Early Production of Cognitive Skill: Evidence from Randomly-Assigned Childcare Prices and Pre-natal Investments

Aaron Sojourner and Juan Chaparro
University of Minnesota

December 2014

Abstract:

This paper estimates an early childhood production function in which age-3 cognitive skill is modeled as a function of innate endowment, pre-natal influences, and the quality of the post-natal care environment. We develop a model of maternal choice of child human-capital investment, consumption, labor, and leisure in which production technology is distinguished from maternal preferences. The model is estimated using data from the Infant Health and Development Study (IHDP), which randomly assigned 985 new mothers and their low-birth weight babies to a control group or an intervention involving access to home visits during the child's first year of life and very high-quality, center-based care while the child was 1 and 2 years old. We treat the intervention as a post-natal investment shock to the price and availability of high-quality, non-maternal care. Data on pre-natal care choices such as smoking, made under a veil of ignorance with respect to child endowment, are used to proxy for maternal tastes. Maternal responses with respect to parenting effort, maternal time use, and other margins are studied. Key estimated parameters of the early cognitive skill production function are the degree of productive complementarity between (1) pre-natal and early post-natal investments (inter-temporal complementarity as in Cunha and Heckman (2007) and (2) maternal and non-maternal post-natal care (intra-temporal complementarity as relevant to assess effects of maternal employment). We find strong complementarity between pre-natal and early post-natal investments in the production of early child cognitive skill. Moreover, we find that maternal and non-maternal care are close substitutes in child human-capital production, with the offer of subsidized maternal care leading to increased maternal-care quality through the quantity and effort margins.

Note: asojourn@umn.edu and chap0441@umn.edu. Many thank to Greg Duncan for his advice throughout this project; Ana Auger, Jeanne Brooks-Gunn, Flavio Cunha, George Farkas and Kyoo Il Kim for help developing these ideas and, especially, to Jade Jenkins for comments and helping us use the ECLS-B data. Research reported in this publication was also supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD065704. Support was also received from the Center for Personalized Prevention Research and the Human Capital Research Collaborative at the University of Minnesota through the 2014 Early Career Investigators Network Summer Fellowship. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Table of Contents

I. Introduction

Emerging evidence from human and animal studies shows that the brain develops critically-important neural structures and functions during pregnancy and early childhood, which in turn shape future cognitive, social, emotional development and health outcomes (Sapolsky, 2004; Knudsen, Heckman, Cameron, & Shonkoff, 2006). Moreover, brain development differs between children born into low-, middle-, and high-socioeconomic status (SES) families. In their study of early brain structure, Hanson et al (2013) find that the relationship between SES and average gray-matter volume is weak in the first year of life. However, large SES-based gaps emerge between ages 1 and 3 as average gray-matter volume becomes strongly and positively correlated with SES.

These structural differences are matched by the variation in behavioral measures of cognitive skills (e.g., IQ and achievement tests) assessed in the early years of children's lives. By age 5, reading and math achievement is strongly correlated with family income (Heckman, 2006; Reardon, 2011). Gaps in cognitive and other skills that exist at that point persist throughout childhood and have strong relationships with adult productivity and life success (Cunha, Heckman, Lochner, & Masterov, 2006). Improving the quality of children's environments at very early ages can raise skill levels and reduce skill gaps in both the short- and long-run (Committee on Integrating the Science of Early Childhood Development , 2000; Ramey, Campbell, & Ramey, 1999; Duncan & Magnuson, 2013). This begs more basic questions regarding the nature and determinants of parental choices during pregnancy and in the early years of life that set children on stronger or weaker skill-building trajectories.

The current paper makes both theoretical and empirical contributions in this direction. First, we propose a model of early childhood cognitive skill formation and maternal pre- and post-natal investment choice that combines features of some existing models (Ribar, 1995; Kimmel & Connelly, 2006; Cunha, Heckman, & Schennach, 2010; Bernal & Keane, 2010; Gelber & Isen, 2013), while adding key innovations including endogenous parenting effort and a framework for analyzing maternal and non-maternal care through a unified lens. The model is of a mother with one child. The child requires some type of care – either maternal or non-maternal – at all times. The mother has a money budget, with expenditures split between non-maternal child care and consumption, and a time budget split among labor-market work, parenting, and other uses. Time spent providing maternal care requires foregoing wages and leisure. Each care type has an endogenous quality level, which is defined by how well it promotes the development of child cognitive skills. Higher quality and larger quantities of non-maternal care can be purchased with money.

For a given mother, increasing maternal-care quantity or quality requires additional parenting effort. On the margin, additional parenting effort is a source of disutility for the mother. Incorporating endogenous parenting effort is a novel contribution of our paper and captures essential economic tradeoffs parents face. The model allows for heterogeneity in maternal tastes, maternal labor-market productivity, and maternal productivity in parenting, including possible correlations between labor market productivity and parenting productivity through both observable and unobservable characteristics. We derive first-order conditions and corner solutions for optimal choices that trade off maternal leisure, consumption, parenting effort, and child-skill development. Maternal responses with respect to parenting effort, maternal time use, and other margins are studied. The model illuminates important economic tradeoffs parents face.

Second, in estimating the model, we extend the empirical skill-production-function literature back to an earlier stage of development, prior to age 3, where the biological and medical literature suggests critical development takes place. We estimate our model using data from the Infant Health and Development Study (IHDP), which randomly assigned 985 new mothers and their low-birth weight babies in eight sites to either a control group or to an intervention involving home visits and very high quality center-based care up through the child's third birthday (Bradley, et al., 1994; Gross, Spiker, & Haynes, 1997). We model the intervention as a post-natal investment shock to the price and availability of high-quality non-maternal care.

Third, drawing on ideas of inter-temporal complementarity (Cunha & Heckman, 2007; Cunha, Heckman, & Schennach, 2010), our model captures the productive relationship between pre-natal and post-natal investments in producing early child cognitive skill while accounting for parental responses to endowment shocks (Cunha & Heckman, 2007; Currie & Almond, 2011; Almond & Mazumder, 2013). Relative to post-natal investment choices, pre-natal choices are made under a veil of ignorance with respect to child endowment (Aizer & Cunha, 2012). Therefore, mothers' pre-natal investment choices – such as number of cigarettes smoked during pregnancy – provide important information about maternal willingness to trade personal consumption utility against utility from future child human capital. Because the IHDP sample is selected on birth weight and gestational age, we go to an outside source – the nationally-representative Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) – to measure how pre-natal investment choices and maternal characteristics predict birth weight and gestational age. With this forecast model, we score each mother-child pair in the ECLS-B on a pre-natal investment index and an index of child endowment, measured as the deviation of the child's realized birth status from its conditional expectation, yielding new nationally-representative estimates of the joint distribution of indexes of pre-natal investment and child endowment. By scoring each mother-child pair in the IHDP with this model, we characterize them in the national distribution in terms of pre-natal investment level and child endowment. We use the mother's pre-natal investment index as a proxy for maternal preference for child human capital in our model of post-natal investment choices. Together with a rich set of observable characteristics -- including maternal education, IQ, age, and family structure -- this gives a strong characterization of maternal and child type. Because we do not directly observe individual wage nor non-maternal child care quality, we use models estimated in outside data sources, the Current Population Survey (CPS) and the National Institute of Child Health and Human Development (NICHD) respectively, and then score the IHDP sample to measure these. In the end, we find strong complementarity between pre-natal and early post-natal investments in the production of early child cognitive skill, which drives reinforcing post-natal investments such that parents of children born with unexpectedly low birth weight tend to invest less post-natally. For comparison, we also estimate the production function using non-linear least squares to relate inputs to outputs, ignoring endogeneity. This model misses the inter-temporal complementarity.

Fourth, our study also provides estimates of the effects of maternal care and its interaction with nonmaternal care on child-skill formation (Bernal & Keane, 2010; Bernal & Keane, 2011; Gelber & Isen, 2013). For each post-natal care type – maternal and nonmaternal – the product of care quality and care quantity produces effective units of care. We allow for intra-temporal complementarity between post-natal maternal and non-maternal care in producing child cognitive skill. Gelber & Isen (2013) found that the offer of subsidized nonmaternal care raised the quality of nonmaternal care and interpret this as evidence of an intra-temporal complementarity between the two types of care. Our theoretical framework and data allow us to test this interpretation against an alternative: the increase in the quality of maternal care is driven by the reduction in the quantity of maternal care and parenting effort required upon receipt of subsidized nonmaternal care. Gelber & Isen mention this alternative interpretation as a possibility but, lacking data on maternal-care quantities, could not examine it directly. Our analysis replicates Gelber & Isen's basic, care-quality finding but supports the alternative interpretation: maternal and nonmaternal care are close substitutes in production and the offer of subsidized nonmaternal care leads to increased maternal-care quality through the quantity and effort margins instead of through complementarity.

Finally, though the reduced-form treatment effect of the IHDP intervention on child cognitive skill is known to vary by maternal education (Brooks-Gunn, Gross, Kraemer, Spiker, & Shapiro, 1992) and income (Duncan & Sojourner, 2013), the channels creating this heterogeneity in effects are unclear. Are they due to differences in child endowments, maternal tastes, maternal productivity in the labor market or productivity in parenting? A series of papers outside economics reported treatment effects on various outcomes in various subsamples such as child cognitive skill and behavior (Brooks-Gunn, Klebanov, Liaw, & Spiker, 1993), quality of the home environment (Bradley, et al., 1994), quality of parenting, maternal employment (Brooks-Gunn, McCormick, Shapiro, Benasich, & Black, 1994), and the use of paid child care (Gross, Spiker, & Haynes, 1997). Though Berlin, Brooks-Gunn, McCarton, & McCormick (1998) and Bradley, Burchinal, & Casey (2010) have studied mechanisms, they have paid less attention to endogeneity than the current paper. This paper contributes to the literature by studying the IHDP evidence through a unified theoretical lens that exploits the presence of a random shock to the budget constraint (Wolpin, 2013; Heckman & Pinto, 2013).

Bernal & Keane (2010) is the closest paper in the literature. It also estimates a production function for early child cognitive skill taking account of many of the same behavioral factors we do through a quasi-structural model with linearized policy functions. However, the IHDP affords two large advantages over the NLSY data on which they rely. First, we have experimental variation in the budget constraint driven by the IHDP design, whereas they rely on nonexperimental variation in work incentives driven by changes in welfare rules across states. Second, we have randomization in an economically-diverse sample. By focusing on variation in welfare rules, their variation occurs only among low-income mothers who tend to be low-education. This is where they must focus exclusively. The diversity of our sample allows for a better understanding of the underlying economic tradeoffs and yields more insight into the merits of universal versus income-targeted care subsidies.

Viewing the evidence through the lens of a structural model will allow for counter-factual policy simulations accounting for parental behavioral responses [TBA: we intend to do soon!]. Compare IHDP (full-day, high-quality care) effects to other policies such as full-day, mediocre care or half-day, high-quality care. Universal versus income-targeted.

The following section describes the structure and properties of our model, including the economics of the relevant tradeoffs parents face. Section III describes how we identify and estimate the model. Results are described in Section IV, followed by a discussion of the findings.

II. Model

This section presents a model of early cognitive skill production and of maternal pre-natal and post-natal investment decisions particularly with respect to the quality and quantity of maternal and non-maternal care. We have incorporated ideas especially from Cunha & Heckman (2007); Cunha, Heckman, & Schennach (2010); Aizer & Cunha (2012); Bernal & Keane (2010); Del Bono, Ermisch, & Francesconi (2012); and Gelber & Isen (2013).

a) The post-natal period

Production function: Suppose early childhood cognitive skill is produced according to:

$$h = \tilde{f}(I_1, h_0, \varepsilon).$$

In particular, allow age-3 IQ to depend on post-natal investment (I_1), the stock of human capital at birth (h_0), and unmeasured, post-natal productive heterogeneity (ε). A basic question is to understand the first-order marginal productivities of post-natal investment ($\tilde{f}_1$) and of human capital at birth ($\tilde{f}_2$).

Further, Cunha & Heckman (2007) focused labor economists on trying to understand the dynamic complementarity of investments. Dynamic (or inter-temporal) complementarity captures how the productivity of current investment depends on the incoming stock of skill, embodying past investment and the innate endowment. In the present context, this key property of the human capital production function is $\tilde{f}_{12} = \frac{\partial^2 \tilde{f}}{\partial I_1 \partial h_0}$. It is interesting because it has a strong influence on the optimal timing of investments.

We focus on understanding the inter-temporal complementarity of early post-natal investments (age 0-2) with child status at birth, proxied by birth weight and embodying innate endowment and pre-natal influences. As recently reviewed by Almond & Mazumder (2013), many researchers have been working to develop evidence on this question. Most use a prenatal shock and a single post-natal investment proxy to go after the sign on the cross-partial derivative and the sign on the parental response to an endowment shock. Typically, there is not enough data on inputs to estimate a production function.

For the econometrician, the main challenge lies in separating production function parameters from behavioral responses driven by child or parental heterogeneity. Attempts to estimate the production function directly may not be credible in strictly observational studies due to the endogeneity of chosen inputs. First, measured post-natal investments (I_1) may be positively correlated with unmeasured investments or with determinants of investment productivity, genetic or otherwise (ε). To help deal with this concern, we rely on the large, randomly-assigned shock to post-natal investment generated by the IHDP treatment assignment. Second, human capital at birth (h_0) may be correlated with ε. Absent further information, it almost certainly is. For instance, mothers who choose higher levels of pre-natal investment will tend to give birth to babies with higher h_0. They will also choose higher levels of post-natal investment, some potentially unobserved (ε), inducing a positive correlation. Another possible channel for correlation is that mothers who see that their child is unexpectedly in bad condition at birth, as proxied by low birth weight and prematurity perhaps, might react to the news of this bad endowment shock by compensating with additional post-natal investment. A negative shock revealed at birth by unexpectedly low h_0 may lead the mother to choose higher unmeasured investments, perhaps because the productivity of those investments is higher or because her aversion to very bad development outcomes raises her marginal utility of improvement. Alternatively, the mother may reinforce a negative endowment shock and reduce pre-natal investment, perhaps because the productivity of those investments is lower. To deal with this issue, we draw on pre-natal investment choices, made under the veil of ignorance about child endowment, to develop a proxy for maternal taste (Aizer & Cunha, 2012). This makes a key determinant of investment levels "observable" and, in combination with more conventional demographic measures, allows us to condition on relevant maternal and child heterogeneity. We assume that the residual unobserved variation is independent of observables.

We also explore the productive relationship between two kinds of post-natal investments: embodied in maternal and non-maternal care. For child skill development, quality of care matters. Every child requires supervision and care for a total of T_c <168 waking hours per week, creating a child time budget (Equation 1 below). This is common across all children. The distribution of developmentally-relevant care quality and type varies. Allowing for the possibility that maternal care is special, we consider two kinds of care: maternal and non-maternal such that maternal care hours (r) plus non-maternal care hours (n) must total T_c. This constitutes the child's time budget constraint. Non-maternal care encompasses many arrangements, such as care by other relatives or purchased child care services. The qualities of maternal care (q^r) and non-maternal care (q^n) also vary. Post-natal investment depends on effective units of maternal and non-maternal care:

$$I_1 \equiv g(q^n n\ ,\ q^r r).$$

We study the productive relationship between maternal and non-maternal care. A long and largely-separate strand of economics, sociology, and human development literature has focused on the effects of maternal employment on child development. Implicitly, this literature divides maternal care as separate in kind from non-maternal care and asks about the relative productivity of maternal (g_2) versus non-maternal care (g_1). There is also the possibility of intra-temporal complementarity between different kinds of contemporaneous inputs ($g_{12} = \frac{\partial^2 g}{\partial(q^n n + q^t t)\partial q^r r}$). Are maternal and non-maternal care perfect substitutes or do they complement each other in production, as Gelber & Isen (2013) conclude? Together, these first- and second-order marginal effects help illuminate the developmental effect of shifting from maternal to non-maternal care.

In the context of the IHDP, a special source of non-maternal care is available to those in the treatment group. We assume that households in the treatment group can use the child development center (CDC) services for up to $\bar{\tau}$ hours per week. Mothers choose how many hours to take up ($t \leq \bar{\tau}$). In the control group, $t = \bar{\tau} = 0$. The quality of free CDC daycare is exogenous and equal to q^t. Effective units of CDC care are $q^t t$. Effective units of non-maternal care become the sum of effective units of CDC care and other care: $q^t t + q^n n$. Effective units of care are a central concept in the model and Table 1 summarizes them. These are the investment inputs of the child's human capital production function. Combining $\tilde{f}$ and g yields the production function (Eq. 4).

Economic model: we develop an economic model of maternal choices to develop intuition about the tradeoffs that families face, as a lens for illuminating mechanisms by which the reduced-form IHDP treatment effect operates, and as a way to deal with potential endogeneity in identification of the production function. The variables and relations are discussed before introducing the model formally.

Each mother also has a time budget constraint (Eq. 2). She divides her time endowment (T_p) between three activities. Maternal child care (r), as previously discussed, is one. Leisure (l) and wage work (L) are the others. The mother can earn a potential wage per hour (w), which is an increasing function of observed human capital (m) and unobserved ability (ω) (Eq. 7). Total maternal income equals labor earnings plus any exogenous non-labor income (Y). Total income can be used to pay for consumption (c) and to purchase child care at cost $\pi q^n n$ (Eq 3). We will allow the potential wage offer (w) and the quality of maternal care (q^r) to be correlated due to observed maternal characteristics, like maternal education, or because of unobserved maternal heterogeneity in ability.

The quality of maternal care (q^r) depends on the mother's human capital (m), unobserved individual heterogeneity in ability (ω), and latent, instantaneous parenting effort (e) (Eq. 8). We assume maternal human capital and ability are given but mothers choose the level of parenting effort they invest. Regarding non-maternal, non-CDC sources of care, mothers choose both how much time to use (n) and the quality of care (q^n). These have a non-negative and exogenous price equal to π per each unit of effective care received. Both the labor-market and parenting productivity models (Eq. 7 & 8) will take specific functional forms in estimation.

Total parenting effort is the product of the instantaneous effort level (e) and effort duration, that is hours of maternal care provided (r) (Eq. 6). This parenting quality-quantity tradeoff has been missing from the economics literature, perhaps because datasets with both parenting time and parenting quality are rare. This captures the idea that high-quality parenting is more difficult to maintain over longer periods than shorter periods. Parenting can be exhausting. We are far more likely to stop reading with a child and plop her in front of a TV after 6 hours of intense, developmentally-appropriate parenting than after 6 minutes.

Some distaste for free CDC services is required to explain incomplete take-up of high-quality, free care, as in Bernal and Keane (2010). This distaste captures individual heterogeneity in felt stigma or logistical challenges posed by night work shifts or the presence of multiple young children, with only one eligible for CDC care.

Maternal preferences are represented by $U(c, l, p, h, t)$. Utility is increasing in consumption (c), leisure time (l), and the child's human capital (h), but it may be decreasing in total parenting effort (p) and time the child spends at the CDC (t).

We now have all the elements to summarize the parental decision process as the following utility maximization problem. Appendix 1 contains a dictionary of all the variables in the model.

$$\underset{c,q^n,n,e,r,l,L,t}{\text{Max}} U(c, l, p, h, t)$$

s.t. $r + n + t = T_c$ (1 – Child's time constraint)

$r + L + l = T_p$ (2 – Parent's time constraint)

$c + \pi q^n n = wL + Y$ (3 – Budget constraint)

$h = f[q^n n + q^t t\ ;\ q^r r\ ;\ h_0, \varepsilon]$ (4 – Child's human capital technology)

$t \leq \bar{\tau}$ (5 – Maximum program take-up time)

$p = er$ (6 – Total parenting effort)

$w = w(m, \omega)$ (7 – Wage offer)

$q^r = q^r(m, \omega, e)$ (8 – Parenting quality technology)

A full income - full consumption budget constraint is obtained by combining equations 1, 2 and 3. This simplifies the constraints and yields the following expression:

$$c + [\pi q^n - w]n + wl = w\left[T_p - T_c\right] + wt + Y \qquad (9)$$

Full income, which corresponds to the right hand side of (9), is derived from non-labor income, total free daycare time valued at the parent's market wage and net parental time endowment, also valued at the market wage. On the other hand, full consumption has three components. The first one is traditional consumption. The second one is total value of other sources of care, like purchased daycare. Focus on the economic cost of this decision, which is $\pi q^n - w$: one additional hour of daycare with quality q^n will cost the parent a total of πq^n monetary units, but this decision will free up one hour of parental time, which has a labor market value of w. The third component of full consumption is leisure time priced at the market wage.

Time allocation decisions are reduced to three variables: hours at the CDC (t), other sources of non-maternal and non-CDC care (n), and maternal leisure time (l). If n and t are observed, we can deduce r (maternal care time) from the child's time constraint (1). And once we know r and l, the optimal labor supply decision (L) can be inferred from the parent's time constraint (2). We can now rewrite the post-natal problem as:

$$\underset{c,q^n,e,n,l,t}{\text{Max}} \; U(c,l,p,h,t)$$

s.t. $c + [\pi q^n - w]n + wl = w\left[T_p - T_c\right] + wt + Y$

$h = f[q^n n \; + \; q^t t \; ; \; q^r(e,m,\omega)[T_c - n - t] \; ; \; h_0, \varepsilon]$

$t \leq \bar{\tau}$

$p = er$

Using two multipliers (λ and μ), the Lagrangian of this problem will be:

$$\begin{aligned}\mathcal{L} = U(c,l,e[T_c - n - t]\,, f[q^n n \; + \; q^t t \; ; \; q^r(e,m,\omega)[T_c - n - t] \; ; \; h_0, \varepsilon], t) \\ + \lambda\left[w\left[T_p - T_c\right] + wt + Y - c - [\pi q^n - w]n - wl\right] + \mu[\bar{\tau} - t]\end{aligned}$$

The following notation for marginal rates of substitution (MRS) denominated in consumption and the marginal product of each input ($f_{t=1,2,3}$) will be useful throughout the rest of the document:

$$U_c = \frac{\partial U}{\partial c} > 0, U_l = \frac{\partial U}{\partial l} > 0, U_p = \frac{\partial U}{\partial p} < 0, U_h = \frac{\partial U}{\partial h} > 0, U_t = \frac{\partial U}{\partial t} < 0, q_e^r = \frac{\partial q^r(e,m,\omega)}{\partial e}$$

$$MRS_{l,c} = \frac{U_l}{U_c} > 0, \;\; MRS_{p,c} = \frac{U_p}{U_c} < 0, \;\; MRS_{h,c} = \frac{U_h}{U_c} > 0, \;\; MRS_{t,c} = \frac{U_t}{U_c} < 0$$

$$f_1 = \frac{\partial f}{\partial[q^n n \; + \; q^t t]}, \;\; f_2 = \frac{\partial f}{\partial[q^r(e,m,\omega)r\,]}, \;\; f_3 = \frac{\partial f}{\partial h_0}$$

Denote the MRS for leisure as $MRS_{l,c}$. Denote the MRS for parenting effort as $MRS_{p,c}$ and note that it is negative. Another negative term is the MRS between the time the child spends at the CDC, $MRS_{t,c}$, which captures possible participation stigma. Finally, the MRS for age-3 cognitive skill is $MRS_{h,c}$. The marginal return of instantaneous parenting effort (e) with respect to the quality of maternal care is q_e^r. Regarding the human capital technology, f_1 is the marginal productivity of non-maternal care, f_2 the marginal productivity of maternal care, and f_3 the marginal productivity of human capital at birth, as proxied by birth weight.

Optimal choices: This section describes properties of the optimal choices formally and discusses the economic tradeoffs behind these decisions. The solution to the post-natal parental problem is given by a vector of eight variables $(\lambda^*, \mu^*, c^*, q^{n*}, e^*, n^*, l^*, t^*)$ which comply with all the Kuhn-Tucker conditions available in Appendix 2. Optimal labor supply (L^*) and optimal parental care (r^*) will be given by:

$$r^* = T_c - n^* - t^* \qquad L^* = T_p - l^* - r^*$$

The following expressions are based on the Kuhn-Tucker conditions, but use the marginal rates of substitution which are more suitable for economic interpretation. These first order conditions focus on solutions where the budget constraint is binding ($U_c = \lambda^* > 0$) and parents do not use all the hours available for them at the CDC ($0 \leq t^* < \bar{\tau}\,; \mu^* = 0$), because this is a predominant characteristic in the IHDP data. We contemplate cases where the mother could decide not use help from other caretakers ($n^* \geq 0$). Finally, for a more transparent presentation of the first order conditions, we will focus only on interior solutions for c^*, q^{n*}, e^* and l^*.

$$\frac{\partial \mathcal{L}}{\partial l}: \;\; MRS_{l,c} = w \tag{A}$$

$$\frac{\partial \mathcal{L}}{\partial t}: \;\; MRS_{h,c}[f_1 q^t - f_2 q^r] + w - MRS_{p,c}\, e \;\leq -MRS_{t,c} \qquad \frac{\partial \mathcal{L}}{\partial t} t = 0 \qquad 0 \leq t < \bar{\tau} \tag{B}$$

$$\frac{\partial \mathcal{L}}{\partial n}: \;\; MRS_{h,c}[f_1 q^n - f_2 q^r] + w - MRS_{p,c}\, e \;\leq \pi q^n \qquad \frac{\partial \mathcal{L}}{\partial n} n = 0 \qquad n \geq 0 \tag{C}$$

$$\frac{\partial \mathcal{L}}{\partial q^n}: \; f_1\, MRS_{h,c} = \pi \tag{D}$$

$$\frac{\partial \mathcal{L}}{\partial e}: \;\; f_2\, q_e^r\, MRS_{h,c} = -MRS_{p,c} \tag{E}$$

Equations (A), (B) and (C) determine all optimal time decisions. Like in any other traditional labor supply model, optimal leisure is given by the equality of the market wage rate and the marginal rate of substitution between leisure and consumption (A).

Equation (B) explains the decision to use the free services from the CDC. Possible marginal benefits are on the left hand side of the inequality. Marginal costs are on the right hand side. The effect of one additional hour at the CDC on the child's human capital will depend on the **quality gap** between maternal and CDC care, which is equal to $f_1 q^t - f_2 q^r$. The first term ($f_1 q^t$) measures the raw marginal effect of CDC time on the child's human capital, but such an event implies that the child spent one less hour with her mother. Therefore, we must subtract the marginal effect of maternal time on the child's human capital ($f_2 q^r$) to determine the final effect. Notice that the quality gap could be either positive or negative, and it is valued using the marginal rate of substitution between human capital and consumption ($MRS_{h,c}$). Use of services from the CDC also imply that the mother could work additional hours paid at the market wage rate w. It is also the case that using the CDC implies less total parental effort (er). This possible relief for the mother is valued using the marginal rate of substitution between parental effort and consumption ($MRS_{p,c}$). Although the CDC offers a free service, there may be an implicit cost generated by participation stigma or by associated logistical challenges. This cost is captured by the marginal rate of substitution between time spent at the CDC and consumption ($MRS_{t,c}$).

Optimal non-maternal and non-CDC care time is given by (C). Note its similarity with the decision rule for use of CDC services. In this case, what matters is the **quality gap** between other caregivers and maternal care, $f_1 q^n - f_2 q^r$. Another difference lies in possible economic costs, measured by πq^n.

Recall that quality of care is endogenous in this model. Quality of non-maternal, non-CDC care (q^n) is determined by (D). (E) explains the decision of optimal parenting effort (e), which is the key choice behind quality of maternal care (q^r). In both cases, the marginal return to additional quality depends on the human capital technology. The marginal productivity of non-maternal care (f_1) measures the benefits of additional quality from this type of caregiver. Extra maternal effort translates into additional human capital in the child depending on the marginal productivity of maternal care ($f_2\ q_e^r$). Both marginal effects must be valued using the marginal rate of substitution between the child's human capital and consumption ($MRS_{h,c}$). Recall that π is the price of one unit of effective care by a caregiver different than the mother or the CDC. The implicit price of maternal effort is measured using the marginal rate of substitution between parental effort and consumption ($MRS_{p,c}$).

b) The pre-natal period

The child's human capital at birth (h_0) is not given. Instead, it is determined by pre-natal investments and the child's endowment. We will denote with b the technology behind human capital at birth. The child's endowment (ϕ) is a random variable not controlled by the mother and not known to the mother during pregnancy. The mother chooses how much to invest pre-natally (I_0). Investments promote healthy development of the fetus at some expense to the mother's utility. For instance, these investments can take the form of refraining from smoking, drinking alcohol excessively, or using recreational drugs, getting timely medical care, or managing diet and exercise to gain weight at a desired rate.

Maternal preferences are defined over pre-natal investment efforts before birth (I_0) and the child's expected human capital (h). Therefore, the child's human capital production function (f) is an additional constraint. The pre-natal parental problem is an expected utility maximization problem:

$$\underset{I_0}{\text{Max}}\ E_\phi[V(I_0, h)]$$

s.t.

$$h_0 = b[I_0\ ,\ \phi] \qquad \text{(10 - Human capital at birth technology)}$$

$$h = f[\ (q^n n + q^t t)\ ,\ q^r r,\ h_0, \varepsilon] \qquad \text{(11 - Post-natal production function)}$$

$$\phi\ \sim F_\phi \qquad \text{(12 - Maternal beliefs about endowment)}$$

During the pre-natal period, all mothers assume $t = 0$ because none know about the IHDP study or treatment. Taking account of how optimal post-natal investment will respond to induced changes in h_0, the first-order condition for the choice of pre-natal investment is:

$$E_\phi\left[V_{I_0} + V_h\left(f_1\left[\frac{\partial q^{n*}}{\partial h_0}n^* + q^{n*}\frac{\partial n^*}{\partial h_0}\right] + f_2\left[q^r_e\frac{\partial e^*}{\partial h_0}r^* + q^r(e^*, m)\frac{\partial r^*}{\partial h_0}\right] + f_3\right)b_{I_0}\right] = 0$$

where $b_{I_0} = \frac{\partial b}{\partial I_0} = \frac{\partial h_0}{\partial I_0}$. Consider two pregnant women who differ only in their marginal distaste for pre-natal investment (V_{I_0}) but face the same future marginal returns to human capital (V_h) and marginal returns to pre-natal investment (b_{I_0}). The woman for whom the marginal cost of pre-natal investment is less steep will choose an optimal pre-natal investment level which is higher than the optimal pre-natal investment level of the other woman.[1]

The main purpose of the pre-natal model is to characterize maternal preference heterogeneity, which is useful when interpreting their post-natal investment choices. In

---

[1] We suppose that $\partial V/\partial I_0 < 0$ and $\partial^2 V/\partial {I_0}^2 < 0$, which indicates that pre-natal investment has increasing marginal utility costs for any parent.

particular, two otherwise-similar mothers who choose different levels of pre-natal investment (I_0) reveal information about the relative values they place on child human capital, $MRS_{h,c}$.

III. Data and Identification

To generate our estimates, we draw data from the Infant Health and Development Program (IHDP), which offered a package of services including free, full-day, Abecedarian-type early education to a randomly chosen subset of 985 children in eight sites scattered around the country (Gross et al. 1997). Eligible babies were born low birth-weight (≤ 2,500 g) and premature (≤ 37 weeks gestation). Eligibility was not restricted by family income, race or ethnicity and a demographically heterogeneous set of children and families enrolled in the study.

The IHDP provided weekly home visits from a paraprofessional during the first year of life and seven to nine hours of daily child care when the child was age 1 and 2. Participating child care centers used a game-based curriculum that emphasized language development. A high-quality evaluation design included random assignment of program services to treatment and control groups and assessment of intelligence quotient (IQ) during and up to 15 years after the completion of the program. Published reports have shown very large impacts of the program on IQ during the program and generally smaller impacts, confined exclusively to the heavier babies, after it ended (Brooks-Gunn et al. 1994; Gross et al. 1997; McCarton et al. 1997; McCormick et al. 2006).

We focus here on the first 3 years, when inputs were best measured. The original sample size in the IHDP study was 985 infants. There are no missing values for some fundamental variables, such as birth weight or time measures. However, the procedures for imputing proxies for missing variables also created some missing values (See Appendix f). The overall pattern of missing values implies that our final sample size is 815 infants from the original study.[2]

Table 2 presents summary statistics for all the variables in the model. We use the Stanford Binet IQ at 36 months (all ages are chronologically corrected, based on due date) as a summary measure of the child's cognitive skill (h). Average IQ in the sample (88.5) is below the national standardized average (100). Average birth weight (h_0) for the entire sample is equal to 1.8 kilograms. Table 3 looks at the detail of these variables across four groups of maternal education, for women in the treatment and the control group. Birth weight does not vary with maternal education or treatment group status.

The inputs into the production function should reflect maternal and non-maternal effective units of care during the first three years of life. Hours per week of maternal care (r) correspond to the average of maternal self-reported hours in the 18-month and 30-month family interviews. Hours of care at the CDCs (t) come from administrative data and is the average

[2] Cases with missing values and, thus, excluded from the analytic sample are similar to other cases with complete values on observable baseline characteristics. Treatment status, birth weight, gestational age, maternal education, race, and ethnicity do not predict the selection of an infant into our sample (Appendix Table AT.1).

weekly attendance over the 2 years it was offered. Hours of care with other care takers is calculated as a residual, using the child's time constraint ($n = T_c - r - t$).[3]

Children in the treatment group attended the CDCs for approximately 18 hours per week and hours of maternal care in the same group do not vary substantially with maternal education. These mothers reported an average of 52.4 hours per week of maternal care. There is a different pattern for mothers in the control group, with an inverse relation between maternal education and hours of maternal care: a mother in the control group who did not finish high school reported on average 65.9 hours per week, whereas mothers with a college degree reported 57.3 hours of care.

We want to make a clear distinction between quantity (r) and quality (q^r) of maternal care. To measure quality, we use the Learning and Literacy component from the Infant-Toddler Home Environment score, which is assumed to be affected by maternal effort oriented towards building cognitive capacity in her child (Linver, Martin and Brooks-Gunn, 2004; Fuligni, Han and Brooks-Gunn, 2004). The IHDP gathered data for the Home Environment scores at 12-month and 36-months.

Table 4 presents the yes-or-no questions available in the data. We created two quality indexes, at 12 and 36 months, by performing factor analysis on the tetrachoric correlation matrix across items at each age, standardizing the first factor within each age, averaging across ages for each individual, and standardizing this average. Measurement units of $q^r$are standard deviations within the IHDP sample.[4]

Quality of maternal care is correlated with maternal education (Table 3). In the control group, the gap between college graduate mothers and high school dropouts is equal to 1.5 standard deviations. Additionally, there is a positive effect of treatment on maternal quality and the effect is larger for mothers with low educational attainment. The quality gap between treatment and control for mothers with no high school degree is 0.22 standard deviations. The same gap for college graduate mothers is just 0.01 standard deviations.

We do not observe wage or potential wage. We assume that potential wage depends on observed and unobserved maternal characteristics. Using estimates from a Heckman selection model of potential wage estimated in a similar Current Population Survey sample using variables available in both the CPS and IHDP samples, we obtain the average predicted wage, $\widehat{w}(m)$, for a mother with given observables (m) (Heckman, 1974; Mulligan & Rubinstein, 2008).[5] Because unobserved potential-wage heterogeneity may be correlated with parenting

---

[3] We suppose T_c = 87.5 hours per week. Based on Inglowstein, et. al. (2003), p. 304, average night time sleep duration for 2 year olds is approximately 11.5 hours. Therefore, the average child would require (24 – 11.5) x 7 hours of direct care per week.

[4] For estimation purposes, we need to avoid negative values in q^r and some other variables in the model. We achieve this by adding up the inverse of the minimum value after standardization. This procedure does not change the standard deviation or the underlying covariances in the data, but guarantees that the new minimum value in the distribution will be equal to 0.

5 Details of the procedure are in Appendix f). Summary statistics on predictors for the CPS and IHDP samples and estimates of the selection model are in Appendix Tables AT.2 and AT.3.

productivity and taste parameters in ways that influence optimal choices, we introduce a productive heterogeneity parameter (ω) that captures differences in productivity across mothers that affects both wages and maternal care quality.

$$w(m,\omega) = \widehat{w}(m)\,\omega \quad \Rightarrow \quad ln[w(m,\omega)] = ln[\widehat{w}(m)] + \ ln(\omega) \qquad \text{(F)}$$

We assume that maternal care quality is the sum of instantaneous maternal effort and a fixed productivity term. The productivity term is a function of observed and unobserved wage determinants. We assume a Cobb-Douglas form. Q captures average productivity.

$$q^r(e,m,\omega) = Q\left[\widehat{w}^{\chi_m}\omega^{(1-\chi_m)}\right] + e \qquad \text{(G)}$$

This approach recognizes that parents differ in their ability to produce higher care quality and that this ability may be correlated with observables and with labor market opportunities. It also recognizes that all parents can produce high or low quality care.

To measure the quality of nonmaternal care, we combine IHDP data on child and family characteristics and on the chosen nonmaternal care settings -- partner, sibling, grandmother, another relative, babysitter, day care home, day care center, someone else and the child's father, if he lives in another home – with data from the NICHD-SECCYD on similar variables and care quality. Details are in Appendix f).

We use the following quasi-linear functional form for parental utility,

$$U(c,l,p,h,t) = c + \gamma_l\ ln[l] + \gamma_p\ ln[\bar{p} - p] + \gamma_h\ ln[h] + \gamma_t\ ln[\bar{\tau} - t]$$

where all γ parameters are strictly positive. The marginal utility of consumption is equal to 1.[6] The marginal utility of leisure and human capital are positive. $\bar{\tau}$ is the a maximum number of hours per week of free childcare services. Likewise, $\bar{p}$ represents an upper bound on total parenting effort. Note that the marginal utility of parenting effort is negative. That is also the case for take up time of free childcare services. Under this assumption, the marginal rates of substitution between each commodity and consumption are the following:

$$MRS_{l,c} = \frac{\gamma_l}{l} > 0 \quad \Rightarrow \quad ln\left[MRS_{l,c}\right] = ln[\gamma_l] - ln[l] \qquad \text{(H)}$$

$$MRS_{p,c} = \frac{-\gamma_p}{\bar{p}-p} < 0 \quad \Rightarrow \quad ln\left[-MRS_{p,c}\right] = ln\left[\gamma_p\right] - ln[\bar{p} - p] \qquad \text{(I)}$$

$$MRS_{h,c} = \frac{\gamma_h}{h} > 0 \quad \Rightarrow \quad ln\left[MRS_{h,c}\right] = ln[\gamma_h] - ln[h] \qquad \text{(J)}$$

[6] We will introduce curvature in utility from consumption. Income effects are fundamental in order to explain the positive effect of CDC care on quality and hours of maternal care. Income effects are also important to explain changes on other endogenous variables. Therefore, we plan to modify the utility function from quasi-linear to linear in log of consumption: $U(c,l,p,h,t) = ln(c) + \gamma_l\ ln[l] + \gamma_p\ ln[\bar{p} - p] + \gamma_h\ ln[h] + \gamma_t\ ln[\bar{\tau} - t]$.

$$MRS_{t,c} = \frac{-\gamma_t}{\bar{\tau}-t} < 0 \qquad \Rightarrow \qquad ln\left[-MRS_{t,c}\right] = ln[\gamma_t] - ln[\bar{\tau} - t] \tag{K}$$

Consider the log of the marginal rate of substitution between the child's human capital and consumption, $ln\left[MRS_{h,c}\right]$. It measures how much the parent values *additional* cognitive development in her child and can be broken down in two parts. The first term, $ln[\gamma_h]$, captures parental preferences. Parents who are willing to undergo substantial consumption sacrifices to invest more resources in her child can be represented through a larger value of γ_h. The second term, $-ln[h]$, is a consequence of decreasing marginal utility for child skill.

The theoretical framework accounts for several sources of unobserved heterogeneity: parental preferences ($\gamma_l, \gamma_t, \gamma_p, \gamma_h$), parental productivity (ω, Q) and child-specific shocks (ε). We can decompose each unobserved variable into a first part correlated with one or more shifters and a second part based on an error term, orthogonal to the shifters:

$$ln(\Omega^s) = \mathrm{Z}^s \kappa^s + \nu^s \tag{L}$$

where $\Omega^s \in \left\{\gamma_l, \gamma_t, \gamma_p, \gamma_h, \omega, Q, \varepsilon\right\}$ represents the unobserved heterogeneous variable. Z^s are the corresponding shifters, accompanied by a vector of coefficients denoted by κ^s. The orthogonal error term is symbolized by $\nu^s \in \{\nu^l, \nu^t, \nu^p, \nu^h, \nu^\omega, \nu^Q, \nu^\varepsilon\}$.

In this framework, two key conditions are required for identification. First, the set of shifters for the marginal utility of child cognitive skill is not exactly the same as the set of shifters for unobserved post-natal productivity ($Z^h \neq Z^\varepsilon$). If so, one could not separate whether a mother chooses higher investment due to differences in tastes or productivity. Using the pre-natal investment index as a proxy that shifts maternal tastes is valuable here. Second, the set of shifters for the marginal utility of leisure cannot be exactly the same as the set of shifters for unobserved productivity in the labor market and in parenting ($Z^l \neq Z^\omega$). This is similar to the standard need for an exclusion restriction to identify a Heckman selection model of wages. Otherwise, we could not separate whether chooses to take more leisure because she has a strong preference for it rather than low productivity in alternative time uses.

Table 5 summarizes the chosen shifters. We assume maternal age is correlated with taste for leisure. Therefore, maternal age is used as the shifter for $ln(\gamma_l)$. Parental distaste for the CDC could be a consequence of coordination problems if the parent has more than one child. For this reason, we have used number of children less than age 5 as the shifter for $ln(\gamma_t)$. Consider now total factor productivity in the quality of maternal care technology (Q). Past parenting experience should influence the ability with which the parent is able to deliver quality care, given her productive characteristics. Thus, total number of parenting years of own children up until age 5 is a shifter candidate for $ln(Q)$. The IHDP data includes a measure of maternal IQ (the PPVT score). We use this key variable to shift unobserved maternal productive heterogeneity, $ln(\omega)$, and child-specific shocks to the human capital technology, $ln(\varepsilon)$.

Finally, consider pre-natal investments. Appendix f) describes how we obtain a nationally-normed measure of pre-natal investments (I_0^*) based on the ECLS-B sample. We use I_0^* as a proxy of mother's tastes for child human capital, $ln(\gamma_h)$, and distaste for parenting effort, $ln(\gamma_p)$.

In general, children in the IHDP received strongly negative endowment shocks when compared to the distribution of shocks in the nationally representative ECLS-B sample.[7] The IHDP sample's average endowment z-score is -2.5 and its average percentile is 6. The median percentile is 3. Mothers in the IHDP tend to make lower levels of pre-natal investment than observationally similar mothers in the national population. The average pre-natal investment z-score is -0.78 and average percentile is 27. The median percentile is 19. Summary statistics in Table 2 are for a shifted measure used in production to ensure all production inputs are positive.

We assume that the post-natal human-capital production function (*f*) corresponds to a two-level, nested translog function. That is, $\tilde{f}$ and g are defined as:

$$\ln(h) = \ln\left(\tilde{f}[I, h_0, \varepsilon]\right)$$

$$\equiv \beta_0 + \beta_1 \ln[I] + \beta_2 \ln[h_0] + \beta_3 \ln[I]^2 + \beta_4 \ln[h_0]^2 + \beta_5 \ln[I]\ln[h_0] + \ln[\varepsilon] \qquad \text{(M)}$$

$$\ln(I) = \ln(g[q^n n + q^t t, q^r r])$$

$$\equiv \alpha_1 \ln[q^n n + q^t t] + \alpha_2 \ln[q^r r] + \alpha_3 \ln[q^n n + q^t t]^2 + \alpha_4 \ln[q^r r]^2$$

$$+\alpha_5 \ln[q^n n + q^t t]\ln[q^r r] \qquad \text{(N)}$$

Translogs can be understood as flexible first-order approximations of an unknown function. In our results thus far, we are shutting down curvature ($\alpha_3 = \alpha_4 = \beta_3 = \beta_4 \equiv 0$). Also, we set $\beta_1 \equiv 1$ and let the α parameters capture the contribution of post-natal investments to h. In both functions, we allow for interaction terms. Most notably, β_5 measures dynamic complementarity: whether the productivity of post-natal investments is affected by the stock of human capital at birth (embodying pre-natal investments and child endowment).

Finally, assume all error terms are distributed joint normal, which allows us to estimate the parameters in the model through maximum likelihood,

[7] Appendix f) contains details about how these variables are measured.

$$[\nu^l, \nu^t, \nu^p, \nu^h, \nu^\omega, \nu^Q, \nu^\varepsilon]' \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \qquad \text{(O)}$$

The vector of parameters (Θ) consists of the free coefficients in the translog human capital technology; the shifter coefficients in the unobserved heterogeneity equations and the parameter in the quality of maternal care technology. Therefore,

$$\Theta \equiv [\alpha_1, \alpha_2, \alpha_5, \beta_0, \beta_2, \beta_5, \kappa^l, \kappa^t, \kappa^p, \kappa^h, \kappa^\omega, \kappa^Q, \kappa^\varepsilon, \chi_m]$$

Our primary approach to estimating the model is maximum likelihood. We also estimate the production function parameters by non-linear least squares for comparison. This method does not account for potential endogeneity of inputs or use the pre-natal information. Appendix e) presents the log likelihood equations.

IV. Results

Model estimates of the production function

Recall the specification of the technology based on two nested translog production functions:

$$\ln(I) = \alpha_0 + \alpha_1 \ln[q^n n + q^t t] + \alpha_2 \ln[q^r r] + \alpha_3 \ln[q^n n + q^t t]^2 + \alpha_4 \ln[q^r r]^2 + \alpha_5 \ln[q^n n + q^t t] \ln[q^r r]$$

$$\ln(h) = \beta_0 + \beta_1 \ln[I] + \beta_2 \ln[h_0] + \beta_3 \ln[I]^2 + \beta_4 \ln[h_0]^2 + \beta_5 \ln[I] \ln[h_0] + \ln[\varepsilon].$$

The first order conditions derived from the post-natal problem, in addition to the distributional assumptions and other sources of unobserved individual heterogeneity, allow us to estimate the parameters of the production function using maximum likelihood estimation (MLE). Table 6 reports such estimates, side-by-side with estimates based on non-linear least squares (NLLS) as a benchmark for comparison. The table reports the coefficients for the production technology ($\alpha_1, \alpha_2, \alpha_5, \beta_0, \beta_2, \beta_5$). We also report the coefficient for each shifter associated with different sources of unobserved individual heterogeneity ($\kappa^l, \kappa^t, \kappa^p, \kappa^h, \kappa^\omega, \kappa^Q, \kappa^\varepsilon$), as well as the parameter from the technology that translates parenting effort into quality of maternal care (χ_m). Table 6 also includes estimates for linear combinations of terms from Σ, which correspond to all results denoted by σ. Additional assumptions about the parameters are $\alpha_0 = 0$, $\alpha_3 = 0$, $\alpha_4 = 0$, $\beta_1 = 1$, $\beta_3 = 0$ and $\beta_4 = 0$.

As expected, total effective units of maternal ($\ln[q^r r]$) and non-maternal care ($\ln[q^n n + q^t t]$) translate into post-natal investment. This is a consequence of positive and significant estimates for α_1 and α_2. The negative and significant coefficient for α_5 can be interpreted as evidence of substitution between maternal and non-maternal care. The most important difference between both models is the estimate for β_5. This parameter corresponds to the interaction between post-natal investment and birth weight. According to the dynamic complementarity hypothesis, the marginal productivity of post-natal investment should increase with the child's human capital stock. Thus, we interpret the positive and significant estimate for β_5 as evidence of dynamic complementarity.

The MLE model generates an estimate for χ_m relatively close to 0.5, which indicates that the correlation between quality of maternal care and potential wage in the labor market is due to observable maternal characteristics, such as educational attainment and experience, as well as unobserved individual productive heterogeneity.

Finally, note that the estimate for κ^h is positive and significant. This parameter measures the relation between our measure of observed pre-natal investment (I_0^*) and maternal tastes for the child's human capital, $ln(\gamma_h)$. As expected, higher levels of pre-natal investment can be considered as a strong signal of maternal preferences in favor of the child's human capital.

First-order marginal effects from the MLE estimates

Marginal effects derived from the model's specification allow for better economic interpretation of these results. Based on the nested translog technology, we can calculate the following marginal effects for all the inputs in the production function (as semi - elasticites):

$$\frac{\partial ln[h]}{\partial h_0} = \left(\frac{\partial ln[h]}{\partial ln[h_0]}\right)\frac{1}{h_0}$$

$$\frac{\partial ln[h]}{\partial n} = \left(\frac{\partial ln[h]}{\partial ln[I]}\right)\left(\frac{\partial ln[I]}{\partial ln[q^n n + q^t t]}\right)\left(\frac{q^n}{q^n n + q^t t}\right)$$

$$\frac{\partial ln[h]}{\partial t} = \left(\frac{\partial ln[h]}{\partial ln[I]}\right)\left(\frac{\partial ln[I]}{\partial ln[q^n n + q^t t]}\right)\left(\frac{q^t}{q^n n + q^t t}\right)$$

$$\frac{\partial ln[h]}{\partial r} = \left(\frac{\partial ln[h]}{\partial ln[I]}\right)\left(\frac{\partial ln[I]}{\partial ln[q^r r]}\right)\left(\frac{q^r}{q^r r}\right)$$

$$\frac{\partial ln[h]}{\partial q^r} = \left(\frac{\partial ln[h]}{\partial ln[I]}\right)\left(\frac{\partial ln[I]}{\partial ln[q^r r]}\right)\left(\frac{r}{q^r r}\right)$$

where,

$$\frac{\partial ln[h]}{\partial ln[h_0]} = \beta_2 + 2\beta_4 ln[h_0] + \beta_5 ln[I]$$

$$\frac{\partial ln[h]}{\partial ln[I]} = \beta_1 + 2\beta_3 ln[I] + \beta_5 ln[h_0]$$

$$\frac{\partial ln[I]}{\partial ln[q^n n + q^t t]} = \alpha_1 + 2\alpha_3 ln[q^n n + q^t t] + \alpha_5 ln[q^r r]$$

$$\frac{\partial ln[I]}{\partial ln[q^r r]} = \alpha_2 + 2\alpha_4 ln[q^r r] + \alpha_5 ln[q^n n + q^t t]$$

The six panels in Figure 1 report the first order marginal effects for n, t, r, q^r, $ln(I)$ and h_0. These lines represent the marginal productivity of each input at the 5th, 25th, 50th, 75th and 95th percentile in the sample. In each graph, all other inputs are fixed at the sample mean. Note that the estimated technology has positive but decreasing marginal effects for all of its underlying inputs.

Table 7 explores the pattern of first order marginal effects by treatment group and maternal education. Each marginal effect is calculated at the average input level for the corresponding cell. The most relevant result from Table 7 is the marginal effect of CDC time for mothers in the control group. These marginal effects measure a counter-factual: the return of the first hour of CDC time for mothers that did not have access to the program. Note that the return is the highest for mothers with no high school degree, 0.0047 log points per hour, compared to 0.0016 log points per hour for mothers with a college degree.

Non-maternal and maternal care times have the same units (hours per week). Therefore, we can use the ratio between marginal effects in the control group to measure the equivalent of one hour of CDC time in terms of one hour of maternal time. One hour of care at the CDC is equivalent to 1.81 hours of maternal care by a mother with no high school degree (= 0.00474 /

0.00262). In the case of high school graduates, the ratio is equal to 1.37 hours (= 0.00357 / 0.00260) and for mothers with some college education, the same ratio is equal to 0.93 hours (= 0.00238 / 0.00256). Finally, consider mothers with a college degree. The marginal effect of one additional hour of their care is larger than the marginal effect of the first hour at the CDC. Thus, the equivalence time measure is less than one and equal to 0.68 hours (= 0.00165 / 0.00242).

Effect of treatment on post-natal investments

Consider now Table 8 and Table 9. Table 8 explores the effect of treatment, pre-natal investment and the endowment shock on different components of post-natal investment available in the data. All regressions control for maternal characteristics such as age, maternal education and race, among other controls. Table 9 reports the same set of regressions for observations by treatment and control group. It also includes one additional column, where the dependent variable is care time at the CDC.

We find evidence of reinforcing post-natal investment with respect to the endowment shock. There is also evidence of reinforcing attendance to the CDC, where parents of children with higher endowment shocks used more hours of free child care (column 8 in Table 9). The pre-natal investment measure (I_0^*) is a good predictor of post-natal investment decisions in the control group. In particular, there is a strong association between pre-natal investment and quality of post-natal maternal care (column 3 in Table 9).

Gelber & Isen (2013) found that parents with kids randomly selected for Head Start eligibility raise the level of parenting quality. They interpret this as evidence of perceived complementarity between parental and non-parental care quality. However, they also recognize that this could be due instead to “changes in parent time with children through impacts on the parents’ time constraint” but lack good measures of parental care quantity to get at this directly.

We reproduce their main empirical finding, that low-income parents whose children are eligible for free child care do increase their parenting quality, but we extend the analysis to incorporate a measure of maternal care quantity. In contrast, we find a decline in effective units of maternal care driven by declines in maternal care hours. Further, we find that treatment does not increase parenting quality among higher-income families.

Robustness

We worried that the one parent-one child model obscures too much given that some families include multiple adults and children. However, restricting the sample to households where the mother is the only adult yields similar results. Evidence is provided in Table 10. The table presents, side by side, the model estimates using the full sample (first column, N = 815) and the subsample of infants whose mothers are single and are the head of the household (second column, N = 428). Also, there is no evidence of an effect of the intervention on fertility or family structure.

The weekly home visiting program during the first year of the child's life was intended to shift maternal beliefs and expectations about child development and parenting. However, we modeled the intervention as affecting the budget constraint and did not include this kind of channel. If the home-visiting program shifted maternal beliefs and knowledge and this translated into different parenting and into different child development in the child's first year, then ignoring this channel would be problematic. However, the evidence suggests this did not occur. The IHDP collected data on maternal beliefs and knowledge about child development using the Concepts of Development Questionnaire and the Knowledge of Infant Development Index when the children were age 12-months. There were no treatment effects on these (Gross, Spiker, & Haynes, 1997). Also, there was no evidence of a treatment effect on the HOME environment inventory nor on the children's cognitive skill at 12-months.[8]

The sample is composed exclusively of children born low birth weight and premature. Some may have suffered developmental compromise and may be subject to different developmental processes than children born under normal conditions. There are a few points to make regarding this issue. First, we characterize the sample with respect to the criteria on which they are selected (birth weight and gestational age at birth) within the context of a nationally-representative birth cohort and with respect to the determinants of these selection variables (maternal characteristics, pre-natal investment choices, and child endowment) and we build these differences into our model. Second, we re-estimate the model excluding extremely premature infants, whose gestational age was 30 weeks or less and we obtain similar results (Table 10, column 3, N = 653). Third, even if one is reluctant to generalize outside the sample's support, the estimates are valuable as informative about children born low birth weight and premature.

Future

Solve the model again relaxing quasi-linear assumption on mother's utility function. Is this all driven by an unmodeled wealth effect? How? Calibrate or estimate? What functional form?

Policy simulations

Indirect inference rather than MLE?

[8] According to Bradley et al (1994) state, "…whatever else the intervention accomplished during the 1st year, it did not produce home environments that demonstrated a significantly higher level of stimulation or support for infants' learning and development…." Gross, Spiker, & Haynes (1997) report no effect on age-1 cognitive skill, as measured by Bayley mental development scale.

Limitations

We ignore the costs of goods as inputs, aside from measuring the quality of care. We believe this is justified at this very early stage of development, although the cost of goods themselves and their ability to substitute for or complement personal care-giver attention may be more important as children age.

MLE assumptions strong.

What threats come from using proxies for expected potential wage, pre-natal investment, child endowment, and nonmaternal care quality? Our model of wage already accounts for the possibility of measurement error through inclusion of the unobserved productivity parameter ω. Pre-natal investment (I_0^*) is not an input to production. It is used only as a shifter of utility parameters. The point is to order individuals' tastes. We have estimated models that add squared terms of I_0^* as a taste shifter wherever I_0^* appears. Results do not change appreciably. Child endowment is not used to estimate the production function. We use birth weight, which is measured directly, instead. Endowment is only used later, in Tables 5 and 6, as a predictor in regressions presented to conveniently express the parental behavioral response to treatment and the endowment shock. If our measure of endowment is noisy, then our estimate here may suffer attenuation bias towards zero. However, we detect a significant, positive relationship between endowment and post-natal investment in the full sample and in the control group. The sign alone is evidence of reinforcing response. The magnitude is less important. The fact that endowment do not predict post-natal investment level in the treatment makes good sense because the offer of free investment swamps the parent's usual behavioral response. The most problematic proxy is for nonmaternal-care quality because this is an input into production. However, note that it enters estimation only for individuals with n > 0, as a product with n, and, for the treatment group, it enters as part of a sum with $q^t t$. A large part of the variation driving the estimation of the production function with respect to nonmaternal care comes from $q^t t$. In a sense, $q^n$is a nuisance which we are trying to "control." However, if this creates attenuation bias, it would reduce the estimated impact of nonmaternal care on post-natal investment and of post-natal investment on the outcome. However, we do see large of effects of each.

V. Conclusion

TBA

VI. References

Aizer, A., & Cunha, F. (2012). *The Production of Human Capital: Endowments, Investments, and Fertility.* Cambridge, Mass.: National Bureau of Economic Research.

Almond, D., & Mazumder, B. (2013). *Fetal Origins and Parental Responses.* Chicago, Ill.: Federal Reserve Bank of Chicago.

Berlin, L. J., Brooks-Gunn, J., & McCarton, C. (1998). The Effectiveness of Early Intervention: Examining Risk Factors and Pathways to Enhanced Development. *Preventative Medicine, 27*, 238-245.

Bernal, R., & Keane, M. P. (2010). Quasi-structural estimation of a model of childcare choices and child cognitive ability production. *Journal of Econometrics*, *156*(1), 164-189.

Bernal, R., & Keane, M. P. (2011). Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers. *Journal of Labor Economics, 29*(3), 459-512.

Bradley, R. H., Burchinal, M. R., & Casey, P. H. (2010). Early Intervention: The Moderating Role of the Home Environment. *Applied Developmental Science, 5*(1), 2-8.

Bradley, R. H., Whiteside, L., Mundfrom, D. J., Casey, P. H., Caldwell, B. M., & Barrett, K. (1994). Impact of the Infant Health and Development Program (IHDP) on the Home Environments of Infants Born Prematurely and with Low Birthweight. *Journal of Educational Psychology, 86*(4), 531-541.

Brooks-Gunn, J., Gross, R. T., Kraemer, H. C., Spiker, D., & Shapiro, S. (1992). Enhancing the Cognitive Outcomes of Low Birth Weight, Premature Infants: For Whom Is the Intervention Most Effective? *Pediatrics, 89*(6), 1209-1215.

Brooks-Gunn, J., Klebanov, P. K., Liaw, F.-r., & Spiker, D. (1993). Enhancing the Development of Low-Birthweight, Premature Infants: Changes in Cognition and Behavior over the First Three Years. *Child Development, 64*(3), 736-753.

Brooks-Gunn, J., McCormick, M. C., Shapiro, S., Benasich, A. A., & Black, G. W. (1994). The Effects of Early Education on Maternal Employment, Public Assistance, and Health Insurance: The Infant Health and Development Program. *American Journal of Public Health, 84*(6), 924-931.

Card, D., DellaVigna, S., & Malimendier, U. (2011). The Role of Theory in Field Experiments. *Journal of Economic Perspectives, 25*(3), 29-62.

Committee on Integrating the Science of Early Childhood Development . (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development.* (D. A. Phillips, & J. P. Shonkoff, Eds.) Washington, D.C.: National Academy Press.

Cunha, F., & Heckman, J. J. (2007). The technolgy of skill formation. *American Economic Review, 97*(2), 31-47.

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

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In E. Hanushek, & F. Welch, *Handbook of Economics of Education* (Vol. 1, pp. 697-812). Boston, Mass.: Elsevier.

Currie, J., & Almond, D. (2011). Human Capital Development before Age Five. In D. Card, & O. Ashenfelter, *Handbook of Labor Economics* (Vol. 4b, pp. 1315-1486). Elsevier.

Del Bono, E., Ermisch, J., & Francesconi, M. (2012). Intrafamily Resource Allocations: A Dynamic Structural Model of Birth Weight. *Journal of Labor Economics, 30*(3), 657-706.

Duncan, G. J., & Magnuson, K. (2013). Investing in preschool programs. *Journal of Economic Perspectives*, *27*(2), 109-132.

Duncan, G., & Sojourner, A. (2013). Can Intensive Early Childhood Intervention Programs Eliminate Income-based Cognitive and Achievement Gaps? *Journal of Human Resources, 48*(4), 945-968.

Gelber, A., & Isen, A. (2013). Children's Schooling and Parents' Behavior: Evidence from the Head Start Impact Study. *Journal of Public Economics, 101*, 25-38.

Gross, R. T., Spiker, D., & Haynes, C. W. (1997). *Helping low birth weight, premature babies : the infant health and development program.* Stanford, Calif.: Stanford University Press.

Hanson, J. L., Hair, N., Shen, D. G., Shi, F., Gilmore, J. H., Wolfe, B. L., & Pollak, S. D. (2013). Family Poverty Affects the Rate of Human Infant Brain Growth. *PLoS one, 8*(12).

Heckman, J. J. (1974). Shadow Prices, Market Wages, and Labor Supply. *Econometrica, 42*(4), 679-694.

Heckman, J. J. (2006). Skill Formation and the Economics of Investing in Disadvantaged Children. *Science, 312*(1900), 1900-1902. doi:10.1126/science.1128898

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *PNAS, 104*(33), 13250-13255.

Heckman, J. J., & Pinto, R. (2013). *Econometric Mediation Analyses: Identifying the Sources of Treatment Effects from Experimentally Estimated Production Technologies with Unmeasured and Mismeasured Inputs.* Cambridge, Mass.: National Bureau of Economic Research.

Kimmel, J., & Connelly, R. (2006). Mothers' Time Choices: Caregiving, Leisure, Home Production, and Paid Work. *Journal of Human Resources*, 644-681.

Knudsen, E. I., Heckman, J. J., Cameron, J. L., & Shonkoff, J. P. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences, 103*(27), 10144-10162.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school.* Washington, D.C.: Economic Policy Institute.

Mulligan, C. B., & Rubinstein, Y. (2008). Selection, Investment, and Women's Relative Wages Over Time. *Quarterly Journal of Economics, 123*(3), 1061-1110.

Olds, D. L., Kitzman, H., Cole, R., Robinson, J., Sidora, K., Luckey, D. W., . . . Holmberg, J. (2004). Effects of Nurse Home-Visiting on Maternal Life Course and Child Development: Age 6 Follow-Up Results of a Randomized Trial. *Pediatrics, 114*(6), 1550-1559.

Ramey, C. T., Campbell, F. A., & Ramey, S. L. (1999). Early Intervention: Successful Pathways to Improving Intellectual Development. *Developmental Neuropsychology, 16*(3), 385-392.

Reardon, S. F. (2011). The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations. In G. J. Duncan, & R. J. Murnane, *Whither Opportunity* (pp. 91-116). New York, N.Y.: Russell Sage Foundation.

Ribar, D. C. (1995). A Structural Model of Child Care and the Labor Supply of Married Women. *Journal of Labor Economics, 13*(3), 558-597.

Sapolsky, R. M. (2004). Social Status and Health in Humans and Other Animals. *Annual Review of Anthropology, 33*, 393-418. doi:10.1146/annurev.anthro.33.070203.144000

Wolpin, K. I. (2013). *The Limits of Inference without Theory.* Cambridge, Mass.: MIT Press.

VII. Tables

Table 1: Possible caretakers and effective units of care provided

		Variables in the model		
	Caretaker	Time with caretaker	Quality of care	Effective units of care provided
Maternal Care	Mother	r	q^r	$q^r r$
Non-maternal care	Free Daycare (CDC)	t	q^t	$q^t t$
	Non-maternal, non-CDC	n	q^n	$q^n n$

Note: total effective units of non-maternal care will be equal to $(q^t t) + (q^n n)$.

Table 2: Summary statistics

	Variable	Mean	Std. Dev.	Min	Max	N
Stanford Binet IQ (Corrected Age) at 36M	h	88.55	20.09	43	147	815
Birth Weight (kgs)	h_0	1.80	0.46	0.54	2.5	815
Hours per week with other caretakers	n	22.22	14.46	0	61	815
Hours per week at CDC	t	7.00	10.46	0	40.52	815
Hours per week of maternal care	r	58.28	14.70	12.5	87.5	815
Learning and Literacy score, Avg. 12m 36m	q^r	2.87	0.99	0.17	4.53	815
Stimulation of Development ORCE, predicted	q^n	3.92	1.00	0	6.27	815
Predicted Wage Offer, US$ of 2012 per hour	w	8.78	5.78	0.39	23.69	815
Hours per week of working time	L	16.85	16.62	0	57	815
Pre-natal Investment	I_0	1.76	0.01	1.66	1.79	815
Endowment shock	ϕ	3.44	1.20	0.54	6.63	805

Table 3: Averages by treatment status and mother's education

	Maternal education level			
Treatment Group	Less than High School	High School graduate	Some College	College graduate
	h_0: birth weight (kgs)			
Treatment	1.85	1.81	1.78	1.77
Control	1.77	1.76	1.87	1.81
	t: hours per week of CDC care			
Treatment	17.96	19.34	17.97	17.33
Control	0	0	0	0
	r: hours per week of maternal care			
Treatment	53.83	52.55	49.73	53.43
Control	65.97	61.56	57.98	57.30
	n: hours per week of non-maternal, non-CDC care			
Treatment	15.71	15.61	19.80	16.74
Control	21.53	25.94	29.52	30.20
	q^r: Learning and Literacy score, Avg. 12m and 36m			
Treatment	2.51	2.93	3.43	3.84
Control	2.29	2.64	3.22	3.83

Table 4: Learning an Literacy components (IT-Home score) available in the IHDP sample

12-month Home Assessment	36-month Home Assessment
At least 10 books are present and visible	Child has toys which teach color, size, shape
Muscle activity toys or equipment	Child has three or more puzzles
Push or pull toys	Child has toys permitting free expression
Parent provides toys for child during visit	Child has toys or games requiring refined movements
Learning equipment appropriate to age: cuddly toys or role playing toys	Child has at least 10 children's books
Learning facilitators: mobile, table and chairs, high chair, play pen	At least 10 books are visible in the apartment
Complex eye-hand coordination toys	Child is encouraged to learn the alphabet
Toys for literature and music	Interior of apartment not dark or perceptually monotonous
Parent reads stories to child at least 3 times weekly	Parent converses with child at least twice during visit
Child has 3 or more books of her own	Child is encouraged to learn spatial relationships
	Child is encouraged to learn to read a few words
	Child has real or toy musical instrument

Based on Linver, Martin and Brooks-Gunn (2004) and Fuligni, Han and Brooks-Gunn (2004).

Table 5: Shifters and sources of unobserved individual heterogeneity in the model

Unobserved heterogeneous variable	Shifter	Maternal age	Number of children < 5	Pre-natal Investment (I_0^*)	Maternal IQ	Total parenting experience
$ln(\gamma_l)$	$Z^l\kappa^l$	X				
$ln(\gamma_t)$	$Z^t\kappa^t$		X			
$ln(\gamma_p)$	$Z^p\kappa^p$			X		
$ln(\gamma_h)$	$Z^h\kappa^h$			X		
$ln(\omega)$	$Z^\omega\kappa^\omega$				X	
$ln(Q)$	$Z^Q\kappa^Q$					X
$ln(\varepsilon)$	$Z^\varepsilon\kappa^\varepsilon$				X	

Table 6: Model estimates

	MLE	NLLS
α_1	0.223***	0.067
	(0.0408)	(0.0714)
α_2	0.268***	0.211***
	(0.0492)	(0.0650)
α_5	-0.037***	0.001
	(0.0068)	(0.0139)
β_0	2.865***	3.046***
	(0.2761)	(0.3459)
β_2	-1.616***	-0.119
	(0.2654)	(0.2986)
β_5	1.109***	0.115
	(0.3427)	(0.2231)
κ^l	0.217***	
	(0.0031)	
κ^t	2.523***	
	(0.0715)	
κ^p	-1.954***	
	(0.0238)	
κ^h	3.845***	
	(0.0433)	
κ^ω	-0.276***	
	(0.0238)	
κ^Q	-0.007***	
	(0.0028)	
κ^ε	0.169***	
	(0.0102)	
χ_m	0.511***	
	(0.0077)	
σ_{q^r}	0.455***	
	(0.0117)	
$\sigma_{(\nu^\varepsilon+\nu^h)}$	0.925***	
	(0.0231)	
$\sigma_{(\nu^p-\nu^\varepsilon-\nu^h)}$	0.500***	
	(0.0127)	
$\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}$	0.589***	
	(0.0154)	
$\sigma_{(\nu^t-\nu^h-\nu^\varepsilon)}$	1.960***	
	(0.0844)	
$\sigma_{(\nu^l-\nu^\omega)}$	1.209***	
	(0.0300)	
N	815	815

MLE: Maximum Likelihood Estimation. NLLS: Non-linear Least Squares
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: First order marginal effects by sub-group

	Maternal education level			
Treatment Group	Less than High School	High School graduate	Some College	College graduate
Marginal effect of birth weight on logarithm of age-3 IQ $(\partial ln(h)/\partial h_0)$				
Treatment	0.04156	0.04957	0.05718	0.06482
	(0.0228)	(0.0233)	(0.0238)	(0.0241)
Control	0.03128	0.04074	0.05125	0.06324
	(0.0239)	(0.0239)	(0.0226)	(0.0235)
Marginal effect of non-maternal care time on logarithm of age-3 IQ $(\partial ln(h)/\partial n)$				
Treatment	0.00145	0.00122	0.00112	0.00099
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Control	0.00279	0.00213	0.00166	0.00126
	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Marginal effect of CDC time on logarithm of age-3 IQ $(\partial ln(h)/\partial t)$				
Treatment	0.00255	0.00211	0.00170	0.00143
	(0.0002)	(0.0002)	(0.0001)	(0.0001)
Control	0.00474	0.00357	0.00238	0.00165
	(0.0004)	(0.0003)	(0.0002)	(0.0001)
Marginal effect of maternal care time on logarithm of age-3 IQ $(\partial ln(h)/\partial r)$				
Treatment	0.00243	0.00240	0.00239	0.00229
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Control	0.00262	0.00260	0.00256	0.00242
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Marginal effect of quality of maternal care on logarithm of age-3 IQ $(\partial ln(h)/\partial q^r)$				
Treatment	0.05201	0.04300	0.03468	0.03192
	(0.0045)	(0.0038)	(0.0031)	(0.0029)
Control	0.07558	0.06061	0.04622	0.03622
	(0.0065)	(0.0053)	(0.0039)	(0.0031)
Marginal effect of total investment on logarithm of age-3 IQ $(\partial ln(h)/\partial \ln(I))$				
Treatment	1.67971	1.66096	1.63899	1.63301
	(0.2101)	(0.2043)	(0.1975)	(0.1956)
Control	1.63309	1.62844	1.69402	1.65745
	(0.1957)	(0.1942)	(0.2145)	(0.2032)

Standard errors in parentheses

Table 8: Effect of treatment on post-natal investments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variable	Total post-natal investment	Effective Maternal care	Quality of maternal care	Maternal care time	Effective Non-maternal care	Quality of non-maternal, non-CDC care	Non-maternal, non-CDC care time
VARIABLES	ln(I)	$\ln(q^r r)$	q^r	r	$\ln(q^n n + q^t t)$	q^n	n
Treatment indicator	0.0221***	-0.0715**	0.215***	-9.438***	0.847***	-0.252***	-8.776***
	(0.00269)	(0.0229)	(0.0310)	(0.879)	(0.0642)	(0.0345)	(1.071)
Pre-natal Investment	0.379**	1.532	5.280**	-41.76	4.724*	9.953***	43.09
	(0.116)	(1.160)	(1.981)	(42.54)	(2.031)	(1.292)	(29.82)
Endowment	0.00579***	-0.00366	-0.00859	-0.213	0.164***	0.543***	-0.0713
	(0.00138)	(0.0103)	(0.0268)	(0.442)	(0.0342)	(0.0185)	(0.475)
Maternal Age	-0.00271**	0.00123	0.000130	0.512	-0.0822*	-0.0194	-0.624
	(0.00108)	(0.0171)	(0.0282)	(0.714)	(0.0371)	(0.0195)	(0.574)
Maternal age, squared	4.11e-05**	-8.85e-05	-1.20e-05	-0.0114	0.00138*	0.000261	0.0140
	(1.56e-05)	(0.000335)	(0.000537)	(0.0130)	(0.000620)	(0.000370)	(0.0108)
Number of own children under age 5	-0.0152**	-0.0494**	-0.178***	2.000**	-0.260**	-0.448***	-1.940*
	(0.00483)	(0.0204)	(0.0457)	(0.769)	(0.0839)	(0.0636)	(0.894)
Age of the youngest own child	-0.00237	0.00324	0.00420	0.297	-0.0493	-0.248***	-0.686
	(0.00261)	(0.0197)	(0.0377)	(0.851)	(0.0742)	(0.0310)	(1.010)
Maternal education: Less than High School	-0.0161**	-0.0619	-0.245**	3.464*	-0.258***	-0.00572	-2.879*
	(0.00579)	(0.0376)	(0.0978)	(1.479)	(0.0665)	(0.0647)	(1.435)
Maternal education: Some College	0.0188***	0.107**	0.418***	-2.964*	0.180*	0.370***	3.233*
	(0.00230)	(0.0323)	(0.0580)	(1.455)	(0.0790)	(0.0403)	(1.437)
Maternal education: College graduate	0.0213***	0.157***	0.593***	-2.597	0.339***	0.472***	2.522
	(0.00396)	(0.0351)	(0.107)	(2.149)	(0.0723)	(0.0799)	(2.471)
Race: African American	-0.0345***	-0.357***	-0.792***	-4.171**	0.00730	-0.790***	3.199**
	(0.00485)	(0.0370)	(0.0845)	(1.366)	(0.0923)	(0.0313)	(1.040)
Race: Hispanic	-0.0187***	-0.258***	-0.559***	-4.545***	0.153	-0.564***	4.458*
	(0.00338)	(0.0168)	(0.0736)	(1.231)	(0.0883)	(0.0955)	(1.900)

Race: Other	-0.0232***	-0.261**	-0.592**	-2.447	0.00518	-0.00203	2.143
	(0.00517)	(0.0905)	(0.173)	(2.402)	(0.131)	(0.0636)	(2.245)
Marital status: Single	-0.0133**	-0.0689*	-0.140*	1.031	-0.156*	-0.308***	-1.376
	(0.00380)	(0.0353)	(0.0679)	(1.096)	(0.0810)	(0.0447)	(1.046)
Marital status: Sep./Div./Wid.	-0.00788	-0.0245	-0.0558	-0.0269	-0.169	-0.474***	-0.156
	(0.00487)	(0.0486)	(0.0997)	(2.858)	(0.200)	(0.0495)	(2.765)
Site Name: ARK	0.00680*	0.112***	0.297***	-1.196**	0.00351	-0.00365	-0.101
	(0.00348)	(0.0187)	(0.0457)	(0.368)	(0.0389)	(0.0328)	(0.385)
Site Name: EIN	0.0316***	0.379***	0.633***	4.060***	-0.234***	0.254***	-4.463***
	(0.00288)	(0.0224)	(0.0518)	(0.692)	(0.0574)	(0.0175)	(0.726)
Site Name: HAR	0.0149**	0.167***	0.394***	-0.341	-0.0513	0.0952*	0.0747
	(0.00543)	(0.0332)	(0.0754)	(0.456)	(0.0527)	(0.0484)	(0.534)
Site Name: PEN	0.0318***	0.247***	0.730***	-5.501***	0.276***	0.0648*	5.264***
	(0.00135)	(0.00988)	(0.0291)	(0.695)	(0.0355)	(0.0297)	(0.699)
Site Name: TEX	0.0243***	0.262***	0.476***	0.879*	-0.0481*	0.130***	-1.744***
	(0.000904)	(0.0104)	(0.0225)	(0.396)	(0.0221)	(0.0120)	(0.331)
Site Name: WAS	0.0229***	0.218***	0.535***	-2.103***	0.0657	0.0730	1.500**
	(0.00553)	(0.0378)	(0.0872)	(0.510)	(0.0535)	(0.0462)	(0.479)
Site Name: YAL	0.0177***	0.266***	0.455***	2.818***	-0.184***	0.0114	-3.159***
	(0.00471)	(0.0313)	(0.0696)	(0.381)	(0.0415)	(0.0425)	(0.499)
Constant	0.897***	2.476	-6.193	129.5*	-2.925	-13.57***	-40.81
	(0.198)	(2.089)	(3.477)	(64.78)	(3.140)	(2.342)	(46.64)
Observations	805	805	805	805	805	805	805
R-squared	0.381	0.280	0.483	0.174	0.258	0.738	0.178

Robust standard errors by site in parentheses. The excluded maternal education category is “High School graduates”. The excluded race and ethnicity category is “Non-Hispanic Whites”. The excluded marital status category is “Married”. The excluded site category is “Miami (MIA)”.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Effect of treatment on post-natal investments by treatment group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable	Total post-natal investment	Effective Maternal care	Quality of maternal care	Maternal care time	Effective Non-maternal care	Quality of non-maternal, non-CDC care	Non-maternal, non-CDC care time	CDC care time
VARIABLES	ln(I)	$\ln(q^r r)$	q^r	r	$\ln(q^n n + q^t t)$	q^n	n	t
				Subsample: treatment group				
Pre-natal Investment	0.0779	-1.010	1.939	-78.81	3.559	5.428***	59.24	19.57
	(0.122)	(1.593)	(3.047)	(66.92)	(3.446)	(1.425)	(33.95)	(48.33)
Endowment	0.00192	-0.0101	-0.0134	-0.317	0.0879**	0.486***	-0.440	0.757*
	(0.00122)	(0.0160)	(0.0394)	(1.126)	(0.0353)	(0.0134)	(1.048)	(0.361)
Observations	310	310	310	310	310	310	310	310
R-squared	0.354	0.316	0.476	0.096	0.121	0.722	0.099	0.065
				Subsample: control group				
Pre-natal Investment	0.626***	3.461**	8.102***	-15.85	5.924	13.33***	15.85	-
	(0.163)	(1.031)	(1.751)	(44.00)	(3.233)	(1.763)	(44.00)	-
Endowment	0.00679**	-0.00460	-0.0178	-0.0684	0.193***	0.570***	0.0684	-
	(0.00211)	(0.0162)	(0.0256)	(0.544)	(0.0446)	(0.0259)	(0.544)	-
Observations	495	495	495	495	495	495	495	-
R-squared	0.399	0.271	0.505	0.119	0.168	0.753	0.119	-

Robust standard errors by site in parentheses. Maternal controls used in Table 8 are included but not reported (maternal age, number of children under age 5, age of youngest child, maternal education, race and ethnicity, marital status and site).
*** p<0.01, ** p<0.05, * p<0.1

Table 10: Robustness analysis, model estimates based on subsamples

	MLE: Full sample	MLE: Single mothers	MLE: Gestational age > 30 weeks
α_1	0.223***	0.119***	0.277***
	(0.0408)	(0.0311)	(0.0712)
α_2	0.268***	0.142***	0.330***
	(0.0492)	(0.0371)	(0.0847)
α_5	-0.037***	-0.019***	-0.0461***
	(0.0068)	(0.0051)	(0.0119)
β_0	2.865***	3.529***	2.605***
	(0.2761)	(0.2100)	(0.4803)
β_2	-1.616***	-1.078***	-1.267***
	(0.2654)	(0.2590)	(0.4204)
β_5	1.109***	1.423**	0.672*
	(0.3427)	(0.5680)	(0.3816)
κ^l	0.217***	0.226***	0.220***
	(0.0031)	(0.0047)	(0.0034)
κ^t	2.523***	2.481***	2.507***
	(0.0715)	(0.0895)	(0.0804)
κ^p	-1.954***	-2.078***	-1.949***
	(0.0238)	(0.0363)	(0.0264)
κ^h	3.845***	4.057***	3.883***
	(0.0433)	(0.0793)	(0.0476)
κ^ω	-0.276***	-0.309***	-0.260***
	(0.0238)	(0.0419)	(0.0263)
κ^Q	-0.007***	-0.006	-0.008***
	(0.0028)	(0.0040)	(0.0031)
κ^ε	0.169***	0.097***	0.169***
	(0.0102)	(0.0198)	(0.0112)
χ_m	0.511***	0.539***	0.516***
	(0.0077)	(0.0160)	(0.0087)
σ_{q^r}	0.455***	0.549***	0.458***
	(0.0117)	(0.0198)	(0.0132)
$\sigma_{(\nu^\varepsilon+\nu^h)}$	0.925***	1.006***	0.909***
	(0.0231)	(0.0347)	(0.0253)
$\sigma_{(\nu^p-\nu^\varepsilon-\nu^h)}$	0.500***	0.548***	0.494***
	(0.0127)	(0.0191)	(0.0141)
$\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}$	0.589***	0.610***	0.589***
	(0.0154)	(0.0229)	(0.0171)
$\sigma_{(\nu^t-\nu^h-\nu^\varepsilon)}$	1.960***	1.783***	1.998***
	(0.0844)	(0.0998)	(0.0954)
$\sigma_{(\nu^l-\nu^\omega)}$	1.209***	1.229***	1.208***
	(0.0300)	(0.0422)	(0.0334)
N	815	428	653

MLE: Maximum Likelihood Estimation.
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

VIII. Figures

Figure 1 - MLE First-order marginal effects (h: Age 3 IQ)

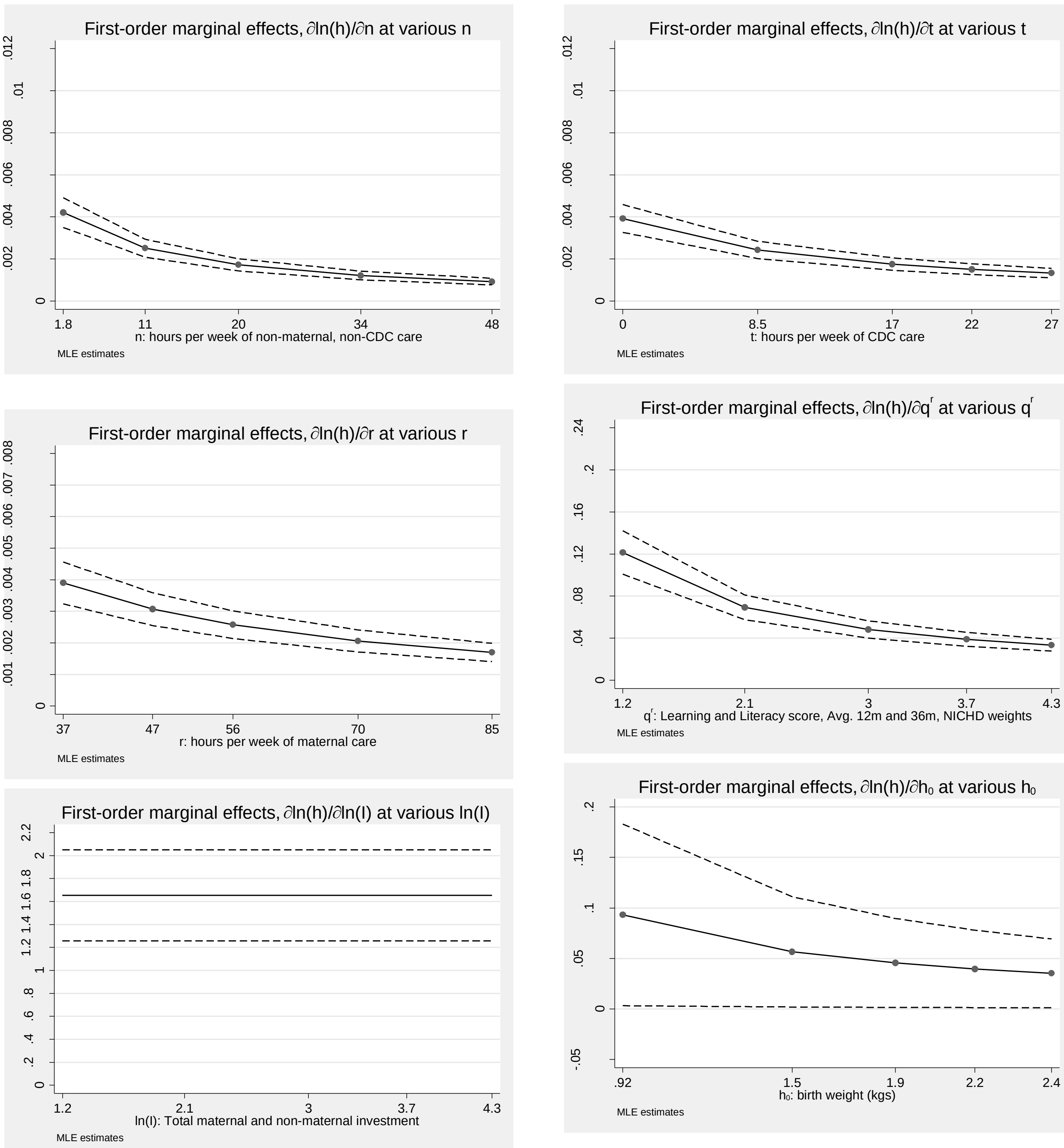
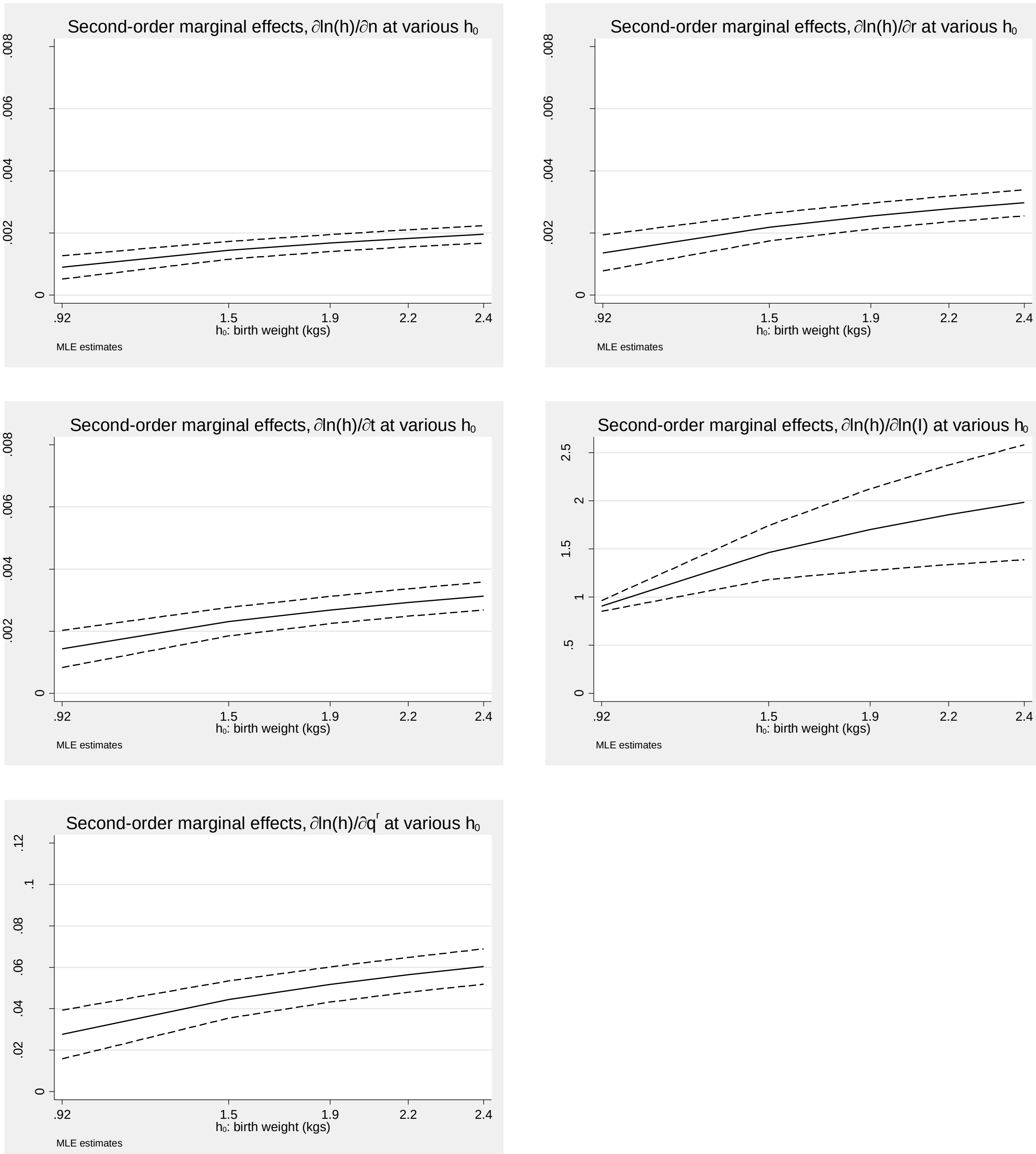


Figure 2 - MLE Second-order marginal effects at different levels of h_0



IX. Appendixes

a) Appendix: variable dictionary for the post-natal parental problem

Decision variables:

- Parental consumption: c
- Maternal care time investment: r
- Hours of care at CDC: t
- Other sources of care (non-maternal and non-CDC, in time units): n
- Quality of other sources of care (per unit of time): q^n
- Maternal effort, which translates into quality of the maternal time investment: e
- Maternal leisure time: l
- Maternal working time: L

Parameters:

- Time endowment, for children and parents: T_c, T_p
- Human capital of the parent: m
- Unobserved individual heterogeneity in ability: ω
- Labor market wage, an increasing function of human capital and ability: $w(m, \omega)$
- Non-labor income: Y
- Price of other sources of care, per unit of effective care: π
- Human capital at birth: h_0
- Maximum number of hours at the CDC: $\bar{\tau}$

Treatment variables:

- For families in the control group, $\bar{\tau} = 0$. For families in the treatment group, $\bar{\tau} > 0$
- Quality of CDC services: $q^t > 0$

Child's human capital technology depends on:

- Total effective units of non-maternal care: $q^n n \; + \; q^t t$
- Total effective units of maternal care: $q^r(e, m, \omega) * r$
- Human capital at birth: h_0

c) Appendix: Kuhn-Tucker conditions for the post-natal parental problem

$$\frac{\partial \mathcal{L}}{\partial \lambda} = w\left[T_p - T_c\right] + wt + Y - c - [\pi q^n - w]n - wl = 0 \qquad \lambda \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial \mu} = \bar{\tau} - t \geq 0 \qquad \frac{\partial \mathcal{L}}{\partial \mu}\mu = 0 \qquad \mu \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial c} = U_c - \lambda \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial c}c = 0 \qquad c \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial q^n} = U_h f_1 n - \lambda \pi n \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial q^n}q^n = 0 \qquad q^n \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial e} = U_p r + U_h f_2 q_e^r r \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial e}e = 0 \qquad e \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial n} = U_h[f_1 q^n - f_2 q^r] - U_p e - \lambda[\pi q^n - w] \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial n}n = 0 \qquad n \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial l} = U_l - \lambda w \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial l}l = 0 \qquad l \geq 0$$

$$\frac{\partial \mathcal{L}}{\partial t} = U_h[f_1 q^t - f_2 q^r] + U_t - U_p e + \lambda w - \mu \leq 0 \qquad \frac{\partial \mathcal{L}}{\partial t}[t - \bar{\tau}] = 0 \qquad 0 \leq t \leq \bar{\tau}$$

d) Appendix: guide to connection between theory and IHDP data

				Year 1			Year 2		Year 3	
Variable in the model	**Symbol**	**Variable in IHDP data**	**Variable name**	**4m**	**8m**	**12 m**	**18 m**	**24 m**	**30 m**	**36m**
Parental care time investment	r	Mother is the PCG	mother_pcg	x	x	x	x	x	x	x
		Hours of maternal care, weekdays (M-F)	hxw_mother _weekdays	-	-	-	x	-	x	-
		Hours of maternal care, weekends (S-S)	hxw_mother _weekend	-	-	-	x	-	x	-
Quality of parental care	$q^r(e,m)$	Parental Warmth (standarized score)	Warmth	-	-	x	-	-	-	x
		Learning and Literacy (standarized score)	Learn_Lit	-	-	x	-	-	-	x
Hours of daycare (free or purchased)	$n+\tau$	Daycare is the PCG	daycare_pcg	x	x	x	x	x	x	x
		Hours per week in daycare, primary or secondary	total_hxw_d care	-	-	-	x	x	x	x
		PCG daycare refers to the CDC?	dayc_36m_C DC_IHDP	-	-	-	-	-	-	x: "Yes" for 97% of treatme nt group.
Parental working time	L	Maternal employment status	employed	x	x	x	x	x	x	x
		Full time vs. Part time employment	full_time	-	-	-	x	x	x	x
		Hours worked per week	hours_x_wee k_worked	-	-	-	x	x	x	x
Total family income	$w(m)L + Y$	Family income, annual	f18v61(1985) f34v64(1986) f52v96(1987)	-	-	x (85)	-	x (86)	-	x (87)
Child's human capital	h	Bayley Mental Scale	bayley	-	-	x	-	x	-	-
		Stan. Bin. IQ	iq	-	-	-	-	-	-	x
Child's human capital at	h_0	Birth Weight (kgm)	bw							
		Head Circumf. (cms)	f3av5							

birth		Birth Length (cms)	f3av4			
		Gest. Age(weeks)	anga			
				Year 1	**Year 2**	**Year 3**
Hours of CDC	t	Number of Days	cdays	-	x	x
		Avg. Number of Hours	chour	-	x	x
		Avg. hours per week	hxw_cdc	-	x	x
		No use of CDC	no_cdc	-	x	x

PCG: Primary Care Giver. SCG: Secondary Care Giver. CDC: Child Development Center (free daycare). **Note:** No data for consumption (c) and leisure (l). Looking for data to capture quality and price of other sources of care (q^n, π) and quality of CDC services (q^t).

e) Appendix: Maximum likelihood estimation

The likelihood of the parameters can be decomposed using Bayes' rule.

$$L(\Theta\,;\, l,t,n,e,q^n) = L(\Theta\,; l \mid t,n,e,q^n) * L(\Theta\,; t \mid n,e,q^n) * L(\Theta\,; n \mid e,q^n) * L(\Theta\,; e \mid q^n) * L(\Theta\,;\, q^n)$$

Taking the log yields:

$$\ell(\Theta\,;\, l,t,n,e,q^n) = \ell(\Theta\,; l \mid t,n,e,q^n) + \ell(\Theta\,; t \mid n,e,q^n) + \ell(\Theta\,; n \mid e,q^n) + \ell(\Theta\,; e \mid q^n) + \ell(\Theta;\, q^n)$$

Each of the components of the log-likelihood will be derived from the first order conditions (equations A, B, C, D and E), combined with the wage equation (equation F), the technology of maternal care (equation G), the functional forms for the marginal rates of substitution (equations H, I, J and K), the shifter equations (equation L), the functional form for human capital technology (equations M and N) and the distributional assumptions (equation O).

From the FOC for quality of non-maternal care, equation D:

$$\ell(\Theta;\, q^n) = ln\left[\phi\left(\frac{ln(\pi) - ln(f_1) - \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} - \mathrm{Z}^{h}\kappa^{h} + ln[h]}{\sigma_{(\nu^{\varepsilon}+\nu^{h})}}\right)\right] - ln\left[\sigma_{(\nu^{\varepsilon}+\nu^{h})}\right] \qquad \text{(Equation E1)}$$

From the FOC for instantaneous parenting effort, equation E:

$$\ell(\Theta\,; e \mid q^n) = ln\left[\phi\left(\frac{ln(f_2) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + \mathrm{Z}^{h}\kappa^{h} - \mathrm{Z}^{p}\kappa^{p} - ln[h] + ln[\bar{p} - p]}{\sigma_{(\nu^{p}-\nu^{\varepsilon}-\nu^{h})}}\right)\right] - ln\left[\sigma_{(\nu^{p}-\nu^{\varepsilon}-\nu^{h})}\right] \qquad \text{(Equation E2)}$$

From the FOC for non-maternal, non-CDC care time, equation C: there are 2 cases.

Case a) $n^* > 0$:

$\ell(\Theta\,; n \mid n > 0, e, q^n)$

$$= ln\left[\phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^{\varepsilon}\kappa^{\varepsilon} - Z^{h}\kappa^{h} - Z^{Q}\kappa^{Q} + \chi_m Z^{\omega}\kappa^{\omega} + ln[h]}{\sigma_{(\nu^{\varepsilon}+\nu^{h}+\nu^{Q}-\chi_m\nu^{\omega})}}\right)\right]$$

$$-ln\left[\sigma_{(\nu^{\varepsilon}+\nu^{h}+\nu^{Q}-\chi_m\nu^{\omega})}\right] \qquad \text{(Equation E3)}$$

Case b) $n^* = 0$:

$\ell(\Theta\,; n \mid n = 0, e, q^n)$

$$= ln\left[1 - \Phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^{\varepsilon}\kappa^{\varepsilon} - Z^{h}\kappa^{h} - Z^{Q}\kappa^{Q} + \chi_m Z^{\omega}\kappa^{\omega} + ln[h]}{\sigma_{(\nu^{\varepsilon}+\nu^{h}+\nu^{Q}-\chi_m\nu^{\omega})}}\right)\right] \qquad \text{(Equation E4)}$$

From the FOC for CDC care time, equation B, there are four cases to consider: a) $n^* > 0\,;\ t^* > 0$; b) $n^* > 0\,;\ t^* = 0$; c) $n^* = 0\,;\ t^* > 0$; and d) $n^* = 0\,;\ t^* = 0$.

Case a) $n^* > 0\,;\ t^* > 0$:

$\ell(\Theta\,; t \mid n > 0, t > 0, e, q^n)$

$$= ln\left[\phi\left(\frac{ln(f_1) + ln(q^t) + Z^{h}\kappa^{h} + Z^{\varepsilon}\kappa^{\varepsilon} - Z^{t}\kappa^{t} - ln[h] + ln[\bar{\tau} - t]}{\sigma_{(\nu^{t}-\nu^{h}-\nu^{\varepsilon})}}\right)\right]$$

$$-ln\left[\sigma_{(\nu^{t}-\nu^{h}-\nu^{\varepsilon})}\right] \qquad \text{(Equation E5)}$$

Case b) $n^* > 0\,;\ t^* = 0$:

$\ell(\Theta\,; t \mid n > 0, t = 0, e, q^n)$

$$= ln\left[1 - \Phi\left(\frac{ln(f_1) + ln(q^t) + Z^{h}\kappa^{h} + Z^{\varepsilon}\kappa^{\varepsilon} - Z^{t}\kappa^{t} - ln[h] + ln[\bar{\tau} - t]}{\sigma_{(\nu^{t}-\nu^{h}-\nu^{\varepsilon})}}\right)\right] \qquad \text{(Equation E6)}$$

Case c) $n^* = 0\,;\ t^* > 0$.[9]

$\ell(\Theta\,; t \mid n = 0, t > 0, e, q^n)$

$$\approx ln\left[\phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^{h}\kappa^{h} - Z^{\varepsilon}\kappa^{\varepsilon} - Z^{Q}\kappa^{Q} + \chi_m Z^{\omega}\kappa^{\omega} + ln[h]}{\sigma_{(\nu^{h}+\nu^{\varepsilon}+\nu^{Q}-\chi_m\nu^{\omega})}}\right)\right]$$

[9] We make an additional assumption here, which is discussed in the appendix: $\pi q^t \approx -MRS_{t,c}$

$$-ln\left[\sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}\right] \qquad \text{(Equation E7)}$$

Case d) $n^* = 0\ ;\ t^* = 0$:

$\ell(\Theta\,;t\mid n=0, t=0, e, q^n)$

$$= ln\left[1-\Phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)]-ln(f_2)-Z^h\kappa^h-Z^\varepsilon\kappa^\varepsilon-Z^Q\kappa^Q+\chi_m Z^\omega\kappa^\omega+ln[h]}{\sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}}\right)\right] \qquad \text{(Equation E8)}$$

Put together all four cases, we obtain the log-likehood

$$\begin{aligned}\ell(\Theta\,;t\mid n,e,q^n) \\ &= \mathbb{1}[n>0,t>0,\bar{\tau}>0]*\ell(\Theta\,;t\mid n>0,t>0,e,q^n) \\ &+ \mathbb{1}[n>0,t=0,\bar{\tau}>0]*\ell(\Theta\,;t\mid n>0,t=0,e,q^n) \\ &+ \mathbb{1}[n=0,t>0,\bar{\tau}>0]*\ell(\Theta\,;t\mid n=0,t>0,e,q^n) \\ &+ \mathbb{1}[n=0,t=0,\bar{\tau}>0]*\ell(\Theta\,;t\mid n=0,t=0,e,q^n)\end{aligned}$$

From the FOC for leisure time, equation A:

$\ell(\Theta\,;l\mid t,n,e,q^n)$

$$= ln\left[\phi\left(\frac{ln[\widehat{w}(m)]+Z^\omega\kappa^\omega-Z^l\kappa^l+ln[l]}{\sigma_{(\nu^l-\nu^\omega)}}\right)\right]-ln\left[\sigma_{(\nu^l-\nu^\omega)}\right] \qquad \text{(Equation E9)}$$

And the following abbreviations were used,

$$\sigma_{(\nu^\varepsilon+\nu^h)} \equiv \sqrt{\sigma_\varepsilon^2+\sigma_h^2+2\sigma_{\varepsilon,h}}$$

$$\sigma_{(\nu^p-\nu^\varepsilon-\nu^h)} \equiv \sqrt{\sigma_p^2+\sigma_\varepsilon^2+\sigma_h^2-2\sigma_{p,\varepsilon}-2\sigma_{p,h}+2\sigma_{\varepsilon,h}}$$

$$\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)} \equiv$$
$$\sqrt{\sigma_\varepsilon^2+\sigma_h^2+\sigma_Q^2+\chi_m^2\sigma_\omega^2+2\sigma_{\varepsilon,h}+2\sigma_{\varepsilon,Q}+2\sigma_{h,Q}-2\chi_m\sigma_{\varepsilon,\omega}-2\chi_m\sigma_{h,\omega}-2\chi_m\sigma_{Q,\omega}}$$

$$\sigma_{(\nu^t-\nu^h-\nu^\varepsilon)} = \sqrt{\sigma_t^2+\sigma_h^2+\sigma_\varepsilon^2-2\sigma_{t,h}-2\sigma_{t,\varepsilon}+2\sigma_{h,\varepsilon}}$$

$$\sigma_{(\nu^l-\nu^\omega)} = \sqrt{\sigma_l^2+\sigma_\omega^2-2\sigma_{l,\omega}}$$

$\ell(\Theta;\ q^n)$: this log-likelihood will come from the first order condition for quality of non-maternal care (Equation D).

$$f_1\ MRS_{h,c} = \pi$$

$$\Rightarrow ln(f_1) + ln\left(MRS_{h,c}\right) = ln(\pi)$$

$$\Rightarrow ln(f_1) + ln(\varepsilon) + ln[\gamma_h] - ln[h] = ln(\pi)$$

$$\Rightarrow ln(f_1) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + \nu^{\varepsilon} + \mathrm{Z}^{h}\kappa^{h} + \nu^{h} - ln[h] = ln(\pi)$$

$$\Rightarrow \nu^{\varepsilon} + \nu^{h} = ln(\pi) - ln(f_1) - \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} - \mathrm{Z}^{h}\kappa^{h} + ln[h]$$

Using the distributional assumptions, $\nu^{\varepsilon} + \nu^{h} \sim \mathcal{N}\left(0, \sigma_{\varepsilon}^2 + \sigma_h^2 + 2\sigma_{\varepsilon,h}\right)$. Therefore,

$$\Rightarrow \ell(\Theta;\ q^n) = ln\left[\phi\left(\frac{ln(\pi) - ln(f_1) - \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} - \mathrm{Z}^{h}\kappa^{h} + ln[h]}{\sigma_{(\nu^{\varepsilon}+\nu^{h})}}\right)\right] - ln\left[\sigma_{(\nu^{\varepsilon}+\nu^{h})}\right]$$

where $\sigma_{(\nu^{\varepsilon}+\nu^{h})} \equiv \sqrt{\sigma_{\varepsilon}^2 + \sigma_h^2 + 2\sigma_{\varepsilon,h}}$

(Equation E1)

$\ell(\Theta\,; e \mid q^n)$: this log-likelihood will be based on the first order condition for optimal maternal effort (Equation E).

$$f_2\ q_e^r\ MRS_{h,c} = -MRS_{p,c}$$

$$\Rightarrow ln(f_2) + ln(q_e^r) + ln\left(MRS_{h,c}\right) = ln\left(-MRS_{p,c}\right)$$

$$\Rightarrow ln(f_2) + ln(\varepsilon) + ln(q_e^r) + ln\left(MRS_{h,c}\right) = ln\left(-MRS_{p,c}\right)$$

$$\Rightarrow ln(f_2) + ln(\varepsilon) + ln(q_e^r) + ln[\gamma_h] - ln[h] = ln\left[\gamma_p\right] - ln[\bar{p} - p]$$

$$\Rightarrow ln(f_2) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + \nu^{\varepsilon} + ln(q_e^r) + \mathrm{Z}^{h}\kappa^{h} + \nu^{h} - ln[h] = \mathrm{Z}^{p}\kappa^{p} + \nu^{p} - ln[\bar{p} - p]$$

$$\Rightarrow ln(f_2) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + ln(q_e^r) + \mathrm{Z}^{h}\kappa^{h} - ln[h] - \mathrm{Z}^{p}\kappa^{p} + ln[\bar{p} - p] = \nu^{p} - \nu^{\varepsilon} - \nu^{h}$$

$$Var(\nu^{p} - \nu^{\varepsilon} - \nu^{h}) = Var(\nu^{p} - \nu^{\varepsilon}) + Var(\nu^{h}) - 2Cov(\nu^{p} - \nu^{\varepsilon}, \nu^{h})$$

$$= Var(\nu^p) + Var(\nu^\varepsilon) - 2Cov(\nu^p, \nu^\varepsilon) + Var(\nu^h) - 2Cov(\nu^p - \nu^\varepsilon, \nu^h)$$

$$= Var(\nu^p) + Var(\nu^\varepsilon) - 2Cov(\nu^p, \nu^\varepsilon) + Var(\nu^h) - 2Cov(\nu^p, \nu^h) + 2Cov(\nu^\varepsilon, \nu^h)$$

$$= \sigma_p^2 + \sigma_\varepsilon^2 + \sigma_h^2 - 2\sigma_{p,\varepsilon} - 2\sigma_{p,h} + 2\sigma_{\varepsilon,h}$$

Then, $\nu^p - \nu^\varepsilon - \nu^h \sim \mathcal{N}\left(0, \sigma_p^2 + \sigma_\varepsilon^2 + \sigma_h^2 - 2\sigma_{p,\varepsilon} - 2\sigma_{p,h} + 2\sigma_{\varepsilon,h}\right)$

Finally, $ln(q_e^r) = 0$ under the functional form assumption for q^r. Thus,

$$ln(f_2) + \mathrm{Z}^\varepsilon\kappa^\varepsilon + \mathrm{Z}^h\kappa^h - \mathrm{Z}^p\kappa^p - ln[h] + ln[\bar{p} - p] \sim \mathcal{N}\left(0, \sigma_p^2 + \sigma_\varepsilon^2 + \sigma_h^2 - 2\sigma_{p,\varepsilon} - 2\sigma_{p,h} + 2\sigma_{\varepsilon,h}\right)$$

$$\Rightarrow \ell(\Theta\,; e \mid q^n) = ln\left[\phi\left(\frac{ln(f_2) + \mathrm{Z}^\varepsilon\kappa^\varepsilon + \mathrm{Z}^h\kappa^h - \mathrm{Z}^p\kappa^p - ln[h] + ln[\bar{p} - p]}{\sigma_{(\nu^p - \nu^\varepsilon - \nu^h)}}\right)\right] - ln\left[\sigma_{(\nu^p - \nu^\varepsilon - \nu^h)}\right]$$

where $\sigma_{(\nu^p - \nu^\varepsilon - \nu^h)} \equiv \sqrt{\sigma_p^2 + \sigma_\varepsilon^2 + \sigma_h^2 - 2\sigma_{p,\varepsilon} - 2\sigma_{p,h} + 2\sigma_{\varepsilon,h}}$

(Equation E2)

$\ell(\Theta\,; n \mid e, q^n)$: the first order condition that defines optimal care time from non-maternal non-CDC caregivers (n), Equation C, is the starting point for this likelihood. We can simplify this equation by using Equations D and E:

$$MRS_{h,c}[f_1 q^n - f_2 q^r] + w(m, \omega) - MRS_{p,c}\, e \le \pi q^n$$

$$\Rightarrow MRS_{h,c} f_1 q^n - MRS_{h,c} f_2 q^r + w(m, \omega) - MRS_{p,c}\, e \le \pi q^n$$

$$\Rightarrow \pi q^n - MRS_{h,c} f_2 q^r + w(m, \omega) - MRS_{p,c}\, e \le \pi q^n \qquad \text{(Using Eq. D)}$$

$$\Rightarrow -MRS_{h,c} f_2 q^r + w(m, \omega) + f_2\, q_e^r\, MRS_{h,c}\, e \le 0 \qquad \text{(Using Eq. E)}$$

$$\Rightarrow f_2\, MRS_{h,c}[q_e^r\, e - q^r] + w(m, \omega) \le 0$$

$$\Rightarrow w(m, \omega) \le f_2\, MRS_{h,c}[q^r - q_e^r\, e] \qquad \text{(Equation F)}$$

We can simplify further by taking logs, using the wage equation, the functional form for q^r and the expression for the marginal rate of substitution between human capital and consumption:

$$\Rightarrow ln[w(m,\omega)] \ \leq ln(f_2) + ln\left(MRS_{h,c}\right) + ln(q^r - q_e^r\ e)$$

$$\Rightarrow ln[\widehat{w}(m)] + ln(\omega) \leq ln(f_2) + ln(\varepsilon) + ln\left(MRS_{h,c}\right) + ln(q^r - q_e^r\ e)$$

If $q^r(e,m,\omega) = Q\left[\widehat{w}^{\chi_m}\omega^{(1-\chi_m)}\right] + e$, then $ln(q^r - q_e^r e) = ln(Q) + \chi_m ln(\widehat{w}) + (1-\chi_m)ln(\omega)$. Therefore,

$$\begin{aligned} ln[\widehat{w}(m)] + ln(\omega) & \\ &\leq ln(f_2) + ln(\varepsilon) + ln\left(MRS_{h,c}\right) + ln(Q) + \chi_m ln(\widehat{w}) \\ &+ (1-\chi_m)ln(\omega) \end{aligned}$$

$$\Rightarrow (1-\chi_m)ln[\widehat{w}(m)] \leq ln(f_2) + ln(\varepsilon) + ln\left(MRS_{h,c}\right) + ln(Q) - \chi_m ln(\omega)$$

$$\begin{aligned} \Rightarrow (1-\chi_m)ln[\widehat{w}(m)] & \\ &\leq ln(f_2) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + \nu^{\varepsilon} + ln[\gamma_h] - ln[h] + \mathrm{Z}^{Q}\kappa^{Q} + \nu^{Q} \\ &- \chi_m[\mathrm{Z}^{\omega}\kappa^{\omega} + \nu^{\omega}] \end{aligned}$$

$$\begin{aligned} \Rightarrow (1-\chi_m)ln[\widehat{w}(m)] & \\ &\leq ln(f_2) + \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} + \nu^{\varepsilon} + \mathrm{Z}^{h}\kappa^{h} + \nu^{h} - ln[h] + \mathrm{Z}^{Q}\kappa^{Q} + \nu^{Q} \\ &- \chi_m\mathrm{Z}^{\omega}\kappa^{\omega} - \chi_m\nu^{\omega} \end{aligned}$$

$$\begin{aligned} \Rightarrow (1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - \mathrm{Z}^{\varepsilon}\kappa^{\varepsilon} - \mathrm{Z}^{h}\kappa^{h} - \mathrm{Z}^{Q}\kappa^{Q} + \chi_m\mathrm{Z}^{\omega}\kappa^{\omega} + ln[h] & \\ \leq \nu^{\varepsilon} + \nu^{h} + \nu^{Q} - \chi_m\nu^{\omega} & \end{aligned}$$

$$\begin{aligned} Var(\nu^{\varepsilon} + \nu^{h} + \nu^{Q} - \chi_m\nu^{\omega}) & \\ &= Var(\nu^{\varepsilon} + \nu^{h} + \nu^{Q}) + \chi_m^2 Var(\nu^{\omega}) \\ &- 2Cov(\nu^{\varepsilon} + \nu^{h} + \nu^{Q}, \chi_m\nu^{\omega}) \end{aligned}$$

$$\begin{aligned} = Var(\nu^{\varepsilon} + \nu^{h}) + Var(\nu^{Q}) + 2Cov(\nu^{\varepsilon} + \nu^{h}, \nu^{Q}) + \chi_m^2 Var(\nu^{\omega}) & \\ - 2Cov(\nu^{\varepsilon} + \nu^{h} + \nu^{Q}, \chi_m\nu^{\omega}) & \end{aligned}$$

$$\begin{aligned} = Var(\nu^{\varepsilon}) + Var(\nu^{h}) + Var(\nu^{Q}) + \chi_m^2 Var(\nu^{\omega}) + 2Cov(\nu^{\varepsilon}, \nu^{h}) & \\ + 2Cov(\nu^{\varepsilon}, \nu^{Q}) + 2Cov(\nu^{h}, \nu^{Q}) - 2\chi_m Cov(\nu^{\varepsilon}, \nu^{\omega}) & \\ - 2\chi_m Cov(\nu^{h}, \nu^{\omega}) - 2\chi_m Cov(\nu^{Q}, \nu^{\omega}) & \end{aligned}$$

$$= \sigma_{\varepsilon}^2 + \sigma_h^2 + \sigma_Q^2 + \chi_m^2\sigma_{\omega}^2 + 2\sigma_{\varepsilon,h} + 2\sigma_{\varepsilon,Q} + 2\sigma_{h,Q} - 2\chi_m\sigma_{\varepsilon,\omega} - 2\chi_m\sigma_{h,\omega} - 2\chi_m\sigma_{Q,\omega}$$

Let $\sqrt{\sigma_\varepsilon^2 + \sigma_h^2 + \sigma_Q^2 + \chi_m^2\sigma_\omega^2 + 2\sigma_{\varepsilon,h} + 2\sigma_{\varepsilon,Q} + 2\sigma_{h,Q} - 2\chi_m\sigma_{\varepsilon,\omega} - 2\chi_m\sigma_{h,\omega} - 2\chi_m\sigma_{Q,\omega}} = \sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}$

Thus, $\nu^\varepsilon + \nu^h + \nu^Q - \chi_m\nu^\omega \sim \mathcal{N}\left(0, \sigma^2_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}\right)$.

We now have to contemplate two cases, due to possible corner solutions: a) $n^* > 0$, b) $n^* = 0$.

Case a) If $n^* > 0$, then $(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^\varepsilon\kappa^\varepsilon - Z^h\kappa^h - Z^Q\kappa^Q + \chi_m Z^\omega\kappa^\omega + ln[h] = \nu^\varepsilon + \nu^h + \nu^Q - \chi_m\nu^\omega$

$$\Rightarrow (1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^\varepsilon\kappa^\varepsilon - Z^h\kappa^h - Z^Q\kappa^Q + \chi_m Z^\omega\kappa^\omega + ln[h] \sim \mathcal{N}\left(0, \sigma^2_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}\right)$$

Therefore,

$$\ell(\Theta\,; n \mid n > 0, e, q^n) = ln\left[\phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^\varepsilon\kappa^\varepsilon - Z^h\kappa^h - Z^Q\kappa^Q + \chi_m Z^\omega\kappa^\omega + ln[h]}{\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}}\right)\right] - ln\left[\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}\right]$$ (Equation E3)

Case b) If $n^* = 0$, $(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^\varepsilon\kappa^\varepsilon - Z^h\kappa^h - Z^Q\kappa^Q + \chi_m Z^\omega\kappa^\omega + ln[h] \leq \nu^\varepsilon + \nu^h + \nu^Q - \chi_m\nu^\omega$

The analysis is similar to a Tobit equation. In that case:

$$\ell(\Theta\,; n \mid n = 0, e, q^n) = ln\left[1 - \Phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - Z^\varepsilon\kappa^\varepsilon - Z^h\kappa^h - Z^Q\kappa^Q + \chi_m Z^\omega\kappa^\omega + ln[h]}{\sigma_{(\nu^\varepsilon+\nu^h+\nu^Q-\chi_m\nu^\omega)}}\right)\right]$$ (Equation E4)

Note that in case b we have to use the CDF of a normal distribution, instead of the PDF.

In conclusion,

$$\ell(\Theta\,; n \mid e, q^n) = \mathbb{1}[n > 0] * \ell(\Theta\,; n \mid n > 0, e, q^n) + \mathbb{1}[n = 0] * \ell(\Theta\,; n \mid n = 0, e, q^n)$$

$\ell(\Theta\,; t \mid n, e, q^n)$: Equation B, the FOC for optimal CDC time, is the basis for these likelihood functions. Only families in the treatment group choose optimal time at the CDC. ($\bar{\tau} > 0$).

$$MRS_{h,c}[f_1 q^t - f_2 q^r] + w(m,\omega) - MRS_{p,c}\, e \;\leq -MRS_{t,c}$$

$$\Rightarrow MRS_{h,c} f_1 q^t - MRS_{h,c} f_2 q^r + w(m,\omega) - MRS_{p,c}\, e \;\leq -MRS_{t,c}$$

$$\Rightarrow MRS_{h,c} f_1 q^t + w(m,\omega) + MRS_{h,c}\, f_2\, [q_e^r\, e - q^r] \leq -MRS_{t,c} \quad \text{(Using Eq. E)}$$

Due to possible corner solutions, there are four cases to consider:

a) $n^* > 0\,;\; t^* > 0$.
b) $n^* > 0\,;\; t^* = 0$.
c) $n^* = 0\,;\; t^* > 0$.
d) $n^* = 0\,;\; t^* = 0$.

Case a) $n^* > 0\,;\; t^* > 0$

In this case, the first order condition for non-maternal non-CDC time (n) holds with equality. Therefore, and using Equation F, $w(m,\omega) = \; f_2\; MRS_{h,c}[q^r - q_e^r\; e]$, we can simplify the expression to:

$$\Rightarrow MRS_{h,c} f_1 q^t = -MRS_{t,c}$$

$$\Rightarrow ln\big(MRS_{h,c}\big) + ln(f_1) + ln(q^t) = ln\big(-MRS_{t,c}\big)$$

$$\Rightarrow ln[\gamma_h] - ln[h] + ln(f_1) + ln(\varepsilon) + ln(q^t) = ln[\gamma_t] - ln[\bar{\tau} - t]$$

$$\Rightarrow \mathrm{Z}^h \kappa^h + \nu^h - ln[h] + ln(f_1) + \mathrm{Z}^\varepsilon \kappa^\varepsilon + \nu^\varepsilon + ln(q^t) = \mathrm{Z}^t \kappa^t + \nu^t - ln[\bar{\tau} - t]$$

$$\Rightarrow ln(f_1) + ln(q^t) + \mathrm{Z}^h \kappa^h + \mathrm{Z}^\varepsilon \kappa^\varepsilon - \mathrm{Z}^t \kappa^t - ln[h] + ln[\bar{\tau} - t] = \nu^t - \nu^h - \nu^\varepsilon$$

We can show that $Var(\nu^t - \nu^h - \nu^\varepsilon) = \sigma_t^2 + \sigma_h^2 + \sigma_\varepsilon^2 - 2\sigma_{t,h} - 2\sigma_{t,\varepsilon} + 2\sigma_{h,\varepsilon}$

Then, $\nu^t - \nu^h - \nu^\varepsilon \sim \mathcal{N}\big(0, \sigma_t^2 + \sigma_h^2 + \sigma_\varepsilon^2 - 2\sigma_{t,h} - 2\sigma_{t,\varepsilon} + 2\sigma_{h,\varepsilon}\big)$

And therefore,

$$\ell(\Theta\,; t \mid n > 0, t > 0, e, q^n) =$$
$$ln\left[\phi\left(\frac{ln(f_1) + ln\big(q^t\big) + \mathrm{Z}^h\kappa^h + \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^t\kappa^t - ln[h] + ln[\bar{\tau} - t]}{\sigma_{(\nu^t - \nu^h - \nu^\varepsilon)}}\right)\right] - ln\left[\sigma_{(\nu^t - \nu^h - \nu^\varepsilon)}\right]$$

(Equation E5)

where $\sigma_{(\nu^t-\nu^h-\nu^\varepsilon)} = \sqrt{\sigma_t^2 + \sigma_h^2 + \sigma_\varepsilon^2 - 2\sigma_{t,h} - 2\sigma_{t,\varepsilon} + 2\sigma_{h,\varepsilon}}$

Case b) $n^* > 0\ ;\ t^* = 0$. In this case,

$$ln(f_1) + ln(q^t) + \mathrm{Z}^h\kappa^h + \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^t\kappa^t - ln[h] + ln[\bar{\tau} - t] \leq \nu^t - \nu^h - \nu^\varepsilon$$

Therefore,

$$\ell(\Theta\,; t \mid n > 0, t = 0, e, q^n) = ln\left[1 - \Phi\left(\frac{ln(f_1)+ln(q^t)+\mathrm{Z}^h\kappa^h+\mathrm{Z}^\varepsilon\kappa^\varepsilon-\mathrm{Z}^t\kappa^t-ln[h]+ln[\bar{\tau}-t]}{\sigma_{(\nu^t-\nu^h-\nu^\varepsilon)}}\right)\right] \text{ (Equation E6)}$$

Case c) $n^* = 0\ ;\ t^* > 0$. To analyze this case, we require the following assumption:

$$\pi q^t \approx -MRS_{t,c}$$

In words, the market value of high-quality care like the service provided by the CDC is an approximate measure of a parent's participation stigma (the marginal rate of substitution between time at the CDC and consumption). Using this new assumption, the relevant FOC becomes:

$$\overbrace{\left[\pi q^t - \left(-MRS_{t,c}\right)\right]}^{\approx 0} + w(m, \omega) = MRS_{h,c}\, f_2\, [q^r - q_e^r\, e]$$

Then, using the expression for q^r and the wage equation,

$$ln[\widehat{w}(m)] + ln(\omega) \approx ln\left(MRS_{h,c}\right) + ln(f_2) + ln(q^r - q_e^r\, e)$$

$$\Rightarrow ln[\widehat{w}(m)] + ln(\omega) \approx ln[\gamma_h] - ln[h] + ln(f_2) + ln(\varepsilon) + ln(Q) + \chi_m ln(\widehat{w}) + (1 - \chi_m) ln(\omega)$$

$$\Rightarrow (1 - \chi_m) ln[\widehat{w}(m)] \approx ln[\gamma_h] - ln[h] + ln(f_2) + ln(\varepsilon) + ln(Q) - \chi_m ln(\omega)$$

$$\Rightarrow (1-\chi_m)ln[\widehat{w}(m)] \approx \mathrm{Z}^h\kappa^h + \nu^h - ln[h] + ln(f_2) + \mathrm{Z}^\varepsilon\kappa^\varepsilon + \nu^\varepsilon + \mathrm{Z}^Q\kappa^Q + \nu^Q - \chi_m(\mathrm{Z}^\omega\kappa^\omega + \nu^\omega)$$

$$\Rightarrow (1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - \mathrm{Z}^h\kappa^h - \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^Q\kappa^Q + \chi_m\mathrm{Z}^\omega\kappa^\omega + ln[h] \approx \nu^h + \nu^\varepsilon + \nu^Q - \chi_m\nu^\omega$$

Recall we have established already the following:

$$Var(\nu^h + \nu^\varepsilon + \nu^Q - \chi_m\nu^\omega) = Var(\nu^h + \nu^\varepsilon + \nu^Q) + Var(\chi_m\nu^\omega) - 2Cov(\nu^h + \nu^\varepsilon + \nu^Q, \chi_m\nu^\omega)$$

$$= Var(\nu^h + \nu^\varepsilon) + Var(\nu^Q) + 2Cov(\nu^h + \nu^\varepsilon, \nu^Q) + Var(\chi_m\nu^\omega) - 2Cov(\nu^h + \nu^\varepsilon + \nu^Q, \chi_m\nu^\omega)$$

$$= Var(\nu^h) + Var(\nu^\varepsilon) + 2Cov(\nu^h, \nu^\varepsilon) + Var(\nu^Q) + 2Cov(\nu^h + \nu^\varepsilon, \nu^Q) + Var(\chi_m\nu^\omega) - 2Cov(\nu^h + \nu^\varepsilon + \nu^Q, \chi_m\nu^\omega)$$

$$= Var(\nu^h) + Var(\nu^\varepsilon) + 2Cov(\nu^h, \nu^\varepsilon) + Var(\nu^Q) + 2Cov(\nu^h, \nu^Q) + 2Cov(\nu^\varepsilon, \nu^Q) + Var(\chi_m\nu^\omega) - 2Cov(\nu^h, \chi_m\nu^\omega) - 2Cov(\nu^\varepsilon, \chi_m\nu^\omega) - 2Cov(\nu^Q, \chi_m\nu^\omega)$$

$$= Var(\nu^h) + Var(\nu^\varepsilon) + Var(\nu^Q) + 2Cov(\nu^h, \nu^\varepsilon) + 2Cov(\nu^h, \nu^Q) + 2Cov(\nu^\varepsilon, \nu^Q) + \chi_m^2 Var(\nu^\omega) - 2\chi_m Cov(\nu^h, \nu^\omega) - 2\chi_m Cov(\nu^\varepsilon, \nu^\omega) - 2\chi_m Cov(\nu^Q, \nu^\omega)$$

$$= \sigma_h^2 + \sigma_\varepsilon^2 + \sigma_Q^2 + \chi_m^2\sigma_\omega^2 + 2\sigma_{h,\varepsilon} + 2\sigma_{h,Q} + 2\sigma_{\varepsilon,Q} - 2\chi_m\sigma_{h,\omega} - 2\chi_m\sigma_{\varepsilon,\omega} - 2\chi_m\sigma_{Q,\omega}$$

Let $\sqrt{\sigma_h^2 + \sigma_\varepsilon^2 + \sigma_Q^2 + \chi_m^2\sigma_\omega^2 + 2\sigma_{h,\varepsilon} + 2\sigma_{h,Q} + 2\sigma_{\varepsilon,Q} - 2\chi_m\sigma_{h,\omega} - 2\chi_m\sigma_{\varepsilon,\omega} - 2\chi_m\sigma_{Q,\omega}} = \sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}$

So in conclusion,

$$\ell(\Theta\,; t \mid n = 0, t > 0, e, q^n) \approx ln\left[\phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - \mathrm{Z}^h\kappa^h - \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^Q\kappa^Q + \chi_m\mathrm{Z}^\omega\kappa^\omega + ln[h]}{\sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}}\right)\right] - ln\left[\sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}\right]$$ (Equation E7)

Case d) $n^* = 0\,;\ t^* = 0$

Based on the analysis for case c), we can state the following:

$$(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - \mathrm{Z}^h\kappa^h - \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^Q\kappa^Q + \chi_m\mathrm{Z}^\omega\kappa^\omega + ln[h] \leq \nu^h + \nu^\varepsilon + \nu^Q - \chi_m\nu^\omega$$

Therefore, the log-likelihood at this corner solution will be the following:

$$\ell(\Theta\,; t \mid n = 0, t = 0, e, q^n) = ln\left[1 - \Phi\left(\frac{(1-\chi_m)ln[\widehat{w}(m)] - ln(f_2) - \mathrm{Z}^h\kappa^h - \mathrm{Z}^\varepsilon\kappa^\varepsilon - \mathrm{Z}^Q\kappa^Q + \chi_m\mathrm{Z}^\omega\kappa^\omega + ln[h]}{\sigma_{(\nu^h+\nu^\varepsilon+\nu^Q-\chi_m\nu^\omega)}}\right)\right] \text{ (Equation E8)}$$

We can now put together all four cases to obtain the complete log-likehood $\ell(\Theta\,; t \mid n, e, q^n)$.

$$\begin{aligned}\ell(\Theta\,; t \mid n, e, q^n) &= \mathbb{I}[n > 0, t > 0, \bar{\tau} > 0] * \ell(\Theta\,; t \mid n > 0, t > 0, e, q^n) \\ &+ \mathbb{I}[n > 0, t = 0, \bar{\tau} > 0] * \ell(\Theta\,; t \mid n > 0, t = 0, e, q^n) \\ &+ \mathbb{I}[n = 0, t > 0, \bar{\tau} > 0] * \ell(\Theta\,; t \mid n = 0, t > 0, e, q^n) \\ &+ \mathbb{I}[n = 0, t = 0, \bar{\tau} > 0] * \ell(\Theta\,; t \mid n = 0, t = 0, e, q^n)\end{aligned}$$

$\ell(\Theta\,; l \mid t, n, e, q^n)$: the first order condition for leisure is the starting point to obtain this log-likelihood.

$$MRS_{l,c} = w(m, \omega)$$

$$\Rightarrow ln\big(MRS_{l,c}\big) = ln[\widehat{w}(m)] + ln(\omega)$$

$$\Rightarrow ln[\gamma_l] - ln[l] = ln[\widehat{w}(m)] + ln(\omega)$$

$$\Rightarrow \mathrm{Z}^l\kappa^l + \nu^l - ln[l] = ln[\widehat{w}(m)] + \mathrm{Z}^\omega\kappa^\omega + \nu^\omega$$

$$\Rightarrow \nu^l - \nu^\omega = ln[\widehat{w}(m)] + \mathrm{Z}^\omega\kappa^\omega - \mathrm{Z}^l\kappa^l + ln[l]$$

Using the distributional assumptions, $\nu^l - \nu^\omega \sim \mathcal{N}\left(0, \sigma^2_{(\nu^l-\nu^\omega)}\right)$

where $\sigma_{(\nu^l-\nu^\omega)} = \sqrt{\sigma_l^2 + \sigma_\omega^2 - 2\sigma_{l,\omega}}$

Therefore,

$$\ell(\Theta\,;l\mid t,n,e,q^{n}) = ln\left[\phi\left(\frac{ln[\widehat{w}(m)]+\mathrm{Z}^{\omega}\kappa^{\omega}-\mathrm{Z}^{l}\kappa^{l}+ln[l]}{\sigma_{(\nu^{l}-\nu^{\omega})}}\right)\right]-ln\left[\sigma_{(\nu^{l}-\nu^{\omega})}\right]$$

(Equation E9)

f) Appendix: measurement of missing variables

This appendix discusses measurement of variables that are theoretically important but not available directly in the IHDP: 1) expected potential wage, 2) pre-natal investment choice and child endowment, and 3) nonmaternal-care quality. In each case, the IHDP contains a lot of variables that should have strong relationships with the missing variable. In each case, fortunately, there is a separate, high-quality, micro-dataset in which both the predictors and the missing variable are measured. We aim to harness this outside information to measure the variables of interest for each individual in the IHDP. The basic approach is estimate a model in the outside dataset and use the estimated parameters to score each IHDP individual using the model's estimated parameters and the IHDP individual's observed values on the predictors. That is, we impute the conditional mean in place of the missing value.

This is different than mean-imputation or multiple-imputation as usually practiced. Usually, the problem is that, within a single dataset, a variable (x) has some individuals with observed values and other individuals with missing values. Let z indicate whether the value is observed for each individual. Typically, other variables (d) have fully-observed values. In this case, researchers often model the relationship between the variable with some missing values and the variables with fully-observed values in the subsample where x is observed ($z=1$). Then, the subsample where x is missing ($z=0$) are scored and this is used to impute missing values and the primary relationship of interest, $E[y|x]$, is then estimated using the full sample. While this can produce unbiased estimates under some conditions, the conditions are often not credible. Some selection process drove some individuals to have missing values and others to have observed values. This selection process might also affect the primary relationship of interest and lead to bias.

Our situation is different. Here, all individuals have missing data on the variables in question. There is no selection into observability. The original IHDP researchers collected data on a huge number of variables but missed a few specific variables that we care about. We are harnessing the outside data to understand the relationship between observables in both datasets and the missing variables of interest. Then, we use the conditional mean prediction as an imputed proxy for the missing values. We will discuss the implications of relying on proxies, which may have classical measurement error, in the robustness section.

Expected potential wage

We draw on Current Population Survey March supplements for 1986-89 from IPUMS. We limit the sample to mothers between the ages of 15 and 55 with at least one child below the age of 5, excluding non-civilians, unpaid family workers, and the self-employed. In terms of cleaning and modeling, we largely follow Mulligan & Rubinstein (2008). However, we include women of color and allow wage offers and employment probabilities to differ by ethnicity. Observed hourly wage is the ratio of last year's total labor income divided by usual hours per week times weeks worked. Wages below \$3.73 and above \$80 in 2012 dollars are trimmed. We use a standard Heckit model of selection into the workforce (L=1) estimated by the 2-step method (Heckman, 1974):

$$\ln(w) = X\beta^w + \theta^w \lambda(Z\delta^w) + \epsilon^w$$

$$\Pr(L = 1|Z) = \Phi(Z\delta^w)$$

Observed variables are described here with omitted categories in *italics*. Wage determinants (X) are indicators of educational attainment (less than high school, *high school only*[10], some college[11], college graduate[12]), ethnicity (*non-Hispanic whites*, African-American, Hispanic, other), marital status (never-married, *married*, separated/widowed/divorced). To capture differences in local market conditions in the IHDP sites in particular, we include an indicator for residence in each of the 8 IHDP site's metropolitan area and an indicator for other SMSA residency (*non-SMSA residency* excluded) and indicators for region and year. We also include a quartic of potential work experience, defined as maximum{0, age - years of completed schooling - 7}, and its interactions with the education indicators. The participation determinants (Z) include all components of X as well as the following variables, which are excluded from the wage equation, measures of the number of children below 5, age of the youngest child, and number of other children in household, and the interaction of these 3 with the marital status indicators. Observations with any demographic variables missing are dropped. This produces estimates of (β^w,δ^w,θ^w). For each mother in the IHDP sample, these estimates are used to impute an expected potential wage, $\widehat{w} = \ln(w_i) = X_i\hat{\beta}^w + \hat{\theta}^w \lambda\left(Z_i\hat{\delta}^w\right)$, treating the estimates as known parameters. In our model, wage is the sum of this and an unobserved productivity parameter (ω).

Pre-natal investment choice and child endowment (I_0^* , ϕ)

Birth outcomes such as birth weight and gestational age at birth are influenced by pre-natal choices of the mother and, therefore, by maternal preferences and constraints. We seek to model mothers' pre-natal choices given their beliefs about the relationship between effort choices and child outcomes. [13]

[10] Women who finished 12th grade, have a high school diploma or equivalent.
[11] Between one and three years of college education.
[12] Four or more years of college education.
[13] Fox et al (1987) report that knowledge about the risks of smoking and drinking were widespread in 1985. A nationally-representative survey of over 20,000 Americans aged 18-44 nationally found that 85% of women reported that smoking during pregnancy increased the risk of low weight at birth and 88% said the same for heavy smoking. In connecting smoking with low birth weight, more respondents say the behavior "definitely increases risk" than said "probably increases risk." They report some bivariate relationship between risk perceptions with age and with education, which adds to the justification for including these in X. Race was related to perception of risks from smoking but not heavy drinking. Little relationship with income was observed.

The pre-natal production function b maps latent maternal pre-natal investments (I_0^*), observed maternal characteristics that would influence fetal development and maternal beliefs (X), and the child's idiosyncratic endowment (ϕ) into h_0. Assume g is linear, though only additive ϕ is necessary for identification. Also, assume ϕ is mean independent of I_0^* conditional on X. This assumption is credible given that I_0^* is chosen pre-natally, before information about child endowment ϕ is known to the mother (Aizer & Cunha, 2012).

$$h_0 = b[I_0^* \, ; \, X, \phi] \; \equiv \pi_0 + \pi_1 I_0^* + \pi_2 X + \phi$$

Because the IHDP sample is selected on explicit, defined thresholds for birth weight and gestational age, studying the relationships between I_0^* and h_0 in the IHDP sample directly might produce misleading conclusions. Therefore, we seek to understand these relationships in the ECLS-B, the nation's first nationally-representative birth cohort consisting of approximately 14,000 children born in 2001. We will normalize Var(I_0^*)=1 with respect to this nationally-representative cohort.

This allows us to characterize the IHDP sample in terms of the national joint distribution of pre-natal investment choices and child endowment shocks $(I_0^* \, , \phi)$. After obtaining estimates from the nationally-representative birth cohort, we score each observation in the IHDP sample with an estimated (I_i^*, ϕ_i). Each mother's choice of I_0^* reveals information about the strength of her preference for future child human capital compared to other, otherwise similar (X) mothers. The endowment measures how different the child's condition is at birth from the expected level given maternal type and pre-natal investment choices. To the extent allowed by our set of observables, we will be able to say whether the IHDP sample is born low birth weight and premature because of (a) low level of pre-natal investment presumably driven by low levels of maternal value placed on child development or (b) bad child endowment shocks orthogonal to maternal type.

To approximate (I_0^*, ϕ), we proxy $h_0$with the two birth outcomes on which the IHDP sample is selected: weight (W) and gestational age (A). In a SUR framework, we regress each of these birth outcomes on a vector of observable pre-natal investment choices (C_0) and on maternal characteristics (X).

$$\begin{pmatrix} W \\ A \end{pmatrix} = \pi_0 + \pi_1 C_0 + \pi_2 X + \begin{pmatrix} \phi_W \\ \phi_A \end{pmatrix}$$

Given our strategy, we limit the analysis to variables that are available in both the IHDP and ECLS-B, which is an extensive list. C_0 includes average number of cigarettes smoked per day during pregnancy and its square, average number of alcoholic drinks consumed per week during pregnancy and its square, an indicator of drug use, maternal weight gain during pregnancy and its square, trimester of first pre-natal care with no pre-natal care coded as 4. The measures of X are ethnicity indicators, marital status, education indicators with high school excluded, indicator for non-singleton pregnancy, indicator for cesarean delivery, maternal weight at conception, and indicator for female baby. Appendix Table AT.4 provides summary statistics from the ECLS-B and IHDP samples on these variables. Estimating this model in the ECLS-B using appropriate weights produces estimates of (π_0, π_1, π_2) displayed in Appendix Table AT.5.

Applying these coefficients to each member of the ECLS-B sample's own values of (C_0,X) produces estimates of $(\hat{\pi}_1^k C_{o,i}, \hat{\pi}_0 + \hat{\pi}_2^k X_i)$ for each individual and each birth outcome k = W,A. What is $\hat{\pi}_1^W C_{o,i}$? It measures the pre-natal investment level chosen by mother-i in terms of birth-weight units where each component of C_0 is weighted by its importance in conditionally predicting birth weight. What is $\hat{\pi}_0 + \hat{\pi}_2^k X_i$? It captures the predicted birth weight associated with a particular maternal type (X) holding pre-natal investment choices fixed. Put another way, we look at how different pre-natal investment choices are associated with different birth weights conditional on mothers' observable type. The distribution of $\hat{\pi}_1^W C_{o,i}$ is a nationally-representative estimate of the full distribution of birth-weight-determining, pre-natal investment levels. We record the percentiles of this distribution, its mean and standard deviation, and transform each individual's measure of investment into a z-score. We do the same for the birth weight residuals (ϕ_{Wi}). Further, we do the same with gestational age: k=A.

Next, we average the z-scores of investment levels for each individual across birth weight and gestational age. We standardize that average so that it has mean 0 and standard deviation 1.[14] This is our proxy for I_0^*. We also record the percentiles of this distribution in the nationally-representative sample. Similarly, we average the standardized residuals from the birth weight and gestational age predictions and standardize the result. This is our proxy for ϕ.

Finally, we use the same scoring procedure for each member of the IHDP sample. This delivers measures of I_0^* and ϕ for each mother and child in the IHDP. These are measured with respect to national norms. Because the translog production function requires positive inputs, the IHDP endowment estimated z-scores are shifted up by a constant just above the magnitude of the minimum value observed.

Quality of nonmaternal care

The IHDP data has very specific information about non-maternal care. The survey asked for the primary and secondary caregivers during a typical week at the 18-month, 24-month, 30-month and 36-month family interviews. The respondent could choose from nine different categories (partner, sibling, grandmother, another relative, babysitter, day care home, day care center, someone else and the child's father, if he lives in another home). However, the IHDP did not directly measure the quality of non-maternal care.

[14] This ensures the two outcomes receive equal weight, even though they are measured in different units.

To get a continuous measure the quality of these care settings, we draw in data from a pioneering study of nonmaternal care quality, the Study of Early Child Care and Youth Development (SECCYD) by the National Institute of Child Health and Human Development (NICHD). The SECCYD collected panel data on child and family characteristics and their use of various care settings. The SECCYD classifies non-maternal caregivers into nine categories: father / partner, grandparent in-home, grandparent out-of-home, other relative in-home, other relative out-of-home, non-relative in-home, non-relative out-of-home, child care center and others. The study included a sample of 1,364 children aged 0 to 3 during 1991 to 1994 in 10 study sites around the country, 2 of which overlap with the IHDP's 8 sites.[15]

For each child and each nonmaternal care setting used, the SECCYD measured care quality using the Observational Record of the Childcare Environment (ORCE) (NICHD, 2003; Vandell, 2004), which is composed of three different types of scores: Behavioral Scales, Qualitative Ratings and measures of Structural Variables. We follow Auger & Burchinal (2013), who suggest that a good measure of the quality of interactions geared toward cognitive stimulus is the ORCE's Qualitative Rating on Stimulation of Development. This rating is available in the SECCYD data at 15, 24 and 36 months (Phase 1).

We estimate a pooled OLS model in the SECCYD data, in which the dependent variable is standardized ORCE Qualitative Rating on Stimulation of Development. The set of predictors must be variables available in both the SECCYD and IHDP datasets. They include child's age, birth order, gender, birth weight (level and square), gestational age at birth (level and square), maternal age at child birth, maternal education (four categories), race, ethnicity, marital status, and study site. As a predictor, we also use the standardized Learning and Literacy score based on components from the HOME score (Linver, Martin & Brooks-Gunn, 2004; Fuligni, Han & Brooks-Gunn, 2004). Finally, we match the nine categories of non-maternal caregivers from the IHDP with the nine categories used in the SECCYD. Thus, the last set of predictors are indicators for the category of the caregiver.

After estimating the linear relationship between mean nonmaternal care quality and the set of predictors in the SECCYD, we score each IHDP child based on the same set of predictors and impute that mean prediction as the IHDP child's measure of nonmaternal-care quality (q^n). Summary statistics for the SECCYD data and model estimates are displayed in Appendix Tables AT.6 and AT.7, respectively. Because the translog production function requires positive inputs, each value is shifted up by a constant just above the magnitude of the minimum value observed.

[15] The 10 sites of the SECCYD – NICHD study are University of Arkansas, UC Irvine, University of Kansas, University of New Hampshire, Penn State University, Temple University, University of Virginia, University of Washington, Western Carolina Center and University of Wisconsin. The sites which overlap with the IHDP study are the University of Arkansas and the University of Washington.

X. Appendix Tables and Figures

AT. 1: Baseline characteristics as predictors of the subsample (N = 815)

	(1)	(2)	(3)	(4)
Treatment indicator	0.00280	0.00727	-0.0228	0.0205
	(0.0291)	(0.0280)	(0.0121)	(0.0296)
Male infant	0.0232	0.0202	-0.0218	0.0314
	(0.0273)	(0.0290)	(0.0154)	(0.0297)
Birth Weight (kgm)	0.0103	0.00357	0.0159	-0.0175
	(0.0357)	(0.0360)	(0.0238)	(0.0250)
Gestational Age (weeks)	0.00329	0.00393	-0.00710	0.00632
	(0.00774)	(0.00783)	(0.00519)	(0.00746)
Maternal Age	0.000132	0.00445	-0.000985	0.00183
	(0.00229)	(0.00246)	(0.000987)	(0.00253)
Maternal educ.: < HS	-0.0480	-0.117**	-0.0325	-0.0338*
	(0.0331)	(0.0393)	(0.0336)	(0.0170)
Maternal educ.: Some College	-0.000504	0.0788	-0.00641	-0.00328
	(0.0446)	(0.0517)	(0.0339)	(0.0363)
Maternal educ.: College graduate	0.0278	0.228**	0.0209	0.0260
	(0.0432)	(0.0655)	(0.0346)	(0.0271)
Race: African American	-0.0182	-0.0291	0.0169	-0.00747
	(0.0476)	(0.0481)	(0.0255)	(0.0458)
Race: Hispanic	-0.0327	-0.0766	-0.0117	-0.0334
	(0.0657)	(0.0671)	(0.0233)	(0.0655)
Race: Other	-0.122	-0.158	-0.0765	-0.0960
	(0.0933)	(0.0960)	(0.0438)	(0.116)
Marital status: Single	-0.0144	-0.0601*	-0.00345	0.00813
	(0.0304)	(0.0299)	(0.0246)	(0.0251)
Marital status: Sep./Div./Wid.	-0.0496	-0.0575	0.0378	-0.0417
	(0.0428)	(0.0432)	(0.0305)	(0.0335)
Predicted Wage Offer		-0.0234**		
		(0.00680)		
Quality of Non-Maternal Care			0.0169	
			(0.0262)	
Prenatal Investment				0.123
				(0.710)
Constant	0.737***	0.855***	1.111***	0.428
	(0.199)	(0.209)	(0.0959)	(1.238)
Observations	985	985	867	948
R-squared	0.025	0.042	0.043	0.017

Linear probability models. The dependent variable is an indicator for observations selected in the subsample (N = 815). Standard errors are clustered at the site level. The excluded maternal education category is “High School graduates”. The excluded race and ethnicity category is “Non-Hispanic Whites”. The excluded marital status category is “Married”. The excluded site category is “Miami (MIA)”. Site coefficients are not reported. *** p<0.01, ** p<0.05, * p<0.1

AT. 2: Summary statistics for variables in the predicted wage model

		CPS			IHDP		
Continuous variables							
		Mean	Std. Dev.	N	Mean	Std. Dev.	N
Hourly Rate of Pay Log, US$ of 2012	Working mothers only	2.60	0.56	18,680	2.09	0.71	542
	All the sample	-	-	-	1.89	0.83	985
Worked Indicator		0.60	0.49	30,889	0.52	0.50	913
Potential experience (years)		9.61	5.60	30,889	6.49	5.28	985
Number of own children under age 5		1.30	0.53	30,889	1.50	0.71	985
Age of youngest own child in household		1.75	1.39	30,889	1.70	0.68	985
Number of own children 5 years old or older		0.77	1.03	30,889	0.46	0.84	985

	CPS		IHDP	
Maternal education				
	Share (%)	N	Share (%)	N
Less than High School	18.4	5,682	40.0	394
High School graduate	43.7	13,505	27.4	270
Some College	19.9	6,157	20.0	197
College graduate	18.0	5,545	12.6	124

	CPS		IHDP	
Race and Ethnicity				
	Share (%)	N	Share (%)	N
Non-Hispanic White	70.4	21,752	33.4	329
African American	11.0	3,383	52.5	517
Hispanic	14.6	4,513	10.7	105
Other	4.0	1,241	3.5	34

	CPS		IHDP	
Marital status				
	Share (%)	N	Share (%)	N
Married	80.8	24,964	46.2	455
Single	8.6	2,661	45.8	451
Sep./Div./Wid.	10.6	3,264	8.0	79

CPS Sample: IPUMS-CPS extract, Minnesota Population Center. 1986-89 March Samples. Women, age 15 to 55, with at least one child under the age of 5. Unpaid family workers and self-employed women not included. Hourly Rate of Pay is equal to the ratio of last year's total labor income divided by usual hours per week times weeks worked. Wages below \$3.73 and above \$80 in 2012 dollars are trimmed. IHDP: Infant Health and Development Program sample. Hourly Rate of Pay for the IHDP sample is the predicted value based on the Heckman selection model presented in Appendix Table AT.3.

AT. 3: Estimates from Heckman selection model in CPS sample

VARIABLES	(1) Ln(hourly wage)	(2) 1[worked]	(3) Mills
Potential experience	0.0612***	0.0648***	
	(0.0115)	(0.0233)	
Potential experience, squared	-0.00150	-0.00751**	
	(0.00165)	(0.00327)	
Potential experience, cubed	-3.10e-05	0.000285*	
	(8.85e-05)	(0.000173)	
Potential experience, ^4	1.09e-06	-3.74e-06	
	(1.54e-06)	(2.96e-06)	
Education: Less HS	0.0981*	-0.759***	
	(0.0541)	(0.0860)	
Education: Some Coll.	0.0700	0.348***	
	(0.0455)	(0.0992)	
Education: Coll. grad.	0.429***	0.515***	
	(0.0540)	(0.130)	
Experience * Less HS indicator	-0.0493**	0.0592*	
	(0.0198)	(0.0337)	
Experience * Some Coll. indicator	0.0532**	-0.0773*	
	(0.0208)	(0.0448)	
Experience * Coll. grad. indicator	0.0249	-0.0662	
	(0.0253)	(0.0614)	
Experience^2 * Less HS indicator	0.00265	-0.00372	
	(0.00241)	(0.00424)	
Experience^2 * Some Coll. indicator	-0.00748**	0.00916	
	(0.00303)	(0.00647)	
Experience^2 * Coll. grad. indicator	-0.00423	0.00521	
	(0.00389)	(0.00967)	
Experience^3 * Less HS indicator	-4.75e-05	0.000113	
	(0.000113)	(0.000204)	
Experience^3 * Some Coll. indicator	0.000374**	-0.000382	
	(0.000164)	(0.000352)	
Experience^3 * Coll. grad. indicator	0.000228	-0.000198	
	(0.000229)	(0.000593)	
Experience^4 * Less HS indicator	7.21e-09	-8.05e-07	
	(1.79e-06)	(3.29e-06)	
Experience^4 * Some Coll. indicator	-5.74e-06**	5.39e-06	
	(2.88e-06)	(6.26e-06)	
Experience^4 * Coll. grad. indicator	-4.47e-06	3.90e-06	
	(4.45e-06)	(1.23e-05)	
Race: African American	-0.0932***	0.230***	
	(0.0132)	(0.0282)	
Race: Hispanic	-0.0712***	-0.0992***	
	(0.0132)	(0.0247)	
Race: Other	-0.0418**	-0.0851**	
	(0.0199)	(0.0389)	
Marital status: Single	-0.0403**	-0.183**	
	(0.0166)	(0.0930)	

Marital status: Sep./Div./Wid.	-0.0964***	-0.173*
	(0.0123)	(0.0963)
Metropolitan area: Boston, MA	0.254***	-0.129*
	(0.0340)	(0.0685)
Metropolitan area: Dallas-Fort Worth, TX	0.273***	-0.00558
	(0.0422)	(0.0863)
Metropolitan area: Little Rock-North Little Rock, AR	0.0805	0.558***
	(0.0589)	(0.149)
Metropolitan area: Miami-Hialeah, FL	0.224***	0.0317
	(0.0391)	(0.0811)
Metropolitan area: New Haven-Meriden, CT	0.270**	-0.146
	(0.113)	(0.227)
Metropolitan area: New York, NY	0.335***	-0.352***
	(0.0298)	(0.0515)
Metropolitan area: Philadelphia, PA/NJ	0.215***	-0.0132
	(0.0324)	(0.0626)
Metropolitan area: Seattle-Everett, WA	0.136**	-0.153
	(0.0551)	(0.110)
Other metropolitan areas	0.164***	-0.0422**
	(0.00898)	(0.0187)
Region and division: New England Division	0.0292	0.0448
	(0.0178)	(0.0362)
Region and division: Middle Atlantic Division	0.0458***	-0.164***
	(0.0162)	(0.0311)
Region and division: West North Central Division	-0.105***	0.276***
	(0.0160)	(0.0334)
Region and division: South Atlantic Division	-0.0495***	0.149***
	(0.0137)	(0.0278)
Region and division: East South Central Division	-0.148***	0.0491
	(0.0198)	(0.0404)
Region and division: West South Central Division	-0.0810***	0.0511
	(0.0161)	(0.0321)
Region and division: Mountain Division	-0.113***	0.188***
	(0.0151)	(0.0309)
Region and division: Pacific Division	0.0726***	0.0558*
	(0.0145)	(0.0288)
Number of own children under age 5 in hh		-0.373***
		(0.0168)
Age of youngest own child in household		0.00242
		(0.00680)
Number of own children 5 years old or older		-0.156***
		(0.00936)
Num. of children < 5 * Single indicator		-0.107*
		(0.0566)
Num. of children < 5 * Sep./Div./Wid. indicator		0.0996*
		(0.0531)
Age youngest child * Single indicator		0.0233
		(0.0212)
Age youngest child * Sep./Div./Wid. indicator		0.0987***
		(0.0197)
Num. of children >= 5 * Single indicator		-0.0777**

		(0.0311)	
Num. of children >= 5 * Sep./Div./Wid. indicator		-0.0635***	
		(0.0230)	
Lambda			-0.300***
			(0.0283)
Constant	2.218***	0.755***	
	(0.0331)	(0.0626)	
Observations	30,889	30,889	30,889

AT. 4: Summary statistics for pre-natal investment model

	ECLS-B			IHDP	
Variables	Mean	Std. Dev.		Mean	Std. Dev.
Conditions at birth					
Weight (kg)	3.3	0.6		1.8	0.4
Gestational age (wk)	38.7	2.4		33.0	2.7
Pre-natal investment choices					
Used drugs	0.04			0.04	0.19
Cigs/day				4.3	7.9
Drinks/wk.				0.4	1.8
Weight gain	35.1	23.1		23.5	13.0
Trimester of care	1.2	0.5		1.3	0.6
No pre-natal care	0.01			0.05	0.21
Fixed characteristics					
Fetus female	0.49			0.51	
Non-singleton fetus	0.03			0.11	
African-American	0.14			0.52	
Hispanic	0.25			0.10	
Other race/ethnicity	0.07			0.03	
Never married	0.26			0.45	
Widowed, div., or separated	0.07			0.08	
Maternal age	28.3	6.33		24.7	6.0
Education < HS	0.20			0.40	
HS < Education < BA	0.27			0.20	
Education = BA+	0.24			0.12	
Child parity	1.03	1.18		1.90	1.17

Note: omitted category is fetus male and singleton and mother non-Hispanic Caucasian, married, and high-school degree only.

AT. 5: Estimates from pre-natal investment model using the nationally-representative ECLS-B sample

	Outcomes: birth conditions	
Predictor variables	Weight (g)	Gestational age (weeks)
Pre-natal investment choices		
Used drugs	-6.80 (1.51)	
Cigarettes/day	-24.7	
$(\text{Cigarettes/day})^2$		
Alcoholic drinks/week		
$(\text{Alcoholic drinks/week})^2$		
Weight gain during pregnancy		
$(\text{Weight gain})^2$		
Trimester of care		
No pre-natal care		
Fixed characteristics		
Fetus female		
Non-singleton fetus		
African-American		
Hispanic		
Other race/ethnicity		
Never married		
Widowed, div., or separated		
Maternal age		
Education < HS		
HS < Education < BA		
Education = BA+		
Child parity		

Note: Estimated coefficients (SE). All coefficients are significant at p=0.002.

AT. 6: Descriptive statistics from the NICHD – SECCYD data

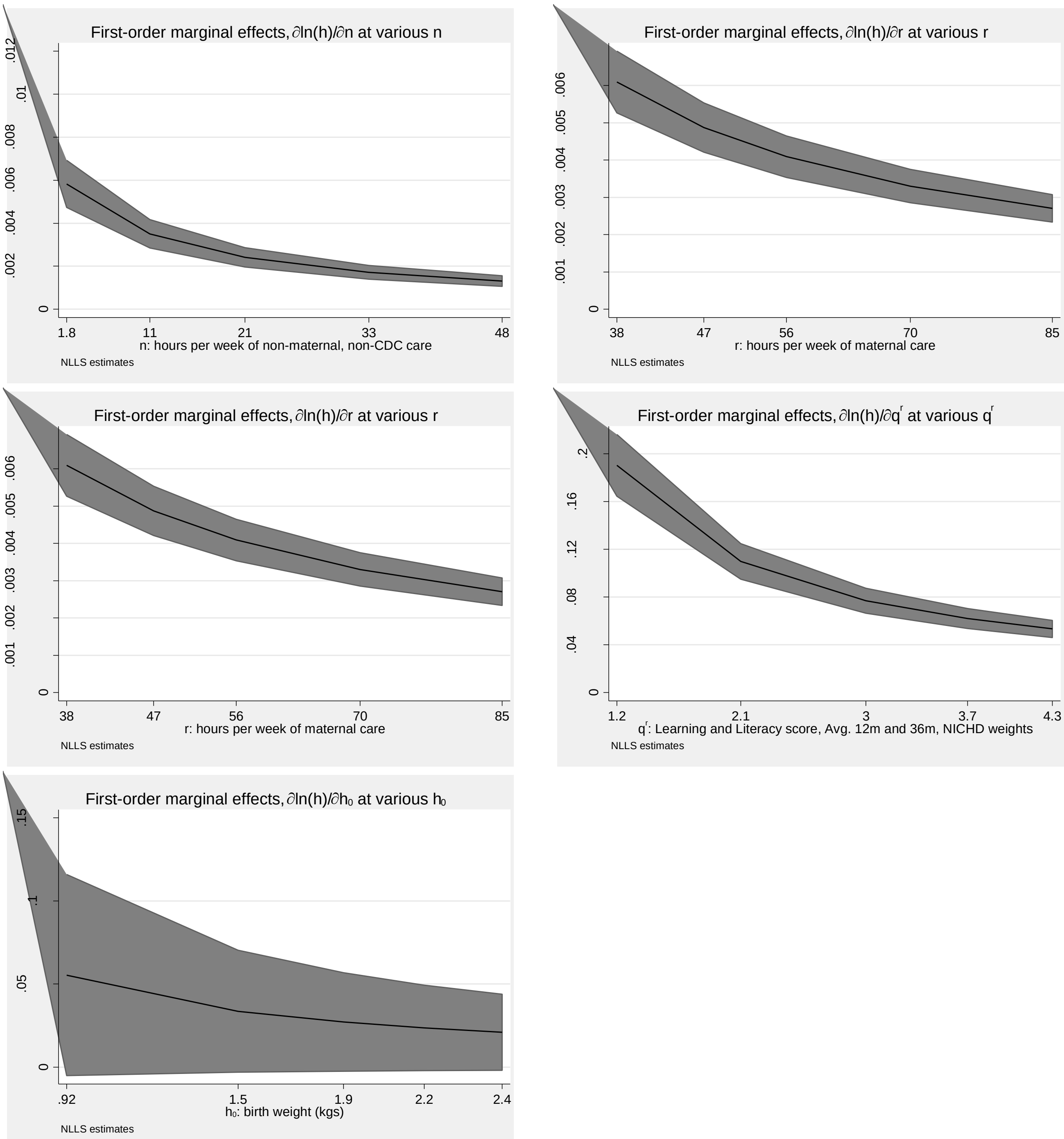
	Mean	Std. Dev.	Min	Max	N
ORCE, Stimulation of Development score	0.00	1.00	-1.39	3.26	1,837
Child's age (months)	25.29	8.64	15	36	1,837
Birth order	1.67	0.81	1	5	1,837
Female indicator	0.49	0.50	0	1	1,837
Child's birth weight (kgs)	3.50	0.51	2	5.34	1,837
Child's gestational age (weeks)	39.27	1.47	33	43	1,837
Mother's age (years)	28.92	5.39	18	46	1,837
Learning and Literacy Score, HOME Inventory	5.02	0.89	0	6.13	1,837
Mother's Education	Percent				
Less than High School	4.9				
High School graduate	17.8				
Some College	35.0				
College graduate	42.4				
Race and Ethnicity	Percent				
Non-Hispanic White	82.6				
African American	10.3				
Hispanic	4.3				
Other	2.7				
Non-Maternal Caregiver	Percent				
Father / Partner	14.8				
Grandparent	10.3				
Another Relative	5.6				
Non-Relative In-Home	10.8				
Day Care Home	27.3				
Child Care Center	31.3				

AT. 7: Model estimates for the quality of non-maternal care in the SECCYD – NICHD data

	(1)
Child’s age indicator, 24 months	0.0305
	(0.0463)
Child’s age indicator, 36 months	0.0940*
	(0.0500)
Child's birth order	-0.125***
	(0.0348)
Female child indicator	0.123**
	(0.0536)
Birth weight (grams)	0.412
	(0.472)
Birth weight squared	-0.0571
	(0.0661)
Child's gestational age	0.884*
	(0.511)
Child's gestational age squared	-0.0114*
	(0.00663)
Mother's age	0.0106*
	(0.00625)
Mother’s education: Less than High School	0.0228
	(0.126)
Mother’s education: Some college	0.0943
	(0.0722)
Mother’s education: College graduate	0.150*
	(0.0802)
Race and ethnicity: African-American	-0.194**
	(0.0861)
Race and ethnicity: Hispanic	-0.104
	(0.143)
Race and ethnicity: Other	0.179
	(0.128)
Marital status: Single	-0.132
	(0.0916)
Marital status: Separated / Divorced / Widowed	-0.260
	(0.179)
Avg. Learning and Literacy score, 15m and 36m	0.142***
	(0.0339)
Non-Maternal Caregiver: Father / Partner	0.336***
	(0.0889)
Non-Maternal Caregiver: Grandparent	0.342***
	(0.0949)
Non-Maternal Caregiver: Another Relative	0.0302
	(0.104)
Non-Maternal Caregiver: Non-Relative In-Home	0.534***
	(0.105)
Non-Maternal Caregiver: Day Care Home	0.138**
	(0.0661)
Constant	-18.94**
	(9.626)
Observations	1,837
R-squared	0.140

Note: the dependent variable is the Observational Rating of the Caregiving Environment (ORCE), Stimulation of Development score. The excluded child's age category is 15 months. The excluded mother's education category is high school graduates. The excluded race and ethnicity category are non-hispanic whites. The excluded marital status category is married women. The excluded non-maternal caregiver category is child care centers. 9 site dummies are included but not reported.

AF.1 Non-linear Least Squares (NLLS) first-order marginal effects



AF. 2: NLLS Second order effects