9-11	29.99	29.70	43.33	43.63
12-15	31.77	32.52	43.76	44.77

(c) Active Time with Child (Avg.)

Child Age	Mother		Father	
	Data	Simulated	Data	Simulated
3-5	21.01	20.85	8.53	8.95
6-8	14.32	15.12	5.28	6.58
9-11	11.18	12.09	4.95	5.53
12-15	6.88	6.99	4.18	3.56

(d) Joint Parental Time with Child (Avg.)

Child Age	Data	Simulated
3-5	10.23	10.34
6-8	9.59	10.42
9-11	10.10	11.22
12-15	7.88	8.92

(e) Child Self-Investment Time (Avg.)

Child Age	Data	Simulated
3-5	0.60	1.39
6-8	1.42	1.30
9-11	3.37	2.68
12-15	6.01	5.49

Notes: Data is actual data. Simulated is the model prediction at estimated parameters. **Source**: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Table 8: Sample Fit of Wages and Non-Labor Income

(a) Hourly Wages	Mother		Father	
	Data	Simulated	Data	Simulated
Average	20.13	21.82	28.79	30.77
Standard deviation	13.32	13.07	18.38	16.24

(b) Weekly Non-labor Income				
	Data	Simulated		
Average	89.41	87.69		
Standard deviation	161.77	161.46		
Fraction with $I > 0$	0.73	0.73		

Notes: Data is actual data. Simulated is the model prediction at estimated parameters. **Source**: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Table 9: Comparing Models with and without Internal CCTs: Unrestricted Cash Transfer at age 15

	Benchmark Costly ICCT		Special Case No ICCT	
	Baseline	Policy	Baseline	Policy
(a) Child Quality				
Level k_{t+1}	10.56	+0.32	10.02	+0.32
Growth rate g_t in %	12.27	+2.90	10.52	+3.37
(b) Investments				
Mother's Active Time τ_1	5.61	+1.33	5.73	+1.35
Father's Active Time τ_2	3.00	+2.10	3.04	+2.14
Joint Parental Time τ_{12}	7.83	+2.52	7.90	+2.55
Child Expenditures e	122.51	+3.13	121.40	+3.13
Child Self-Investment Time τ_c	6.59	-0.88	2.71	-1.01
(c) Labor Supply and Leisure				
Mother's Hours Work h_1	29.03	-7.87	28.50	-8.01
Mother's Leisure l_1	70.57	+2.81	70.91	+2.78
Father's Hours Work h_2	45.07	-5.28	44.58	-5.38
Father's Leisure l_2	55.37	+3.80	55.75	+3.80
Child's Leisure l_c	51.61	-0.50	55.27	-0.54
(d) Consumption and Welfare				
Parents' Consumption c	1695.07	+3.15	1707.52	+3.16
Child Consumption x	509.68	+3.10	471.55	+3.13
Parents' Utility $\widetilde{u}_{p,t}$	4.40	+0.49	4.40	+0.49
Parents' Utility $\widetilde{u}_{p,t+1}$	4.45	+0.02	4.45	+0.02
Child's Utility $u_{c,t}$	3.67	+0.12	3.67	+0.12
Child's Utility $u_{c,t+1}$	3.75	+0.04	3.74	+0.04
(e) Internal CCT Use				
Prop. Using Internal CCT	0.72	+0.00	-	-
Reward Elasticity r	0.17	+0.05	-	-

Notes: This table summarizes the effects of an unrestricted cash transfer of 170 dollars per week to all households at age $t = 15$. Columns 1-2 present results for the (benchmark) model where households choose whether to use an ICCT at a given cost. Columns 3-4 present results for the special case where ICCTs are infinitely costly. For each model, the "Baseline' column shows averages taken across all households and simulation rounds, at child age $t = 15$, in the absence of policy. The "Policy' column contains average percentage deviations from the respective baseline column. All experiments were done using $R = 10$ simulated data sets.

Table 10: Comparing Models with and without Internal CCTs: Restricted Cash Transfer at age 15

	Benchmark Costly ICCT		Special Case No ICCT	
	Baseline	Policy	Baseline	Policy
(a) Child Quality				
Level k_{t+1}	10.56	+1.13	10.02	+0.92
Growth rate g_t in %	12.27	+12.55	10.52	+11.32
(b) Investments				
Mother's Active Time τ_1	5.61	+2.46	5.73	+0.82
Father's Active Time τ_2	3.00	+2.51	3.04	+1.27
Joint Parental Time τ_{12}	7.83	+2.28	7.90	+1.51
Child Expenditures e	122.51	+50.63	121.40	+51.67
Child Self-Investment Time τ_c	6.59	+4.21	2.71	-0.68
(c) Labor Supply and Leisure				
Mother's Hours Work h_1	29.03	-5.60	28.50	-4.34
Mother's Leisure l_1	70.57	+1.86	70.91	+1.51
Father's Hours Work h_2	45.07	-3.69	44.58	-2.91
Father's Leisure l_2	55.37	+2.54	55.75	+2.02
Child's Leisure l_c	51.61	-1.21	55.27	-0.33
(d) Consumption and Welfare				
Parents' Consumption c	1695.07	+2.47	1707.52	+2.02
Child Consumption x	509.68	-2.44	471.55	+2.12
Parents' Utility $\widetilde{u}_{p,t}$	4.40	+0.23	4.40	+0.27
Parents' Utility $\widetilde{u}_{p,t+1}$	4.45	+0.07	4.45	+0.05
Child's Utility $u_{c,t}$	3.67	-0.29	3.67	+0.07
Child's Utility $u_{c,t+1}$	3.75	+0.13	3.74	+0.11
(e) Internal CCT Use				
Prop. Using Internal CCT	0.72	+0.84	-	-
Reward Elasticity r	0.17	+11.68	-	-

Notes: This table summarizes the effects of a restricted cash transfer of 170 dollars per week to all households at age $t = 15$. Columns 1-2 present results for the (benchmark) model where households choose whether to use an ICCT at a given cost. Columns 3-4 present results for the special case where ICCTs are infinitely costly. For each model, the "Baseline' column shows averages taken across all households and simulation rounds, at child age $t = 15$, in the absence of policy. The "Policy' column contains average percentage deviations from the respective baseline column. All experiments were done using $R = 10$ simulated data sets.

Table 11: Comparing Models with and without Internal CCTs: External CCT at age 15

	Benchmark Costly ICCT		Special Case No ICCT	
	Baseline	Policy	Baseline	Policy
(a) Child Quality				
Level k_{t+1}	10.56	+3.55	10.02	+2.92
Growth rate g_t in %	12.27	+42.26	10.52	+38.74
(b) Investments				
Mother's Active Time τ_1	5.61	+20.88	5.73	+21.70
Father's Active Time τ_2	3.00	+21.83	3.04	+23.20
Joint Parental Time τ_{12}	7.83	+23.23	7.90	+24.91
Child Expenditures e	122.51	+25.97	121.40	+32.19
Child Self-Investment Time τ_c	6.59	+30.62	2.71	-4.84
(c) Labor Supply and Leisure				
Mother's Hours Work h_1	29.03	-2.72	28.50	-2.67
Mother's Leisure l_1	70.57	-3.32	70.91	-3.83
Father's Hours Work h_2	45.07	-1.63	44.58	-1.50
Father's Leisure l_2	55.37	-3.26	55.75	-3.90
Child's Leisure l_c	51.61	-11.17	55.27	-7.32
(d) Consumption and Welfare				
Parents' Consumption c	1695.07	-3.01	1707.52	-3.80
Child Consumption x	509.68	-5.10	471.55	-3.64
Parents' Utility $\widetilde{u}_{p,t}$	4.40	-1.01	4.40	-0.81
Parents' Utility $\widetilde{u}_{p,t+1}$	4.45	+0.60	4.45	+0.64
Child's Utility $u_{c,t}$	3.67	-1.62	3.67	-1.09
Child's Utility $u_{c,t+1}$	3.75	+0.63	3.74	+0.46
(e) Internal CCT Use				
Prop. Using Internal CCT	0.72	+21.70	-	-
Reward Elasticity r	0.17	+103.59	-	-

Notes: This table summarizes the effects of an external conditional cash transfer which offers a reward of 250 dollars per week at age 16, conditional on the growth rate in child quality between age 15 and 16 being at least 20 percent for the benchmark model, and 17.5 percent for the no-ICCT model. Columns 1-2 present results for the (benchmark) model where households choose whether to use an ICCT at a given cost. Columns 3-4 present results for the special case where ICCTs are infinitely costly. For each model, the "Baseline' column shows averages taken across all households and simulation rounds, at child age $t = 15$, in the absence of policy. The "Policy' column contains average percentage deviations from the respective baseline column. All experiments were done using $R = 10$ simulated data sets.

Table 12: External CCTs: Comparing Optimal Policies at Age 15

	(0)	(1)	(2)	(3)	(4)
	Baseline	**ECCT Policy**			
Avg. Cost per Household	-	36.0	72.2	115.4	146.3
(a) Policy Parameters					
Target growth rate ρ^{CCT}	-	0.185	0.195	0.200	0.200
Reward ϕ^{CCT}	-	75	135	195	225
(b) Child Capital					
Growth Rate g_t, Average	0.123	0.144	0.155	0.165	0.171
Growth Rate g_t, St.dev.	0.041	0.050	0.052	0.052	0.049
Level k_{t+1}, Average	10.562	10.734	10.814	10.878	10.915
Level k_{t+1}, St.dev.	5.493	5.503	5.480	5.454	5.437
(c) Take-up Rate					
Fraction receiving transfer: $Pr(k_{t+1}^{**} \geq \overline{k}_{t+1})$	-	0.480	0.535	0.592	0.650
Fraction affected by policy: $Pr(k_{t+1}^{**} \geq \overline{k}_{t+1} > k_{t+1}^{*})$	-	0.421	0.500	0.564	0.623
(d) Fraction using ICCT at time t					
All Households	0.720	0.781	0.814	0.845	0.864
Baseline, given $k_{t+1}^{**} \geq \overline{k}_{t+1} > k_{t+1}^{*}$	-	0.816	0.782	0.757	0.745
Policy, given $k_{t+1}^{**} \geq \overline{k}_{t+1} > k_{t+1}^{*}$	-	0.967	0.978	0.980	0.979
(e) ICCT reward elasticity (r_t) at time t					
All Households	0.172	0.241	0.282	0.317	0.339
Given $r_t > 0$	0.239	0.309	0.347	0.375	0.392

Notes: All results are obtained for a growth-based External CCT policy administered when each child is $t = 15$ years old, where the parents receive a transfer of ϕ^{CCT} if the child's optimal level of latent human capital at age $t+1$ under the ECCT policy (denoted by k_{t+1}^{**}) is at least $(1+\rho^{CCT})$ times the baseline level at age t, k_t. The optimal $t+1$ level of child quality in the baseline (in the absence of an ECCT policy) is denoted by k_{t+1}^{*}.

Appendices

A   Model Solution

A.1   Final period: $t = M$

For both the parents and the child, the decision rules are solved using backward recursion beginning from the end of the development process, time M.

Child's problem.   In period M, the child makes her choice of $\tau_{c,M}$ given the time and budget allocation of the parents, $a_{p,M}$, and, if the parents are using an Internal CCT, the contract specified by $\{r_M, b_M\}$. Then the child's period M problem is

$$V_{c,M}(\Gamma_M | a_{p,M}, r_M, b_M) = \max_{\tau_{c,M} | a_{p,M}, r_M, b_M} \quad \lambda_1 \ln(\tilde{T}_M - \tau_{p,M} - \tau_{c,M}) + \lambda_2 \ln x_M + \lambda_3 \ln k_M$$

$$+ \beta_{c,M} \psi_{c,M+1} E(\ln k_{M+1} | \tau_{c,M}, a_{p,M}, r_M, b_M)$$

where $\tilde{T}_M$ is the child's time endowment after subtracting exogenous school time s_M. We can substitute out these two components:

$$
\begin{aligned}
E(\ln k_{M+1} | \tau_{c,M}, a_{p,M}, r_M, b_M) &= \ln R_M + \delta_{1,M} \ln \tau_{1,M} + \delta_{2,M} \ln \tau_{2,M} + \delta_{3,M} \ln \tau_{12,M} \\
&\quad + \delta_{4,M} \ln e_M + \delta_{5,M} \ln \tau_{c,M} + \delta_{6,M} \ln k_M, \\
\ln x_M &= b_M + r_M \ln \tau_{c,M}
\end{aligned}
$$

Since we assume that all parameters, including Total Factor Productivity R_M, are known at the time of the period M decisions, there is no uncertainty present in the production technology, allowing us to drop the expectation operator. The optimal decision of the child is given by

$$
\begin{aligned}
\tau_{c,M}^*(\tau_{p,M}, r_M) &= \frac{\lambda_2 r_M + \Delta_{c,M}}{\lambda_1 + \lambda_2 r_M + \Delta_{c,M}}(\tilde{T}_M - \tau_{p,M}) \qquad \text{(A-1)} \\
&= \gamma_M(r_M)(\tilde{T}_M - \tau_{p,M})
\end{aligned}
$$

where $\Delta_{c,M} \equiv \beta_{c,M} \psi_{c,M+1} \delta_{5,M}$. Given the properties of the production, utility and reward functions, the choice of time in investment is independent of all of the parents' decisions with the exception of (1) the total time they spend interacting with the children, $\tau_{p,M}$, the effect of which is to reduce the child's effective time endowment, and (2) the child's "wage" rate r_M, which corresponds to the elasticity of child consumption with respect to child study time. The fact that b_M drops out will prove useful in deriving some of the results below. Note that when $r_M = 0$, this solution simplifies to the special case in which the parents make a fixed transfer of x_M to the child that is not tied to the child's investment time. Clearly, the solution to the child's problem is increasing in r_M and the child can be induced to spend virtually all of its time in investment as r_M

becomes arbitrarily large. However, this could never be optimal since (1) the child has an incentive compatibility (IC) constraint that must be satisfied whenever the parents use an ICCT scheme, and (2) even in the absence of an IC constraint, the parents would not want their child to have zero leisure as long as they are altruistic ($\varphi > 0$).

Parents' problem. Given the child's reaction function $\tau_{c,M}^*(\tau_{p,M}, r_M)$, the parents solve the following problem:

$$V_{p,M}(\Gamma_M) = \max_{a_{p,M}, r_M, b_M} \tilde{u}_p(l_{1,M}, l_{2,M}, c_M, k_M, l_{c,M}, x_M) + \beta_p \psi_{p,M+1} \ln(k_{M+1})$$

$$+\mu_M \Big(\lambda_1 \ln(l_{c,M}) + \lambda_2 \ln(x_M) + \lambda_3 \ln(k_M) + \beta_{c,M} \psi_{c,M+1} \ln(k_{M+1}) - V_{c,M}(\Gamma_M | a_{p,M}^0) \Big) \text{(A-2)}$$

where $\mu_M \geq 0$ is the Lagrange multiplier on the child's IC constraint, and $V_{c,M}(\Gamma_M | a_{p,M}^0)$ denotes the child's outside option, i.e. the indirect value function evaluated at the parents' choices in the absence of an ICCT. We can substitute out (1) c_M for the period M budget constraint, (2) $l_{1,M}$, $l_{2,M}$ and $l_{c,M}$ for the individual time constraints, (3) $\ln k_{M+1}$ for the production technology, and (4) $\tau_{c,M}$ for the child's optimal reaction function derived in the previous paragraph. In order to simplify the first order conditions with respect to the remaining choices $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, x_t, r_t, b_t\}$, note that the parents jointly choose the triple $\{x_M, r_M, b_M\}$ subject to the reward function. Rearranging this equation yields

$$\begin{aligned} b_M &= \ln(x_M) - r_M \ln(\tau_{c,M}^*(\tau_{p,M}, r_M)) \\ &= \ln(x_M) - r_M \ln(\gamma_M(r_M)) - r_M \ln(\tilde{T}_M - \tau_{p,M}) \end{aligned}$$

Conditional on $\{x_M, r_M, \tau_{1,M}, \tau_{2,M}, \tau_{12,M}\}$, this will pin down the optimal choice of b_M. Taking first order conditions with respect to $\{e_M, x_M\}$ and using the budget constraint yields the following solutions for the expenditures (conditional on labor supply choices and the multiplier on the child's IC constraint):

$$c_M^\star = \frac{\tilde{\alpha}_3}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \qquad \text{(A-3)}$$

$$e_M^\star = \frac{\beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M \beta_{c,M} \psi_{c,M+1} \delta_{4,M}}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \qquad \text{(A-4)}$$

$$x_M^\star = \frac{\tilde{\alpha}_6 + \mu_M \lambda_2}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,M+1} \delta_{4,M} + \mu_M(\lambda_2 + \beta_{c,M} \psi_{c,M+1} \delta_{4,M})} Y_M \qquad \text{(A-5)}$$

where $Y_M = w_{1,M} h_{1,M} + w_{2,M} h_{2,M} + I_M$ and $\mu_M \geq 0$. It is easy to see that the fraction of income spent on the parents' private consumption, c_M, is strictly decreasing in μ_M. Conversely, the fraction spent on the child's consumption, x_M, is strictly increasing in μ_M under a weak condition on the primitives:

$$\frac{\partial x_M^* / Y_M}{\partial \mu_M} > 0 \iff \beta_p \psi_{p,M+1} - \varphi \beta_{c,M} \psi_{c,M+1} > \frac{-\tilde{\alpha}_3}{\delta_{4,M}}$$

Since the parents in our model are not perfectly altruistic ($\varphi < 1$), are more patient than children ($\beta_p > \beta_{c,t}$) and usually (but not always in our model) care more about future child quality than the child ($\psi_{p,t+1} > \psi_{c,t+1}$), the left-hand side of this inequality should typically be positive, making this condition satisfied.[30]

After substituting out the (conditional) optimal choices $\{c_M^\star, e_M^\star, x_M^\star\}$ and the child's reaction function in the value function $V_{p,M}$, and after dropping some constant terms, we are left with a "residual" maximization problem which can be decomposed into two separate maximization problems, conditional on the Lagrange multiplier μ_M:

$$W_{p,M} = \max \nu_{1,M}(h_{1,M}, h_{2,M}, \tau_{1,M}, \tau_{2,M}, \tau_{12,M}; \mu_M) + \nu_{2,M}(r_M; \mu_M)$$

where the second component is given by

$$\nu_{2,M}(r_M; \mu_M) \equiv \left(\tilde{\alpha}_5 + \mu_M \lambda_1\right) \ln(1 - \gamma_M(r_M)) + \left(\Delta_{p,M} + \mu_M \Delta_{c,M}\right) \ln(\gamma_M(r_M)). \quad \text{(A-6)}$$

where $\Delta_{c,M} \equiv \beta_{c,M} \psi_{c,M+1} \delta_{5,M}$ and $\Delta_{p,M} \equiv \beta_p \psi_{p,M+1} \delta_{5,M}$.

Unconstrained optimum. Denote the unconstrained optimal time, budget and ICCT choices by the vector $\{a_{p,M}^{unc}, r_M^{unc}, b_M^{unc}\}$. If the multiplier μ_M equals 0, the parents' optimization problem becomes truly separable, since the two components, $\nu_{1,M}(a_{p,M})$ and $\nu_{2,M}(r_M)$, no longer have any common components. Therefore, none of the unconstrained optimal time and budget choices (summarized by $a_M^{unc} = \{h_{1,M}, h_{2,M}, \tau_{1,M}, \tau_{2,M}, \tau_{12,M}, c_M, e_M, x_M\}$) will depend on the parents' choice of r_M. Indeed, under our functional form assumptions, it must be the case that

$$a_{p,M}^{unc} = a_{p,M}^0 = \hat{a}_{p,M}$$

where $a_{p,M}^0$ denotes the optimal time and budget allocation in the no-ICCT Stackelberg equilibrium (where $r_M = 0$), and where $\hat{a}_{p,M}$ denotes the optimal time and budget allocation in the dictatorial model, where parents (hypothetically) choose the child's study time directly. Conditional on parental labor supply, we can find the optimal budget allocation by plugging in $\mu_M = 0$ into Equations (A-3)-(A-5). The remaining optimal choices can be found by maximizing the sub-function $\nu_{1,M}(a_{p,M})$ analytically.

It is convenient that the second sub-function, $\nu_{2,M}(r_M)$ shown in (A-6) only depends on the reward elasticity, r_M, due to the functional form of the ICCT reward function. The parents' ability to implement an ICCT (assuming a slack incentive constraint for the child) allows them to perfectly align the child's incentives with their own altruistic preferences by implementing the following ICCT contract:

$$
\begin{aligned}
r_M^{unc} &= \frac{\Delta_{p,M} - \varphi \Delta_{c,M}}{\lambda_2 \varphi}, \\
b_M^{unc} &= \ln(x_M^{unc}) - r_M^{unc} \ln(\tau_{c,M}(\tau_{p,M}^{unc}, r_M^{unc}))
\end{aligned}
$$

---

[30]Given our model estimates and random simulation draws, this condition always holds in our analysis.

where r_M^{unc} follows from differentiating $\nu_{2,M}(r_M)$, and b_M^{unc} follows from the ICCT reward function. Note that given this contract, the child's optimal response coincides with what the parents would choose themselves if they were the household dictator (i.e. $\tau_{c,M}^{unc} = \hat{\tau}_{c,M}$).

By complementary slackness, we should finally verify whether the child's IC constraint in the parents' problem (A-2) is indeed slack, given the parents' unconstrained choice vector. As stated in Proposition 1, we know that whenever the parents are using an ICCT with $r_t \neq 0$, the child's incentive constraint must be binding in equilibrium, thereby effectively ruling out the unconstrained equilibrium we have just derived. For completeness, we also provide a more formal proof of Proposition 1 for a general period $t \in \{1, ..., M\}$.

First, we define the child's outside option, $V_{c,t}(\Gamma_t | a_{p,t}^0)$, as her indirect value when the parents are not using an Internal CCT scheme, i.e. when $a_{p,t} = a_{p,t}^0$ and $r_t = 0$. We prove by contradiction, i.e. by assuming that the child's IC constraint will be slack in the Stackelberg equilibrium with ICCT and $\mu_t = 0$. From before, we know that the optimal reaction function of the child is given by Equation (A-1). After plugging in this reaction function and the parents' optimal choices, we can write the child's indirect value function as follows:

$$V_{c,t}(\Gamma_t | a_{p,t}^{unc}, r_t^{unc}, b_t^{unc}) = \lambda_1 \ln(l_{c,t}^{unc}) + \lambda_2 \ln(x_t^{unc}) + \lambda_3(k_t) + \beta_{c,t} \psi_{c,t+1} \ln(k_{t+1}(a_{p,t}^{unc}, r_t^{unc}))$$

where the child's leisure $l_{c,t}^{unc} = (1 - \gamma_t(r_t^{unc}))(\tilde{T}_t - \tau_{p,t}^{unc})$. Importantly, in this indirect utility function, the child's consumption level no longer depends on the child's study time. Indeed, even though the parents are offering the child an incentive scheme (or reward function) given by $x_t(\tau_{c,t}; r_t, b_t)$, the child realizes that irrespective of how much she studies, the parents can always implement their first-best value of child consumption (given by $x_t^{unc} = x_t^0 = \hat{x}_t$) by simply readjusting (or reneging on) the value of b_t^{unc} after the child has chosen how much time to devote to studying. This lack of commitment on behalf of the parents would make the child unwilling to participate in the incentive scheme and deviate back to the no-ICCT Stackelberg equilibrium. Indeed, given our previous result that $a_{p,t}^{unc} = a_{p,t}^0$, we can simplify child's IC constraint as follows:

$$V_{c,t}(\Gamma_t | a_{p,t}^{unc}, r_t^{unc}, b_t^{unc}) \geq V_{c,t}(\Gamma_t | a_{p,t}^0)$$
$$\iff \lambda_1 \ln(l_{c,t}^{unc}) + \beta_{c,t} \psi_{c,t+1} \delta_{5,t} \ln(\tau_{c,t}^{unc}) \geq \lambda_1 \ln(l_{c,t}^0) + \beta_{c,t} \psi_{c,t+1} \delta_{5,t} \ln(\tau_{c,t}^0)$$
$$\iff \lambda_1 \ln(1 - \gamma_t(r_t^{unc})) + \Delta_{c,t} \ln(\gamma_t(r_t^{unc})) \geq \lambda_1 \ln(1 - \gamma_t(0)) + \Delta_{c,t} \ln(\gamma_t(0))$$

First, note that this inequality is binding if and only if $r_t^{unc} = 0$, which is only optimal in the knife-edge case where $\Delta_{p,t} = \varphi\Delta_{c,t}$.[31] Second, while the right-hand side of the

---

[31]Under the relatively weak assumption on the primitives that $\Delta_{p,t} > \varphi\Delta_{c,t}$, parents prefer to positively incentivize their children, i.e. to set $r_t > 0$. Although our model does not rule out that some parents may prefer to implement negative incentive schemes ($r_t < 0$), it is never the case given our parameter estimates and random simulation draws. Moreover, even in those cases where $r_t^{unc} < 0$, the child would still prefer to deviate back to the no-ICCT equilibrium by studying *more* than what the parents prefer.

above inequality does not depend on r_t, we can show that the left-hand side is strictly *decreasing* in r_t, by taking the partial derivative:

$$
\frac{\partial V_{c,t}(\Gamma_t | a_{p,t}^{unc}, r_t, b_t)}{\partial r_t} = \frac{\partial \gamma_t(r_t)}{\partial r_t} \left[\frac{\Delta_{c,t}}{\gamma_t(r_t)} - \frac{\lambda_1}{1 - \gamma_t(r_t)} \right]
$$

$$
= \frac{\partial \gamma_t(r_t)}{\partial r_t} \left[\frac{-\lambda_2 r_t}{\gamma_t(r_t)} \right]
$$

where we implicitly use the results that the parents' time and budget allocation $(a_{p,t}^{unc})$ does not vary with r_t, and that the child's action does not depend on b_t. Since $\gamma_t(r_t)$ is strictly increasing in r_t, this partial derivative is zero only when $r_t = 0$, strictly negative whenever $r_t > 0$, and strictly positive whevener $r_t < 0$. This implies that compared to the no-ICCT equilibrium where $r_t = 0$, the child's value function (evaluated at $a_{p,t}^{unc}$) is globally maximized at $r_t = 0$, and strictly decreases whenever $r_t \neq 0$. This means the child's IC constraint is violated whenever $\mu_t = 0$, which rules out the unconstrained ICCT equilibrium. Therefore, the IC constraint is always binding in the ICCT equilibrium.

Constrained optimum. Although we cannot solve for r_M in closed form, we can find the optimal value conditional on μ_M, by differentiating $V_{p,M}$ (see Equation (A-2)) with respect to r_M. Assuming for now that there will be an interior solution (i.e. $r_M \neq 0$), we obtain:

$$
\frac{dV_{p,M}(\Gamma_M | a_{p,M}, r_M, b_M; \mu_M)}{dr_M} = \frac{\partial V_{p,M}}{\partial r_M} + \frac{\partial V_{p,M}}{\partial \mu_M} \frac{\partial \mu_M}{\partial r_M} = 0 \qquad (A\text{-}7)
$$

where we have imposed that $\frac{\partial a_{p,M}}{\partial r_M} = \frac{\partial b_M}{\partial r_M} = 0$ due to the optimality principle. We know that $\frac{\partial V_{p,M}}{\partial \mu_M} \leq 0$, since the presence of the child's incentive constraint must decrease the parents' value relative to the unconstrained equilibrium, which coincides with the parents' first-best outcome. Moreover, from Proposition 1, we know that the incentive constraint becomes binding whenever $r_M \neq 0$. Since $\mu_M = 0$ only if $r_M = 0$, this implies that the partial derivative $\frac{\partial \mu_M}{\partial r_M}$ is positive when $r_M^{unc} > 0$, and negative when $r_M^{unc} < 0$. Thus, the second component in (A-7) must be negative whenever $r_M^* > 0$, and positive whenever $r_M^* < 0$. By optimality, the first component must have the opposite sign as the second component. Given our previous discussion of the parents' constrained problem and the expression given in Equation (A-6), we can derive this first component as follows:

$$
\frac{\partial V_{p,M}}{\partial r_M} = \frac{\partial \nu_{2,M}}{\partial r_M} = \frac{d\gamma_M(r_M)}{dr_M} \left[\frac{\Delta_{p,M} + \mu_M \Delta_{c,M}}{\gamma_M(r_M)} - \frac{\tilde{\alpha}_5 + \mu_M \lambda_1}{1 - \gamma_M(r_M)} \right]
$$

$$
= \frac{d\gamma_M(r_M)}{dr_M} \left[\frac{\Delta_{p,M} - \varphi \Delta_{c,M} - \lambda_2 r_M(\varphi + \mu_M)}{\gamma_M(r_M)} \right]
$$

where we used the fact that $\tilde{\alpha}_5 = \varphi\lambda_1$. Since $\gamma_M(r_M)$ is strictly increasing in r_M, we can derive a bound on the constrained optimal value r_M^*:

$$\frac{dV_{p,M}}{dr_M} = 0 \iff |r_M^*| \leq \frac{|\Delta_{p,M} - \varphi\Delta_{c,M}|}{\lambda_2(\varphi + \mu_M)} < |r_M^{unc}| = \frac{|\Delta_{p,M} - \varphi\Delta_{c,M}|}{\lambda_2\varphi}$$

where φ is the parents' altruism parameter. Note that as μ_M approaches 0, r_M^* converges to r_M^{unc}. By plugging this bound into the child's reaction function, the corresponding bounds on the child's optimal fraction of study time are:

$$r_M^* > 0 \iff \frac{\Delta_{c,M}}{\lambda_1 + \Delta_{c,M}} < \gamma_M(r_M^*) \leq \frac{\Delta_{p,M} + \mu_M\Delta_{c,M}}{\tilde{\alpha}_5 + \Delta_{p,M} + \mu_M(\lambda_1 + \Delta_{c,M}))} < \frac{\Delta_{p,M}}{\tilde{\alpha}_5 + \Delta_{p,M}}$$

where all inequality signs reverse for the (rare) cases where $r_M^* < 0$, i.e. when $\Delta_{p,M} < \varphi\Delta_{c,M}$. Without an explicit expression for μ_M, we cannot characterize the constrained ICCT optimum any further.

Conditions for Costly ICCT use. Since implementing an Internal CCT is, in our most general model, costly for the parents, it may be optimal to not use one, by setting $r_M = 0$. The parents will choose to use an ICCT when the welfare gain from the *constrained* equilibrium exceeds the utility cost ω_M, i.e. under the following necessary and sufficient condition:

$$r_M^* \neq 0 \iff V_{p,M}(\Gamma_M|a_{p,M}^*, r_M^*, b_M^*) - V_{p,M}(\Gamma_M|a_{p,M}^0) \geq \omega_M$$

where $a_{p,M}^*$ and $a_{p,M}^0$ denote the parents' optimal time and budget allocations in the constrained ICCT equilibrium and the no-ICCT equilibrium, respectively. Since the child's incentive constraint in the ICCT equilibrium is always binding, the optimal parental choices will change whenever $r_M \neq 0$ (i.e. $a_{p,M}^* \neq a_{p,M}^0$), preventing us from simplifying this expression any further. In the empirical implementation, we use a numerical solver to evaluate this necessary and sufficient condition for every household at every child age.

We have previously argued that the *unconstrained* parents' problem is separable into two parts, where only the second component, $\nu_2(r_M, \mu_M)$ (see (A-6)) depends on the parents' ICCT parameter. This insight allows us to derive the following necessary (but not sufficient) condition for the parents' choice whether to use an ICCT:

$$r_M^* \neq 0 \implies \nu_{2,M}(r_M = r_M^{unc}, \mu_M = 0) - \nu_{2,M}(r_M = 0, \mu_M = 0) \geq \omega_M$$
$$\iff \Delta_{p,M}\ln\left(\frac{\gamma_M(r_M^{unc})}{\gamma_M(0)}\right) \geq \tilde{\alpha}_5\ln\left(\frac{1 - \gamma_M(0)}{1 - \gamma_M(r_M^{unc})}\right) + \omega_M$$

where the closed form for r_M^{unc} is known, and where $\gamma_M(0) = \frac{\Delta_{c,M}}{\lambda_1 + \Delta_{c,M}}$ is the child's optimal fraction of study time in the no-ICCT equilibrium, which is strictly smaller than the fraction of study time in the unconstrained ICCT equilibrium, $\gamma_M(r_M^{unc}) =$

$\frac{\Delta_{p,M}}{\tilde{\alpha}_5 + \Delta_{p,M}}$. Consider the most common case where the parents would choose $r_M^{unc} > 0$. Then, the left-hand side of the second line can be interpreted as the parents' benefit of implementing an internal CCT which, by raising the child's study time, will increase child capital in the next period. The right-hand side can be interpreted as the parents' utility cost of implementing the ICCT, comprising both the direct cost, ω_M, as well as a utility loss through the reduction in the child's leisure time, which is weighted by the parents' value of child leisure, $\tilde{\alpha}_5 = \varphi\lambda_1$. Conversely, in the case where $r_M^{unc} < 0$, the parents would like to reduce the child's study time, which now has a utility benefit in terms of leisure, and a utility cost in terms of lost capital. Intuitively, the inequality will be satisfied if the parents' optimal reward elasticity is sufficiently different from 0, either positively or negatively. However, since this necessary condition does not include the additional utility loss the parents must incur due to the child requiring some additional compensation in the constrained equilibrium, it is not sufficient.

Optimal choices. Given the functional form assumptions and the presence of the child's incentive constraint, we cannot find closed form solutions for any of the parental choices $\{h_{1,M}, h_{2,M}, \tau_{1,N}, \tau_{2,M}, \tau_{12,M}, c_M, x_M, e_M, r_M, b_M\}$. In the computational exercise, we will use a numerical solver to find the optimal choice vector, taking into account the possible corner solutions for labor supply. Enforcing the child's incentive constraint involves first solving the unconstrained parents' problem, which (1) allows us to verify that, in accordance with Proposition 1, the child's IC is violated whenever $r_M^{unc} \neq 0$, and (2) provides us with a good initial guess before numerically solving the harder constrained problem where the incentive constraint is imposed at equality. Appendix C contains more details on this estimation procedure.

A.2 Remaining periods: $t = 1, ..., M-1$

The solution has exactly the same characteristics in the general period t case. The only adjustments to the solution occur with respect to the variables $\psi_{j,t}$, $j = c, p$, which measure the future impacts of improvements in child quality in period t and the remaining periods in the development process. The time-varying characteristics that appear in the solution include the production function parameters, the realizations of wages and non-labor income in period t, and the discount factor of the child, which is monotonically increasing in t. Thus the t-period solution is as follows:

No-ICCT Stackelberg Equilibrium. First, we solve the household problem assuming the parents are not using an ICCT, such that $r_t = 0$. We denote the total vector of optimal parental choices in the no-ICCT equilibrium as $a_{p,t}^0$.

1. Condition on a choice vector of $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}\}$, including potential corners for labor supply. Given these values, the household income in period t is $Y_t = w_{1,t}h_{1,t} + w_{2,t}h_{2,t} + I_t$. Total parental time is defined as $\tau_{p,t} = \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$.

2. The optimal expenditures in the no-ICCT Stackelberg equilibrium (conditional on labor supply) are given by

$$c_t^0 = \frac{\tilde{\alpha}_3}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$

$$e_t^0 = \frac{\beta_p \psi_{p,t+1} \delta_{4,t}}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$

$$x_t^0 = \frac{\tilde{\alpha}_6}{\tilde{\alpha}_3 + \tilde{\alpha}_6 + \beta_p \psi_{p,t+1} \delta_{4,t}} Y_t$$

where

$$
\begin{aligned}
\psi_{p,M+1} &\equiv \xi_p \alpha_4, \\
\psi_{p,t} &\equiv \tilde{\alpha}_4 + \beta_p \delta_{6,t} \psi_{p,t+1}, \ \ t = 1, ..., M.
\end{aligned}
$$

The optimal study time of the child in the absence of an ICCT is given by:

$$\tau_{c,t}^0(\tau_{p,t}, r_t = 0) = \frac{\Delta_{c,t}}{\lambda_1 + \Delta_{c,t}}(\tilde{T}_t - \tau_{p,t})$$

where

$$
\begin{aligned}
\Delta_{c,t} &\equiv \beta_{c,t} \psi_{c,t+1} \delta_{5,t}, \ \ t = 1, ..., M, \\
\psi_{c,M+1} &\equiv \xi_c \lambda_3, \\
\psi_{c,t} &\equiv \lambda_3 + \beta_{c,t} \delta_{6,t} \psi_{c,t+1}, \ \ t = 1, ..., M.
\end{aligned}
$$

3. By using the time constraints and the production technology function, we find the leisure of each individual $(l_{1,t}^0, l_{2,t}^0, l_{c,t}^0)$ and future child capital, k_{t+1}^0. This allows us to define the parental value function:

$$V_{p,t}(\Gamma_t, a_{p,t}^0) = \tilde{u}_p(l_{1,t}^0, l_{2,t}^0, c_t^0, k_t, l_{c,t}^0, x_t^0) + \beta_p \psi_{p,t+1} \ln(k_{t+1}^0)$$

We use a numerical solver to maximize this function with respect to the remaining choices for which we cannot find closed form solutions: $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}\}$.

4. Finally, we evaluate the child's value function at the no-ICCT Stackelberg equilibrium to define the child's outside option:

$$V_{c,t}(\Gamma_t, a_{p,t}^0) = \lambda_1 \ln(l_{c,t}^0) + \lambda_2 \ln(x_t^0) + \lambda_3 \ln(k_t) + \beta_{c,t} \psi_{c,t+1} \ln(k_{t+1}^0)$$

Constrained ICCT Equilibrium. Now, we solve the household's problem if the parents are using an ICCT, summarized by the reward function $x_t(\tau_{c,t}; r_t, b_t)$. If the parents choose a strictly positive reward elasticiticy $(r_t > 0)$, we know by Proposition 1 that the child's incentive compatibility constraint must be binding. Although some parents in our model might theoretically prefer to set a negative reward elasticity (see

above), this is never the case for our estimates and random simulation draws. Therefore, we can abstract from those cases in the empirical implementation. We denote the vector of optimal parental time, budget and ICCT choices in this equilibrium by $\{a_{p,t}^*, r_t^*, b_t^*\}$. In the constrained optimum, we no longer have closed form solutions for any of the parental choices. To simplify the numerical solution, we first find the unconstrained parents' optimum, which is almost identical to the no-ICCT equilibrium, except now (1) the parents choose $r_t^{unc} = \frac{\Delta_{p,t} - \varphi \Delta_{c,t}}{\lambda_2 \varphi} > 0$, and (2) consequently, the child studies more. Using this as an initial guess, we then solve the constrained problem as follows:

1. Condition on a choice vector of $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, r_t\}$, including potential corners for labor supply, and restricting r_t to be strictly positive. Given these values, the household income in period t is $Y_t = w_{1,t} h_{1,t} + w_{2,t} h_{2,t} + I_t$. Total parental time is defined as $\tau_{p,t} = \tau_{1,t} + \tau_{2,t} + \tau_{12,t}$.

2. The optimal reaction of the child is given by

$$\begin{aligned}
\tau_{c,t}^*(\tau_{p,t}, r_t) &= \frac{\lambda_2 r_t + \Delta_{c,t}}{\lambda_1 + \lambda_2 r_t + \Delta_{c,t}}(\tilde{T}_t - \tau_{p,t}) \\
&= \gamma_t(r_t)(\tilde{T}_t - \tau_{p,t})
\end{aligned}$$

3. By using the time constraints, we find each individual's leisure $(l_{1,t}^*, l_{2,t}^*, l_{c,t}^*)$. Since we know all the inputs $\{\tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, \tau_{c,t}, k_t\}$, we can also find future child quality, k_{t+1}^*. This allows us to invert the child's binding IC constraint, to find the amount of child consumption needed to make the child indifferent:

$$\ln(x_t^*) = \frac{1}{\lambda_2}\Big(V_{c,t}(\Gamma_t, a_{p,t}^0) - \lambda_1 \ln(l_{c,t}^*) - \lambda_3 \ln(k_t) - \beta_{c,t}\psi_{c,t+1}\ln(k_{t+1}^*)\Big)$$

Finally, parental consumption c_t^* follows from the budget constraint, and b_t^* fan be backed out from the ICCT reward function:

$$b_t^* = \ln(x_t^*) - r_t \ln(\tau_{c,t}^*)$$

4. The parents' value (not including the ICCT cost) can then be defined as:

$$V_{p,t}(\Gamma_t, a_{p,t}^*, r_t, b_t^*) = \tilde{u}_p(l_{1,t}^*, l_{2,t}^*, c_t^*, k_t, l_{c,t}^*, x_t^*) + \beta_p \psi_{p,t+1} \ln(k_{t+1}^*)$$

We use a numerical solver to to maximize this function with respect to the remaining choices for which we cannot find closed form solutions: $\{h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t}, e_t, r_t\}$.

5. By construction, the child is indifferent between the two equilibria. The parents will implement the ICCT equilibrium if and only if

$$V_{p,t}(\Gamma_t, a_{p,t}^*, r_t^*, b_t^*) - \omega_t \geq V_{p,t}(\Gamma_t, a_{p,t}^0)$$

where ω_t is the per-period utility cost of implementing the ICCT.

Appendix C contains more details on the estimation procedure.

B Data Appendix

B.1 Sample criteria

All our results are based on a selected sample of households that satisfy the following criteria:

(1) All households are intact over the observed period (i.e. only stable two-parent households).

(2) Households have either one or two children. We select only one child from each household (see below).

(3) All children are biological; no adopted children, no step-parents.

(4) All selected children are at least three years old in 1997, because we need a valid initial Letter Word (LW) score observation.

(5) All selected children have an observed LW score in 1997 and in 2002. Some of these also have an observed LW score in 2007 as well, although it is not required.

(6) If a household has two eligible siblings satisfying requirements (4) and (5), we select the youngest sibling by default. This has two potential advantages: parental labor supply is probably more responsive to the age of the youngest sibling than the age of the oldest sibling, and we also have a higher chance of observing the youngest sibling in 2007, which enriches the total sample.

(7) We only keep data rows for which the selected child's age is between 0 and 16.

This sample selection approach results in a final sample of $N = 247$ children or households. We have exactly 17 data rows per child, and we load the following variables after cleaning the data in Stata (not all of which are used in the code): (1) household identifier, (2) year, (3) number of child, (4) mother's age, (5) father's age, (6) family size, (7) mother's education, (8) mother's weekly labor, (9) mother's hourly wage, (10) father's weekly labor, (11) father's hourly wage, (12) weekly non-labor income, (13) child's age, (14) Letter Word raw score, (15) father's education, (16) joint parental active time, (17) mother's active time, (18) father's active time, (19) total school time, (20) regular school time, (21) other school time, (22) child's effective time endowment, (23) child's age in 1997.

B.2 Censoring and truncation

Actual data. Obvious reporting errors in the parental wage and labor supply data were resolved in the following way. For a given spouse in a given year, we replace the reported labor income and labor supply by missing values if (1) the reported labor income is positive but the reported labor hours are 0, (2) if the reported labor hours are positive but the labor income is 0, or (3) if either reported labor hours or labor income is missing.

If the non-labor income in any given year (calculated as the residual yearly income after subtracting both spouses' labor income) was either negative or above 1000 dollars

per week, we replace all the corresponding hourly wage, labor hours and non-labor income data by missing values for that year.

If the labor supply for a given spouse was above 80 hours per week, we truncate that observation at 80. If an hourly wage rate for a given spouse was either less than $5 per hour or more than $150 per hour, we replace that observation by a missing value. However, we keep all the other information pertaining to that household.

Simulated data. All simulated data are being censored in exactly the same way as the original data. Hence, if the original data contain a missing value or a censored observation for some variable at some child age, then the simulated data will have a missing value in the corresponding cell (i.e. in all R corresponding cells, since we simulate $R > 1$ data sets). Similarly, whenever the simulations yields a corner solution for labor supply, we censor the corresponding simulated wage. However, we do not censor extreme simulated wage draws (i.e. below $5 or above $150 per hour).

Given our estimation procedure for the non-labor income process, simulated non-labor income draws cannot be negative. In the event that they exceed $1000 per week, we truncate that draw at $1000. Note that we cannot replace these extreme draws by a missing value (as we did for the actual data), since we always need a real-numbered (non-missing) value of non-labor income to simulate household choices in each period.

B.3 School time

We believe the reported school time data from the CDS to be relatively noisy, as can be seen in Table B-1, which shows the distribution of reported school time at each child age t. Given the implausibly wide data range of these reported school times, we only use the median of these reported values (conditional on child age t), and use that as a measure to define the child's effective time endowment at age t as $T_{c,t} = 112 - med(s_t)$. To construct school time s_t, we use combined CDS data from 1997, 2002 and 2007, and define total school time as the sum of "regular" school time and "other" school time. These two subcomponents were constructed based on the following CDS time categories:

1. Regular school time: All time use with activity code

   - 5090: Student (full-time); attending classes; school if full-time student.
   - 5091: Daycare/nursery school for children not in school.
   - 5092-5093: School field trips inside/outside of regular school hours.

2. Other school time: all activities taking place at school with activity code

   - 5190-5193: Other classes, courses, lectures, being tutored.
   - 5680: Daycare/nursery before or after school only.

- 6130-6138: Attending a before or after school club (math, science, drama, debate, band, ...).

Detailed descriptive statistics of these schooling components are available upon request. Finally, we note that time spent with babysitters, time spent at daycare before or after school, or time spent in home care from a non-household member (CDS activity code 4870) is not counted as school time.

Table B-1: Total School Time s_t by Child Age

	Mean	Std.	Min	P25	**Median**	P75	Max	NrZeros	NrObs
$t = 3$	11.250	17.866	0.000	0.000	0.000	26.042	47.083	6	9
$t = 4$	9.845	16.152	0.000	0.000	0.000	15.833	55.000	23	36
$t = 5$	13.725	16.830	0.000	0.000	0.000	29.583	56.250	21	40
$t = 6$	24.534	17.404	0.000	0.000	32.500	37.083	47.083	9	34
$t = 7$	31.739	9.956	0.000	32.500	33.333	35.000	45.417	1	23
$t = 8$	28.274	14.201	0.000	30.833	33.458	35.000	48.333	6	38
$t = 9$	31.842	12.055	0.000	30.833	34.167	37.812	50.000	5	59
$t = 10$	31.719	10.948	0.000	31.667	33.750	36.250	47.500	4	62
$t = 11$	30.629	14.027	0.000	30.771	34.583	38.750	56.833	6	53
$t = 12$	32.826	16.841	0.000	33.333	38.333	42.500	53.750	4	22
$t = 13$	31.948	12.409	0.000	32.500	34.583	37.604	45.833	4	37
$t = 14$	37.560	14.888	0.000	35.000	37.125	43.750	74.833	5	54
$t = 15$	35.908	15.003	0.000	34.375	37.500	43.750	60.833	6	56
$t = 16$	33.638	17.220	0.000	31.042	37.917	45.000	65.000	8	49

C Estimation, Identification, and Computation Details

C.1 Computation of Model Solution

Given some vector of model parameters, we next describe the model solution algorithm for a given household i in the dataset. For each household there is a vector of observable characteristics X, including parental age (at birth) and parental education levels. In addition, we observe a measure of the child's cognitive skills at some child initial age (where the age of the initial test score observation can vary across children).

For each household in the dataset, and starting at the initial child age, we draw $r = 1, \ldots, R$ wage offer and non-labor income shocks, test score measure shocks, preferences. For each simulation draw, the model solution takes the following steps:

1. Solve for the latent cognitive skills given the draw.

2. Parental labor supply falls into 1 of 4 possible cases:

   (a) $h_{1,t} > 0, h_{2,t} > 0$
   (b) $h_{1,t} = 0, h_{2,t} > 0$
   (c) $h_{1,t} > 0, h_{2,t} = 0$
   (d) $h_{1,t} = 0, h_{2,t} = 0$

   For each of the four labor supply cases, we numerically solve the optimal time allocation vector $(h_{1,t}, h_{2,t}, \tau_{1,t}, \tau_{2,t}, \tau_{12,t})$ and, for the ICCT model, also for (e_t, r_t). For the case where both mother and father have positive labor hours, there are 5 free choice variables in the no-ICCT model, and 7 choice variables in the ICCT model. We use the Newton-Raphson algorithm to solve for the utility maximizing choices. We constrain each choice appropriately using the logit transformaton:

   $$q_i = \frac{exp(p_i)}{1 + exp(p_i)} \in (0,1),$$

   and search over the $p_i \in (-\infty, \infty)$ parameters for $i = 1, \ldots, 7$.

   For each q_i point, we define the choice variables sequentially as

   (a) Total parental investment time $\tau_{p,t} = q_1(\tilde{T}_t - s_t)$.
   (b) Mother's active time $\tau_{1,t} = q_2 \tau_{p,t}$.
   (c) Father's active time $\tau_{2,t} = q_3(\tau_{p,t} - \tau_{1,t})$.
   (d) Mother's labor time $h_{1,t} = q_4(T - \tau_{1,t} - \tau_{12,t})$
   (e) Father's labor time $h_{2,t} = q_5(T - \tau_{2,t} - \tau_{12,t})$

(f) For the ICCT model: child expenditures $e_t = q_6 Y_t$, where $Y_t = w_{1,t} h_{1,t} + w_{2,t} h_{2,t} + I_t$. For the no-ICCT model, we have a closed-form solution for e_t as a function of $(h_{1,t}, h_{2,t})$.

(g) For the ICCT model: reward elasticity $r_t = q_7 r_{max}$, where $r_{max} = 20$. For the no-ICCT model, we set $r_t = 0$.

This ensures that all parental time choices (including joint parental time $\tau_{12,t} = \tau_{p,t} - \tau_{1,t} - \tau_{2,t}$) are strictly positive and satisfy the time and budget constraints. For each time allocation choice, we compute $\tau_{c,t}(\tau_{p,t}, r_t)$ using the child's reaction function.

3. In the no-ICCT model, we find x_t, c_t and e_t using the closed form solutions derived above. After defining k_{t+1}, we can define the child's outside option, $V_{c,t}^0(\Gamma_t | a_{p,t}^0)$. In the constrained ICCT model, we numerically solve for e_t, so we can (1) use the technology function to define k_{t+1} conditional on all inputs, (2) find x_t by inverting the child's binding incentive compatibility constraint conditional on the outside option (see also Appendix A), and (3) find c_t through the budget constraint.

4. We solve for the utility maximizing choices for all possible labor supply cases and retain the highest utility choices for both the no-ICCT and ICCT models. In the benchmark model with endogenous costly ICCT choice, we retain those choices which maximize the parents' value (net of the ICCT cost ω_t).

5. With the optimal choices computed, we use the rth measurement shock to compute the measure $\tilde{k}_{r,t+1}$. Then, we reiterate by updating t to $t+1$ and latent capital k_t to k_{t+1}.

C.2 Identification

In this sub-section, we provide more details on several of the more involved identification issues.

Production Technology: Measurement Error In order to focus on key issues, consider a simplified version of our production technology, where $\ln k = \ln R + \delta \ln \tau$, with k representing latent cognitive ability, R is TFP, and τ is an observed input with associated parameter δ. Consider the following conditional mean of the observed test score k^*, given some level of the observed input.

$$E(k^* | \tau) = NQ \frac{\exp(\lambda_0 + \lambda_1 \ln R + \lambda_1 \delta \ln \tau)}{1 + \exp(\lambda_0 + \lambda_1 \ln R + \lambda_1 \delta \ln \tau)} \tag{C-1}$$

We observe the left-hand side of this expression in the data, and the right-hand side is a function of the primitives we would like to identify.

It is clear from this expression that we cannot separately identify the production function primitives (R, δ) from the measurement parameters (λ_0, λ_1). This is a generic

problem of indeterminacy due to the fact that latent child quality/skill k does not have any natural units. Identification requires some normalization to fix the location and scale of the latent variable. We normalize $\lambda_{0t} = 0$ and $\lambda_{1t} = 1$ for all t, and proceed to identify the production primitives up to this normalization.

First, consider evaluating this conditional expectation at the point $\tau = 1$, so that

$$E(k^*|\tau = 1) = NQ\frac{R}{1 + R}$$

Given the number of test questions NQ, we identify the TFP term R. Next, we identify δ from the difference in mean test scores for two values of the input $\tau \in \{1, b\}$, for $b \neq 1$:

$$E(k^*|\tau = b) - E(k^*|\tau = 1) = NQ\{\frac{\exp(\ln R + \delta \ln b)}{1 + \exp(\ln R + \delta \ln b)} - \frac{R}{1 + R}\}$$

We can extend this approach to any number of multiple observed inputs.

Production Technology: Unobserved Expenditures In our data, in contrast to the time inputs, child expenditures are not observed directly (the PSID-CDS data provides some expenditure data but is likely incomplete). To identify the productivity of the unobserved child expenditure input, we require a different identification strategy from the one we utilized for the observed time inputs. Consider two households with the same observed time inputs, but who differ in their household income (due to differences in labor or non-labor income). Given child expenditures are a normal good, this implies that the higher income household has larger expenditures on children. Expanding our simplified production function notation to include an expenditure input e and observed household income Y, we can construct the following conditional moment of the observed test scores:

$$E(k^*|\tau, Y) = NQ\frac{\exp(\ln R + \delta_\tau \ln \tau + \delta_e E(\ln e|Y))}{1 + \exp(\ln R + \delta_\tau \ln \tau + \delta_e E(\ln e|Y))}$$

$E(\ln e|Y)$ is the expected (log) expenditure for a household of income Y. Building on the analysis above, comparing households with different observed incomes then allows us to identify this term $\delta_e E(\ln e|Y)$ for any Y in the support of our data.

Our task is then to separately identify the productivity parameter δ_e from the unobserved average level of expenditure by income $E(\ln e|Y)$. We separately identify these two components using the model structure, in particular the restrictions implied by the budget constraint and from observed household choices. From the solution to our model, the optimal expenditure on children is given by $e = \Delta_e Y$, where $\Delta_e \in (0, 1)$ is the income share spent on children, a non-linear function of the primitive household preferences and technology. Δ_e is identified jointly with the other household parameters, with the key parameters comprising this share parameter (i.e. the household preference for consumption relative to the taste for child skills) identified from the observed household time allocation (i.e. time with children and labor supply).

Production Technology: Latent Skill Distribution We also need to identify the distribution of latent skills because they serve as input to the production process, not only as an output. For each child we observe measures of child quality for at least two different ages. We use the first measure of child quality as an initial condition. However, to solve the model and identify the production technology, we require an initial level of latent child quality k_t, not the measure k_t^*.

Given the measurement error assumptions, the probability of answering a question correctly p is distributed according to the Beta distribution, with parameters $(1 + k_t^*, (NQ - k_t^*) + 1)$, where k_t^* is the observed number of correct answers out of the $NQ = 57$ items. For any given realization of p (given k_t^*), $p = \tilde{p}$, we then invert the normalized measurement equation (2) to obtain a realized value of latent child quality:

$$k_t = \frac{\tilde{p}}{1 - \tilde{p}}.$$

Repeatedly drawing from the Beta distribution given the observed measure then provides a simulated distribution of latent child quality values. From these initial values of k_t, we then begin the construction of each sample path, recursively substituting the latent k_t values and other endogenous inputs determining latent k_{t+1}. When we get to the period of the second measurement, at which time the child is of age $t' > t$, the observed test score is a draw from a Binomial distribution with parameters $(NQ, p(k_{t'}))$, as described above.

C.3 Estimator

For the same household i, this process is repeated S times, so that in the end we have $S \times N$ sample paths. Using the simulated data set, we then compute the analogous simulated sample characteristics to those determined from the actual data sample. The characteristics of any simulated sample are determined by Ω, the vector of all primitive parameters that characterize the model, and the actual vector of pseudo-random number draws made in generating the sample paths. Denote the simulated sample characteristics generated under the parameter vector Ω by $\tilde{M}_S(\Omega)$. The Method of Simulated Moments (MSM) estimator of Ω is then given by

$$\hat{\Omega}_{S,N,W} = \arg\min_{\Omega}(M_N - \tilde{M}_S(\Omega))'W_N(M_N - \tilde{M}_S(\Omega)),$$

where W_N is a symmetric, positive-definite weighting matrix.[32] Given random sampling from the population of married households with a given number of children (one or two, in our case), we have $\text{plim}_{N\to\infty} M_N = M$. The weighting matrix, W_N, is simply the

---

[32]Simulation in our context is used to solve the computationally intensive integration problem. Our choice of MSM vs. an alternative simulation estimator, for example simulated maximum likelihood (SMLE) is due the greater flexibility that the MSM estimator offers in combining data from multiple sources with different sampling schemes.

inverse of the covariance matrix of M_N, which is estimated by resampling the data.[33] Given that the simulated moments are non-linear functions of the simulated draws so that $\tilde{M}_S$ is biased for fixed S, for consistency of the MSM estimator we require that S also grow indefinitely large. Let the true value of the parameter vector characterizing the model be denoted by Ω_0. Then $\text{plim}_{S\to\infty} \tilde{M}_{S,N}(\Omega_0) = M_N(\Omega_0)$. Given identification and these regularity conditions,

$$\text{plim}_{N\to\infty,S\to\infty} \tilde{\Omega}_{S,N,W} = \Omega \text{ for any positive definite } W.$$

Since W_N is positive definite by construction, our estimator Ω_{S,N,W_N} is consistent as well. We have not utilized the asymptotically optimal weighting matrix in this case due to the computational cost and issues regarding the differentiability of the objective function given the crude simulator we use. This does not seem to be a major concern since virtually all of the parameters are precisely estimated with the exception of those which we know from our earlier discussion to be tenuously identified in a data set that is the size of ours.

---

[33]We computed the M_N^g vector for each of Q resamples of the original N data points, and the covariance matrix of M_N is given by

$$W_N = \left(Q^{-1} \sum_{g=1}^{G} (M_N^g - M_N)(M_N^g - M_N)' \right)^{-1}.$$

The number of draws, Q, was set at 200.

D   Additional Tables and Figures

Figure D-1: Distribution of Child Self-investment Time by Age
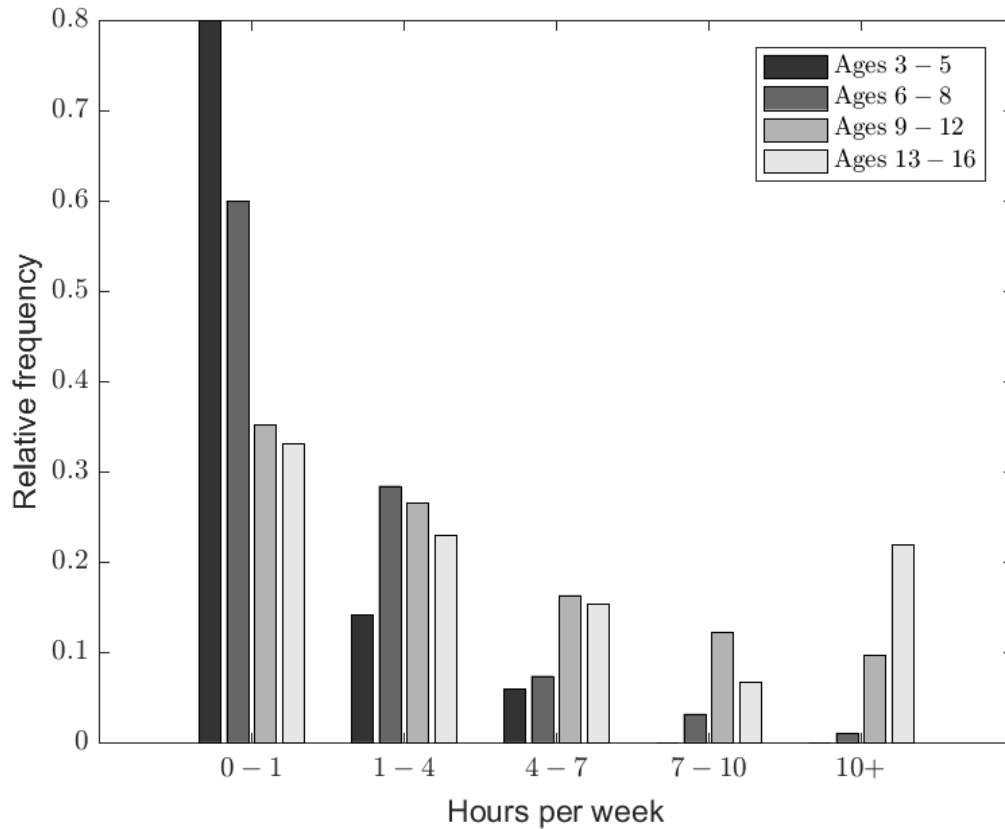
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews.
Notes: Within each child age category, the vertical bars represent the fraction of households whose reported child self-investment time was between $0-1$ hours, $1-4$ hours, $4-7$ hours, $7-10$ hours, or more than 10 hours per week.

Figure D-2: Boxplots of Child Self-investment Time by Age

(a) Hours per week
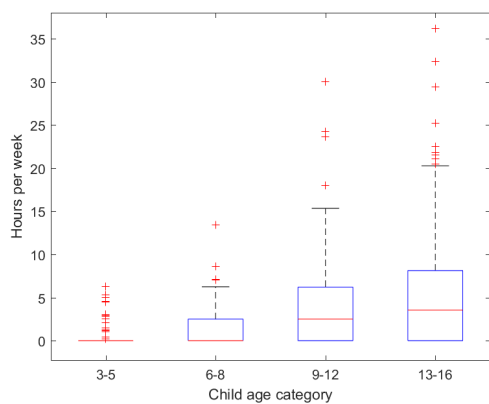
(b) Fraction of total investment time
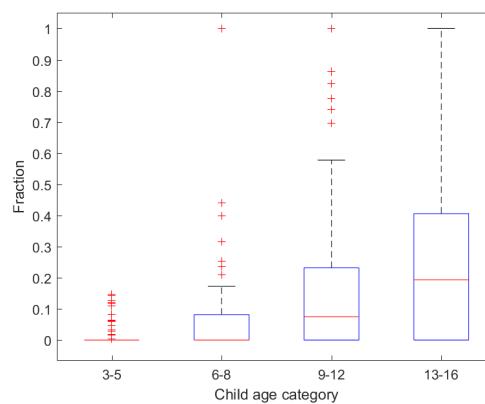
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews.
Notes: The left panel plots the distribution of the reported weekly child study time for each child age category. The right panel shows child study time as a fraction of total investment time, defined as the sum of child study time and all active time with either or both of the parents.

Figure D-3: The Effect of Household Income on Productive Time Inputs and Test Scores
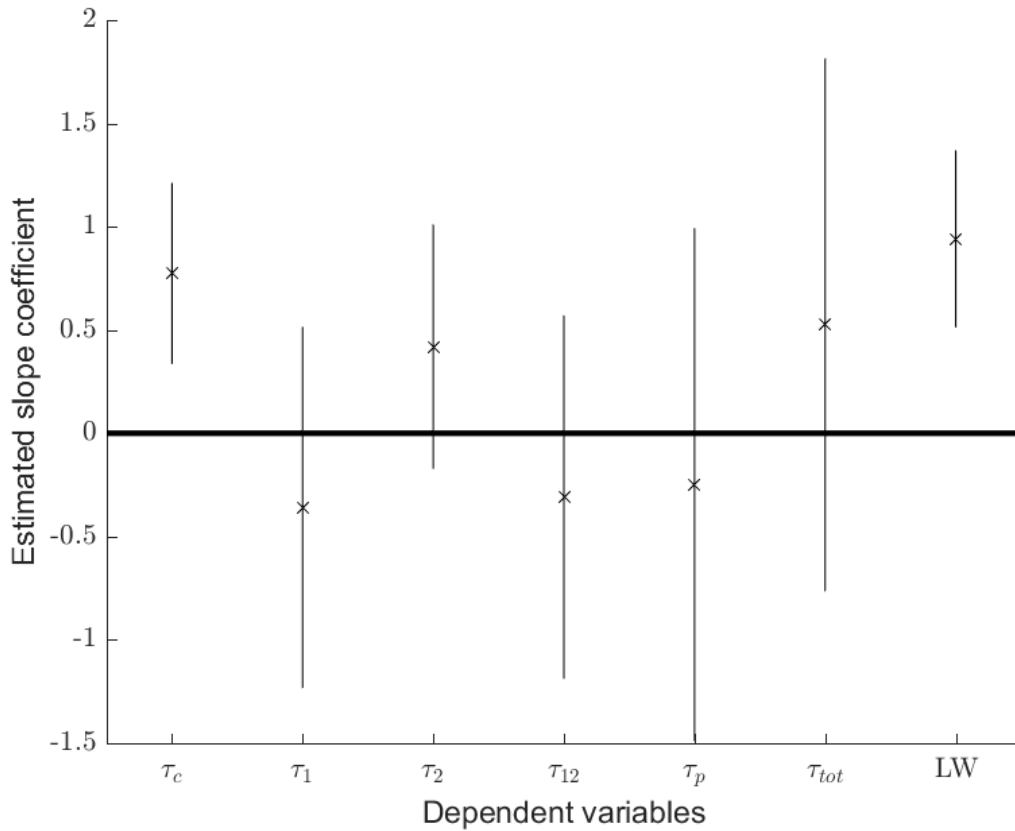
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Notes: We regress various weekly time inputs and test scores on weekly household income (in thousands of dollars, averaged across all observed years). All regressions also include child age fixed effects. We plot the estimated slope coefficients on income and their corresponding 95% confidence intervals. The dependent variables are (from left to right): (1) the child's self-investment time, τ_c, (2) mother's active time, τ_1, (3) father's active time, τ_2, (4) joint parental time, τ_{12}, (5) total parental time, $\tau_p = \tau_1 + \tau_2 + \tau_{12}$, (6) total investment time, $\tau_{tot} = \tau_c + \tau_p$ and (7) the child's raw Letter Word score, LW.

Figure D-4: Simulated and Actual Average Child's Letter Word Score
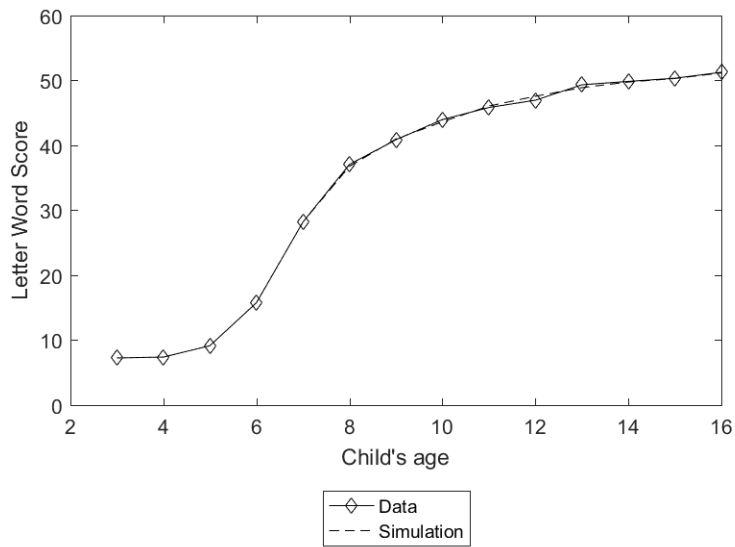
Notes: Data is actual data. Simulated is the model prediction at estimated parameters given above.
Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010.

Figure D-5: Parental Labor Supply and LFP by Child Age

(a) Working Mother's Labor Supply
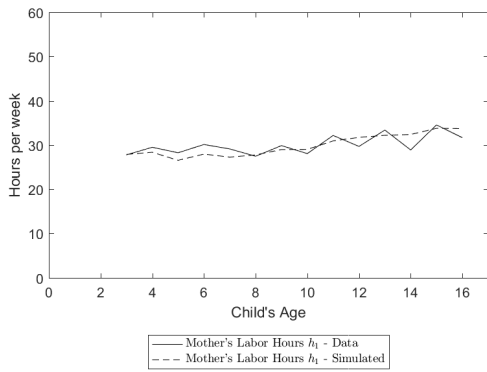
(b) Working Father's Labor Supply
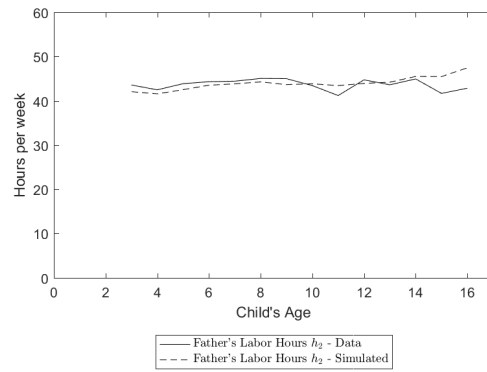
(c) Mother's Labor Force Participation
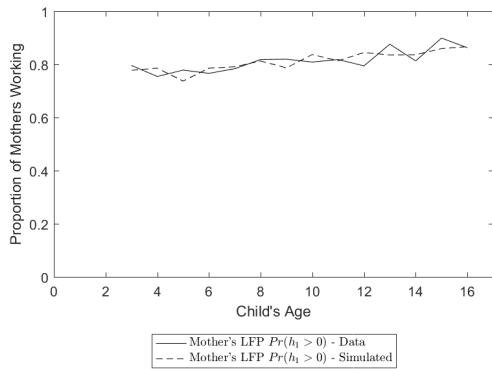
(d) Father's Labor Force Participation

99
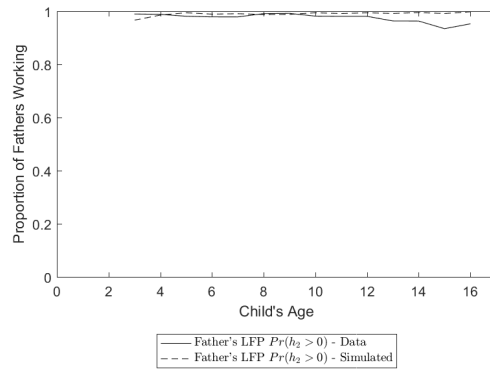
Figure D-6: Productive Time Inputs by Child Age

(a) Mother's Active Time
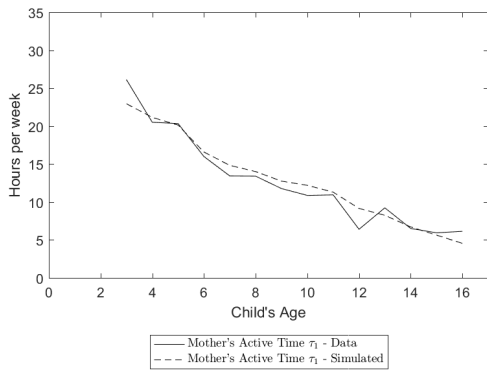
(b) Father's Active Time
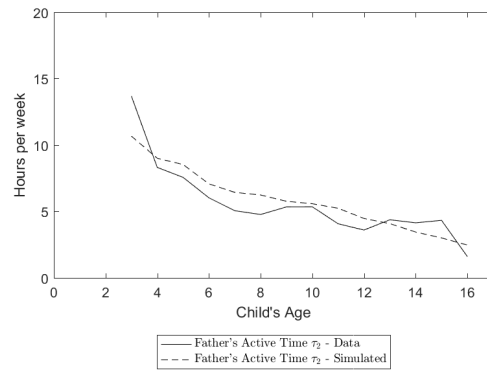
(c) Joint Parental Time
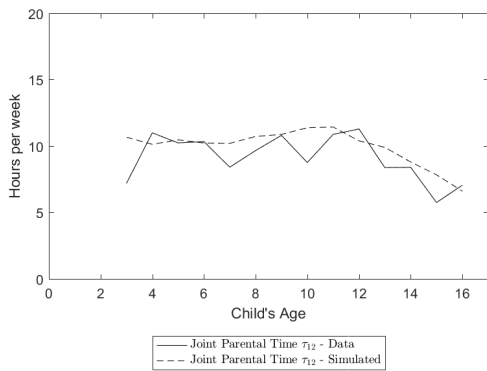
(d) Child's Self-investment Time
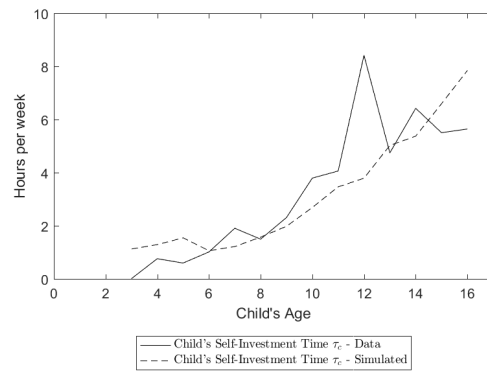
Figure D-7: Parental Hourly Wages by Parental Age and Education

(a) Mother's Hourly Wage, by Age
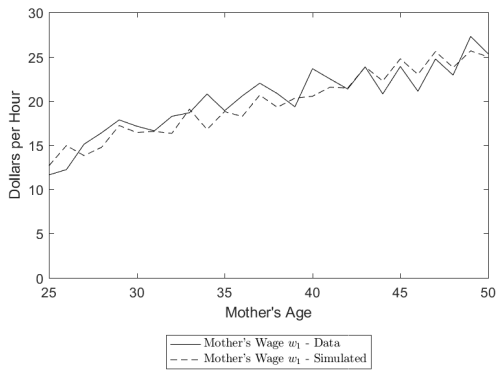
(b) Mother's Hourly Wage, by Education
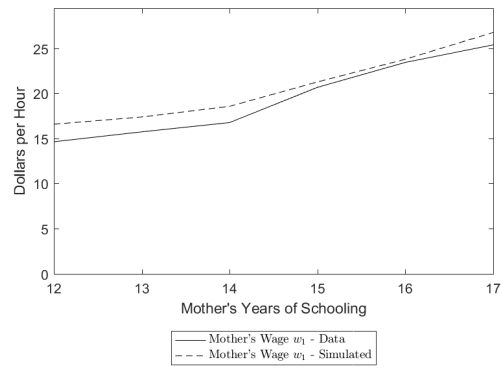
(c) Father's Hourly Wage, by Age
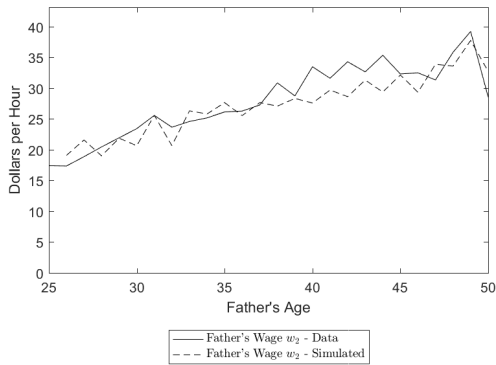
(d) Father's Hourly Wage, by Education

101
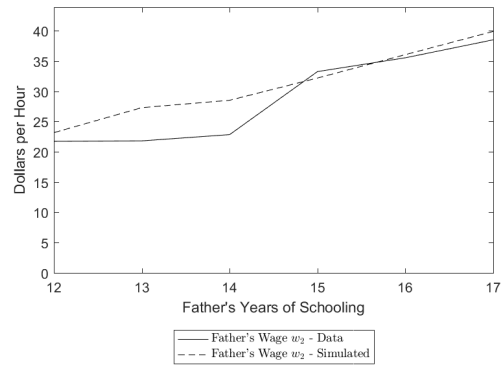
Figure D-8: Weekly Non-Labor Income by Parents' Age and Education

(a) All, by Father's Age
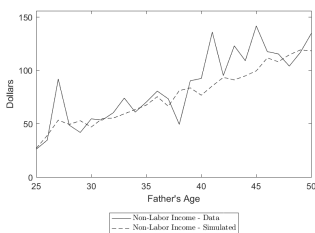
(b) Positives, by Father's Age
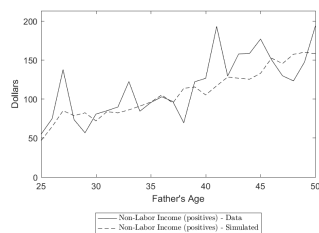
(c) Fraction > 0, by Father's Age
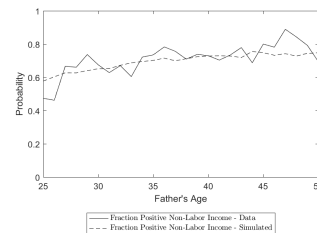
(d) All, by Mother's Age
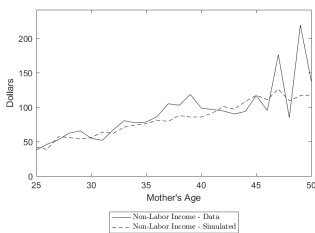
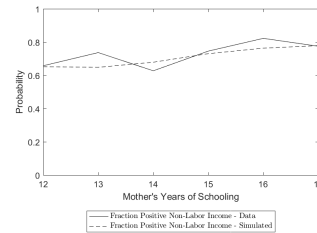(e) Positives, by Mother's Age
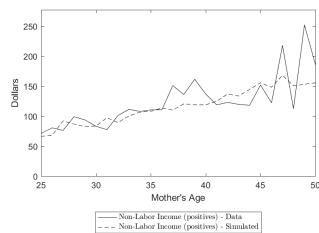
(f) Fraction > 0, by Mother's Age
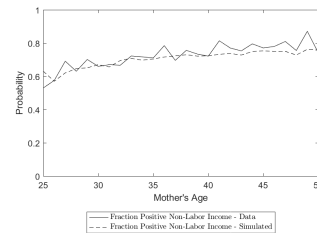
(g) All, by Father's Educ.
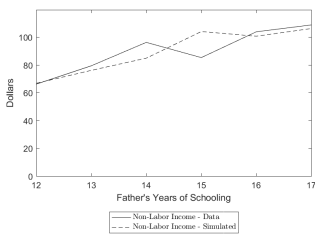
(h) Positives, by Father's Educ.
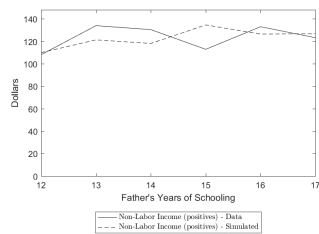
(i) Fraction > 0, by Father's Educ.
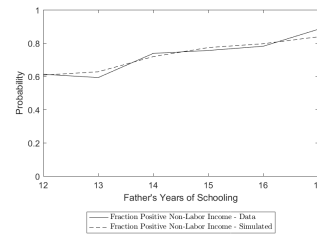
(j) All, by Mother's Educ.
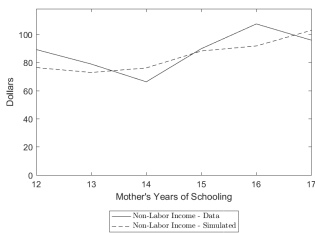
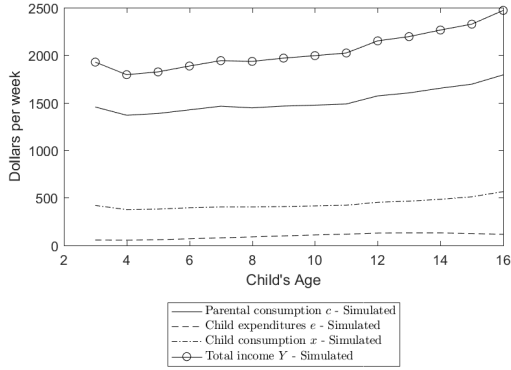(k) Positives, by Mother's Educ.
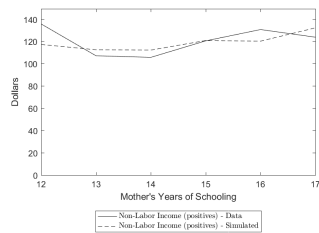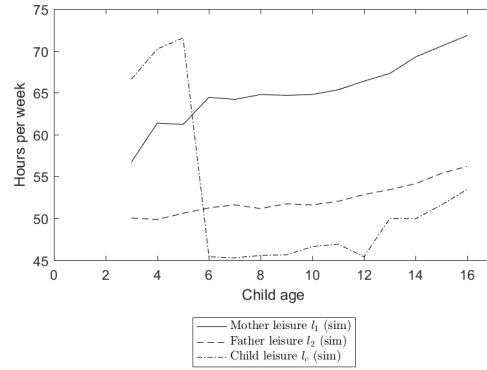
(l) Fraction >0, by Mother's Educ.

102

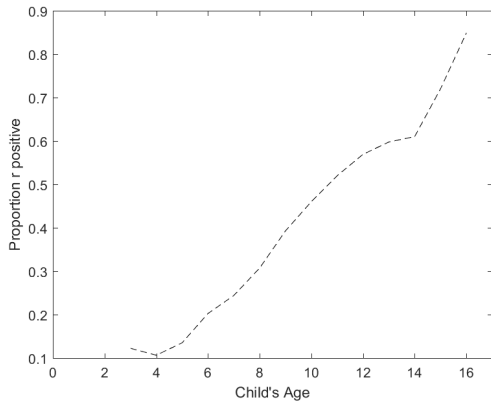# Figure D-9: Expenditures, Leisure and Internal CCT Use by Child Age

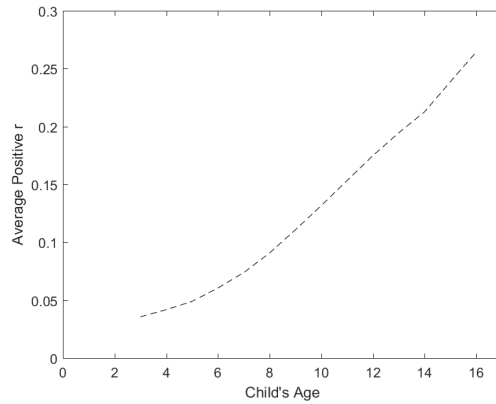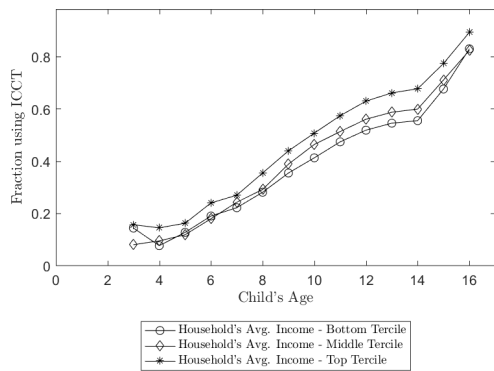### (a) Consumption, Expenditures and Income

(b) Household Leisure Time

(c) Fraction of Households using Internal CCT

(d) Avg. Reward Elasticity r (if $r > 0$)

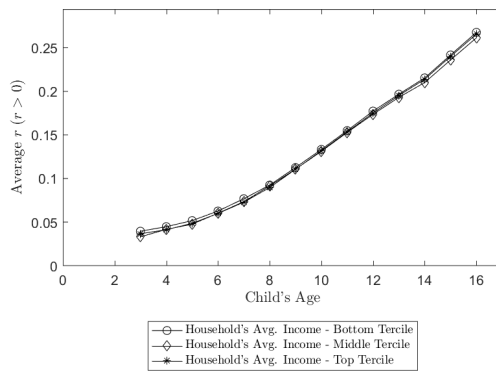### (e) Fraction using ICCT by Income

(f) Avg. Reward Elasticity by Income

Table D-1: Data Correlations

Child Ages	9-12	13-16
Letter Word Score, Child time	0.264	0.128
	(0.069)	(0.071)
Letter Word Score, Mother's Educ.	0.245	0.265
	(0.068)	(0.068)
Letter Word Score, Father's Educ.	0.301	0.342
	(0.067)	(0.066)
Letter Word Score, HH Income	0.325	0.287
	(0.076)	(0.077)
Child time, Mother's Educ.	0.078	0.160
	(0.072)	(0.071)
Child time, Father's Educ.	0.095	0.283
	(0.071)	(0.069)
Child time, HH Income	0.145	0.280
	(0.082)	(0.079)

Source: PSID-CDS combined sample from 1997, 2002 and 2007 interviews and PSID core data between 1986 and 2010. To alleviate the missing data problem at young child ages, "HH income" is defined as the average total household income within each relevant child age bin. Standard Errors of the correlations are between brackets, and are defined as $SE_r = \sqrt{\frac{1-r^2}{n-2}}$.

Table D-2: Technology Parameter Estimates

		Estimate	**SE**
Mother's Active Time (δ_1)	Intercept $\gamma_{1,0}$	-0.244	(0.03502)
	Slope $\gamma_{1,1}$	-0.252	(0.00683)
	Mother's Educ. $\gamma_{1,2}$	-0.001	(0.00007)
Father's Active Time (δ_2)	Intercept $\gamma_{2,0}$	-1.662	(0.04973)
	Slope $\gamma_{2,1}$	-0.239	(0.00846)
	Father's Educ. $\gamma_{2,2}$	0.042	(0.00276)
Joint Parental Time (δ_3)	Intercept $\gamma_{3,0}$	-1.259	(0.06380)
	Slope $\gamma_{3,1}$	-0.133	(0.00127)
	Mother's Educ. $\gamma_{3,2}$	0.020	(0.00219)
	Father's Educ. $\gamma_{3,3}$	0.018	(0.00100)
Child Expenditures (δ_4)	Intercept $\gamma_{4,0}$	-4.219	(0.17291)
	Slope $\gamma_{4,1}$	-0.053	(0.00118)
Child's Self-Investment Time (δ_5)	Intercept $\gamma_{5,0}$	-7.930	(0.13529)
	Slope $\gamma_{5,1}$	0.249	(0.00942)
Last Period's Child Quality (δ_6)	Intercept $\gamma_{6,0}$	-1.644	(0.01502)
	Slope $\gamma_{6,1}$	0.264	(0.00170)
Total Factor Productivity (R_t)	$\gamma_{7,0}$	0.47365	(0.00677)
	$\gamma_{7,1}$	1.01128	(0.00414)
	$\gamma_{7,2}$	1.44493	(0.13486)
	$\gamma_{7,3}$	8.24483	(0.10487)

Notes: Productivity parameters take the form $\delta_{i,t} = 0.01 + 0.99 \frac{\exp(\gamma_{i,0} + \gamma_{i,1}(t-1))}{\exp(\gamma_{i,0} + \gamma_{i,1}(t-1))}$, for all $i = 1, ..., 6$ and $t = 1, ..., 16$. Total Factor Productivity parameters take the form $R_t = \gamma_{7,0} + \frac{\gamma_{7,1} - \gamma_{7,0}}{1 + \exp(-\gamma_{7,2}(t - \gamma_{7,3}))}$. SEs are standard errors computed using a cluster bootstrap sampling each household with replacement.