NBER WORKING PAPER SERIES

EARLY CHILDHOOD CARE AND COGNITIVE DEVELOPMENT

Juan Chaparro
Aaron Sojourner
Matthew J. Wiswall

Working Paper 26813
http://www.nber.org/papers/w26813

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2020

Thanks to Flavio Cunha, Liz Davis, Greg Duncan, Drew Griffen, Rob Grunewald, Jeff Smith, Chris Taber, Deb Vandell and communities at Cornell, Purdue, Ohio State, UC Berkeley, UC Irvine, UMinnesota, UWisconsin IRP, and Yale for valuable discussion. Chaparro and Sojourner received funding from the Eunice Kennedy Shriver NICHD under Award P01HD065704, the Center for Personalized Prevention Research, and the Human Capital Research Collaborative. Wiswall received support from National Science Foundation grant 1357636. Content is solely the authors' responsibility and doesn't represent the National Institutes of Health's official views. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Early Childhood Care and Cognitive Development
Juan Chaparro, Aaron Sojourner, and Matthew J. Wiswall
NBER Working Paper No. 26813
February 2020
JEL No. J13

ABSTRACT

This paper combines multiple sources of information on early childhood development in a unified model for analysis of a wide range of early childhood policy interventions. We develop a model of child care in which households decide both the quantities and qualities of maternal and non-maternal care along with maternal labor supply. The model introduces a novel parenting-effort channel, whereby child care subsidies that permit less parenting may enable better parenting. To estimate the model, we combine observational data with experimental data from the Infant Health and Development Program (IHDP) which randomly assigned free child care when the child was 1 and 2 years old. We estimate a cognitive skill production function and household preferences, giving insight into mechanisms driving the ex post heterogeneous effects of the IHDP intervention, accounting for alternative care substitutes available to the control group and spillovers of the child care offer across the household's decisions. We also estimate ex ante effects of counterfactual policies such as an offer of lower-quality care, requiring a co-pay for subsidized care, raising the maternal wage offer, or a cash transfer. Finally, we use the model to rationalize existing evidence from outside the US on the effects of universal child care programs.

Juan Chaparro
School of Economics and Finance
Universidad EAFIT
Carrera 49 # 7 Sur - 50
Medellin, Colombia, South America
jcchaparrc@eafit.edu.co

Aaron Sojourner
Carlson School of Management
321 19th Ave S, 3-300
Minneapolis, MN 55409
asojourn@umn.edu

Matthew J. Wiswall
Department of Economics
University of Wisconsin-Madison
1180 Observatory Drive
Madison, WI 53706
and NBER
matt.wiswall@gmail.com

1 Introduction

This paper develops a framework to combine various sources of information on early childhood development, including observational data and experiments, into a unified model which can be used to analyze a wide range of early childhood policy interventions. Our empirical analysis focuses on a period of human development, the first 3 years of life, that is critical to cognitive development: this is the period when development by family socioeconomic status (SES) differs the most.[1] However, in the U.S., child care and education during these younger ages receives substantially less government funding than at later ages, primarily due to large expenditures on public K-12 schooling.[2]

Our analysis starts with a model of early childhood development in which households with heterogeneous structures, incomes, skills, and preferences make a rich set of choices regarding the quantity and quality of maternal and nonmaternal child care. Children's cognitive skills develop through the combination of these care experiences according to a production technology that allows for dynamic complementarity between the child's initial human capital endowment and multiple investments in care quality in the first years of life. To estimate the model, we combine information from a large, randomized-control, early-childhood care experiment from the 1980s, the Infant Health and Development Program (IHDP), with several nationally-representative, observational datasets. The combination of observational and experimental evidence identifies key parameters of the model that would not be identifiable from either of the sources alone. The estimated model allows us to understand the mechanisms underlying heterogeneous, *ex post* effects of the IHDP treatment, and provides a credible basis for estimating important economic parameters and simulating *ex ante* effects of a variety of early childhood policies.

The paper makes four main contributions. First, we extend the literature on models of child development by integrating new dimensions of choice into a tractable framework, offering new insights into the tradeoffs parents face. Following in the path of previous structural models of child development including Blau and Hagy [1998], Bernal and Keane [2010], Del Boca et al. [2014], and Griffen [2019], we model mothers' decisions about labor supply jointly with how much time to spend providing maternal care and how much nonmaternal care to use. In

[1] In the first three years of life, children's brains develop critically-important neural structures and functions, which shape lifelong development and outcomes [Knudsen et al., 2006]. The relationship between SES and average gray-matter volume is weak in the first year of life, but large SES-based gaps emerge between ages 1 and 3 as average gray-matter volume becomes strongly and positively correlated with SES [Hanson et al., 2013]. By age 5, reading and math achievement is strongly correlated with family income [Heckman, 2006; Reardon, 2011; Figlio et al., 2014]. Gaps in cognitive skills existing at that point persist throughout childhood without further widening and have strong relationships with adult productivity and life success [Cunha et al., 2006; Council of Economic Advisers, 2016b]. The early childhood period matters for human development.

[2] Children from birth through age 4 are covered by publicly-financed early care and education programs for an average of only 5 hours per week during standard business hours, equivalent to 9 percent of children's time during this period. In contrast, kids age 6 to 12 spend just over half these hours in publicly-financed care and education [Council of Economic Advisers, 2016a]. And, eligible children are more than 11 times as likely to be served by the federal Head Start program each year when age 3-4 than when age 0-2, via the much smaller Early Head Start program (Authors' calculations from enrollment figures in Elango et al. [2015]). See also Hotz and Wiswall [2019].

contrast to these studies, which primarily model the quantity of child care time, one of our main contributions is to allow for a rich space for heterogeneous and endogenous child care quality as well.[3] Several other studies, including Cunha et al. [2010], Agostinelli and Wiswall [2020], and Attanasio et al. [forthcoming], use various measures, such as number of books in the home and number of times parents read to their children, to capture some notion of quality and quantity of parental investments, but without explicit models of household decision making or child care.

Our model is of one child and a unitary household decision maker, which we refer to as the mother because data indicate that mothers provide the vast majority of child care hours. In our framework, a household can purchase higher quality child care on the private market, allowing higher-income households to provide higher quality care for their children. We also allow the quality of maternal care to vary, depending on maternal characteristics (human capital), household structure (number and ages of children), and through an endogenous "parenting effort" choice. In this way, maternal care quality and labor market activity can be correlated. And, our model captures the fact that parenting and other forms of caregiving are demanding, effortful activities. To our knowledge, modeling effort is novel in the economics of parenting but related to common models of worker productivity. This model offers a new channel by which child care subsidies can increase maternal-care quality; by reducing the number of maternal-care hours, subsidies can lead mothers to provide higher quality maternal care.[4] We identify the magnitude of this relationship leveraging the IHDP experiment and measures of maternal-care quantity and quality to causally identify the effect of maternal-care hours on maternal-care quality.

Second, our model's rich choice set and combination of experimental and observational data provides a general framework to understand standard experimental treatment effect estimates. Our paper stands at the intersection of two strands of the child development literature: one focuses on evaluating existing programs, either smaller demonstration programs (like the famous Perry and Abcedarian programs) or larger scale programs like Head Start, typically through random-control trials (RCT) [Elango et al., 2015; Duncan and Magnuson, 2013], and the other strand that uses structural models of child care choice, as discussed above. As in Kline and Walter's study of the Head Start Impact Study RCT [Kline and Walters, 2016], we show how the effects of program take-up on child development critically depends on the counterfactual alternative child care arrangement: offers of subsidized care have larger effects on households which would otherwise provide lower-quality care to their children.[5] We generalize the Kline and

[3] Blau and Hagy [1998] model demand for quantities of different child care modes, and child care attributes (caregiver-to-child ratio) which are possibly correlated with quality. Del Boca et al. [2014] models the number of hours of "passive" and "active" time from mothers and father separately, and these different kinds of time have different productivities in producing child cognitive skills. See a recent extension in this framework [Mullins, 2019]. Griffen [2019] assumes that for every 6 month period, non-Head Start eligible households receive one non-maternal childcare offer (a quality and price pair) and can choose the number of hours to take-up (none, part-time, or full-time). Quality of mother's care is exogenous, but mothers can endogenously affect how much children are exposed to this home quality by adjusting child-care hours. A number of other recent papers consider related issues but in a simpler child care decision space, see for example Guner et al. [2018], Caucutt and Lochner [2017], Abott [2019], and Daruich [2018].

[4] This notion is already captured in the psychology literature. Parenting less may create psychological space to parent better and avoid parenting burnout [Roskam et al., 2018].

[5] In this way, this issue relates to the larger literature on program evaluation when the programs have "close"

Walters approach, which estimates a model of the choice between 3 homogeneous care types. Our model allows for multiple alternative arrangements which vary continuously in intensity and quality, and allows multiple channels through which program take-up itself endogenously affects the household, through effects on labor supply, income, and expenditures on private care, and through changing parenting effort by providing participating mothers child care relief. By combining the standard data on program effects on child outcomes with rich information about effects on the household more generally, we identify various spillovers by which program take-up affects child development.

This paper builds on a long literature studying the IHDP's *ex post* average treatment effects on maternal and child outcomes and how effects differed conditional on child birth weight, maternal education, and other observables.[6] Although selected on low birth weight and early birth, unlike many other experimental studies whose study population consisted of only disadvantaged children, the IHDP sample is economically diverse. One of the key findings from this experiment is that this intervention had large positive effects on child IQ at age 3 for the low SES children, but almost no effect on the high SES children. And, these effects on child cognitive skill fade down but remain positive, substantial, and significant for many years after the end of the treatment for children from low-income families in the higher birthweight subsample [Duncan and Sojourner, 2013]. Our study leverages the model as a unified lens through which to interpret the complex patterns of *ex post* treatment effects [Wolpin, 2013], in particular the larger effects on children from lower SES households.

Third, we integrate experimental and observational data in a novel way, and the random-assignment experiment at the heart of the IHDP provides valuable leverage to evaluate the credibility of our identifying conditions. We develop an identification strategy for the effect of maternal hours on parenting quality using the random assignment of the treatment offer as an instrument for maternal care hours. We provide an identification argument for preference heterogeneity without strong parametric assumptions using observed continuous choices. Finally, we validate the structural model showing that we can replicate the experiment's control-group data by counterfactually removing the treatment offer from the treatment group. This is similar to a typical validation exercise of counterfactually providing the treatment to the control group [Todd and Wolpin, 2006; Griffen, 2019]. We do the converse because the program specific parameters are not identified in the control sample.

The final contribution is that our estimated model allows *ex ante* simulation of the effects of many hotly-debated, child care and family subsidy policies. The model represents the IHDP treatment as one possible child care subsidy policy with a particular quality and price (free). Our analytic framework facilitates counterfactual policy simulations of alternative in-kind child care programs, varying the quality of the program offer and price (co-pay), and then simulating

substitutes, see for example [Heckman et al., 2000].

[6]Gross et al. [1997] describes the experiment and initial analysis of effects through age-3. A series of studies in child development and medicine reported treatment effects on various outcomes such as child cognitive skill and behavior [Brooks-Gunn et al., 1993], quality of the parenting and the home environment [Bradley et al., 1994], quality of parenting and maternal employment [Brooks-Gunn et al., 1994], and the use of paid child care [Gross et al., 1997]. Berlin et al. [1998] studied mechanisms focusing especially on heterogeneity along demographic lines. They find that child cognitive effects are larger among those who take-up more care. Duncan and Sojourner [2013] and Chaparro and Sojourner [2015] look at heterogeneous effects without the lens of a structural model.

endogenous take-up of the counterfactual program offers and their heterogeneous effects on child development, maternal labor supply, maternal-care quality, and demand for nonmaternal care. Our framework also allows simulation of counterfactual policies that replace the offer of a specific subsidized program with different types of unrestricted income transfers, allowing households to choose their child care as desired.

We highlight four key empirical results. First, we find a substantial quantitative importance of the "exhaustion effect," and estimate that reducing maternal-care hours increases average maternal-care quality. This channel is important. The program offer raised measured maternal quality by about 1/4 of a standard deviation. And, using the estimated model, if we do not allow adjustment on the effort margin, the average predicted treatment effect of the program offer falls by about 10 percent. Second, we find that the offer of access to high-quality, in-kind care boosts average child cognitive skill by more than 4 times as much as an equivalent-cost cash transfer to the family, though effects of each are bigger among mothers with less education. Third, the quality of the subsidized care offered is of first-order importance. Controlling for the quantity and quality of other care used in our general framework, we estimate that the IHDP program care was of very high quality and estimate how counterfactual effects on child skill, maternal time allocation, and other choices would differ depending on the quality of the program-care offered. For instance, reducing offered program-care quality to the average nonmaternal-care quality observed in the control group eliminates most of the average effect on child skill. Finally, given observed heterogeneity in program effects, the composition of the program take-up group governs the overall effect of the program on children. For households with good alternative care, taking up a low-quality program would have negative effects on child skill development. We show how this finding can rationalize why some universal expansions of subsidized child care, such as in Québec, have negative effects [Baker et al., 2015, 2008].

The remainder of our paper is organized as follows. The next section describes the data and some main empirical findings. The following sections describe the model and our identification and estimation strategy. The final sections discuss results and counterfactual policies. We conclude by discussing some limitations of our approach and directions for future research.

2 Setting

The Infant Health and Development Program (IHDP) was an experimental intervention carried out in eight U.S. cities during the late 1980s to support the healthy development of babies born low birthweight (≤ 2.5 kg) and premature (≤ 37 weeks gestational age at birth). Researchers recruited families into the study in hospitals at an eligible child's birth and randomly assigned each family to a treatment (one-third chance) or control group.[7] The treatment group was offered free, high-quality center-based child care service between the ages of 12- and 36-months. Child care services were provided at a researcher-established Child Development Center (CDC) in each city for up to nine hours per day (45 hours maximum each week) over those 2 years. Caregivers followed an Abecedarian-type game-based curriculum that emphasized language

[7]Recruitment occurred in eight university hospitals in Boston, Dallas, Little Rock, Miami, New Haven, New York City, Philadelphia, and Seattle.

development [Gross et al., 1997], and the staff was paid substantially more than typical day care teachers.[8] Unlike many early childhood interventions, family income was not a selection criteria for participation and an economically diverse set of families participated. The IHDP also collected rich data on the quantity and quality of maternal and non-maternal care, the allocation of mothers' time and each child's skill development.[9]

We focus on the IHDP's 880 singleton infants.[10] The remainder of this section describes the sample's baseline characteristics at the time of the child's birth and intent-to-treat effects of the program offer, both overall and by maternal-education subgroup. Although results are similar to those described in the prior literature, we present them here as context for readers unfamiliar with the IHDP and as motivation for the structural model we present next. Birth weight averages 1.8 kilograms (4 pounds) and gestational age at birth averages 33 weeks. The sample includes never-married (50 percent), married (42 percent), and previously-married mothers (8 percent). On average, mothers are 25 years old at the time of the child's birth. Forty-three percent had not completed high school, 27 percent had only a high school diploma, 17 percent had some college but no bachelor's degree, and 13 percent had at least a bachelor's degree. Most mothers are either African-American (52 percent) or non-Hispanic white (34 percent). Panel A of Table 1 presents basic descriptive statistics broken out by treatment status and a balance test on each observable. Overall, treatment groups are balanced at baseline (p=0.79 of joint test).

Panel B of Table 1 reports the means by treatment group and average intent-to-treat effects on children's cognitive skill, time allocation, and quality of care. Overall, the child cognitive skill effect at the end of the intervention (age 36-months) is 9.5 IQ points, about two-thirds of a nationally-normed standard deviation.[11]

Underneath this average effect, child cognitive-skill effects differ greatly depending on maternal education (Table 2). There is no significant cognitive-skill effect among children of mothers with at least a bachelor's degree (2.6 points, $p = 0.48$), but children of mothers with less formal

[8] Gross et al. [1997]: "whereas most day care teachers earned between $12,000 and $18,000 per year, IHDP child development teachers earned between $15,000 and $22,000 (1986-1988$).... higher salaries were essential to the project's success in hiring and keeping qualified and experienced staff."

[9] The treatment group was also offered weekly home visits between birth and age 12-months, which was intended to shift maternal beliefs and expectations about child development and parenting. Our model ignores this element of the treatment and focuses only on the child care and education offer. This choice could be problematic if the home-visiting program shifted maternal beliefs and knowledge. However, plentiful evidence suggests this did not occur. The IHDP collected data on maternal beliefs and knowledge 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 measures [Gross et al., 1997]. Also, there was no evidence of a treatment effect on the HOME environment inventory nor on child skill at 12-months. 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 et al. [1997] report no effect on age-1 cognitive skill, as measured by the Bayley mental development scale. Home visits were not conducted by nurses, which research has shown limits their effectiveness [Olds et al., 2002].

[10] The IHDP included another 105 twin pairs but tradeoffs and production technologies may be quite different here. We take account of family size and ages of other children in the household in our modeling of the maternal/home care quality.

[11] Child cognitive skill is measured by Stanford Binet IQ at 36 months, with mean 100 and standard deviation 15 in the national population. Child age is chronologically-corrected based on due date.

education obtain large cognitive-development benefit from the intervention. Among children of mothers with only a high-school degree, treatment increases IQ by a full standard deviation (16.4 points, $p < 0.01$), and, for children whose mothers did not complete high school, the treatment lifts child IQ by about two-thirds of a standard deviation (9.1 points, $p < 0.01$).[12]

One might suspect that large differences in effects on child skill by maternal education group derive from differences in take-up of the IHDP's offer of free CDC care, but this is not the case. Average hours of program care taken up do not vary significantly by maternal education. On average, treatment-group children whose mothers did not have a high-school degree spent 17.7 hours weekly in CDC care. This is neither statistically nor substantially different than the 19.7, 18.2, and 16.2 hours weekly for children of mothers with high-school only, some college, and bachelor's degree, respectively. The treatment and its take-up are the same across maternal education group but its effects on children's development are very different. Understanding differences in the re-allocation of time and resources across family types in reaction to the treatment offer is a primary motivation for the present study.

Mothers with less education use program care largely as a substitute for maternal care but those with more education tend to use it to substitute for other, nonmaternal care. In the control group, the children of mothers with less education spend more time in maternal care than other children do. Given free access to a CDC, mothers without a high-school degree reduce average weekly maternal care by 13.3 hours ($p < 0.01$) and reduce other nonmaternal care less (-4.4 hours, $p = 0.01$). At the other end of the maternal-education spectrum, mothers with a bachelor's degree do not substantially reduce maternal-care hours (treatment effect = -1.9 hours weekly, $p = 0.53$) but cut back sharply on other nonmaternal care (-14.3 hours, $p < 0.01$). There is no significant effect on maternal hours spent earning in the labor market overall, although there are marginal increases of 3-4 earning hours among mothers with less education.

Treatment effects on maternal-care quality also differ sharply by maternal education. Among mothers with less education – who cut back maternal care hours, increase labor supply, and whose children experience large positive effects on cognitive skill – treatment raises maternal-care quality by over a third of a (control-group) standard deviation. In contrast, treatment has no evident effect on maternal-care quality among mothers with a bachelor's degree, those whose maternal-care and earning hours change little and whose children do not experience a cognitive-skill effect.

Non-monotonicity of child-skill effects in maternal education is consistent with the operation of countervailing channels. For instance, as maternal education rises, the child-skill returns to substituting CDC care for maternal care may fall, diminishing the incentive to replace maternal care with CDC care. However, hourly earnings from maternal earning increase with maternal education, strengthening the incentive to replace maternal care with CDC care. These factors

[12]The effects on child cognitive skill fade down but remain positive, substantial, and significant for many years after the end of the treatment for children from low-income families in the higher birthweight subsample (2,000-2,500 grams) [Duncan and Sojourner, 2013]. As discussed there, these results suggest great promise for offers of high-quality child care to close skill gaps between children from higher- and low-income families from either targeted or universal programs, especially given evidence that the skill-development process is smooth in birth weight above 800 grams (Figlio et al. [2014]: Fig. 5).

could also be confounded with maternal preferences for child skill, "leisure" time, parenting time, or with the productivity of care quality in generating child skill. In the following sections, our model and empirical analysis serve as a unified lens through which to interpret this rich, *ex post*, reduced-form evidence, to disentangle mechanisms, and to harness empirical variation to answer additional questions.

3 Model

Our model embeds early cognitive-skill production within household resource-allocation choices. The framework imposes a child care time constraint and features a rich space of child care options that households can select to meet this constraint while balancing consequences for the family budget, maternal time-use, and child development. It represents care-subsidy policies in a way that is useful for expressing key tradeoffs parents face in their child's first years of life, interpreting experimental evidence, and predicting counterfactual policies.

3.1 Environment

The model is of one child and a unitary household decision maker, which we refer to as the mother.[13] It is a model of resource allocation for a child aged 12 to 36-months, during the period when the treatment group is eligible for free care in the IHDP program. A child enters the period endowed with a level of baseline skill [Cunha and Heckman, 2008; Aizer and Cunha, 2012]. Given this, and several other household characteristics, the mother jointly decides her labor supply and what combination of child care options to use. Through a child-development production technology, the child care inputs and the initial level of child skill produce the child's level of skill at 36 months. The mother's own labor income combined with non-labor income, including labor income from a spouse or transfers from a non-resident father, constrain the household's purchase of non-maternal care. Each mother's utility depends on consumption, hours and effort parenting the child, "leisure" hours, and the child's skill at the end of the period. The model allows for heterogeneity in the mother's skills in parenting and in the labor market, the mother's preferences, the child's initial endowments of skills, in family composition (number and ages of other children, if any) and marital status (never, currently, or formerly), and the father's contribution to family income, if any.

3.2 Child Care

Our model's rich child care decision space distinguishes it from previous work. Each type of care is characterized by a quantity, quality, and cost. Maternal-care quality depends on the mother's skills and a parenting effort decision. Program care has a fixed, externally-determined

[13] A single agent makes decisions. Whether we label the decision maker as the "parent," "household," "mother," or "father" does not fundamentally affect the modeling framework. The preferences we specify and ultimately estimate from observed choices are not necessarily the mother's or the father's preferences *per se*, but could be determined by some prior stage intra-household interaction between parents or between the mother and a non-father spouse.

quality and a maximum hours available, which depends on treatment status. If program care is available, the mother chooses how many hours to take up. Program care is free but we allow for individual heterogeneity in (dis)taste for its use, allowing an additional channel to understand incomplete program take-up. Other non-maternal care is available for purchase, with hourly cost increasing in quality.[14]

3.2.1 Quantities of Care

A young child requires supervision and care for a total of T_C waking hours per week, creating a child time constraint. The weekly time budget for every child is

$$\tau_m + \tau_p + \tau_n = T_C \tag{1}$$

where $\tau_m \geq 0$ is hours of maternal care, τ_p hours of program care with $\tau_p \in [0, \bar{\tau}_p]$, and $\tau_n \geq 0$ is non-maternal, non-program care.[15] Households in the treatment group can choose up to $\bar{\tau}_p = 45$ hours per week. In the control group, this care is not offered so $\bar{\tau}_p = 0$.

3.2.2 Qualities and Costs of Care

Care differs in quality, where quality is defined with respect to its effect on child-skill development, as defined by the child skill production function, discussed below. There is a quality level for each of the three types of care: q_m denotes the quality of maternal care, q_p the quality of program care, and q_n the quality of other care. Each quality level is strictly positive: $q_j > 0$ for $j \in \{m, p, n\}$.

Maternal Care Maternal-care quality is both heterogeneous and endogenous. Mothers vary in their parenting ability and choose how much instantaneous parenting effort to exert. The quality of maternal care, q_m, is determined by

$$\ln q_m = X'\delta_{q_m} + \ln e \tag{2}$$

where X is a vector of maternal and household characteristics, including the mother's stock of human capital measured using education and an IQ score, the numbers of other children aged less than five years and aged five years or more in the household, and the mother's age. δ_{q_m} is the associated vector of quality-productivity parameters. This specification expresses the idea that parents can differ in their ability to produce higher care quality and that this ability can be correlated, via the mother's human capital, with earning power in the labor market. Maternal-care quality can also depend on the number of other children the mother has and their ages to capture her human-capital in parenting accumulated from experience raising

[14] Because in the period of the IHDP there is very little to no other government programs available for 12-36 month old children, we do not model an alternative government programs. Note that Head Start, a major federal program which exists during this period, is almost entirely available only to older children.

[15] This includes care at all other providers, such as the father or any other relative, childcare centers or home-based providers, neighbors, or nannies.

other children, possible spillovers of (in)attention in multiple-child families (congestion effects), or other sources of heterogeneity correlated with fertility.

The maternal-effort choice variable is $e \in (0, \infty)$, with $e \to 0$ indicating lowest effort. Effort is expressed directly in terms of maternal-care quality. Maternal effort in parenting affects maternal-care quality, and ultimately child development through the skill-production process described below.[16] The costs of maternal care are the time and effort required to produce a given quantity-quality combination, (τ_m, q_m), with shadow prices described below.

The inclusion of this parenting effort choice is a novel feature of the model. It borrows a convention from many principal-agent models in personnel economics and elsewhere such that productivity depends on observable worker characteristics as well as a latent, choice-of-effort variable. Allowing for maternal effort recognizes that it may be possible for all types of parents to produce high- or low-quality care. For a given mother, producing a higher level of maternal-care quality requires higher effort. Mothers with different levels of human capital and parenting experience (X) may require different levels of effort to produce the same quality.

Non-Maternal Care Non-maternal, non-program care is available at any quality level, q_n, and is chosen by the mother. The hourly price is $\pi q_n > 0$, with exogenous $\pi > 0$ is the price per "effective unit" (quality x quantity), $q_n \tau_n$, of such care. Total expenditure on this care is $\pi q_n \tau_n$.[17]

Program Care The quality of program care, q_p, is fixed by program designers and out of the mother's control. If offered program care, a mother chooses how much to use (τ_p). In the program we study, program care is free to the parents, and we experiment with counterfactual non-zero prices (co-pay) below. In addition to possible direct monetary costs, individuals also have heterogeneous utility costs (or benefits) to using program care, as specified below. Modeling the program as a particular type of non-maternal care available at a different cost and quality gives us leverage to use the experiment to model how families would respond to offers of subsidized care at counterfactual quality and price levels.

[16] Our specification is equivalent to $\ln q_m = X'\delta_{q_m} + \delta_e \ln e$, where we normalize the productivity of effort parameter $\delta_e = 1$. Effort e is then in "efficiency" units in terms of maternal-care quality.

[17] We return to the issue of the market availability of care of different qualities below and discuss how assumptions regarding the child care market affect the interpretation of our findings. We also discuss how we proxy for the price of care when care is provided by relatives and does not have an observed price. An alternative formulation of non-maternal care is found in Griffen [2019] where it is assumed households exogenously receive a stochastic draw of a child care quality and price pair. In that model, the exogenous quality and price draws depend on household characteristics (e.g. maternal education), thus allowing the model to fit the pattern of a positive correlation between child care quality and expenditure with family income. Our model takes a more traditional approach, treating non-maternal child care as a good purchased in a market, and child care is endogenously chosen given the household's preferences, productivity, and resources available. Several other papers model endogenous expenditures on children generally, but do not link these costs to the time spent caring for children, e.g. [Del Boca et al., 2014].

3.3 Child Skill Production Technology

Child skill is denoted $h_t > 0$ for periods $t = 0, 1$. h_0 is the initial stock of skills (at age 12 months), and h_1 is the level produced at the end of the investment period (at age 36 months). We specify the technology of early-childhood cognitive skill h_1 as a function of the initial skill level h_0, the quantities and the qualities of the three types of child care inputs, and a productivity shock η_h:

$$\ln h_1 = \delta_0(h_0) + \delta_m(h_0)\frac{\tau_m}{T_c}\ln q_m + \delta_n(h_0)\frac{\tau_n}{T_c}\ln q_n + \delta_p(h_0)\frac{\tau_p}{T_c}\ln q_p + \eta_h, \tag{3}$$

where the δ parameters express the productivity of the various inputs for age-3 skill. $\delta_0(h_0)$ is a general function of h_0 representing the contribution of the initial endowment. The level of "investment" in children is determined by both the qualities and quantities of care received by the child. Modeling investment as the combination of care experiences in this way is substantially different from many previous models which have generally focused on either care quality or quantities, given the kind of data available, but not both together.[18]

Our specification of the skill-production function incorporates two key generalities. First, the investment contribution of each type of care $\delta_j(h_0)\ln q_j$ is weighted by the fraction of time the child spends in that care: $(\tau_j/T_C) \in [0,1]$ for $j \in \{m, p, n\}$. Given the child's time constraint, $(\tau_m/T_C) + (\tau_n/T_C) + (\tau_p/T_C) = 1$, these time-allocation fractions can be considered endogenous "factor shares" and provide the relative "weighting" over the child care quality inputs. For example, if a child spends all of her time with her mother $(\tau_m/T_C = 1)$, so (log) child investment is simply the (log) quality of this mother's maternal care $\delta_m(h_0)\ln q_m$. Therefore, unlike some standard alternatives, this specification allows for well-defined "corner solutions" in child investments. In general, this function embodies the idea that improving the quality of each type of child care improves the overall level of child care investment in proportion to the endogenously-chosen fraction of time spent in that type of child care. To see this, note that the elasticity of child skill with respect to an improvement in maternal-care quality is:

$$\frac{\partial \ln h_1}{\partial \ln q_j} = (\tau_j/T_C)\delta_j(h_0) \text{ for } j \in \{m, p, n\}. \tag{4}$$

For maternal time, for example, our model implies that the productivity of increasing maternal care quality on child skill development is increasing in the fraction of time the mother actually spends with the child (τ_m/T_C). Both the quantity and quality inputs are endogenously and jointly determined in our model.

[18] For instance, Del Boca et al. [2014] model parental time and expenditures but not on particular types of care and without an endogenous effort channel. Cunha et al. [2010] measure the quality of care at home but not elsewhere nor the quantities of time spent in different care settings. Bernal and Keane [2010] measure if a child is in maternal care or not as the complement to whether the mother is earning full-time, part-time, or not at all. They do not measure variation in the quality of non-maternal care nor have direct measures of quality of maternal care. They proxy for it using maternal human capital. Kline and Walters [2016] measure if a child is in any non-parental care but not each household's number of hours or level of quality. The closest formulation to ours is found in Griffen [2019], which models investment, as we do here, as a weighted average of the childcare quality received by the child.

A second key feature of our technology specification is that $\delta_j(h_0)$ for $j \in \{m, p, n\}$ productivity parameters of each care type are functions of the child's initial skill, or endowment h_0. The elasticity of child-skill output with respect to any input (4) can depend on h_0. $\delta_j(h_0)$ increasing in h_0 would imply a higher return on investment for children with high levels of baseline skill than for children with low levels of baseline skill, a positive dynamic complementarity [Cunha and Heckman, 2008; Aizer and Cunha, 2012; Almond and Mazumder, 2013; Cunha et al., 2010; Attanasio et al., forthcoming, 2019]. $\delta_j(h_0)$ decreasing in h_0 would imply that investment productivity declines in initial skill, a negative complementarity. We let our data speak and estimate the relationship between investment productivity and initial skill without imposing a particular sign on the relationship.[19]

3.4 Other Household Constraints

Beyond the child time constraint (1) and the production technologies (2-3), there are two other household constraints: the maternal-time constraint (5) and the budget constraint (6). The mother divides her time endowment (T_M) between three activities: maternal child care (τ_m), earning wages through hours in the labor market (L), and leisure or other household production (l).[20]

$$\tau_m + L + l = T_M \tag{5}$$

For every hour in the labor market, L, the mother earns wage w.[21] Total maternal income equals labor earnings plus any non-labor income (Y). Non-labor income includes spousal income, transfer income from the government, and any capital income. Total income can be used to pay for consumption (c) and to purchase child care via expenditure $\pi q_n \tau_n$.

$$c + \pi q_n \tau_n = (1 - \nu(w, L, Y, X))(wL + Y) \tag{6}$$

where $\nu(w, L, Y, X) \in [0, 1]$ is the income tax rate, which depends on the labor income, non-labor income (primarily spouse's income for married women), and household characteristics X (marital status and number of children).[22]

[19] A limitation of the standard constant-returns-to-scale CES production function in this context, $h_1 = (\gamma I^\sigma + (1-\gamma h_0^\sigma)^(1/\sigma)$ with I being some investment, is that complementarities across inputs (I and h_0) are constrained to be nonnegative, $\frac{\partial h_1}{\partial I \partial h_0} \geq 0$. The CES form in the child development context assumes that the marginal productivity of investment is nondecreasing in the child's initial endowment. Although this is of course sensible in standard production technologies (e.g. capital and labor inputs producing some output good), this may not be appropriate in the child development context, where one of the inputs is the existing stock of skill. In contrast, our specification allows this complementarity to be either positive or negative. See Agostinelli and Wiswall [2020] for evidence of negative complementarities in child cognitive development at ages five and older, and Hoynes et al. [2014] for evidence of higher impacts of Head Start on low skill children.

[20] τ_m includes all time the mother is caring for the focal child while the child is awake, including time also engaging in other activities or with other people simultaneously. "Leisure" here includes home production and caring for other children if the focal child is not in her care.

[21] Our model also abstracts away from dynamic impacts of career interruptions, which can impact parents' incentives to take time away from work to provide care [Anderson et al., 2002; Adda et al., 2017].

[22] The tax function $\nu(w, L, Y, X)$ – an approximation of federal, state, and local taxes – is taken from Guner

3.5 Preferences

We assume the following form for mother/household/parental utility:

$$U = \ln c + \gamma_l \ln l + \gamma_{h_1} \ln h_1 + \gamma_{\tau_m} \ln(\tau_m + 1) - \gamma_{e,1} e \tau_m^{\gamma_{e,2}} + \gamma_p \tau_p. \tag{7}$$

We normalize utility parameters with respect to log consumption, which is assumed to have a marginal utility of 1. To conserve on notation, we do not write household level subscripts on the preference parameters, but as described below in the estimation section, we allow a rich level of individual heterogeneity in the full set of preferences given by the vector $\gamma = (\gamma_l, \gamma_{h_1}, \gamma_{\tau_m}, \gamma_{e,1}, \gamma_{e,2}, \gamma_p)$.

$\gamma_l \geq 0$ represents the preference for leisure. Importantly, given that "leisure" here includes caring for other children, this preference could depend on the number and ages of other children in the household. We discuss the heterogeneity of preferences in the next section. $\gamma_{h_1} \geq 0$ the preference for the child's human capital. We assume the household has a preference for the child's human capital directly. The value of γ_{h_1} for a household can represent a combination of altruistic preferences over the child's future consumption and the household's belief about how human capital at this early period generates the child's future labor market earnings and consumption.

The specification of the utility from maternal time τ_m and maternal effort (e) is the least standard aspect of our preference specification. It allows for the mother either to enjoy maternal time ($\gamma_{\tau_m} > 0$) or to have a distaste for it ($\gamma_{\tau_m} < 0$), beyond the utility implication of foregone leisure.

Utility from maternal time may also depend on effort expended in producing maternal care. Effort $e > 0$ is expressed as per each hour of maternal time τ_m. We write the marginal utility cost of maternal effort as depending on the level of effort and the time spent in care-giving: $c'(e) = \gamma_{e,1} \tau_m^{\gamma_{e,2}}$, with $\gamma_{e,1} \geq 0$, assuming a (weak) distaste for effort. With $\tau_m = 0$ (no maternal care), then there is no effort cost. If the mother devotes the minimum effort to her child's care ($e \to 0$), then the effort cost approaches 0, and the direct utility from any maternal time approaches $\gamma_{\tau_m} \ln(\tau_m + 1)$.

Parents' tradeoff between parenting quality and quantity has been missing from the economics literature, perhaps because datasets with both parenting time and parenting quality are rare.[23] This model captures the idea that high-quality parenting is more difficult to maintain over longer periods than shorter periods. Parenting can be exhausting. We are 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. A potentially non-linear effort cost in parenting time allows for the possibility that effort in care-giving becomes increasingly costly as parenting hours increase. The marginal cost of additional effort is $\gamma_{e,1} \tau_m^{\gamma_{e,2}}$. With $\gamma_{e,2} > 0$, the marginal cost of effort is increasing in maternal time, representing an "exhaustion effect" in maternal care. If $\gamma_{e,2} = 0$, then there is no exhaustion effect and the marginal cost of effort is invariant

et al. [2018]. $\nu(L) = \alpha + \beta \ln(52(Lw + Y)/53{,}006) + 0.045$, with α and β depending on the household marital status and number of children.

[23] The family-size literature focuses on parents trading off the quantity of children against their quality but not the hours and quality of parenting a particular child.

to the hours of maternal care provided. In the special case $\gamma_{e,1} = 0$ and $\gamma_{\tau_m} = 0$, the only "utility cost" to maternal time is foregone leisure, and is essentially the same preference specification as in some previous work, for example [Del Boca et al., 2014]. In the next section, we devote considerable attention to analyzing how we can use our available data and the exogenous experimental variation to identify these more general preferences for maternal care.

Finally, including hours of program care selected (τ_p) as an element of the utility function implies possible taste or distaste for program care. This is necessary to rationalize the choice of some households to pass up available high-quality, free care and to pay for lower-quality care, similar to Moffitt [1983] and Bernal and Keane [2010]. Individual heterogeneity in distaste (γ_p) expresses variation in, for instance, felt stigma, sensitivity to administrative hassle, and logistical challenges posed by nonstandard work hours or by having multiple young children with only one eligible for program care.

3.6 Timing and Decision Making

The mother first observes the child's initial level of human capital, her wage offer and non-labor income, and the productivity and price of various inputs, and then jointly decides her labor supply, time and effort in maternal care, and a quantity and quality of non-maternal care. This yields the level of child skill h_1.

The full vector of choices is $\Gamma = (c, q_n, \tau_n, e, \tau_m, l, L, \tau_p)$. The decision problem is:

$$\max_{\Gamma} U(c, l, h_1, e, \tau_m, \tau_p) \tag{8}$$

subject to the (i) child's time constraint (1), (ii) production technologies (2-3), (iii) maternal time constraint (5), (iv) budget constraint (6).

The model has five sources of heterogeneity: i) child's initial human-capital endowment h_0, ii) household preferences γ, iii) non-labor income, iv) mother's human capital and associated maternal care quality and wage offers, and v) household composition (mother's marital status and the number and age of other children).

There are three inter-related choices faced in the model: (i) the mother's time and effort allocation decision (between market work, care for the child with a chosen effort intensity, and all other uses), (ii) the mother's allocation of income between expenditures on non-maternal care and consumption, and (iii) the portfolio of child care types: maternal time, non-maternal care of varying quality, and subsidized program care, if offered.

3.7 Interior Solutions

To provide some intuition for the tradeoffs in the model and to set up our analysis of identification in the next section, we describe interior solutions for all choice variables. However, our full model-solution algorithm allows for corner solutions in time allocation (i.e. mother not working ($L = 0$), no program participation ($\tau_p = 0$), take-up of all program hours ($\tau_p = \bar{p}$), and exclusive maternal care ($\tau_m = T_C$)). Substituting the main constraints (5) and (6) into the utility function (7) yields the following optimization:

$$\max_{L,q_n,e,\tau_m,\tau_p} \ln[(1-\nu(L))(Lw+Y)-\pi q_n(T_C-\tau_m-\tau_p)]+\gamma_l \ln(T_M-L-\tau_m)$$
$$+\gamma_h[\delta_m(h_0)\frac{\tau_m}{T_C}(X'\delta_{q_m}+\ln e)+\delta_n(h_0)\frac{T_C-\tau_m-\tau_p}{T_C}\ln q_n+\delta_p(h_0)\frac{\tau_p}{T_C}\ln q_n]$$
$$+\gamma_{\tau_m}\ln(\tau_m+1)-\gamma_{e,1}e\tau_m^{\gamma_{e,2}}+\gamma_p\tau_p$$

where we have implicitly conditioned on observable characteristics and ignored some elements of the production technology that are not relevant for the optimal choices. The case we consider here is when program care is offered (treatment group). In the control group, this care is not offered so $\tau_p = 0$.

First-order conditions for $(L, q_n, e, \tau_m, \tau_p)$ illuminate the marginal costs and benefits balanced in each choice. For an interior solution to labor supply, we have a standard condition:

$$L\colon \quad \frac{w[1-\nu'(L)L-\nu(L)]}{c}-\frac{\gamma_l}{l}=0 \tag{9}$$

where $\nu'(L)$ is the derivative of the tax function with respect to labor supply. The optimal amount of maternal time allocated to earning in the labor market balances the marginal benefit from an after-tax consumption boost against the shadow value of lost leisure. The interior condition need not apply. We allow for corner solutions in earning ($L = 0$).

For purchased non-maternal quality, the FOC is given by

$$q_n\colon \quad -\frac{\pi\tau_n}{c}+\gamma_h\delta_n(h_0)(\tau_n/T_c)\frac{1}{q_n}=0. \tag{10}$$

The optimal level of nonmaternal-care quality balances the mother's marginal cost in foregone consumption from spending on higher-quality nonmaternal care against the marginal benefit to the mother of additional child skill generated by that marginal increase in nonmaternal-care quality.

The maternal effort FOC, given by,

$$e\colon \quad -\gamma_{e,1}\tau_m^{\gamma_{e,2}}+\gamma_h\delta_m(h_0)(\tau_m/T_c)\frac{1}{e}=0. \tag{11}$$

expresses how the optimal level of parenting effort balances the marginal cost of additional effort against the marginal benefit of additional child skill generated through improved maternal-care quality. Both the marginal cost and benefit depend on the quantity of endogenous maternal-care time.

The interior-solution FOC for maternal time is given by

$$\tau_m\colon \quad \frac{\pi q_n}{c}-\frac{\gamma_l}{l}+\frac{\gamma_{\tau_m}}{\tau_m+1}-\gamma_{e,1}\gamma_{e,2}e\tau_m^{\gamma_{e,2}-1}+\gamma_h(1/T_c)(\delta_m(h_0)\ln q_m-\delta_n(h_0)\ln q_n)=0. \tag{12}$$

The optimal amount of maternal time devoted to directly providing child care balances marginal costs and benefits across five channels. First, a marginal hour in maternal care generates a benefit by raising consumption by the saved cost of nonmaternal care. Second, an hour of maternal care costs an hour of maternal leisure. Third, there is the direct preference for maternal care: utility may be affected, positively or negatively according to the sign of γ_{τ_m}, by the addition of an hour of maternal care. Fourth, because effort is expended across all maternal hours, marginal maternal care time requires additional dis-utility from additional effort. Fifth, maternal time has an impact on child skill development. Whether the impact is positive or negative depends on whether, on the margin, maternal care is more or less productive than the nonmaternal care it substitutes for: $\delta_m(h_0) \ln q_m - \delta_n(h_0) \ln q_n$.

Finally, the interior FOC for program time, when program care is offered, is:

$$\tau_p\text{:} \quad \frac{\pi q_n}{c} + \gamma_h(1/T_c)(\delta_p(h_0) - \delta_n(h_0) \ln q_n) + \gamma_p = 0 \tag{13}$$

The optimal amount of program care time balances costs and benefits across three channels. First, adding an hour of program care boosts maternal consumption by the cost of an hour of nonmaternal care that would otherwise be required. Second, it impacts child skill development: whether the impact is positive or negative depends on whether, on the margin, program care is more or less productive of child skill than nonmaternal care. In either case, the effect on take-up is magnified by the strength of the mother's preference for child skill. Third, there is a direct taste or dis-taste for the program.

This FOC expresses some basic insights into what drives families' reactions to an offer of program care. They will take up more hours of offered care if the quality of offered program care is higher. Even if offered care is of low quality, some households may substitute away from more productive forms of care to save parenting time and effort or nonmaternal care expenses. In that case, the offer of subsidized care would make those children *worse* off, although the decision maker–even an altruistic one–would be better off [Peltzman, 1973; Herbst and Tekin, 2010].

The model clarifies how the choice to allocate child time across maternal, nonmaternal, and program care requires balancing each type's particular combination of costs and benefits. It allows these to vary across family types in a way that sensibly reflects differences in maternal productivity, nonlabor income, preferences, and child endowment. It clarifies reasons that the same offer of subsidized care induces very different responses from families with different preferences and constraints.

3.8 Full Solution using Mixed Numerical/Analytic Methods

We conclude this Section by describing our full solution method for the model, including the possibility of corner solutions. Our functional forms facilitate a mixed numerical and analytic solution to the model. For each possible time allocation $(L, \tau_m, \tau_n, \tau_p)$, we solve for the optimal maternal effort choice e and the optimal non-maternal quality expenditure q_n. Given the specifications described above, the solutions for e and q_n must be at interior solutions, with q_n undefined if the child spends no time in non-maternal care and q_m undefined if no maternal effort. Importantly, our numerical solution to the problem explicitly allows "corner solutions". There are a number of "corner solutions" possible: households can choose no maternal labor supply ($L = 0$), no non-maternal care ($\tau_n = 0$), no use of the program care ($\tau_p = 0$), or take-up of all program hours ($\tau_p = \bar{p}$). However, for each of these corner solutions, there is still a well-defined interior solution for the maternal effort and quality of care purchased, when applicable. This mixed numerical and analytic solution is both tractable and allows for important model generalities.

Our solution algorithm proceeds as follows. For each possible time allocation configuration $(L, \tau_m, \tau_n, \tau_p)$, the "conditional" (on time allocation) solution for maternal effort is found by solving (11):

$$e^*(\tau_m) = \frac{\gamma_h \Delta_e}{\gamma_{e,1} \tau_m^{\gamma_{e,2}}} \tag{14}$$

where $\Delta_e = \delta_m(h_0)\frac{\tau_m}{T_c}$.

Likewise, the conditional solution for maternal quality from (10) is

$$q_n^*(L, \tau_n) = R(L) \frac{\gamma_h \Delta_{q_n}}{\pi \tau_n (1 + \gamma_h \Delta_{q_n})} \tag{15}$$

where $R(L)$ is the level of after-tax income (given labor supply L): $R(L) = (1 - \nu)(wL + Y)$ and $\Delta_{q_n} = \delta_n(h_0)\frac{\tau_n}{T_c}$.

Our mixed numerical-analytic solution to the model (8) uses a grid over the time allocation choices (imposing the appropriate time constraints, depending on whether the program care is offered) and computes $e^*(\tau_m)$ and $q_n^*(L, \tau_n)$ for each feasible time allocation choice on the grid. For each combination of choices on the grid, we compute the utility (7), and the optimal choice is the highest utility choice vector on the grid. The specification of the utility function and the direct tradeoffs in time allocation imposed by the time and budget constraints ensure a unique solution to the problem. This mixed numerical-analytic algorithm not only easily allows for corner solutions, it is also computationally fast.

4 Identification and Estimation

In this section, we analyze the identification of the model and describe our multiple-step estimation procedure. Given the importance of credible identification of model primitives for the counterfactual exercises to follow, we devote considerable space to this analysis. In general, our identification and estimation algorithm relies on two sources of identification: i) rich data including measures of quality and quantity of child care, and ii) exogenous variation from the IHDP RCT experiment. Further details about the data sources is provided in the Appendix.

We need to estimate two sets of parameters. One characterizes the production technologies and resources available to each household, comprising the main constraints in the model. The second characterizes the distribution of preferences. We discuss each in turn, and the identification concepts we use are constructive, leading to a transparent multiple-step estimator.

4.1 Child Skill Production Technology

Our identification of the production technology (3) relies on the unusually rich data collected in the IHDP experiment, the availability of exogenous variation due to the experimental design, and supplemental data from other datasets. We use the 36-month IQ score as the measure of $\ln h_1$ and the 12-month Bayley mental development index score as the measure of $\ln h_0$, standardized in the control group.[24] The quantity of time in maternal, non-maternal, and IHDP program care (τ_m, τ_n, τ_p) are observed.

We measure endogenous maternal-care quality (q_m) with direct observations of mother-child interactions and the home environment. We focus on a subset of the age-36 month Home Observation for Measuring the Environment (HOME) that prior research finds to have a strong theoretical and empirical relationship to child cognitive skill growth, the 12-item subscale measuring parental Support for Learning and Literacy [Fuligni et al., 2004]. We measure $\ln q_m$ as the sum of the 12 binary Learning and Literacy items, standardized in the control group. The $\delta_m(h_0)$ parameter relates maternal-care quality to age-3 child cognitive skill.

We measure non-maternal care quality (q_n) using hourly expenditures on such care, supplemented by a model of observation-based scores when expenditure data are missing. We normalize the non-maternal care price (π) to 1, implying that q_n is in units of expenditure on child care per hour. Given this normalization, q_n is directly measured from the ratio of reported expenditures on and hours in non-parental child care. For households who do not report paying for child care directly, we proxy for the quality of care by exploiting data from the National Institute for Child Health & Human Development's Study of Early child care and Youth Development (SECCYD). It covers a distinct, nonexperimental sample that overlaps in calendar time with the IHDP sample. Both the IHDP and SECCYD include rich family background variables and data on primary nonmaternal care category (paternal/partner, grandparent, other relative, nonrelative in child's own home, nonrelative day-care home, or nonrelative day-care center). The SECCYD contains a direct measure of care quality based on observed caregiver-child interactions for the child's primary nonmaternal care provider. To construct a proxy for nonmaternal care quality in the IHDP subsample that does not pay directly for nonmaternal care, we use the rich set of covariates available in both datasets and construct an expenditure proxy, essentially an estimate of the effective market price of the nonmaternal care the household uses. More details are provided in the Appendix.

Finally, for program-care quality (q_p), we identify it directly from the variation in cognitive outcomes h_1 between the randomized treatment and control groups. We normalize $\ln q_p = 1$ and identify the effective program quality $\delta_p(h_0)$ jointly with the remaining parameters, treating program quality as homogeneous, but allowing its productivity to vary by the child's initial human capital h_0.

Putting all these elements together and adding i subscripts to denote child-i level variables, we write (3) as

[24] We also estimated models using various combinations of weight, gestational age, length, and head circumference at birth, which all produced similar results. [Aizer and Cunha, 2012] also focus on the 12-month Bayley mental development index as their preferred measure of initial endowment.

$$\ln h_{1i} = \delta_{h_0}(h_{0i}) + \delta_m(h_{0i})\frac{\tau_{mi}}{T_c}\ln q_{mi} + \delta_n(h_{0i})\frac{\tau_{ni}}{T_c}\ln q_{ni} + \delta_p(h_{0i})\frac{\tau_{pi}}{T_c} + \eta_{hi}, \tag{16}$$

For any measured level of initial child skill h_0, we identify the δ parameters in (16) given variation in the maternal, non-maternal, and program care inputs. The model structure implies that optimal decisions do not depend on the realization of the η_{hi} shock, hence we identify these production function parameters directly from the IHDP data without jointly solving the household behavior model.[25] Although direct identification of this part of the production technology is credible given the rich data available, as we discuss next, omitted-variable/endogeneity issues become critical in identifying other aspects of the production technology, and we exploit the experimental variation for identification there.

4.2 Maternal-Care Quality Function

As discussed above, maternal-care quality is measured using items from the age 36-month HOME inventory. Adding household level indexes i, the maternal-care quality function (2) becomes

$$\ln q_{mi} = X_i'\delta_{q_m} + \nu_i \tag{17}$$

where, given that maternal effort e is unobserved, the residual is $\nu_i \equiv \ln e_i$. We have two inextricably-linked tasks. One, identify the δ_{q_m} productivity parameters (e.g. the productivity of maternal human capital) in producing maternal-care quality. Two, identify the level of unobserved parenting effort (e) each mother is deploying. Because maternal characteristics generally relate to her endogenous effort decision (X is correlated with ν), we cannot credibly identify δ_{q_m} from a regression of $\ln q_m$ on X. For example, college-educated mothers have higher average q_m than mothers with less education. To what extent is this due to a difference in productive human capital ($X'\delta_{q_m}$) or to different levels of endogenous effort e_i?

We address this identification challenge by exploiting the exogenous variation from the IHDP experiment. Identification and the estimation procedure takes 3 steps:

Step 1: Identify the "exhaustion" effect ($\gamma_{e,2}$). Re-arranging the analytic solution for effort (14) implies that log effort for any household i is:

$$\ln e_i = \ln \gamma_{hi} + \ln \delta_m(h_{0i}) - \ln T_c - \ln \gamma_{e,1,i} + (1 - \gamma_{e,2}) \ln \tau_{mi} \tag{18}$$

Substituting into the maternal-care quality function (2) yields:

[25] The key assumption is the additive log separability of both the production and utility functions, implying that decisions on the margin depend on the δ parameters and the level of h_0 but not η. Given this, the OLS estimator consistently estimates δ parameters. In addition, the Appendix reports two robustness exercises: (i) we re-estimated the production technology also including a vector of household characteristics, and (ii) we re-estimated the production technology using an IV estimator, using the randomized treatment offer and this offer interacted with household characteristics as instruments. In both cases, we find the main counterfactual results are similar to those using the baseline estimation.

$$\ln q_{mi} = X_i'\delta_{q_m} + \alpha \ln \tau_{mi} + \chi_i \tag{19}$$

where $\alpha = 1-\gamma_{e,2}$ and χ_i is the sum of equation (18)'s remaining terms: $\chi_i = \ln \gamma_{hi} + \ln \delta_m(h_{0i}) - \ln T_C - \ln \gamma_{e,1,i}$. We assume $\gamma_{e,2}$ is homogeneous but allow $\gamma_{e,1,i}$ to be heterogeneous, imposing a common exhaustion effect across all households but allowing the marginal cost of effort to vary freely across households. With $(\ln q_{mi}, X_i, \tau_{mi})$ observed, we could obtain a naive OLS estimator of α and therefore $\gamma_{e,2} = -\alpha + 1$ from a regression following (19). However, this is an inconsistent estimator because maternal time τ_{mi} is likely correlated with the residual χ_i: maternal-care hours depend on heterogeneous preferences and productivities in χ_i.

To solve this endogenous-regressors problem, we use the experimental randomization as an instrument for maternal-care hours. Consider conditioning on the randomly-assigned treatment (offer of IHDP program care), treat $\in \{0, 1\}$. The difference between the average treatment effects on maternal-care quality and hours identifies $\gamma_{e,2}$:

$$\begin{aligned} &E(\ln q_m|\text{treat} = 1) - E(\ln q_m|\text{treat} = 0) \\ &\qquad = (1 - \gamma_{e,2})[E(\ln \tau_m|\text{treat} = 1) - E(\ln \tau_m|\text{treat} = 0)] \end{aligned}$$

This result follows because the experimental randomization is balanced across the X variables and is independent of remaining preference and productivity characteristics, i.e. $(X, \chi) \perp$ treat. This exclusion restriction assumes that the only way the program affects maternal-care quality is through the maternal-care hours adjustment and the subsequent psychological-relief channel.[26]

Step 2: Identify δ_{q_m}. Note that the main equation (19) has at least several endogenous variables: each element of X_i (mother and household characteristics) are likely correlated with the residual, representing unobserved parenting effort. Although we have a plausible instrument for maternal time, and can consistently estimate the "exhaustion effect" as described above, the IHDP experiment does not offer any additional exogenous variables we can use to instrument for the maternal and household characteristics in X_i.

Our strategy to identify the δ_{q_m} parameters is to use the first step estimation of the exhaustion effect, identified using the IHDP experiment, combined with a rich set of conditioning

[26] In particular, consider some alternatives. First, it assumes parents do not learn how to parent better from the program elements intended to improve parenting, such as the home visits during the first year and bimonthly parent meetings in years 2 and 3. Second, it assumes parents do not learn how to parent better from interactions with the CDC teachers during 12 to 36 months. Third, it assumes parents do not learn how to parent better directly from their young children. We review a few key factors that lend credibility of this assumption. First, as will be discussed in Table 3, there is no treatment effect on maternal-care quality at age 12-months, after the year of offered home visits but before the start of program-care access. Home visits were also not conducted by nurses, which research has shown limits their effectiveness [Olds et al., 2002]. Second, the IHDP provided free transportation to the program-care centers. This encouraged take up but reduced parent-teacher interaction. Third, as described by the IHDP principal investigators in Gross et al. [1997], "Significant main effects did not emerge on any of the measures of maternal knowledge or concepts of child development at 12, 24, or 36 months." Finally, recall that the maternal education groups experiencing the largest increases in maternal-care quality are the same ones experiencing the largest reductions in maternal-care hours, even though take-up of program care is about the same across the groups. We explore this in Figure 1. Behavior changed but not beliefs or knowledge.

variables included in the IHDP data. We first purge maternal quality from the contribution of maternal hours using the IV estimate of α (19) from the first stage: $\ln q_{mi} - \hat{\alpha} \ln \tau_{mi}$. We then regress this exhaustion effect purged measure of maternal quality on the X_i characteristics of interest and a set of pre-determined variables W_i to capture additional variation in effort. The optimal effort equation (18) from the model solution suggests relevant W_i variables that relate to the effort choice and are pre-determined before the experiment: initial skill level h_{0i} (as these relate to the productivity of effort), and a number of other variables, that could plausibly relate to preferences for maternal effort including a control-group normalized index of pre-natal maternal behaviors such as smoking, drinking, and drug use during the pregnancy, and quarter of first prenatal exam. We estimate δ_{q_m} using OLS applied to the following equation:

$$(\ln q_{mi} - \hat{\alpha} \ln \tau_{mi}) = X_i' \delta_{q_m} + W_i' \iota + \zeta_i$$

A sufficient condition for identification of δ_{q_m} is then $E(\zeta_i | X_i, W_i) = 0$.

Step 3: Identifying effort: e**.** The residual or "unexplained" level of each household's maternal-care quality identifies the level of effort expended:[27]

$$\ln e_i = \ln q_{mi} - X_i' \delta_{q_m}.$$

Note that our specification for maternal quality assumes there is no measurement error: any residual variation in the maternal quality measure represents unobserved effort, not measurement error. Although this is a strong assumption–as are any of the common assumptions that the data at hand measures exactly the variable of interest–without access to multiple measures of maternal quality, it would be difficult to identify separately measurement error from unobserved effort.[28]

4.3 Preferences

Our goal is to identify the preference vector $(\gamma_l, \gamma_{\tau_m}, \gamma_{e,1}, \gamma_{e,2}, \gamma_h, \gamma_p)$, up to the normalization $\gamma_c = 1$. Above, we used the measure of maternal-care quality and experimental variation to identify a common effort-disutility "curvature" parameter $(\gamma_{e,2})$, representing the "exhaustion

[27] Note the X_i vector contains no constant. We implicitly assume that $X_i = 0_{\dim\{X\}}$. A hypothetical mother with 0 years of schooling, age 0, with no other children, would have no human capital, and all maternal quality for this mother is due to effort. This is a kind of identification-at-infinity argument to identify the constant [Heckman and Navarro-Lozano, 2004].

[28] To see this difficulty, write maternal quality as $\ln q_{mi} = X_i' \delta_{q_m} + \ln e_i + \omega_i$, where ω_i is a measurement error term. Even under strong classical measurement error assumptions (ω_i is unbiased/mean-zero and independent of X_i and e_i), we still could not separately identify the distribution of "real" unobserved effort from measurement error without some additional parametric restrictions. For example, the observed conditional variance of maternal quality under the classical measurement error assumptions is given by $V(\ln q_m | X) = V(\ln e | X) + V(\omega)$, and it is not possible to identify separately the variance due to unobserved effort and that due to measurement error. In the Appendix, we report a robustness exercise where we re-estimate model parameters assuming classical Normal measurement error calibrated so that 50 percent of total variation is measurement error. The main counterfactual results under this alternative specification remain similar to the baseline results.

effect of parenting. This leaves 5 remaining preference parameters, which we allow to be heterogeneous across households.

Identification at the Interior We begin by considering non-parametric identification of the preferences if all households were at an interior solution. Although a minority of households are in fact at corner solutions, this analysis provides both a constructive component of identification for households at an interior, and accessible intuition for the sources of identification, which is typically quite murky in a structural model.

For each household i, we observe labor supply L_i, maternal time τ_{mi}, maternal-care quality q_{mi}, non-maternal quality q_{ni}, and maternal effort e_i. As discussed above, maternal-care quality is identified from direct observer measures of the mother-child interactions and the home environment, non-maternal quality is identified from per hour expenditures on non-maternal care, and maternal effort is identified from the "unexplained" residual maternal-care quality. Parameters of the maternal-care and child-skill production technology, (2) and (3), are assumed identified based on the analysis above. For this initial identification analysis, we also assume wage offers w_i and non-labor income Y_i for each household are observed.[29] We identify the 5-vector of each household's preferences $(\gamma_{hi}, \gamma_{e,1,i}, \gamma_{li}, \gamma_{\tau_m,i}, \gamma_{pi})$ from this 5-vector of household-level choice data $(L_i, \tau_{mi}, \tau_{pi}, q_{ni}, e_i)$. We discuss each element in turn.

First, inverting the non-maternal care quality expenditure solution (15) yields:

$$\gamma_{hi} = \frac{q_{ni}\pi\tau_{ni}}{\Delta_{q_{ni}}(R_i(L_i) - q_{ni}\pi\tau_{ni})} \tag{20}$$

where the price of non-maternal care is normalized to $\pi = 1$, $R_i(L_i)$ is after-tax income as a function of labor supply, and $\Delta_{q_{ni}} \equiv \delta_n(h_{0i})\frac{\tau_{ni}}{T_c}$. Higher expenditures on non-maternal quality (q_n), given household income $R(L)$, implies a stronger preference for child skill, i.e. a higher γ_h.

Similarly, inverting the optimal e solution (14), we solve for the effort preference that rationalizes the observed data:

$$\gamma_{e,1,i} = \frac{\gamma_{hi}\delta_m(h_{0i})\frac{\tau_{mi}}{T_c}}{e_i\tau_{mi}^{\gamma_{e,2}}} \tag{21}$$

Given the preference for child skill and maternal-care time, a higher effort level implies a smaller dis-taste for parenting effort $(\gamma_{e,1})$.

Similarly, the preference for leisure is found from the FOC for an interior solution (9):

$$\gamma_{li} = \frac{l_i}{c_i}w_i[1 - \nu'(L_i)L_i - \nu(L_i)] \tag{22}$$

where household i consumption c_i is derived from the income and expenditures data. Preference for leisure is identified from the ratio of leisure observed relative to consumption, given the effect on the after-tax wage rate: mothers observed with a higher ratio of leisure hours relative to consumption given their wage and tax rate are inferred to have a higher preference for leisure.

[29] We discuss the data that we use to identify these below.

Preference for maternal-care time is found by inverting its FOC (12):

$$\gamma_{\tau_{mi}} = (\tau_{mi}+1)\{\frac{\gamma_{li}}{l} - \frac{\pi q_{ni}}{c_i} + \gamma_{e,1,i}\gamma_{e,2}e_i\tau_{mi}^{\gamma_{e,2}-1} - \gamma_{hi}(1/T_C)(\delta_m(h_{0i})\ln q_{mi} - \delta_n(h_{0i})\ln q_{ni})\} \quad (23)$$

This expression indicates that a household's preference for maternal-care time increases in the observed amount of maternal-care time and declines in its productivity relative to endogenous nonmaternal care time.

Finally, the preference for program participation is identified by inverting the FOC for program-care hours (13):

$$\gamma_{pi} = -\frac{\pi q_{ni}}{R(L_i) - \pi q_{ni}(T_C - \tau_{mi} - \tau_{pi})} - \gamma_{hi}(1/T_C)(\delta_p(h_{0i}) - \delta_n(h_{0i})\ln q_{ni}) \quad (24)$$

where preference for program care increases in the observed use of the program given the relative quality of program care versus alternative care.

Implementation To implement this identification strategy as a practical estimator, we face two challenges. First, some households are not at the interior of the choice set, and we cannot non-parametrically identify their preferences.[30] Second, even for households that are in the interior and preferences are non-parametrically identified, exactly identifying the 5-vector of preference parameters for each sample household from a 5-vector of observed data for each household risks over-fitting, which may compromise the counterfactual analysis.

We solve these challenges by imposing some additional parametric structure on the distribution of preferences. We assume preferences are distributed jointly Normal (imposing the appropriate sign restrictions), with the mean of the joint distribution depending on a rich set of observable household characteristics, including marital status, maternal cognitive skill and education, numbers and ages of other children in the household, initial child skill at age 12 months, and pre-natal characteristics. We estimate this preference distribution using the control sample only and, given random assignment, assume the treatment sample shares the same (conditional on observables) distribution of preferences. Given their observable characteristics, the preferences for any given household, including those at corner solutions, are assumed to be a draw from this estimated distribution. The Appendix, particularly Appendix Table B-3, provides additional details.

4.4 Hourly Wage and Non-Labor Income

The IHDP does not contain direct measures of wages and non-labor income but measures strong determinants of them. To enrich the IHDP data, we draw on a parallel sample from the

[30] There are four possible corner solutions in our model: no labor supply ($L = 0$), all maternal time ($\tau_m = T_C$), no program time ($\tau_p = 0$), maximum hours program time ($\tau_p = \bar{\tau}_p$). The other corner solutions of no maternal time and only non-maternal time do not occur in the sample and we can safely ignore them. In our sample, about 8 percent of the IHDP sample has full-time maternal care, 40 percent have $L = 0$, 16 percent have $\tau_p = 0$ in treatment group, and no one is at maximum program hours (averaged across 2 years offered).

National Longitudinal Survey of Youth 1979 (NLSY). We construct a similar NLSY subsample to the IHDP, including only mothers who have children under 3, born 1983-1987, and whose children were born low-birth-weight and premature so they meet the IHDP eligibility criteria. Using this IHDP-like sample for the NLSY, we then estimate models of log hourly wages and log weekly non-labor income using predictor variables common to both datasets: maternal education, age, marital status, and maternal IQ score (PPVT in IHDP and AFQT in NSLY but both standardized to national norms). Given the inclusion of education and IQ scores, we argue that the bias from ignoring endogenous labor supply is minimal. Based on the estimated models (Appendix Table B-1), we impute a predicted wage offer and non-labor income for each IHDP sample member.

4.5 Estimation Algorithm

Our estimation algorithm has two main steps. In Step 1, we estimate the production technologies and resource functions. Given these, we then estimate preferences in Step 2. This has two valuable advantages and one small disadvantage over joint estimation of all parameters. First, it substantially reduces computation time. Second, identification of each subset of parameters is more clearly defined by isolating the sources of empirical variation key to their identification. The disadvantage is the familiar one, a loss of efficiency due to not imposing all model restrictions jointly. However, the precision of our parameter estimates is generally high so, in practice, we believe this is not a substantial concern. We conduct inference using the non-parametric bootstrap. We re-sample the data (each household) with replacement and, for every bootstrap sample, re-compute all parameter estimates at each step in sequence. Repeating each step of the estimation ensures that our standard errors accurately reflect the sampling variation at each step of the estimator.

5 Results

Before analyzing the policy simulations in the next section, in this Section, we first discuss estimates of key parameters and results of validity tests. We discuss model parameters directly, and in terms of implied average elasticities. We focus on the unique features of our model, and leave estimates of more standard features (e.g. wage offers) for the Appendix. We conclude this section by discussing the model fit of our estimates and an out-sample validity test.

5.1 Maternal-Care Quality

A central contribution of the paper is introducing a parenting-effort choice, which influences observed maternal-care quality, and proposing it as a channel by which child care subsidies can spillover into affecting parenting care. Allowing parents to parent less may allow them to parent better. When we take the model to the data, we use a specific measure of maternal-care quality, based on the Learning and Literacy index from the HOME inventory at 36-months [Fuligni et al., 2004; Chaparro and Sojourner, 2015]. Do differences in this really capture differences in maternal effort?

Our first indication that maternal-care time is tightly related to maternal-care quality comes from the experimental results reported in Table 2. For mothers with less than a bachelor's degree, there is both an increase in measured maternal-care quality (ranging from 1/4 to 1/3 of a standard deviation across the lower maternal-education subgroups) and a substantial and statistically significant reduction in maternal-care hours (ranging from about 9 to 13 hours per week). In contrast, for the high-education group, there is no statistically significant treatment effect on either maternal hours or quality. To get a more-refined view of whether the types of mothers who experienced larger negative treatment effects on maternal-care time were also the types that experienced larger positive treatment effects on maternal-care quality, we generalize this approach. We add several more baseline covariates (mother's test score, number of other children, and marital status) to maternal education and estimate conditional treatment effects on maternal-care time and quality. Figure 1 plots a binscatter of estimated treatment effects on these two variables conditional on these covariates. There is a strong, negative relationship. Types of mothers where treatment induces large reductions in maternal-care time also have large increases in maternal-care quality. Those with little change in maternal-care time have correspondingly little to no change in maternal-care quality.

We further explore the relationship between maternal-care time and quality using a number of placebo tests. Table 3's Panel A presents estimates of regressions of various measures of maternal-care quality on an indicator of the randomized treatment offer. Column (1) uses the 36-month Learning and Literacy subscale of the HOME inventory, our main measure of maternal quality ($\ln q_m$) as the dependent variable. Controlling for maternal education, cognitive test scores, and age, and numbers and ages of other children, we estimate that the treatment offer increases the Learning and Literacy index by about 0.25 standard deviations on average. To check whether this effect size reflects the treatment, rather than some failure of randomization, Column (2) uses the Learning and Literacy index measured at 12-months, before the offer of program care. We do not find a meaningful nor statistically significant effect of treatment, providing evidence that there were not systematic differences in maternal-care quality before access to program care began.[31] As an additional check, Column (3) repeats the analysis in Column (1), but using each household's difference in scores between 36 and 12 months as the dependent variable (a mother fixed-effect model). The estimated treatment effect is very similar to that using the 36-month level in Column (1).

Another concern is that the treatment effect we estimate may not identify the effect of parenting effort, but instead represent a wealth-effect from receiving free child care. Items in the Learning and Literacy index represent a mix of effort (e.g., child is encouraged to learn to read a few words) and low-cost goods strongly complementary to effort (e.g., child has at least 10 children's books). To test for a confounding wealth effect, we look for items in the whole 55-item, 36-month HOME inventory that measure goods that do not represent maternal effort nor complement effort. Four measures fit this category: home building appears safe, outside play environment appears safe, the neighborhood is aesthetically pleasing, and the house has at least 100 square feet of living space per person. If families have more resources from saved child care expenses or increased earnings, they could provide a better living situation, requiring

[31] This is also evidence that the home-visiting component of the treatment prior to 12 months did not affect maternal-care quality. See above for a more detailed discussion of this.

more financial resources but not more active parenting effort. We sum these four indicators and standardize in the control sample. The estimated treatment effect on this "not-effort" index is not significant (Column 4), again suggesting the channel for the estimated effect on maternal quality is through effort not wealth.

Building on the analysis in Panel A, Panel B of Table 3 presents the estimates for the primitives of the maternal quality model, using the randomized offer of the CDC care as an instrument for log maternal care hours. Consistent with the descriptive patterns described above, we find that the offer of CDC care is a strong instrument for maternal care hours, with the treatment offer estimated to reduce maternal hours by 16.5 percent. The 2SLS IV estimate indicates that a 1 percent increase in maternal hours reduces maternal quality by 0.015 standard deviations.[32] As described in the previous section, the IV estimate (α from equation 19), implies a convex cost to effort parameter of $\gamma_{e,2} = -\alpha + 1 = 2.5$. Our estimates indicate that more maternal-care hours make effort increasingly costly for mothers, reducing the effective quality per hour of maternal care.[33]

Finally, the top panel of Table 4 reports estimates of the remaining maternal-care quality production technology parameters (2). We estimate that a 1 standard deviation increase in the mother's own cognitive test score improves maternal-care quality (36-month Learning & Literacy score) by 0.31 standard deviations, consistent with maternal human capital as an important source of maternal-care quality heterogeneity.[34] We also estimate that maternal maturity has positive returns for productivity in maternal care. All else equal, a 30-year-old mother would produce about 0.20 standard deviations higher maternal quality than a 20-year-old mother. Mothers with more than 1 child in the household are estimated to produce lower-quality maternal care for the focal child. This is consistent with a congestion effect in maternal-care production that dominates any positive returns from parenting experience. This is also consistent with older siblings reducing the average quality of parental care by dividing parents' attention and negatively impacting younger siblings' cognitive-skill development [Kristensen and Bjerkedal, 2007; Price, 2008; Breining et al., 2019].

[32] Repeating the 2SLS estimation using maternal hours in levels, not logs, produces an estimate that a 1 hour per week increase in maternal hours reduces maternal quality by a statistically significant 0.026 standard deviations.

[33] Our evidence is consistent with prior work showing that subsidized child care can affect parenting. Love et al. [2005] found that parents with kids randomly selected for Early Head Start (age 0-2) eligibility raise the level of parenting quality and Gelber and Isen [2013] finds the same for Head Start (age 3-4). Gelber and Isen recognize that it could be due to changes in parent time with children through impacts on the parent's time constraint but lack good measures of parental care quantity to get at this directly. Chaparro and Sojourner [2015] report that this occurs in the IHDP subsample with low maternal wages and add that it doesn't in the higher-wage subsample.

[34] Note that with the inclusion of a maternal cognitive skill measures, maternal education is no longer a statistically significant source of care quality. These two measures of maternal human capital are strongly positively correlated, and dropping the test score measure leaves a strong and statistically significant effect of maternal education.

5.2 Cognitive Skill Production

The bottom panels of Table 4 report our estimates for the parameters of the child cognitive skill production function (3). The output of the production process is age-3 IQ, and parameters are interpretable in terms of that output. Age-3 IQ is in points, with a representative national mean of 100 and standard deviation of 15. We split our sample by initial endowment $\ln h_0$ (measured by 12-month Bayley mental development index) and allow for different parameters between children with below-median and above-median initial skill. This permits either complementarity or substitutability between initial endowment and early investments [Cunha and Heckman, 2008]. We do not find statistically-significant differences in productivity parameter estimates across children with above- versus below-median initial skill, suggesting little evidence of heterogeneity in the productivity of care inputs with respect to the level of initial skill and little support for dynamic complementarity at this developmental stage and for this sample.[35]

For the contribution of initial skill, we use a linear-in-logs function in initial skill with separate slopes and intercepts in each sub-sample (below or above median h_0): $\delta_{h_0}(h_{0i}) = \delta_{0,h_0} + \delta_{1,h_0} \ln h_{0i}$. We estimate that children's initial, 12-month level of skill (h_0) is productive for age-3 cognitive skill. Echoing previous studies [Aizer and Cunha, 2012] and [Cunha and Heckman, 2008], initial skill measured even at very early ages, is self-productive, i.e. "skill begets skill."

Remaining parameters describe the productivity of each type of care (maternal, non-maternal, IHDP program) in producing age-3 IQ. Two issues need to be considered in interpreting parameters. First, a *ceteris paribus* analysis is not strictly possible. An hour increase in one care type must be balanced by an hour reduction in another care type. Second, each productivity parameter provides the productivity given some quality level, and quality is not measured in the same way for each care type. And, our specification of the technology is a general non-linear one, as we allow for explicit complementarities or interactions between quality and quantity.

Beginning with non-maternal care, given the normalization on the price of non-maternal care ($\pi = 1$)–implying that quality is equal to hourly expenditure–δ_n indicates that an additional 1 percent increase in non-maternal expenditure raises IQ scores by δ_n/T_C for every hour of non-maternal care, where $T_C = 87.5$ (the total weekly hours of care each child receives). Given the estimates of δ_n around 16, a 1 percent increase in non-maternal care quality, i.e. hourly expenditure, increases IQ by about 0.18 points (0.012 standard deviations). For maternal care, log care quality is measured using (control-group standardized) scores on the Learning & Literacy dimension of the HOME instrument (as discussed above). The estimate of δ_m of around 10 implies that a 1 percent increase in maternal quality per hour would produce an increase in age-3 IQ of 0.1 points (0.007 standard deviations). We stress that caution should be used in interpreting these numbers as they involve manipulating endogenous variables; below we consider counterfactual changes in child care costs and other exogenous variables and the resulting effects on age-3 IQ, taking into account all endogenous behaviors.

[35] Comparing the two groups, none of the parameters are statistically different individually at the 10% level nor are they different jointly. We caution that given our sample sizes, this is not the ideal dataset to test thoroughly for dynamic complementarities. As a robustness check, we also split the sample at a higher Bayley cutoff, classifying Bayley scores at or above 0.5 standard deviations above our sample mean as high h_0. This yields similar results.

For IHDP program care, we assume that the program has a homogeneous quality (although potentially different effects on households of different child skill levels), and identify this quality level using the experiment and observed age-3 IQ. We can express the estimated quality of the program care in terms of the USD equivalent non-maternal care that would achieve the same level of age-3 skill. We solve for this equivalent hourly care expenditure x from $\delta_n \ln x = \delta_p$, or $x = \exp(\delta_p/\delta_n)$. The estimate of δ_p around 43 and δ_n around 16 implies that the quality of IHDP care is equivalent to private market care costing about $16.7 per hour (in 2018$), well above the average quality of non-program care families purchase in our data.[36] As shown in the Appendix, our estimate is above the 99th percentile for the IHDP sample (but below the highest value observed in this sample), and at around the 90th percentile of the hourly care cost distribution using a more recent 2012 sample. In comparison to actual program costs, combining the average of 18.2 hours of care used per week and the IHDP's weekly care cost of about $336 per child based on administrative spending records from the original study researchers [Gross et al., 1997], the hourly cost of the CDC is estimated at $18.46, somewhat higher than the market price we estimate.

5.3 Implied Elasticities

Beyond reporting model parameters directly, we can also compute elasticities implied by model primitives. Table 5 reports the average elasticity of endogenous outcomes with respect to three margins: an increase in the maternal wage offer (w), the hourly cost of nonmaternal care (π), and nonlabor income (Y). For example, a 1 percent increase in the maternal wage offer is estimated to cause a 0.07 percent increase in child cognitive skill (Column 2). A 1 percent increase in hourly child care price holding quality fixed reduces child cognitive skill by 0.08 percent.[37] Child cognitive skill responds very little to family income on average. In addition to these average elasticities, the model allows general heterogeneous responses. Appendix Table B-6 reports analogous elasticities separately by maternal-education subsample. Elasticities of child skill and maternal-care hours do not vary much by maternal education groups but elasticities of earning hours, maternal-care quality, and nonmaternal-care quality vary substantially, reflecting the importance of analyzing sub-groups separately.

Some of these elasticities can be compared with the literature, though many others are novel. We estimate that a 1 percent increase in child care prices reduces average maternal labor hours by 0.39 percent. This is similar to contemporaneous estimates which place this elasticity in a range from slightly positive to around -0.4.[38] There are relatively few studies of how child care costs relate to non-maternal care demand. We estimate a -0.43 average elasticity of nonmaternal care hours with respect to child care costs. This is similar but higher

[36] For above median families, the quality is estimated at $22 and for below median at $11.4, implying an average cost of 0.5*$22 + 0.5*$11.4 = $16.7.

[37] Recall that, in estimation, we normalized the price of nonmaternal care to $\pi = 1$. Changing π to 1.01 makes equivalent-quality care cost 1% more to families.

[38] Estimates vary by demographic groups and time period, along with methodology. A relatively recent summary of the literature [Morrissey, 2017] indicates that the estimates range from -0.025 to -1.1, with most studies providing estimates around -0.05 to -0.25. This review concludes that more recent estimates are smaller than previous ones using data from the 1990s.

than the comparable estimate of -0.21 (based on Blau and Hagy [1998]; reported on page 75 of Blau [2001]). Finally, our estimate of a 0.32 elasticity of nonmaternal-care quality demanded with respect to family income is much higher than the 0.001 Blau [2001] reports–essentially no demand for quality of child care. That study uses observable child care characteristics (such as the child-teacher ratio and teacher education level) as measures of quality rather than expenditures and observer-based ratings, as we do here. Our estimate of the labor supply elasticity with respect to the maternal wage averages 2.8 and appears higher among mothers with less formal education, echoing Keane and Wolpin [2010].[39]

5.4 Sample Fit and Validity Check

Table 6 displays within-sample fit for the control and treatment groups. The estimated model fits the mean and standard deviation of age-3 IQ, child time allocations, labor supply, maternal quality, and non-maternal care expenditures well. Two relatively minor issues stand out for our within-sample fit for the treatment group: The model under-predicts the treatment group mean age-3 IQ scores by about 1.25 IQ points, and over-predicts maternal hours (correspondingly under-predicts non-maternal hours) by about 5 hours per week. We match the treatment group's mean take-up of the program nearly exactly, about 18 hours in data and simulation.

In Columns (5)-(6) of Panel A, we also perform an out-of-sample validity check. The motivation for this exercise is past practice validating structural models by assessing the extent to which they can predict experimental results. Todd and Wolpin [2006], for example, estimate their structural model using control-group data only, and, then, counterfactually simulate the effects of the Mexican Progressa subsidy program using their model's estimates. The close match of the structural model's predicted effects to the actual observed treatment-group outcomes adds credibility to their model specification and estimates.

We cannot do this exact type of validation test here because we cannot identify the model within the control group. The control group never had access to IHDP program care and we make no *a priori* assumptions comparing the IHDP program care to care available to the control group, so we cannot identify the quality of program care nor the program-specific distaste from control-group data alone (we freely estimate both of these features using the treatment group data). Instead, we conduct a validation test where we put the treatment group counterfactually in the control condition (i.e., remove the treatment offer), simulate their choices, and report how well the model predicts what the treatment group would do as captured by what the control group did do. Thanks to the IHDP's random assignment of families to treatment group, we expect the same distribution of preferences and productivities in the control-condition treatment group as in the observed control group. Columns 5 and 6 present the means and standard deviations of the treatment group's simulated endogenous variables in the control condition. The similarity of columns 1-2 versus 5-6 expresses how well the model predicts families' choices and their impacts on children's cognitive skill. For instance, compare the observed mean (SD)

[39]We estimate 3.2, 3.6, 1.7, and 2.4 for those with no high school degree, high-school only, some college, and at least a bachelor's, respectively (Appendix Table B-6). This echoes estimated elasticities among 6 latent skill-preference types of women with ascending skill of: 9.2, 4.6, 3.9, 3.7, 1.2 and 0.6 from Keane and Wolpin [2010].

of the control group's child IQ of 86.6 (19.2) to the treatment-group's simulated mean (SD) of 87.2 (21.2) in the counterfactual control condition.

5.5 Mechanisms

Prior work on the effects of subsidized child care with economically-diverse samples in the U.S. [Bartik et al., 2012; Duncan and Sojourner, 2013; Cascio and Schanzenbach, 2013] and Europe [Fort et al., 2019; Drange and Havnes, 2019] finds heterogeneous reduced-form effects with respect to family income. However, the extent to which this heterogeneity is due to family income *per se* or to correlated factors–such as maternal wages and human capital, child endowment, non-maternal care quality, maternal-care quality, or preferences–remains unclear. Our estimated model allows us to generate relevant, new evidence.

To illuminate the mechanisms driving the IHDP effects, we simulate what would happen if the economic environment were different and report how average treatment effects of the program offer would change. If all mothers have nonlabor income equal to the average nonlabor income of mothers with at least a bachelor's degree (BA+) but nothing else changes, the average effect of the IHDP program offer on child cognitive skill would fall by more than half (Table 7: Columns (1) & (2)). (Recall that "nonlabor" income here is mainly the mother's spousal labor income.) This quantifies the share of treatment-effect heterogeneity by maternal education due to nonlabor income *per se*, rather than due to wage, endowment, preferences or other factors also correlated with education. If all mothers additionally have the average wage offer of BA+ mothers, the ATE of the program offer on child skill would essentially vanish, with the effect size declining to 0.5 points (Column 3). At least part of the treatment effect re-emerges once we counterfacually provide all households the BA+'s mothers' average productivity in maternal care and child endowment. The final experiment in particular demonstrates some quantitative importance of dynamic complementarity: even when households have higher resources (wage offers and non-labor income), we estimate that the effect of the program offer would be non-trivially positive if they also have a high level of child skill at 12 months. Lower rows display analogous results for endogenous variables besides child skill.

6 Policy Simulations

In this Section, we use our estimated model to simulate the effects of a set of counterfactual policies, and describe the lessons that emerge. First, we compare the distribution of effects from the existing program offer to two alternative programs: (a) a cost-equivalent cash transfer to households, and (b) a 25 percent increase in wages. Second, we change characteristics of the care subsidy from those observed in the IHDP experiment, simulate how different kinds of families react to these counterfactual programs, and estimate their effects on children. In summary, we consider 5 types of changes to the IHDP-CDC program:

- Replace the program with a cost-equivalent cash transfer
- Replace the program with a 25% maternal wage subsidy

- Vary program-care quality
- Vary program subsidy level (co-pay)
- Vary program targeting and take-up

Table 8 reports results for some of the main program changes we consider, including effects on average child IQ, time use, and care quantities and qualities. We next discuss each policy in turn.

6.1 Replacing IHDP with a Cash Transfer or Wage Subsidy

We first consider replacing the offer of IHDP-CDC care with an equally-costly cash transfer. In this scenario, each household receives the average weekly cost of the program $336 in 2018$ as an unrestricted non-labor income transfer (this is an annual transfer of $17,472). As discussed in Del Boca et al. [2014], unrestricted income transfers have the advantage of providing the highest welfare for the household (in these models, the parents), and, in this context, providing broad-based relief to all families, even to those who have a high stigma/hassle cost to a subsidized care program.[40] Disadvantages to unrestricted income transfers from the perspective of influencing child development is that the parents will "consume" part of the transfer and only part of it flows to children.

Figure 2 presents the distribution of treatment effects on age-3 IQ for the baseline IHDP program offer versus the cash transfer. The offer of the IHDP program has a much higher average effect on children's age-3 IQ. The equally-costly cash transfer policy boosts child cognitive skill by an average of only 0.1 standard deviations (1.5 points) compared to nearly 0.5 standard deviations (7.4 points) simulated effect of the program offer. In Table 8, comparing the cash-transfer (Column 3) versus the control condition (Column 1) indicates that households use the extra non-labor income to reduce maternal labor supply by about 10 hours per week, which they divide between a very small increase in maternal-care hours and a large increase in other uses of maternal time (caring for other children or leisure). The quality of maternal care falls slightly, consistent with an effort-channel response to more maternal care. But the largest change and the main source of the higher age-3 IQ for children is the increase in the quality of nonmaternal care purchased, from about $3 per hour to about $4.70. Average expenditure on child care ($q_n \tau_n$) then increases from $83 to $120 per week. As with IHDP program effects, the effect of a cash transfer is smaller among children of mothers with at least some college than among those with less formal education (Appendix Table B-7).[41] As in Del Boca et al. [2014], an unrestricted cash transfer has non-trivial positive average effects on children. But, as we demonstrate here, the effects are small relative to the in-kind offer of high-quality care, given that parents allocate only some of the resources to the children.

[40] From the public perspective, an additional advantage is that unrestricted transfers have lower administrative costs than in-kind care offers because there is no need for costly income verification.

[41] It should also be noted that the cash-transfer policy has a non-trivial, negative effect on some children (Figure 2). The income effect on labor supply causes some mothers to increase their maternal-care hours, which shifts some children to worse care environments.

Raising mothers' hourly wages by 25 percent costs only a quarter as much as the cash-transfer but delivers a similar effect on child skill.[42] The wage subsidy does have different effects on time use and care qualities. On average, mothers raise earning hours by about 5 hours (30 percent), half of that increase coming from reducing maternal-care and half from other time uses ("leisure"). In response to this policy, households on average increase maternal-care quality due to the effort channel and nonmaternal-care quality due to extra income.

6.2 Vary Program-Care Quality

The IHDP offered families access to high-quality, in-kind care, and produced large, positive average treatment effects. Next, we counterfactually vary the quality of the program care offered. For any given value of the δ_p quality parameter, we re-solve the model, including all of the endogenous household choices. Take-up of the program care offer is endogenous in our model and, in general, changes in the quality of the offered care affect take-up. As we have throughout, we assume households know this quality, and can choose any number of hours of the program offer (from 0 to 45 hours).

The top panel of Figure 3 displays how the ATE (average intent-to-treat effect) on child IQ from an offer of free program care varies with the quality of program care offered. The horizontal axis is program-care quality expressed as the dollar-per-hour price of equal-quality nonmaternal care, from near \$0 quality to \$25, with the baseline estimate for the IHDP program-care quality around \$16.7 per hour of care. The vertical axis displays the ATE on age-3 IQ of the program offer at these various quality levels.

As the quality of program care increases, the ATE increases. At the lowest program-qualities, the ATE starts near zero. Few households decide to take-up a low-quality program and the effect is small for the few households that do. As quality improves, the ATE increases, at an accelerating rate, as more households take-up the program and the effect on their children is larger. The horizontal, dashed line shows the experimental sample's ATE estimated directly from the data. The intersection of this line with our model's simulated ATE estimates at different program-care quality levels graphically illustrates how we identify the IHDP's program quality level.

If offered program-care quality were reduced to equal the control group's observed average nonmaternal-care quality (from about \$16.7 to \$3.14), the simulated average effect on child skill falls proportionally more, from 7.4 age-3 IQ points to less than 1 point (Table 8: Columns (5)-(1)). The bottom panel of Figure 3 shows how program-care quality affects the ATE in each maternal education group. We see a similar pattern within each subsample and smaller gains for children of mothers with the most formal education at every offered quality level. Figure 4 describes how the same program-care quality changes affect maternal- and child-time allocations. As program quality increases, take-up of the program increases substantially. There is a smaller negative effect on labor supply and maternal care and a larger reduction in other nonmaternal care.

[42] Although the model includes taxes in calculating each household's budget set, we do not consider taxes here. The 25-percent wage subsidy is pre-tax and the policy cost is the cost of the subsidy, not considering additional tax revenue generated by the additional labor income due to the policy change.

6.3 Universal program

Various communities around the world have implemented universal, in-kind, public child care programs and some have advocated adopting this approach in the U.S. Our results emphasize that the effects of expanding to create such a program depend critically on two factors: the quality of program care and, given the heterogeneity in effects, who is newly induced to take up the program. The latter depends critically on the nature of programs existing prior to a change. New compliers can come only from those not participating already.[43]

Aside from cost and quality, other characteristics of the program affect families' willingness to take-up the program. We next consider how varying the desirability of the program would change the program's overall effects. To bound the effects, we first consider an extreme case where take-up of the program was 100% and all households participated for the maximum 45 hours.[44] Table 8 shows that full take-up of the program (at the estimated baseline quality level) would increase the ATE on age-3 IQ from about 7 IQ points for the voluntary take-up offer in baseline to about 20 IQ points under this counterfactual. Full take-up almost eliminates the use of other nonmaternal care and reduces average maternal care by about 20 hours (33 percent). Earning hours increase by about an hour, but most of mothers' freed hours go to other uses. An important spillover from the full take-up of the program and the substantial reduction in maternal care time is an increase in effort and maternal quality by about 0.6 standard deviations. These effects contribute to the large increase in age-3 IQ from the program.

Figure 5 maintains the full take-up restriction but varies the quality of the program, as in the previous exercise, shown in Figure 3. Two findings are particularly salient. First, for low-quality programs, the overall ATE is *negative* with involuntary full-take-up. In contrast, as shown in the previous exercise in Figure 3, low-quality programs do not do as much "damage" to skill development with voluntary take up as they might otherwise do because households endogenously avoid these programs. The break down by education is particularly stark as children in higher maternal-education households are particularly hurt by full take-up of a low-quality program. These are precisely the households that have high-quality, alternative care and would tend to avoid this program.

Second, as quality increases with full take-up, there are diminishing returns on age-3 IQ, relative to the case with voluntary take-up in Figure 3. This highlights the endogeneity of program take-up: as quality increases, more households are induced to participate, and this increases the ATE of the program offer.[45]

---

[43] At the time of the IHDP, there were almost no public child care subsidies for children under 3 in the U.S. As discussed in the introduction, now there are subsidies via Early Head Start and CCDF vouchers for low-income families. However, funding limitations mean that these cover only a small minority of eligible children. The Child & Dependent Care Tax Credit did exist and provided small care subsidies to higher-income families. Nonrefundability of this tax credit means it's essentially worthless to low-income families.

[44] This counterfactual is not strictly enforceable of course, but represents the upper bound from program changes that would somehow make the program more attractive to households (reducing hassle and stigma costs, and making the program more convenient to households generally) and bringing in households that strongly prefer maternal care over program care. A similar exercise is conducted by Kline and Walters [2016] in the context of their discrete-choice model of Head Start and 2 alternative forms of care.

[45] This does not necessarily change the average treatment effect on the treated (ATT), i.e. those who take-up the program.

These issues of quality and take-up come to the fore in understanding results from various large-scale child care programs. One prominent example is the Québec universal child care program. Starting in the 1990s, the Canadian province of Québec provided universal care to children of all ages, with an out-of-pocket cost capped at CAN$5 per day. Using differences between Québec and other Canadian provinces, Baker et al. [2008] and Baker et al. [2015] find that this program had substantial deleterious effects on child outcomes, such as anxiety, aggressiveness, motor and social skills, and health. However it appears that the types of care subsidized under this program were of lower quality than the programs considered in the U.S.[46] In addition, the take-up of the Québec program was twice as large for educated mothers (some college or more) than for less educated mothers (high school or less), likely because low income households already had access to subsidized care at the time the program began. Therefore, it is possible that the Québec program caused many households to switch from higher-quality care at home or in the market to lower-quality care in the public centers. Baker et al. [2015] conclude that in the case of the Québec program: "Our findings for young children clearly contrast with those of the Perry, Abecedarian, and Head Start studies. These latter programs both provide higher quality care and are targeted at less advantaged children."

Using our estimated model, we attempt to see if we can replicate a similar result. Figure 6 graphs the age-3 IQ ATE from a simulation varying the fraction of college-educated compliers to a low quality ($2 per hour) full take-up program. As we increase the fraction of college educated compliers from 0 to 1, the overall ATE of the program declines, and if the fraction of college-educated compliers exceeds about 2/3 (at least twice as many college graduates than non-college graduates), the overall ATE becomes negative, mirroring the results of the Québec program.

6.4 Vary Subsidized Price

We next consider varying the subsidy level of the program by requiring that parents pay $1 per hour of program care. Compared to the offer of the program for free, instituting this co-pay reduces average hours by about two-thirds, from 18 hours to 6.6 hours per week. The ATE on age-3 IQ is also substantially reduced by over half. Interestingly, even though hours of take-up fall, imposing the co-pay increases maternal earning hours slightly as mothers work more to pay for the program.

6.5 Limitations

Though the mechanisms and tradeoffs highlighted by our model remain at the center of families' choices and sensible policymaking, care is required in drawing out lessons for today. Doing so would require extrapolation in a few dimensions where we want to highlight concerns.

First, the IHDP sample was selected to include only children born low birth weight and premature. Although our sample excludes children born with extreme adverse health conditions and includes primarily children close to the eligibility cutoffs, to the extent that skill-

[46]Based on a audit study of the centers involved, Japel et al. [2005] report that a majority, 61%, had only a "minimal" educational component, and an additional 12% were of "inadequate quality."

development technology, parental preferences, demographics, or correlations with wage and income systematically differ in this subpopulation relative to the general population, then estimates drawn from here may be valid for this subpopulation but less relevant for the general population. As discussed above, we tested for differences in sample statistics using a sample of children close to the eligibility cutoffs and found only minor differences between this subsample and our full sample. In addition, we allow rich patterns of heterogeneity by initial skill differences, both in the skill production technology and preferences. We have also tested for differences in skill production technology by birth weight and gestation age, in addition to age-1 measured skill, and find no statistically significant differences.

On the other hand, our sample is economically diverse because it was selected on a different margin, and can provide many important insights about the role of family resources in child care because of this. In contrast, many papers in the literature on structural estimation of child care choices with credible identification include only children from low-income families such as studies of Perry Preschool [Heckman et al., 2010], welfare reform [Bernal and Keane, 2010], and Head Start [Kline and Walters, 2016]. Others include diverse samples but lack quasi-experimental features Cunha et al. [2010]; Caucutt et al. [2017]; Del Boca et al. [2014].

A second potential limitation is that the IHDP program we study is an offer of care between 12- and 36-months. Different care offers that start earlier and extend later would likely facilitate stronger maternal connection to the labor market and could encourage even higher take-up.

Third, the IHDP parents' choices took place in the late 1980s and much about the economic and policy environment has changed. For instance, the IHDP occurred before any substantial public investment in age 0-2 care. Since then, the Child Care and Development Block Grant's Child Care Development Fund (1990) and Early Head Start (1995) program started subsidizing care for children from low-income families in this age range. Wages, nonlabor income, and prices of nonmaternal child care also evolved over this period (See Hotz and Wiswall [2019] for a recent review). Further, our estimates ignore how scalability might change costs and benefits. Concerns like this apply to many studies.

Fourth, child cognitive skill at age-3 is our measure of child development but other outcomes matter [Kautz et al., 2014]. The IHDP data include other margins of child development at this endpoint of the treatment, such as behavior problems and health. Given that the prior IHDP literature found smaller effects there, we focus on cognitive skill.[47] This is essentially a static model of a point-in-time decision process focused on the earliest ages of development. It should inform and complement research examining longer-term dynamics in the child development process [Cunha et al., 2010; Del Boca et al., 2014; Agostinelli and Wiswall, 2020]. There is a

[47]The IHDP data also include follow-up measures of development across domains at older ages up to 18. However, data on inputs in these intervening years is relatively weak. This model could be extended and made dynamic to cover later ages but, lacking data of similar quality on endogenous inputs, we refrain from that exercise. Duncan and Sojourner [2013] shows that, among the 2,000-2,500 grams at birth subsample, the positive age-3 effect fades down and not out among children from low-income families to 0.59 standard deviations at age 5 and 0.50 at age 8. At age-18, the cognitive skill effect is no longer statistically significant (a third of the sample attrited) but the point estimate remains substantial at 0.36. The smaller age-3 effect for those from higher-income families fades out by and after age-5. No data past age 18 was collected on the IHDP sample but studies of the original Abecedarian project and replications other than the IHDP show positive impacts in adulthood [Campbell et al., 2008, 2014].

large literature on the relationship between surrogate outcomes in childhood and later outcomes of deeper interest [Elango et al., 2015].

7 Conclusion

Each child has only one first 3 years. The quality of the care environments they experience over this short time has life-long consequences. In the first few years of children's lives, parents struggle under the dual burden of care and earning with relatively less support from the community than they will get as their children age. The policy and economic environment shape the choices available. This paper illuminates key tradeoffs parents of young children face, predicts how different kinds of families would react to changes in their environment, and how it would impact their children's development and uses of parents' resources.

Our approach synthesizes empirical evidence across multiple margins of parental choice and leverages it to illuminate mechanisms and predict counteractual policy effects. Our theoretical model highlights tradeoffs that much of the previous literature has not deeply considered – (1) quality versus quantity of maternal care and (2) quality and quantity of non-maternal care – while also integrating more familiar tradeoffs such as how early skill levels affect the productivity of later investment. In particular, adding a choice of parenting effort introduces a new channel into the economic literature by which child care subsidies (and other programs) might influence child development. Allowing parents to parent less can allow them to parent better. We quantify the importance of this endogenous parenting effort channel, finding that it accounts for a substantial part of the treatment effect on child skill and a larger share for children of mothers with less formal education.

Our comparison of alternative policies to invest in families with young children shows that in-kind provision of high-quality care has some advantages over alternatives. The direct impacts on child-skill development are large, as it replaces lower-quality care in the child's time budget, and it induces indirect positive effects on the quality of nonprogram care, both maternal through a relief channel and other nonmaternal through a budget channel. In our simulations, an equivalent-cost cash transfer has a much smaller effect on child skill because little of it goes to improving nonmaternal care quality and there is no positive spillover to maternal-care quality. Relative to the cash-transfer, a 25-percent maternal-wage subsidy has similar effects on child skill at a quarter the cost. Offering free mediocre-quality care has an even-smaller skill effect at even lower cost. Finally, we rationalize evidence on negative average effects of universal programs, showing that low quality programs taken up by advantaged households could lead to overall negative effects.

References

B. Abott. Incomplete markets, the substitutability of time and money, and parental investments in children. *Working Paper*, 2019.

J. Adda, C. Dustmann, and K. Stevens. The career costs of children. *Journal of Political Economy*, 125(2):293–337, 2017.

F. Agostinelli and M. Wiswall. Estimating the technology of children's skill formation. *Working Paper*, 2020.

A. Aizer and F. Cunha. The production of human capital: Endowments, investments and fertility. Technical report, National Bureau of Economic Research, 2012.

D. Almond and B. Mazumder. Fetal origins and parental responses. Technical report, FRB of Chicago Working Paper, 2013.

D. J. Anderson, M. Binder, and K. Krause. The motherhood wage penalty: Which mothers pay it and why? *American economic review*, 92(2):354–358, 2002.

S. E. Anderson, R. A. Gooze, S. Lemeshow, and R. C. Whitaker. Quality of early maternal–child relationship and risk of adolescent obesity. *Pediatrics*, 129(1):132, 2012.

O. Attanasio, C. Meghir, and E. Nix. Human capital development and parental investment in india. Technical report, National Bureau of Economic Research, 2019.

O. Attanasio, S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina. Estimating the production function for human capital: Results from a randomized control trial in colombia. *American Economic Review*, forthcoming.

A. Auger and M. Burchinal. Cognitive stimulation and emotional support in preschool environments and childrens subsequent development. Technical report, University of California Irvine working paper, 2013.

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

M. Baker, J. Gruber, and K. Milligan. Non-cognitive deficits and young adult outcomes: The long-run impacts of a universal child care program. Technical report, National Bureau of Economic Research, 2015.

T. J. Bartik, W. Gormley, and S. Adelstein. Earnings benefits of tulsa's pre-k program for different income groups. *Economics of Education Review*, 31(6):1143–1161, 2012.

L. J. Berlin, J. Brooks-Gunn, C. McCarton, and M. C. McCormick. The effectiveness of early intervention: examining risk factors and pathways to enhanced development. *Preventive medicine*, 27(2):238–245, 1998.

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

D. M. Blau. *Child Care Problem: An Economic Analysis*. Russell Sage Foundation, 2001.

D. M. Blau and A. P. Hagy. The demand for quality in child care. *Journal of Political Economy*, 106(1):104–146, 1998.

R. H. Bradley, L. Whiteside, D. J. Mundfrom, P. H. Casey, B. M. Caldwell, and K. Barrett. 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, 1994.

S. Breining, J. Doyle, D. N. Figlio, K. Karbownik, and J. Roth. Birth order and delinquency: Evidence from denmark and florida. *Journal of Labor Economics*, 2019.

J. Brooks-Gunn, P. K. Klebanov, F.-r. Liaw, and D. Spiker. Enhancing the development of low-birthweight, premature infants: Changes in cognition and behavior over the first three years. *Child development*, 64(3):736–753, 1993.

J. Brooks-Gunn, M. C. McCormick, S. Shapiro, A. Benasich, and G. W. Black. The effects of early education intervention on maternal employment, public assistance, and health insurance: the infant health and development program. *American Journal of Public Health*, 84 (6):924–931, 1994.

F. Campbell, G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. Pungello, and Y. Pan. Early childhood investments substantially boost adult health. *Science*, 343(6178):1478–1485, 2014.

F. A. Campbell, B. H. Wasik, E. Pungello, M. Burchinal, O. Barbarin, K. Kainz, J. J. Sparling, and C. T. Ramey. Young adult outcomes of the abecedarian and care early childhood educational interventions. *Early Childhood Research Quarterly*, 23(4):452–466, 2008.

E. U. Cascio and D. W. Schanzenbach. The impacts of expanding access to high-quality preschool education. Technical report, National Bureau of Economic Research, 2013.

E. Caucutt and L. Lochner. Early and late human capital investments, borrowing constraints, and the family, 2017.

E. M. Caucutt, L. Lochner, and Y. Park. Correlation, consumption, confusion, or constraints: Why do poor children perform so poorly? *Scandinavian Journal of Economics*, 119(1): 102–147, 2017.

J. Chaparro and A. J. Sojourner. Same program, different outcomes: Understanding differential effects from access to free, high-quality early care, 2015.

Council of Economic Advisers. Issue brief: The disconnect between resources and needs when investing in children, 2016a.

Council of Economic Advisers. Economic report of the president. chapter 4: Inequality in early childhood and effective public policy interventions, 2016b.

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

F. Cunha, J. J. Heckman, L. Lochner, and D. V. Masterov. Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education*, 1:697–812, 2006.

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

D. Daruich. The macroeconomic consequences of early childhood development policies. *Working Paper*, 2018.

D. Del Boca, C. Flinn, and M. Wiswall. Household choices and child development. *Review of Economic Studies*, 81(1):137–185, 2014.

N. Drange and T. Havnes. Early child care and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics*, 37(2):581–620, 2019.

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

G. J. Duncan and A. J. Sojourner. Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? *Journal of Human Resources*, 48(4): 945–968, 2013.

S. Elango, J. L. García, J. J. Heckman, and A. Hojman. Early childhood education. Technical report, National Bureau of Economic Research, 2015.

D. Figlio, J. Guryan, K. Karbownik, and J. Roth. The effects of poor neonatal health on children's cognitive development. *American Economic Review*, 104(12):3921–3955, 2014.

M. Fort, A. Ichino, and G. Zanella. Cognitive and non-cognitive costs of daycare 0–2 for children in advantaged families. *Jornal of Political Economy*, 2019.

A. S. Fuligni, W.-J. Han, and J. Brooks-Gunn. The infant-toddler home in the 2nd and 3rd years of life. *Parenting*, 4(2-3):139–159, 2004.

A. Gelber and A. Isen. Children's schooling and parents' behavior: Evidence from the head start impact study. *Journal of public Economics*, 101:25–38, 2013.

A. S. Griffen. Evaluating the effects of child care policies on childrens cognitive development and maternal labor supply. *Journal of Human Resources*, 54(3):604–655, 2019.

R. T. Gross, D. Spiker, and C. W. Haynes. *Helping low birth weight, premature babies: The Infant Health and Development Program*. Stanford University Press, 1997.

N. Guner, R. Kaygusuz, and G. Ventura. Child-related transfers, household labor supply and welfare. *Working Paper*, 2018.

J. L. Hanson, N. Hair, D. G. Shen, F. Shi, J. H. Gilmore, B. L. Wolfe, and S. D. Pollak. Family poverty affects the rate of human infant brain growth. *PloS one*, 8(12):e80954, 2013.

J. Heckman and S. Navarro-Lozano. Using matching, instrumental variables, and control functions to estimate economic choice models. *Review of Economics and statistics*, 86(1):30–57, 2004.

J. Heckman, N. Hohmann, J. Smith, and M. Khoo. Substitution and dropout bias in social experiments: A study of an influential social experiment. *Quarterly Journal of Economics*, 115(2):651–694, 2000.

J. J. Heckman. Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902, 2006.

J. J. Heckman, S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz. The rate of return to the highscope perry preschool program. *Journal of public Economics*, 94(1):114–128, 2010.

C. M. Herbst and E. Tekin. Child care subsidies and child development. *Economics of Education review*, 29(4):618–638, 2010.

V. J. Hotz and M. Wiswall. Child care and child care policy: Existing policies, their effects, and reforms. *The ANNALS of the American Academy of Political and Social Science*, 686 (1):310–338, 2019.

H. Hoynes, M. P. Bitler, and T. Domina. Experimental evidence on distributional effects of head start. Technical report, National Bureau of Economic Research, 2014.

I. Iglowstein, O. G. Jenni, L. Molinari, and R. H. Largo. Sleep duration from infancy to adolescence: reference values and generational trends. *Pediatrics*, 111(2):302–307, 2003.

C. Japel, R. E. Tremblay, and S. Côté. Quality counts! *Choices*, 11(5), 2005.

T. Kautz, J. J. Heckman, R. Diris, B. Ter Weel, and L. Borghans. Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success. Technical report, National Bureau of Economic Research, 2014.

M. P. Keane and K. I. Wolpin. The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women. *International Economic Review*, 51(3):851–892, 2010.

E. E. Kisker, S. L. Hofferth, D. A. Phillips, and E. Farquhar. A profile of child care settings: Early education and care in 1990. Technical report, Mathematica Policy Research for the U.S. Department of Education, 1991.

P. Kline and C. R. Walters. Evaluating public programs with close substitutes: The case of head start. *Quarterly Journal of Economics*, 131(4):1795–1848, 2016.

E. I. Knudsen, J. J. Heckman, J. L. Cameron, and J. P. Shonkoff. Economic, neurobiological, and behavioral perspectives on building americas future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155–10162, 2006.

P. Kristensen and T. Bjerkedal. Explaining the relation between birth order and intelligence. *Science*, 316(5832):1717–1717, 2007.

M. R. Linver, A. Martin, and J. Brooks-Gunn. Measuring infants' home environment: the it-home for infants between birth and 12 months in four national data sets. *Parenting*, 4 (2-3):115–137, 2004.

J. M. Love, E. E. Kisker, C. Ross, H. Raikes, J. Constantine, K. Boller, J. Brooks-Gunn, R. Chazan-Cohen, L. B. Tarullo, C. Brady-Smith, et al. The effectiveness of early head start for 3-year-old children and their parents: lessons for policy and programs. *Developmental psychology*, 41(6):885, 2005.

R. Moffitt. An economic model of welfare stigma. *American Economic Review*, 73(5):1023–35, 1983.

T. W. Morrissey. Child care and parent labor force participation: a review of the research literature. *Review of Economics of the Household*, 15(1):1–24, 2017.

J. Mullins. Designing cash transfers in the presence of childrens human capital formation. *Working Paper*, 2019.

N. E. C. C. R. Network et al. *Child care and child development: Results from the NICHD study of early child care and youth development*. Guilford Press, 2005.

D. L. Olds, J. Robinson, R. OBrien, D. W. Luckey, L. M. Pettitt, C. R. Henderson, R. K. Ng, K. L. Sheff, J. Korfmacher, S. Hiatt, et al. Home visiting by paraprofessionals and by nurses: a randomized, controlled trial. *Pediatrics*, 110(3):486–496, 2002.

S. Peltzman. The effect of government subsidies-in-kind on private expenditures: The case of higher education. *Journal of political Economy*, 81(1):1–27, 1973.

J. Price. Parent-child quality time does birth order matter? *Journal of human resources*, 43 (1):240–265, 2008.

S. F. Reardon. The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. *Whither opportunity*, pages 91–116, 2011.

I. Roskam, M.-E. Brianda, and M. Mikolajczak. A step forward in the conceptualization and measurement of parental burnout: The parental burnout assessment (pba). *Frontiers in psychology*, 9, 2018.

L. S. Sosinsky, H. Lord, and E. Zigler. For-profit/nonprofit differences in center-based child care quality: Results from the national institute of child health and human development study of early child care and youth development. *Journal of Applied Developmental Psychology*, 28 (5-6):390–410, 2007.

P. E. Todd and K. I. Wolpin. Assessing the impact of a school subsidy program in mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American economic review*, 96(5):1384–1417, 2006.

D. L. Vandell. Early child care: The known and the unknown. *Merrill-Palmer Quarterly (1982-)*, pages 387–414, 2004.

D. L. Vandell, J. Belsky, M. Burchinal, L. Steinberg, N. Vandergrift, and N. E. C. C. R. Network. Do effects of early child care extend to age 15 years? results from the nichd study of early child care and youth development. *Child development*, 81(3):737–756, 2010.

K. I. Wolpin. *The limits of inference without theory*. MIT Press, 2013.

8 Tables and Figures

Table 1: Descriptive Statistics and Balance Tests

	Treatment	Control	Difference	P-Value
Panel A: Baseline characteristics				
Birth weight, kg.	1.82	1.79	0.03	0.46
Gestational Age, weeks	33.03	33.06	-0.03	0.90
Cog. skill at 12-months	0.06	0.07	-0.02	0.84
Maternal age, years	25.57	25.87	-0.30	0.54
Maternal cognitive skill	-0.00	0.01	-0.01	0.86
1(married)	0.43	0.51	-0.08	0.04
1(never married)	0.49	0.42	0.07	0.07
1(prev. married)	0.08	0.07	0.01	0.63
1(Less than high school)	0.39	0.35	0.04	0.37
1(High school grad)	0.29	0.29	-0.01	0.87
1(Some college)	0.20	0.22	-0.02	0.53
1(College graduate)	0.13	0.14	-0.01	0.77
# of kids $<$ age 5	1.43	1.44	-0.00	0.95
# of kids $\geq$ age 5	0.48	0.48	0.00	0.95
Panel B: Choice variables				
Child cognitive skill at 36 months ($\ln h_1$)	96.10	86.59	9.51	0.00
Hours per week of program care (τ_p)	18.18	0.00	18.18	0.00
Hours per week of maternal care (τ_m)	52.40	62.07	-9.67	0.00
Hours per week other caretakers (τ_n)	16.92	25.43	-8.52	0.00
Earning hours (L)	17.97	16.68	1.29	0.35
Maternal-care quality ($\ln q_m$)	0.32	0.09	0.23	0.00
Quality non-maternal care (q_n)	3.08	3.10	-0.02	0.88

Note: All calculations were made for singleton-born infants in the IHDP sample (N = 880; 334 in the treatment group and 546 in the control group). Due to missing data, cognitive skill at 12-months (N = 796) and maternal IQ (N = 721) statistics were calculated on smaller sub-samples. F-test of joint significance of baseline characteristics: $F(12, 689) = 0.65$; Prob $> F = 0.795$. Quality of maternal care is based on the sum of 13 binary Learning & Literacy items from the HOME questionnaire at 36-months. It has been standardized within sample, using the mean and standard deviation of the control group, on the full sample before sample restrictions. Quality of non-maternal care is in equivalent USD per hour of care. See Appendix for more details.

Table 2: Experimental Average Treatment Effects by Maternal Education

	Maternal Education Sub-Sample							
	Less than HS		HS grad.		Some coll.		Bachelors+	
Child cognitive skill at 36 months ($\ln h_1$)	9.08	(0.00)	16.43	(0.00)	7.60	(0.03)	2.58	(0.48)
Hours per week of program care (τ_p)	17.71	(0.00)	19.69	(0.00)	18.23	(0.00)	16.16	(0.00)
Hours per week of maternal care (τ_m)	-13.32	(0.00)	-10.08	(0.00)	-8.56	(0.00)	-1.86	(0.53)
Hours per week other caretakers (τ_n)	-4.39	(0.01)	-9.61	(0.00)	-9.67	(0.00)	-14.30	(0.00)
Earning hours (L)	3.31	(0.08)	3.93	(0.12)	-4.41	(0.14)	2.01	(0.61)
Maternal-care quality ($\ln q_m$)	0.33	(0.01)	0.35	(0.02)	0.25	(0.07)	-0.11	(0.32)
Quality non-maternal care (q_n)	-0.03	(0.82)	0.18	(0.25)	0.20	(0.52)	-0.50	(0.56)
Number of observations	356		240		176		108	

Note: All calculations were made for singleton-born infants in the IHDP sample. The cells report estimates of Average Treatment Effects (ATE). For any outcome Y and each educational sub-group X, we report the difference $E(Y|\text{treatment}, X) - E(Y|\text{control}, X)$. The corresponding p value is to the right in parentheses. Hours per week in IHDP program care correspond to averages among members of the treatment group. Four educational sub-groups: less than a high school degree (Less than HS), high-school graduates (HS grad.), some years of college (Some coll.), and at least a bachelor's degree (Bachelors+).

Table 3: Maternal-Care Quality Validation and Exhaustion-Effect Estimation

Panel A: Maternal-Care Quality Validation				
	(1)	(2)	(3)	(4)
Dep. Variable:	36-Month L&L score	12-Month L&L (Placebo)	(36 - 12) Month L&L Difference	36-Month Non-Effort (Placebo)
1(treatment)	0.253***	0.016	0.256***	-0.066
	(0.063)	(0.065)	(0.072)	(0.071)
N	652	687	633	652
Panel B: Exhaustion-Effect Estimation				
Estimator:	Naive OLS	IV 1st Stage	IV 2nd Stage	
Dep. Variable:	36-Month L&L	Log Maternal Hours	36-Month L&L	
Log Hours	-0.157		-1.536***	
	(0.123)		(0.415)	
1(treatment)		-0.165***		
		(0.019)		
N	652	652	652	

Note: Robust SEs in parentheses. Significant at: * 10% ** 5% *** 1%. All models include same controls as the maternal-care quality production function: maternal cognitive skill, education, and age, number of other children aged five or younger, and those over five years of age.
Panel A: Regressions of various measures of maternal-care quality on indicator of randomly-assigned treatment group. Outcomes by model:
(1) The 36-month Learning and Literacy score (L&L), which is our primary maternal-care quality measure ($log(q_m)$).
(2) 12-month L&L: Placebo because program-care offer started at 12-months.
(3) Within-individual change in L&L between 36- and 12-months.
(4) Index of the 4 non-effort items from 36-month HOME inventory: Placebo.
Outcomes 1, 2, and 4 are standardized in control group. 3 is difference between 1 and 2.

Panel B: Estimates of the exhaustion effect, as described in the text (19). 36-month L&L $\ln q_m$ is the primary maternal-care quality measure as described in the text. Column 1 is the simple OLS regression of $\ln q_m$ on log weekly maternal-care hours $\ln \tau_m$. Column 3 is the 2SLS estimate using the randomly-assigned treatment group indicator as the instrument for log maternal hours. Column 2 reports the first-stage from that 2SLS estimator.

Table 4: Estimated Production-Function Parameters

	Estimate	(SE)
Panel A: Production of maternal-care quality		
Maternal cognitive skill	0.31	(0.057)
Maternal years of education	0.01	(0.027)
# of kids $<$ age 5	-0.17	(0.058)
# of kids $\geq$ age 5	-0.19	(0.049)
Maternal age, years	0.02	(0.008)
Panel B: Child cognitive skill for above-median endowment		
δ_{h_0} (Self-Productivity)	6.84	(1.631)
δ_m (Maternal Care)	11.90	(1.210)
δ_n (Non-Maternal Care)	14.84	(3.507)
δ_p (Program Time)	45.90	(6.400)
Panel C: Child cognitive skill for below-median endowment		
δ_{h_0} (Self-Productivity)	7.21	(1.206)
δ_m (Maternal Care)	8.86	(1.444)
δ_n (Non-Maternal Care)	16.57	(4.539)
δ_p (Program Time)	40.32	(7.547)

Notes: This table reports estimates of the maternal-care quality and child skill production functions. Panel A reports on the maternal-care quality function (2). Panels B and C report the child-skill function (3), where parameters can differ by whether the child is above or below the sample median endowment, the 12-month Bayley mental development index score. Standard errors from a cluster bootstrap procedure over all estimation steps.

Table 5: Average Elasticities

	Baseline Level (1)	Wage Offer Elasticity (2)	Care Cost Elasticity (3)	Non-Lab. Inc. Elasticity (4)
Child cog. skill, 36-months ($\ln h_1$)	87.45	0.07	-0.08	0.00
Maternal-care hours (τ_m)	60.96	-0.14	0.11	0.05
Nonmaternal-care hours (τ_n)	26.54	0.58	-0.37	-0.11
Earning hours (L)	16.89	2.54	-0.35	-2.06
Maternal-care quality ($\ln q_m$)	0.16	1.78	-0.63	-0.24
Nonmaternal-care quality (q_n)	3.34	0.66	-0.92	0.32
Effort (e)	1.25	0.23	-0.16	-0.08

Notes: All statistics are averages of simulated endogenous variables in the treatment sample. Column 1: level in the control condition. Column 2: Elasticity with respect to a 1 percent increase in the wage offer. Column 3: Elasticity with respect to a 1 percent increase in the price of non-maternal care. Column 4: Elasticity with respect to a 1 percent increase in non-labor income.

Table 6: In-Sample Fit and Out-of-Sample Validation

	(1) Mean	(2) SD	(3) Mean	(4) SD	(5) Mean	(6) SD
Panel A: Control Condition						
Type of outcome:	Data		Simulated		Simulated	
Sample:	Control		Control		Treated No Program	
Program Hours (τ_p)	0.00	0.00	0.00	0.00	0.00	0.00
Non-Program Hours ($\tau_m + \tau_n$)	87.50	0.00	87.50	0.00	87.50	0.00
Child cognitive skill, 36-months ($\ln h_1$)	86.59	19.16	87.45	19.84	86.70	20.02
Maternal-care hours (τ_m)	62.07	15.54	60.96	14.49	62.07	14.38
Nonmaternal-care hours (τ_n)	25.43	15.54	26.54	14.49	25.43	14.38
Earning hours (L)	16.68	17.05	16.89	16.01	16.74	15.56
Maternal-care quality ($\ln q_m$)	0.09	0.99	0.16	0.97	0.04	1.01
Nonmaternal-care quality (q_n)	3.10	1.93	3.34	2.11	3.00	1.79
Panel B: Treatment Condition						
Type of outcome:	Data		Simulated			
Sample:	Treatment		Treatment			
Program Hours (τ_p)	18.18	8.84	18.52	9.31		
Non-Program Hours ($\tau_m + \tau_n$)	69.32	8.84	68.98	9.31		
Child cognitive skill, 36-months ($\ln h_1$)	96.10	18.38	94.01	19.83		
Maternal-care hours (τ_m)	52.40	11.64	58.03	10.45		
Nonmaternal-care hours (τ_n)	16.92	11.00	10.95	12.22		
Earning hours (L)	17.97	16.11	15.62	15.32		
Maternal-care quality ($\ln q_m$)	0.32	0.90	0.12	1.03		
Nonmaternal-care quality (q_n)	3.08	2.15	3.20	1.94		

Notes: In panel A, columns 1 & 2 are summary statistics from the control group sample's observed endogenous variables. Columns 3 & 4 are from their simulated observations and comparing them to columns 1 & 2 measure within-sample fit. Columns 5 & 6 are from a simulation where we counter-factually assign the treatment-group sample to the control condition by removing the offer of program care. Comparing these to columns 1 & 2 provides an out-of-sample validation. In panel B, columns 1 & 2 are summary statistics from the treatment-group sample's observed variables. Columns 3 & 4 are from their simulated observations and comparing these to columns 1 & 2 measure within-sample fit.

Table 7: Average Treatment Effect of Program Offer Under Counterfactual Conditions

	Baseline (1)	Change Non-Labor Inc. (2)	+ Change Wage (3)	+ Change Mat. Product. (4)	+ Change Endowment (5)
Child cognitive skill, 36-months ($\ln h_1$)	7.31	3.52	0.29	0.85	3.63
Program Hours (τ_p)	18.52	9.29	9.29	9.73	15.89
Maternal-care hours (τ_m)	-4.04	-3.20	0.85	0.93	-0.14
Nonmaternal-care hours (τ_n)	-14.48	-6.09	-10.14	-10.65	-15.75
Earning hours (L)	-1.12	-0.04	-10.27	-9.75	-9.77
Maternal-care quality ($\ln q_m$)	0.09	0.07	-0.04	-0.04	-0.01
Nonmaternal-care quality (q_n)	0.20	0.26	-0.96	-0.90	-0.67
Effort (e)	0.15	0.12	-0.02	-0.01	0.03

Notes: Each cell presents an average treatment effect: $E(y|treat = 1) - E(y|treat = 0)$ for the y variable listed in the rows. Column (1) is the original, baseline estimate. Column (2) changes the non-labor income of households to be the average non-labor income of the BA+ sub-sample. Each column progressively adds additional changes. Column (3) changes both non-labor income and the wage offer to match that of the BA+ sub-sample. Column (4) changes the maternal/household characteristics X in the maternal quality function to match the BA+ sub-sample. Column (5) changes the initial skill endowment to match the average endowment of the BA+ sub-sample.

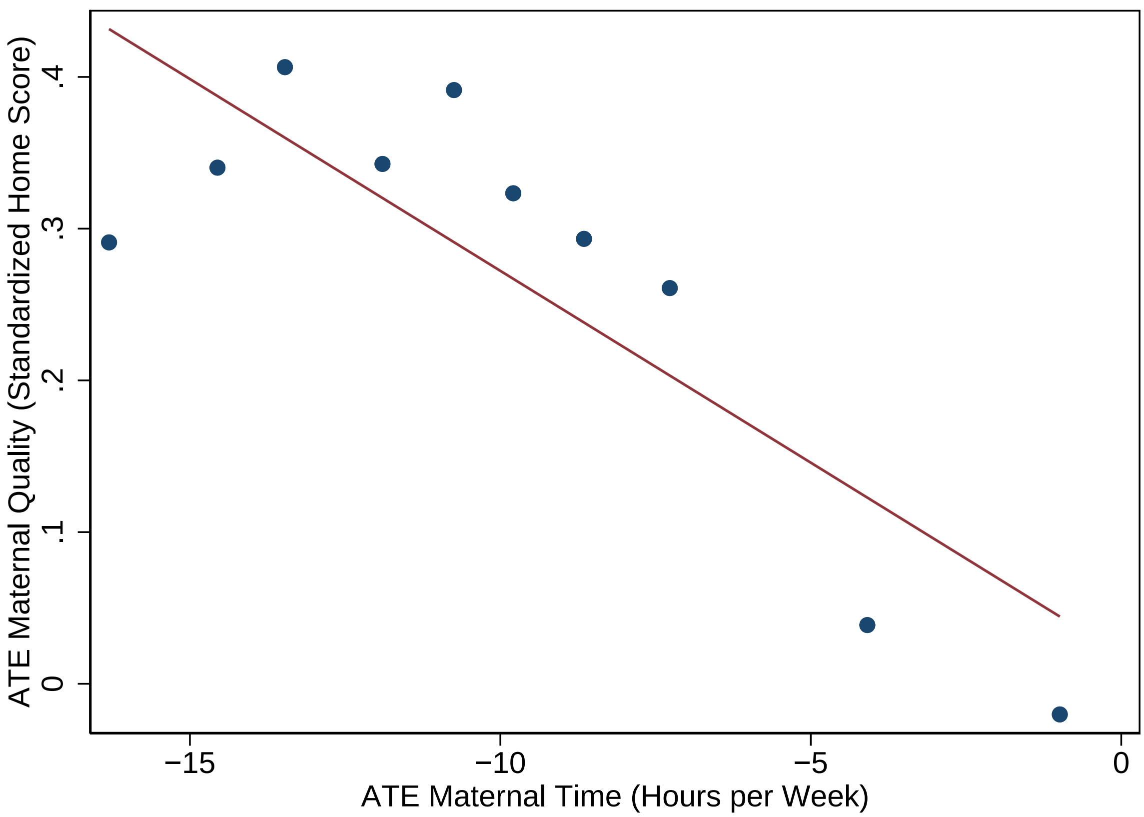
Table 8: Outcomes under Counterfactual Policy Simulations

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ Prog. Qual. (5)	Full Take Up (6)	$1/hour Co-pay (7)
Program Hours (τ_p)	0.00	18.52	0.00	0.00	3.16	45.00	6.69
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.98	87.50	87.50	84.34	42.50	80.81
Child cognitive skill, 36-months ($\ln h_1$)	86.70	94.01	88.08	88.18	87.42	107.33	90.00
Maternal-care hours (τ_m)	62.07	58.03	62.70	60.13	60.33	41.57	58.50
Nonmaternal-care hours (τ_n)	25.43	10.95	24.80	27.38	24.00	0.93	22.31
Earning hours (L)	16.74	15.62	6.77	20.60	17.03	18.45	17.44
Maternal-care quality ($\ln q_m$)	0.04	0.12	0.02	0.09	0.07	0.61	0.11
Nonmaternal-care quality (q_n)	3.00	3.20	4.40	3.50	3.05	3.61	3.08
Effort (e)	1.20	1.34	1.18	1.26	1.25	2.30	1.32
Policy cost, weekly $/child	0	336	336	84	10	831	117

Note: cells present averages of simulated endogenous variables in the treatment sample.

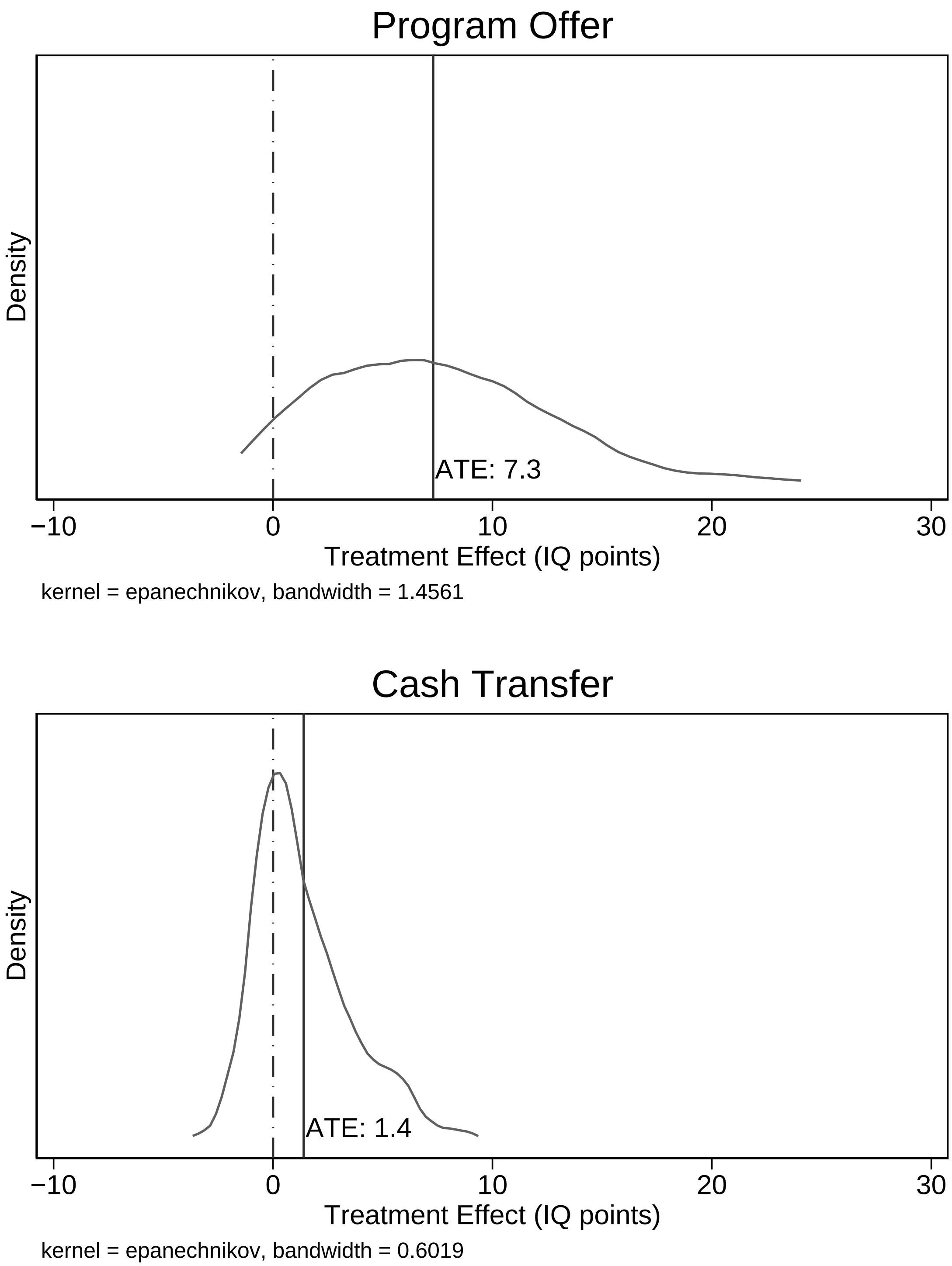
1. *Control*: the control state of the experiment, in which program care is not offered.
2. *Treatment*: the treatment state, free access to IHDP program care is offered.
3. *Cash Transfer*: cash transfer to each household equal to the IHDP's per capita cost. No care subsidy.
4. *Wage +25%*: boosts each mother's wage by 25%. No care subsidy.
5. ↓ *Program-Care Quality*: similar to IHDP's program-care offer but quality of program care reduced to equal the control-group's average nonmaternal-care quality, $3.14 hour.
6. *Full Take-up*: forces all households to take-up treatment program care 45 hours weekly and simulates choices they would make on all other margins.
7. *Co-pay*: like treatment care offer but program care requires a $1 per hour co-pay.

Figure 1: Binscatter of Treatment Effect on Log Maternal Quality vs. Treatment Effect on Maternal Time



Notes: Each dot is the average treatment effect (of program offer) on (log) maternal quality for each of 10 decile bins of the treatment of effect on maternal time. The solid line is the linear regression fit. Log maternal quality is measured using standardized home scores, as described in the text. We estimate the predicted treatment effect for each sample member i using a regression of $Y_i = \{\ln q_{mi}, \tau_{mi}\}$ on a vector of covariates X_i– indicators for mother's education (drop-out, high school graduate, some college, college graduate), mother's test score, number of children aged less than 5 and greater than 5, and indicators for mother's marital status (married and divorced)–by program offer separately. The predicted value is $\widehat{Y}_{\text{offer},i} = X_i'\widehat{\beta}_{Y,\text{offer}}$ for offer $= \{0, 1\}$. The estimated treatment effect for each sample member i is then $\widehat{Y}_{1,i} - \widehat{Y}_{0,i}$.

Figure 2: Distribution of Treatment Effects on Child Skill of Program Care Offer and Equivalent-Cost Cash Transfer



Notes: solid (dashed) vertical line at average (0) treatment effect.

Figure 3: Child Cognitive Skill: Average Intent-to-Treat Effects of Program Care Offer

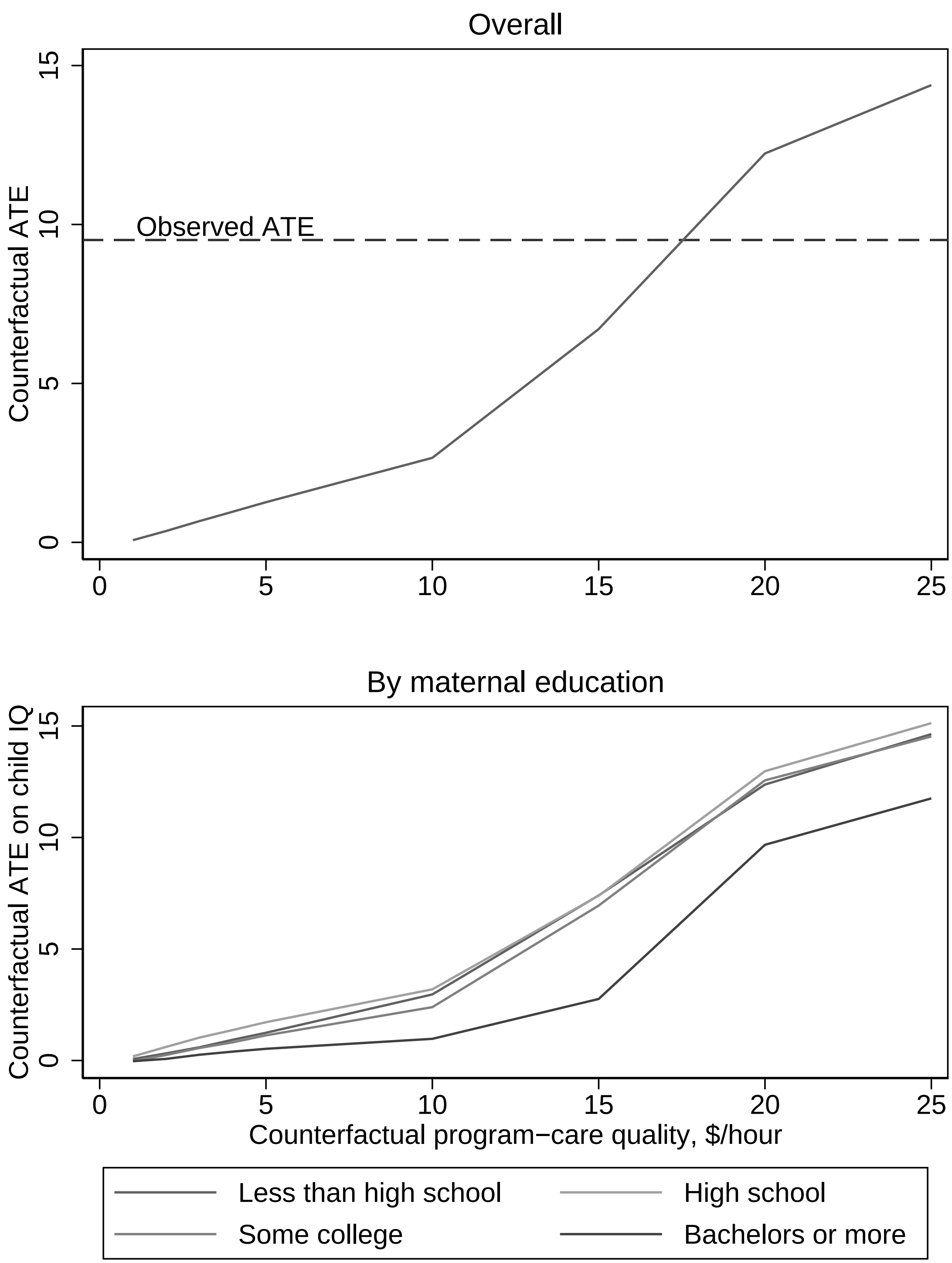


Figure 4: Maternal- and Child-Time Allocations: Average Intent-to-Treat Effects of Program Care Offer by Counterfactual Program Quality

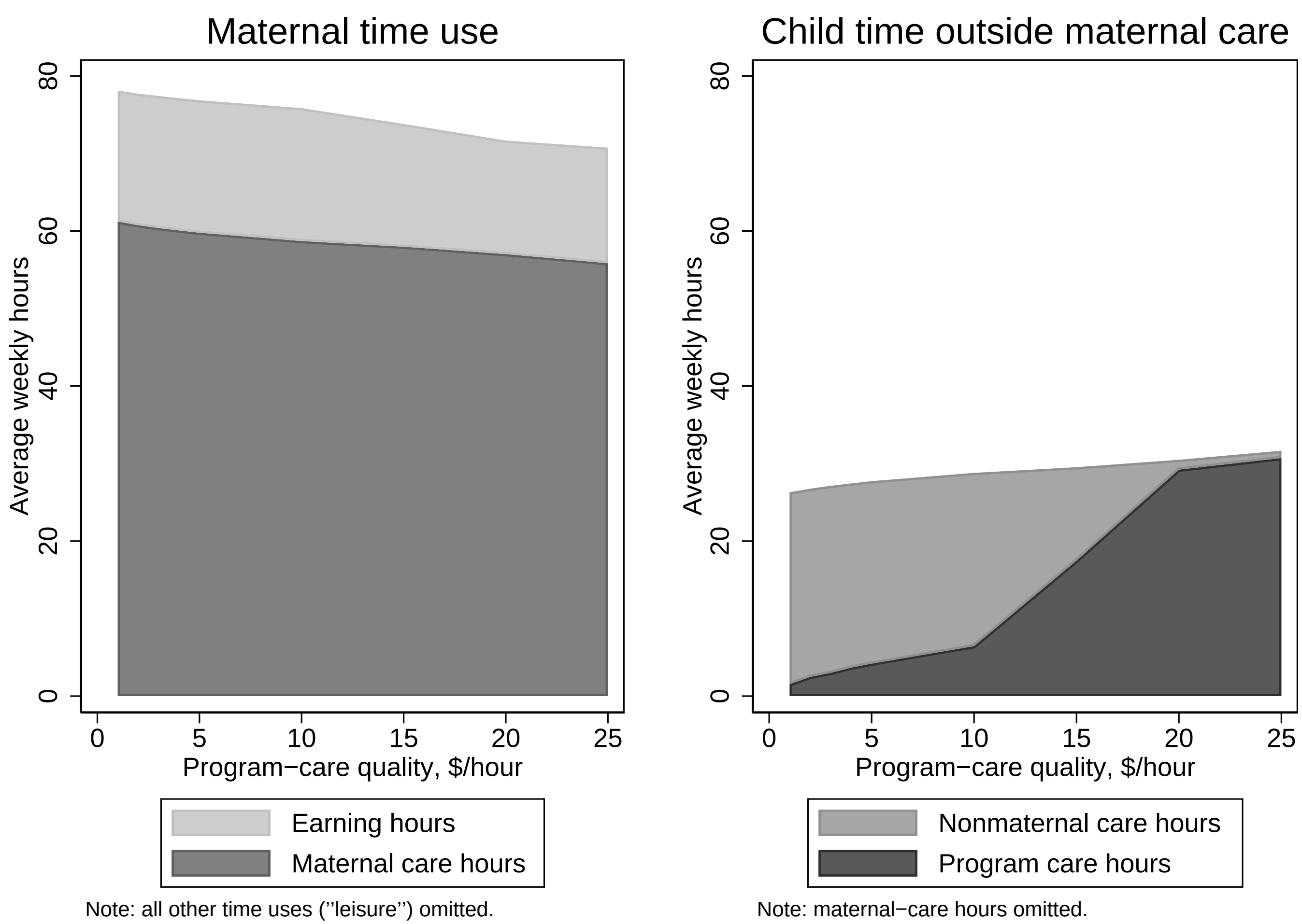


Figure 5: Full Take-Up: Average Intent-to-Treat Effects of Program Care on Child Cognitive Skill as a Function of Counterfactual Program Quality if All Children Attended Full Time, Overall and by Maternal Education

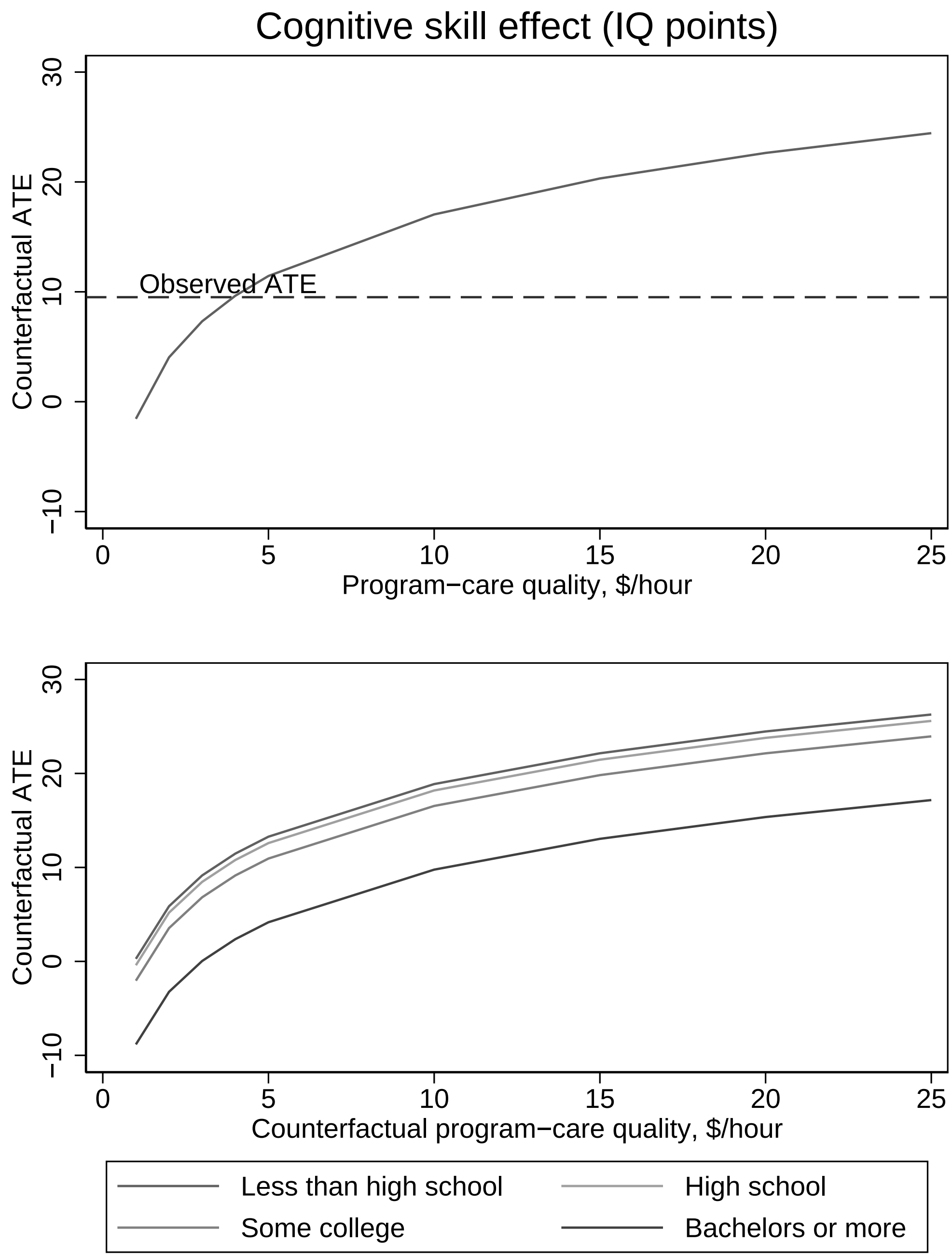
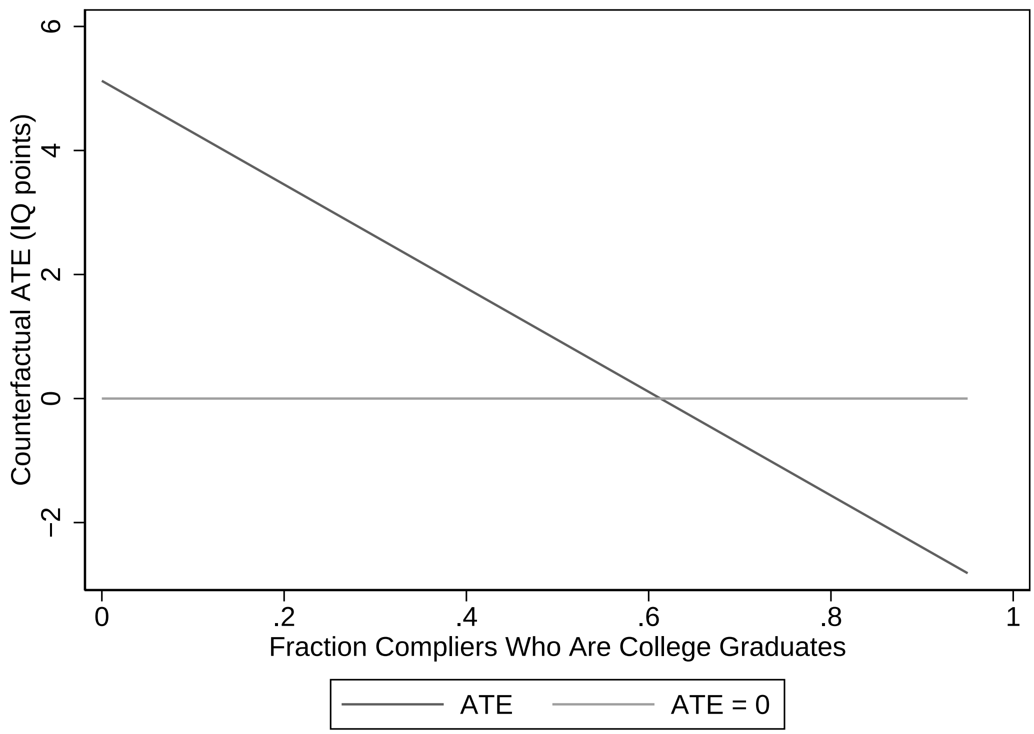


Figure 6: Changing the Compliers: ATE of Low-Quality Program Care as a Function of the Share of Compliers with at Least a Bachelor's Degree



Notes: Program care quality is assumed to be equivalent to $2 per hour. In this simulation, full-take up of the program is assumed. Horizontal axis measures fraction of household compliers that are college graduates, from 0 to 1. Vertical axis measures overall ATE of program on all households. The horizontal line indicates the the 0 ATE. With over 60 percent of households complying, the overall ATE of the program becomes negative.

Online appendices for Chaparro, Sojourner & Wiswall

A Data Appendix

A.1 IHDP Data

The Infant Health and Development Program data are publicly available through the Inter-University Consortium for Political and Social Research (ICPSR #9795 and #23580). The following variables come from the Primary Analysis Dataset: treatment status, birth weight, site, gestational age, length at birth, head circumference, maternal age at birth, marital status at birth, maternal education, maternal race, number of home visits per year, number of days at the CDC per year, average number of daily hours at the CDC per year during the intervention, Bayley test at 12 months and Stanford Binet IQ at 36 months.

Hours per weekday of maternal care correspond to the average of maternal self-reported hours in the 18-month and 30-month family interviews. Hours of care at the CDC come from administrative data and are the average weekly attendance over the 2 years during which the program was offered. Hours of care with other care takers is calculated as a residual, using the child's time constraint. Based on Iglowstein et al. [2003], 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 = 87.5 hours of direct care per week.

A.2 Study of Early Child Care and Youth Development (SECCYD) Data on Nonmaternal Care Quality

One of the problems we face in modeling non-maternal care is that some of this care was not purchased on the private market (e.g., provided by relatives such as grandparents). For only about a quarter of the IHDP sample do we observe positive expenditures on child care. For the IHDP cases without hourly expenditures on nonmaternal care, we leverage the observer-based measure of nonmaternal care quality in another dataset, the Study of Early Child Care and Youth Development (SECCYD). We use this dataset to provide a highly detailed imputation of non-maternal care quality and market price for the IHDP households.

The SECCYD collected panel data on child and family characteristics and their use of various care settings from a sample of 1,364 children aged 0 to 3 from 1991 to 1994 in ten study sites around the country, two of which overlap with the IHDPs eight sites. The two samples are quite parallel in their target populations and purposes, although the SECCYD sample was not limited to low birth weight and premature births. We do use those variables a predictors, as described below. The SECCYD was produced by the National Institute of Child Health and Human Development (NICHD) and is a widely-used observational dataset [Network et al., 2005; Sosinsky et al., 2007; Vandell et al., 2010; Anderson et al., 2012].

Our imputation method uses the overlap in a rich set of variables found in both the IHDP and SECCYD datasets. Our imputation method takes 3 steps.

Step 1: First, we model nonmaternal care quality in the SECCYD sample as a function of a set of child and household variables common to both datasets. For each child and each nonmaternal care setting used, the SECCYD measured care quality using the Observational Record of the Child care Environment (ORCE) [Vandell, 2004; Vandell et al., 2010], which is composed of three different types of scores: Behavioral Scales, Qualitative Ratings and measures of Structural Variables. We follow Auger and Burchinal [2013], who show that a good measure of the quality of interactions geared toward cognitive stimulus is the ORCEs Qualitative Rating on Stimulation of Development. This rating is available in the SECCYD data at 15, 24 and 36 months (Phase 1). We pool these observations together to estimate a model of nonmaternal care quality. The types of care are: father/partner, grandparent, other relative, non-relative in child's home (sitter or nanny), non-relative in day care home, and child care center. Table A-1 reports the means and standard deviations of nonmaternal care quality in the SECCYD by type with this measure standardized within sample (mean 0, standard deviation 1). Notably, average care quality is highest among a non-relative in the child's home and lowest for care from a relative other than the father/partner or grandparent (e.g., older sibling or aunt).

Table A-1: Care Quality by Primary Nonmaternal Care Type in SECCYD Sample

Father/partner	0.11
	(1.10)
Grandparent	0.11
	(0.99)
Other relative	-0.33
	(0.85)
Non-relative in child home	0.39
	(1.12)
Non-relative day care home	-0.016
	(0.98)
Child care center	-0.14
	(0.91)

Note: mean (SD). Observations total 1820.

The relevant, parallel predictors of non-maternal quality available in both datasets that we include in our model of nonmaternal-care quality are the following:

- child: gestational age and weight at birth (levels and squares of each), birth order, gender, and age at the time that caregiver quality is observed,
- maternal: age, education level with high-school only omitted, race/ethnicity with non-Hispanic white omitted and marital status at child's birth with married omitted, and maternal-care quality when child is 12-months old as measured by the sum of the 10 Learning and Literacy components of the HOME inventory ([Linver et al., 2004]),

- primary nonmaternal (and non-CDC in IHDP) care type: harmonizing across the SECCYD and IHDP classifications produced 6 categories. They are: father/partner, grandparent, other relative (e.g., sibling, aunt, or cousin), nonrelative in child's own home (e.g., sitter or nanny), family day care in nonrelative provider's home (the omitted category), and child care center.

To measure primary nonmaternal, non-CDC provider type in the IHDP, we focused on the 36-month family interview because it alone explicitly distinguished child care centers between the CDC and others. It asks who usually provided the most care during the six months between the child's age 30- and 36-months during weekdays. In a series of three sets of questions, the survey elicits information about primary, secondary, and tertiary care providers. There is also a fourth set of questions, which asks about the primary provider now, at 36-months. We look for the primary nonmaternal, non-CDC provider using the three sets of questions and then the fourth set if no nonmaternal, non-CDC provider is named in the first three. This yields a primary nonmaternal, non-CDC provider for most, but not all, of the IHDP sample.

Because of the missing information on non-maternal child care type for some IHDP households, we estimate two models in the SECCYD sample (Table A-2). The first includes all of the predictors listed above, and this is used to predict nonmaternal care quality for the IHDP subsample with observed values on all predictors. The second model excludes primary nonmaternal care type and is used to predict nonmaternal care quality for the IHDP subsample in which primary nonmaternal care type is missing.

Step 2: Second, with the coefficients from this model estimated from the SECCYD and with parallel predictors in the IHDP, we score each IHDP child's predicted nonmaternal care quality based on the maternal, child, and nonmaternal care characteristics listed in the first step. This prediction provides an imputed level of non-maternal quality for each IHDP sample household.

Table A-2: Models of Nonmaternal-Care Quality in SECCYD Sample

	(1)	(2)
Father/partner	0.198*	
	(2.21)	
Grandparent	0.193*	
	(1.99)	
Other relative	-0.114	
	(-1.09)	
Non-relative in child home	0.402***	
	(3.76)	
Child care center	-0.145*	
	(-2.18)	
Less than high school	0.0536	0.0415
	(0.40)	(0.32)
Some college	0.117	0.121
	(1.59)	(1.67)
College grad	0.169*	0.177*
	(2.10)	(2.19)
African-American	-0.226**	-0.257**
	(-2.64)	(-3.01)
Hispanic	-0.129	-0.128
	(-0.91)	(-0.94)
Other	0.164	0.166
	(1.24)	(1.22)
Never married	-0.114	-0.143
	(-1.22)	(-1.61)
Formerly married	-0.257	-0.303
	(-1.46)	(-1.69)
Maternal-care quality index	0.0759***	0.0700***

	(3.63)	(3.36)
Child birth order	-0.146***	-0.123***
	(-4.21)	(-3.67)
Girl child	0.128*	0.135*
	(2.38)	(2.42)
Child birth weight	0.460	0.513
	(0.95)	(1.07)
(Birth weight)^2	-0.0624	-0.0720
	(-0.92)	(-1.08)
Gestational age at birth	0.911	0.698
	(1.78)	(1.28)
(Gest. age)^2	-0.0118	-0.00903
	(-1.77)	(-1.28)
Maternal age	0.0119	0.0114
	(1.91)	(1.80)
At 24 months	0.0305	0.00304
	(0.65)	(0.06)
At 36 months	0.103*	0.0414
	(2.04)	(0.82)
Constant	-19.39*	-15.23
	(-2.01)	(-1.47)
Observations	1820	1820

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Step 3: Using the IHDP sub-sample that pays for care, we regress observed log expenditures per hour on a quadratic function of predicted nonmaternal care quality. We use this regression to predict the market price for care without observed expenditure. Finally, because it appears that total expenditure is too low (likely because some households are only paying for some of their non-maternal care), we adjust the per hour price of non-maternal care to match the average price of center-based care reported in a contemporary study Kisker et al. [1991].

Comparison to NSECE In Table A-4 we compute various statistics on the distribution of hourly prices for care, computed for the IHDP sample (using SECCYD data imputation as described above) and using another, much more recent survey of child care providers, the 2012 National Survey of Early Care and Education (NSECE). The data consists of about 14,000 center and home-based care providers. In this table, we compute the maximum hourly price of care (listed price, before any discounts or subsidies) charged to families of two-year-old children, weighting each provider by sample weights and child enrollment. Hourly rates are expressed in 2018$. The data indicate that their has been secular rise in the price of early child care, in particular, a much larger right-tail in the distribution of hourly rates for the more recent period.

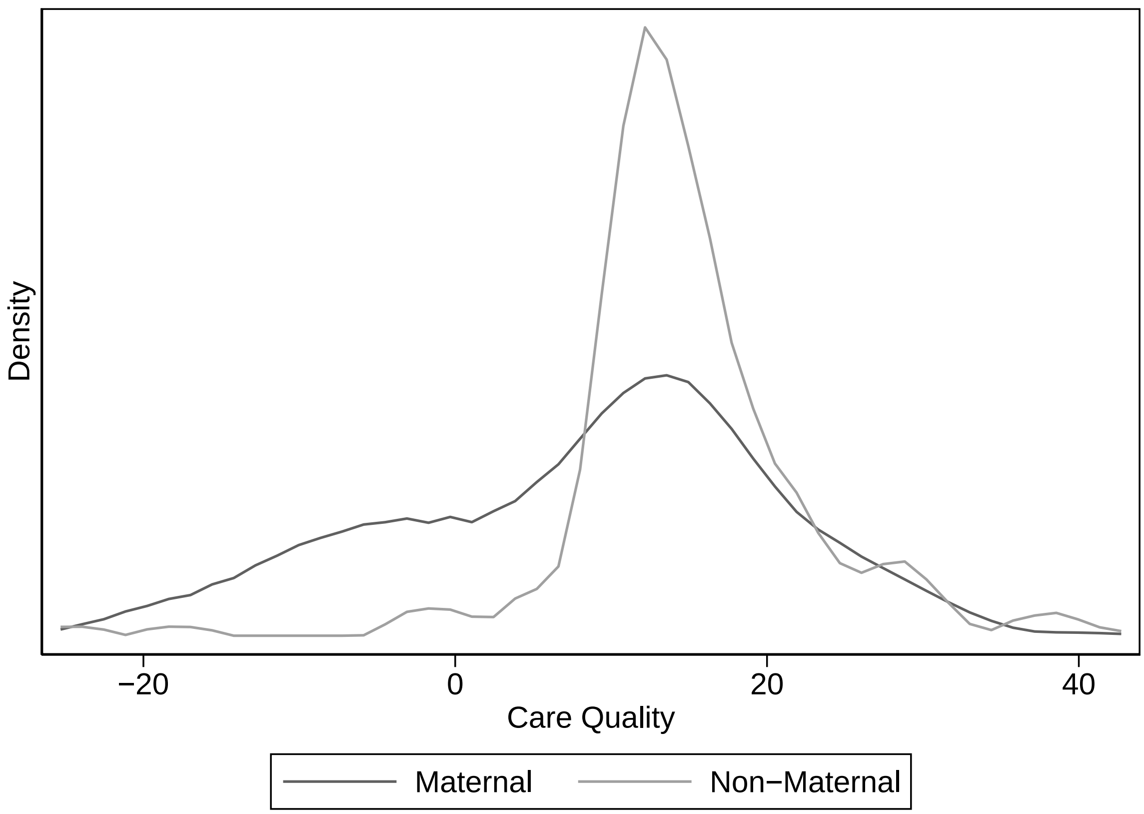
Table A-4: Distribution of Child Care Hourly Rates

	IHDP	NSECE
mean	3.09	6.29
p1	0.51	0.56
p10	1.90	2.05
p25	2.11	2.73
p5	1.71	1.71
p50	2.56	4.10
p75	3.41	6.84
p90	4.93	15.12
p95	6.32	19.15
p99	11.37	27.35
max	21.11	44.35

Notes: NSECE uses combined center- and home-based hourly rates, weighted by enrollment. All figures in 2018$.

B Additional Figures and Tables

Figure B-1: Distribution of Care Qualities Given Production-Function Estimates



Notes: This figure displays the distribution of care qualities using the parameter estimates of the skill production function. Maternal-care quality is $\delta_m \ln q_m$ and nonmaternal-care quality is $\delta_n \ln q_n$.

Table B-1: Estimated Parameters of Wage Offer and Non-Labor Income Functions

	Estimate	(SE)
Panel A: Wage Offer		
Intercept	0.73	(0.184)
Maternal education, years	0.06	(0.013)
Maternal age, years	0.05	(0.007)
Maternal cognitive skill	0.06	(0.022)
Panel B: Non-Labor Income		
Intercept	4.46	(0.298)
Maternal education, years	0.05	(0.016)
1(Married)	1.03	(0.140)
1(Never married)	0.14	(0.153)
Maternal age, years	0.01	(0.012)

Notes: Standard errors are from a cluster-bootstrap procedure over all estimation steps.

Figure B-2: Distribution of Wage Offers

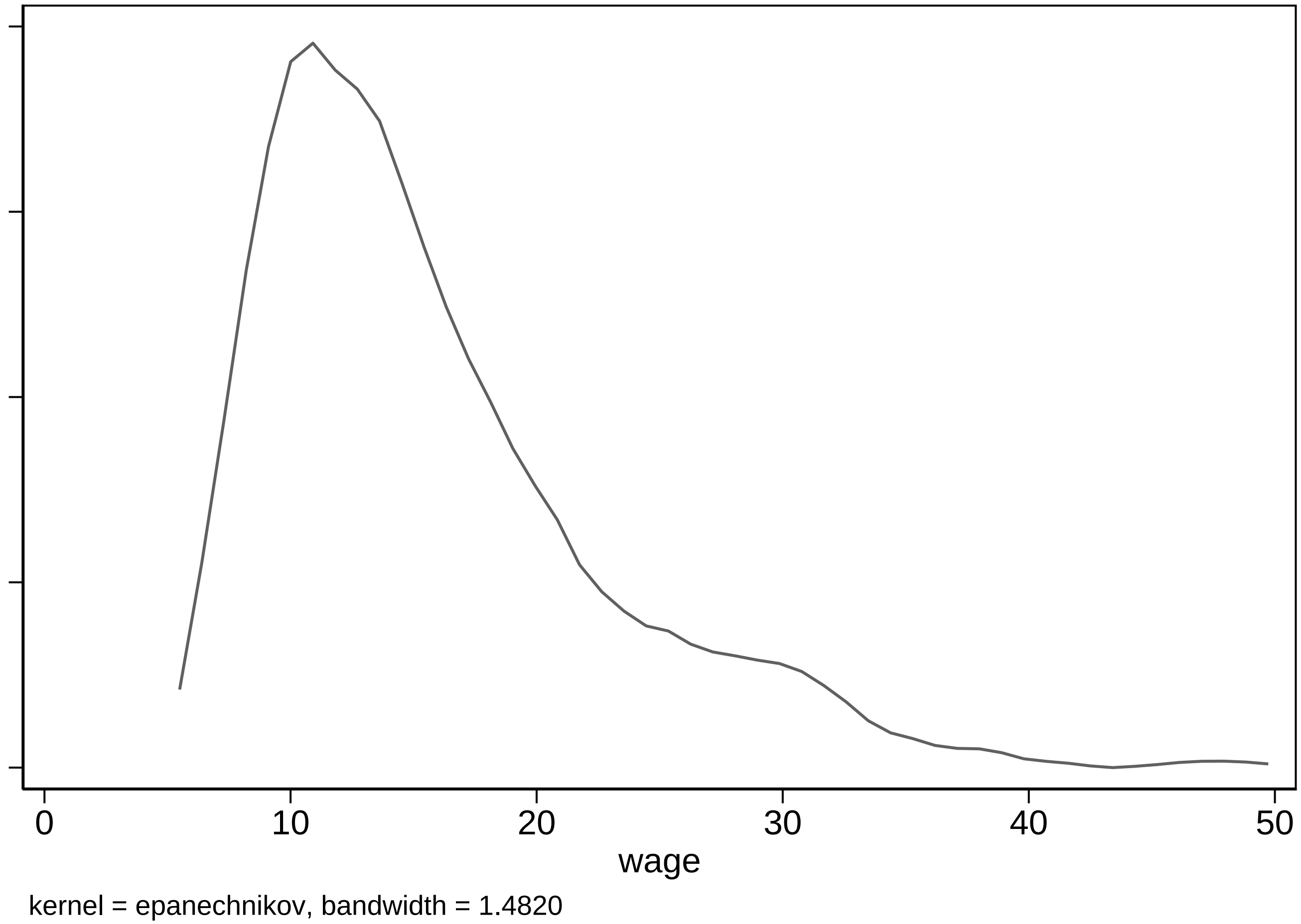


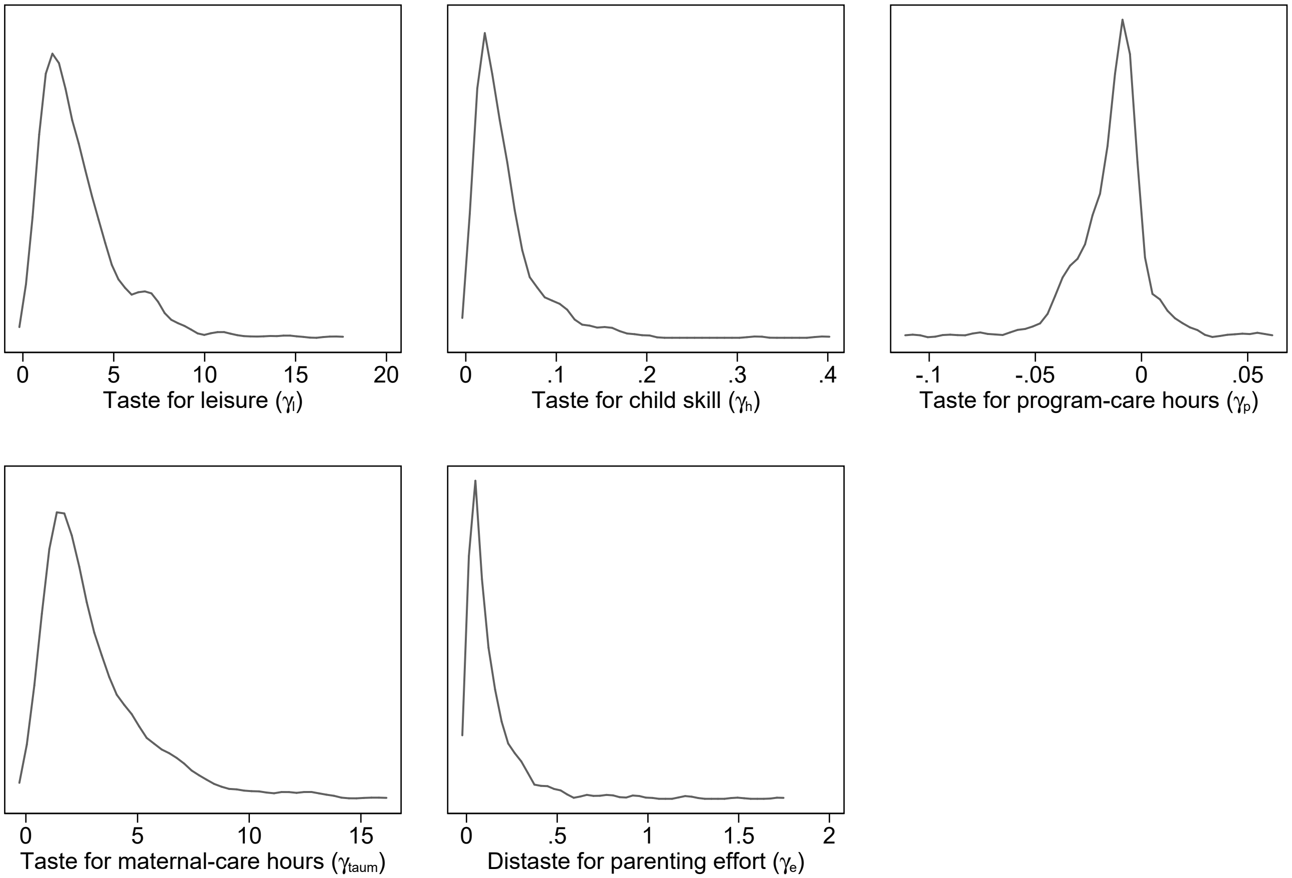
Figure B-3: Distribution of Maternal Labor Supply

Table B-2: Distribution of Estimated Preference Parameters

	Mean	SD
Preference for Leisure (γ_l)	3.099	2.200
Preference for Parenting Time (γ_{τ_m})	3.118	2.543
Disutility of Parenting Effort ($\gamma_{e,1}$)	0.151	0.226
Preference for Child Cognitive Skill (γ_h)	0.044	0.040
Preference for Program Care (γ_p)	-0.012	0.014

Notes: This table reports sample means and standard deviations of the estimated marginal distributions of preferences. Distaste for effort γ_{e1} is re-scaled by 10,000.

Figure B-4: Estimated Marginal Distributions of Preference Parameters



Note: taste for program-care hours (γ_p) estimated only for treatment group.

Table B-3: Relation Between Preferences and Observable Household Characteristics

	(1)	(2)	(3)	(4)	(5)
Dep. Var.	γ_l	γ_h	γ_{τ_m}	$\gamma_{e,1}$	γ_p
Maternal cognitive skill	0.025	0.003	0.020	0.014	0.001
	(0.101)	(0.002)	(0.122)	(0.012)	(0.001)
Mat. age, years	0.040*	-0.001*	0.014	-0.001	0.000
	(0.018)	(0.000)	(0.021)	(0.002)	(0.000)
1(Mat. Educ = high-school grad)	-0.554**	-0.021***	-1.084***	-0.070**	0.004
	(0.190)	(0.004)	(0.228)	(0.023)	(0.002)
1(Mat. Educ = some college)	-1.171***	-0.025***	-1.581***	-0.068*	0.005
	(0.235)	(0.005)	(0.283)	(0.028)	(0.003)
1(Mat. Educ = BA+)	-0.012	-0.006	-0.548	0.004	0.000
	(0.345)	(0.007)	(0.415)	(0.041)	(0.004)
# other kids < 5 years old	0.390***	-0.001	0.518***	0.010	0.001
	(0.112)	(0.002)	(0.135)	(0.013)	(0.001)
# kids 5 or older	0.474***	-0.000	0.261*	-0.015	0.001
	(0.095)	(0.002)	(0.114)	(0.011)	(0.001)
1(Mother married)	-2.065***	-0.022***	-1.654***	-0.068	0.003
	(0.307)	(0.006)	(0.369)	(0.036)	(0.004)
1(Mother never-married)	0.156	0.000	0.331	0.051	-0.000
	(0.310)	(0.006)	(0.373)	(0.037)	(0.004)
Child cognitive skill at 12-months	0.005	0.001	0.145	0.003	-0.001
	(0.078)	(0.002)	(0.094)	(0.009)	(0.001)
Prenatal investment index	-0.057	-0.003**	-0.254**	-0.025**	-0.001
	(0.064)	(0.001)	(0.077)	(0.008)	(0.001)
N	618	618	618	618	240

Table B-4: Sample Fit by Maternal Education (Control Group)

	Data Control		Simulated Control		Simulated Treated No Program	
	Mean (1)	SD (2)	Mean (3)	SD (4)	Mean (5)	SD (6)
Panel A: Less than high school						
Program Hours (τ_p)	0.00	0.00	0.00	0.00	0.00	0.00
Non-Program Hours ($\tau_m + \tau_n$)	87.50	0.00	87.50	0.00	87.50	0.00
Child cognitive skill, 36-months ($\ln h_1$)	78.93	13.60	81.09	15.97	78.47	16.20
Maternal-care hours (τ_m)	67.01	14.33	64.65	14.72	64.16	14.70
Nonmaternal-care hours (τ_n)	20.49	14.33	22.85	14.72	23.34	14.70
Earning hours (L)	9.13	12.99	11.31	12.93	12.52	14.34
Maternal-care quality ($\ln q_m$)	-0.38	0.94	-0.27	0.94	-0.32	0.95
Nonmaternal-care quality (q_n)	2.55	1.28	2.94	1.43	2.63	1.51
Panel B: High school only						
Program Hours (τ_p)	0.00	0.00	0.00	0.00	0.00	0.00
Non-Program Hours ($\tau_m + \tau_n$)	87.50	0.00	87.50	0.00	87.50	0.00
Child cognitive skill, 36-months ($\ln h_1$)	81.47	17.50	81.70	18.00	86.69	18.47
Maternal-care hours (τ_m)	61.90	15.08	61.30	13.73	63.46	14.34
Nonmaternal-care hours (τ_n)	25.60	15.08	26.20	13.73	24.04	14.34
Earning hours (L)	16.00	17.75	14.88	13.76	14.91	13.75
Maternal-care quality ($\ln q_m$)	-0.06	0.95	0.02	0.85	-0.09	0.80
Nonmaternal-care quality (q_n)	2.51	0.91	2.62	1.32	2.61	1.28
Panel C: Some college						
Program Hours (τ_p)	0.00	0.00	0.00	0.00	0.00	0.00
Non-Program Hours ($\tau_m + \tau_n$)	87.50	0.00	87.50	0.00	87.50	0.00
Child cognitive skill, 36-months ($\ln h_1$)	91.49	19.31	91.84	19.42	88.24	18.83
Maternal-care hours (τ_m)	58.07	16.76	57.44	13.70	60.38	12.47
Nonmaternal-care hours (τ_n)	29.43	16.76	30.06	13.70	27.12	12.47
Earning hours (L)	26.20	15.77	25.38	17.37	25.94	15.69
Maternal-care quality ($\ln q_m$)	0.42	0.79	0.48	0.83	0.24	0.99
Nonmaternal-care quality (q_n)	3.22	1.70	3.24	1.55	2.66	0.98
Panel D: College graduates						
Program Hours (τ_p)	0.00	0.00	0.00	0.00	0.00	0.00
Non-Program Hours ($\tau_m + \tau_n$)	87.50	0.00	87.50	0.00	87.50	0.00
Child cognitive skill, 36-months ($\ln h_1$)	109.29	14.44	108.96	16.46	109.09	18.36
Maternal-care hours (τ_m)	56.11	13.81	56.37	14.48	55.23	14.43
Nonmaternal-care hours (τ_n)	31.39	13.81	31.13	14.48	32.27	14.43
Earning hours (L)	22.41	17.55	22.10	18.46	19.55	16.94
Maternal-care quality ($\ln q_m$)	1.10	0.43	1.06	0.69	1.08	0.88
Nonmaternal-care quality (q_n)	5.60	3.01	6.02	3.34	5.49	2.48

Table B-5: Sample fit by Maternal Education (Treated Group)

	Data		Simulated	
	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)
Panel A: Less than high school				
Program Hours (τ_p)	17.71	7.32	18.75	7.78
Non-Program Hours ($\tau_m + \tau_n$)	69.79	7.32	68.75	7.78
Child cognitive skill, 36-months ($\ln h_1$)	88.01	14.00	86.24	16.60
Maternal-care hours (τ_m)	53.68	12.46	59.85	10.52
Nonmaternal-care hours (τ_n)	16.11	10.71	8.90	11.43
Earning hours (L)	12.44	15.24	11.04	13.34
Maternal-care quality ($\ln q_m$)	-0.05	0.86	-0.23	0.98
Nonmaternal-care quality (q_n)	2.51	0.58	2.82	1.65
Panel B: High school only				
Program Hours (τ_p)	19.69	8.29	18.84	9.07
Non-Program Hours ($\tau_m + \tau_n$)	67.81	8.29	68.66	9.07
Child cognitive skill, 36-months ($\ln h_1$)	97.90	17.37	94.61	18.33
Maternal-care hours (τ_m)	51.82	10.28	58.33	9.75
Nonmaternal-care hours (τ_n)	15.99	9.72	10.33	11.95
Earning hours (L)	19.93	14.44	14.07	13.75
Maternal-care quality ($\ln q_m$)	0.29	0.95	0.02	0.84
Nonmaternal-care quality (q_n)	2.69	1.20	2.79	1.42
Panel C: Some college				
Program Hours (τ_p)	18.23	10.52	18.87	10.97
Non-Program Hours ($\tau_m + \tau_n$)	69.27	10.52	68.63	10.97
Child cognitive skill, 36-months ($\ln h_1$)	99.09	19.31	95.68	20.03
Maternal-care hours (τ_m)	49.51	11.79	57.34	10.09
Nonmaternal-care hours (τ_n)	19.76	12.10	11.29	12.08
Earning hours (L)	21.79	16.57	25.40	16.24
Maternal-care quality ($\ln q_m$)	0.67	0.70	0.31	1.00
Nonmaternal-care quality (q_n)	3.42	1.70	2.81	1.13
Panel D: College graduates				
Program Hours (τ_p)	16.16	11.09	16.58	11.42
Non-Program Hours ($\tau_m + \tau_n$)	71.34	11.09	70.92	11.42
Child cognitive skill, 36-months ($\ln h_1$)	111.87	18.81	113.42	18.19
Maternal-care hours (τ_m)	54.24	11.33	52.90	11.02
Nonmaternal-care hours (τ_n)	17.10	12.53	18.02	13.26
Earning hours (L)	24.42	17.11	17.97	16.26
Maternal-care quality ($\ln q_m$)	0.99	0.55	1.13	0.92
Nonmaternal-care quality (q_n)	5.11	4.73	5.83	2.64

Table B-6: Average Elasticities by Maternal Education

	Baseline Level (1)	Wage Offer Elasticity (2)	Child Care Cost Elasticity (3)	Non-Labor Inc. Elasticity (4)
Panel A: Less than high school				
Child cognitive skill, 36-months ($\ln h_1$)	81.09	0.07	-0.08	0.02
Maternal-care hours (τ_m)	64.65	-0.13	0.10	0.01
Nonmaternal-care hours (τ_n)	22.85	0.67	-0.47	0.00
Earning hours (L)	11.31	3.06	-0.50	-2.51
Maternal-care quality ($\ln q_m$)	-0.27	-0.01	-0.04	0.02
Effort (e)	1.11	0.21	-0.16	-0.02
Nonmaternal-care quality (q_n)	2.94	0.58	-0.91	0.42
Panel B: High school only				
Child cognitive skill, 36-months ($\ln h_1$)	81.70	0.08	-0.08	0.00
Maternal-care hours (τ_m)	61.30	-0.14	0.10	0.06
Nonmaternal-care hours (τ_n)	26.20	0.63	-0.35	-0.09
Earning hours (L)	14.88	2.67	-0.32	-2.15
Maternal-care quality ($\ln q_m$)	0.02	3.82	0.01	0.04
Effort (e)	1.21	0.22	-0.15	-0.09
Nonmaternal-care quality (q_n)	2.62	0.64	-0.93	0.33
Panel C: Some college				
Child cognitive skill, 36-months ($\ln h_1$)	91.84	0.08	-0.08	-0.01
Maternal-care hours (τ_m)	57.44	-0.16	0.11	0.10
Nonmaternal-care hours (τ_n)	30.06	0.49	-0.31	-0.29
Earning hours (L)	25.38	1.76	-0.24	-1.63
Maternal-care quality ($\ln q_m$)	0.48	0.56	-0.60	-0.54
Effort (e)	1.44	0.25	-0.17	-0.15
Nonmaternal-care quality (q_n)	3.24	0.79	-0.94	0.20
Panel D: College graduates				
Child cognitive skill, 36-months ($\ln h_1$)	108.96	0.07	-0.07	-0.00
Maternal-care hours (τ_m)	56.37	-0.15	0.12	0.07
Nonmaternal-care hours (τ_n)	31.13	0.37	-0.26	-0.14
Earning hours (L)	22.10	2.60	-0.27	-1.76
Maternal-care quality ($\ln q_m$)	1.06	0.42	0.31	-0.05
Effort (e)	1.39	0.24	-0.18	-0.11
Nonmaternal-care quality (q_n)	6.02	0.71	-0.92	0.27

Note: All statistics are averages of simulated endogenous variables in the treatment sample. Column 1: in the control condition. Column 2: Elasticity with respect to a 1 percent increase in the wage offer. Column 3: Elasticity with respect to a 1 percent increase in the price of non-maternal care. Column 4: Elasticity with respect to a 1 percent increase in non-labor income.

Table B-7: Outcomes under Counterfactual Policy Simulations by Maternal Education

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ Prog. Qual. (5)	Full Take Up (6)	$1/hour Co-pay (7)
Panel A: Less than high school							
Program Hours (τ_p)	0.00	18.75	0.00	0.00	2.67	45.00	6.80
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.75	87.50	87.50	84.83	42.50	80.70
Child cognitive skill, 36-months ($\ln h_1$)	78.47	86.24	80.87	79.87	79.14	100.68	81.91
Maternal-care hours (τ_m)	64.16	59.85	63.43	62.32	62.74	41.65	60.67
Nonmaternal-care hours (τ_n)	23.34	8.90	24.07	25.18	22.09	0.85	20.04
Earning hours (L)	12.52	11.04	3.18	16.09	12.65	13.25	13.00
Maternal-care quality ($\ln q_m$)	-0.32	-0.23	-0.31	-0.28	-0.30	0.30	-0.25
Nonmaternal-care quality (q_n)	2.63	2.82	4.60	3.01	2.66	3.18	2.68
Effort (e)	1.08	1.23	1.11	1.13	1.13	2.20	1.20
Policy cost, weekly $/child	0	327	327	45	8	831	119
Panel B: High school only							
Program Hours (τ_p)	0.00	18.84	0.00	0.00	4.20	45.00	8.46
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.66	87.50	87.50	83.30	42.50	79.04
Child cognitive skill, 36-months ($\ln h_1$)	86.69	94.61	87.90	88.13	87.76	108.51	90.83
Maternal-care hours (τ_m)	63.46	58.33	64.29	61.54	60.77	41.65	58.80
Nonmaternal-care hours (τ_n)	24.04	10.33	23.21	25.96	22.53	0.85	20.24
Earning hours (L)	14.91	14.07	4.23	18.83	15.41	17.09	15.74
Maternal-care quality ($\ln q_m$)	-0.09	0.02	-0.11	-0.04	-0.03	0.52	0.01
Nonmaternal-care quality (q_n)	2.61	2.79	3.88	3.06	2.67	3.13	2.70
Effort (e)	1.05	1.22	1.05	1.10	1.12	2.17	1.20
Policy cost, weekly $/child	0	364	364	62	13	831	148
Panel C: Some college							
Program Hours (τ_p)	0.00	18.87	0.00	0.00	3.40	45.00	6.45
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.63	87.50	87.50	84.10	42.50	81.05
Child cognitive skill, 36-months ($\ln h_1$)	88.24	95.68	88.21	89.88	88.84	108.57	91.42
Maternal-care hours (τ_m)	60.38	57.34	63.38	58.17	59.36	41.55	57.53

Continued on next page

Table B-7 – *Continued from previous page*

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ prog. quality (5)	Full Take Up (6)	$1/hour Co-pay (7)
Nonmaternal-care hours (τ_n)	27.12	11.29	24.12	29.33	24.73	0.95	23.52
Earning hours (L)	25.94	25.40	13.89	30.09	26.19	30.43	26.87
Maternal-care quality ($\ln q_m$)	0.24	0.31	0.16	0.30	0.26	0.78	0.31
Nonmaternal-care quality (q_n)	2.66	2.81	3.34	3.19	2.72	3.18	2.74
Effort (e)	1.34	1.46	1.23	1.42	1.37	2.37	1.45
Policy cost, weekly $/child	0	336	336	143	11	831	113
Panel D: College graduates							
Program Hours (τ_p)	0.00	16.58	0.00	0.00	1.97	45.00	2.81
Non-Program Hours ($\tau_m + \tau_n$)	87.50	70.92	87.50	87.50	85.53	42.50	84.69
Child cognitive skill, 36-months ($\ln h_1$)	109.09	113.42	109.89	110.65	109.37	122.76	110.31
Maternal-care hours (τ_m)	55.23	52.90	55.94	53.35	53.61	41.19	52.77
Nonmaternal-care hours (τ_n)	32.27	18.02	31.56	34.15	31.92	1.31	31.92
Earning hours (L)	19.55	17.97	12.39	23.68	19.94	18.94	20.26
Maternal-care quality ($\ln q_m$)	1.08	1.13	1.06	1.14	1.11	1.49	1.14
Nonmaternal-care quality (q_n)	5.49	5.83	6.56	6.41	5.57	6.62	5.60
Effort (e)	1.66	1.78	1.63	1.74	1.72	2.77	1.78
Policy cost, weekly $/child	0	298	298	157	6	831	49

Note: Cells present averages of simulated endogenous variables in the treatment sample. See notes in Table 8 for definitions of policy experiments in columns.

Table B-8: Experimental Average Treatment Effects by Maternal Education (Above-Median Bayley)

	Maternal Education Sub-Sample							
	Less than HS		HS grad.		Some coll.		Bachelors+	
Child cognitive skill at 36 months ($\ln h_1$)	8.63	(0.00)	13.64	(0.00)	12.59	(0.01)	9.33	(0.01)
Hours per week of program care (τ_p)	17.97	(0.00)	20.65	(0.00)	21.15	(0.00)	18.73	(0.00)
Hours per week of maternal care (τ_m)	-11.02	(0.00)	-9.01	(0.00)	-12.12	(0.00)	0.64	(0.86)
Hours per week other caretakers (τ_n)	-6.95	(0.02)	-11.64	(0.00)	-9.02	(0.02)	-19.37	(0.00)
Earning hours (L)	0.37	(0.90)	7.28	(0.05)	-1.16	(0.77)	-0.92	(0.85)
Maternal-care quality ($\ln q_m$)	0.32	(0.10)	0.16	(0.41)	0.31	(0.06)	-0.00	(0.99)
Quality non-maternal care (q_n)	-0.31	(0.28)	-0.03	(0.87)	0.35	(0.34)	0.20	(0.84)

Note: This sample includes only observations with above-median Bayley scores. The cells report Average Treatment Effects (ATE) estimates. For any outcome Y and each education sub-group X, we report the difference $E(Y|treatment, X) - E(Y|control, X)$. The corresponding p-value is to the right in parentheses. Hours per week at IHDP program care correspond to averages for members of the treatment group. Four education sub-groups: less than a high school degree (Less than HS), high-school graduates (HS grad.), some years of college (Some coll.), and at least a bachelor's degree (Bachelors+).

Table B-9: Experimental Average Treatment Effects by Maternal Education (Below-Median Bayley)

	Maternal Education Sub-Sample							
	Less than HS		HS grad.		Some coll.		Bachelors+	
Child cognitive skill at 36 months ($\ln h_1$)	9.65	(0.00)	18.23	(0.00)	3.30	(0.48)	-7.70	(0.29)
Hours per week of program care (τ_p)	17.52	(0.00)	18.65	(0.00)	14.61	(0.00)	12.09	(0.00)
Hours per week of maternal care (τ_m)	-15.10	(0.00)	-11.16	(0.00)	-3.60	(0.40)	-6.44	(0.23)
Hours per week other caretakers (τ_n)	-2.42	(0.27)	-7.49	(0.01)	-11.01	(0.02)	-5.66	(0.27)
Earning hours (L)	5.56	(0.02)	0.63	(0.86)	-9.23	(0.04)	8.62	(0.21)
Maternal-care quality ($\ln q_m$)	0.35	(0.03)	0.51	(0.02)	0.23	(0.35)	-0.27	(0.13)
Quality non-maternal care (q_n)	0.19	(0.11)	0.40	(0.09)	0.01	(0.98)	-1.73	(0.29)

Note: This sample includes only observations with below-median Bayley scores. The cells report Average Treatment Effects (ATE) estimates. For any outcome Y and each education sub-group X, we report the difference $E(Y|treatment, X) - E(Y|control, X)$. The corresponding p-value is to the right in parentheses. Hours per week at IHDP program care correspond to averages for members of the treatment group. Four education sub-groups: less than a high school degree (Less than HS), high-school graduates (HS grad.), some years of college (Some coll.), and at least a bachelor's degree (Bachelors+).

Table B-10: Experimental Average Treatment Effects by Maternal Education (Exclude Low Bayley)

	Maternal Education Sub-Sample							
	Less than HS		HS grad.		Some coll.		Bachelors+	
Child cognitive skill at 36 months ($\ln h_1$)	9.65	(0.00)	18.23	(0.00)	3.30	(0.48)	-7.70	(0.29)
Hours per week of program care (τ_p)	17.52	(0.00)	18.65	(0.00)	14.61	(0.00)	12.09	(0.00)
Hours per week of maternal care (τ_m)	-15.10	(0.00)	-11.16	(0.00)	-3.60	(0.40)	-6.44	(0.23)
Hours per week other caretakers (τ_n)	-2.42	(0.27)	-7.49	(0.01)	-11.01	(0.02)	-5.66	(0.27)
Earning hours (L)	5.56	(0.02)	0.63	(0.86)	-9.23	(0.04)	8.62	(0.21)
Maternal-care quality ($\ln q_m$)	0.35	(0.03)	0.51	(0.02)	0.23	(0.35)	-0.27	(0.13)
Quality non-maternal care (q_n)	0.19	(0.11)	0.40	(0.09)	0.01	(0.98)	-1.73	(0.29)

Note: This sample excludes 77 observations with Bayley scores below 1 standard deviation below the mean. The cells report Average Treatment Effects (ATE) estimates. For any outcome Y and each education sub-group X, we report the difference $E(Y|treatment, X) - E(Y|control, X)$. The corresponding p-value is to the right in parentheses. Hours per week at IHDP program care correspond to averages for members of the treatment group. Four education sub-groups: less than a high school degree (Less than HS), high-school graduates (HS grad.), some years of college (Some coll.), and at least a bachelor's degree (Bachelors+).

C Robustness to Alternative Specifications

C.1 Alternative Skill-Production Technology: Including Household Covariates

We first consider an alternative to the skill-production technology (16), including a vector of household and mother characteristics (X_i):

$$\ln h_{1i} = \delta_{h_0}(h_{0i}) + \delta_m(h_{0i})\frac{\tau_{mi}}{T_c}\ln q_{mi} + \delta_n(h_{0i})\frac{\tau_{ni}}{T_c}\ln q_{ni} + \delta_p(h_{0i})\frac{\tau_{pi}}{T_c} + X_i'\beta + \eta_{hi}, \qquad \text{(C-1)}$$

where the X_i vector includes mother's education, mother's test score, mother's age, mother's marital status, number of children under 5 years of age, and number of children 5 and older.

Table C-1 reports estimates of the counterfactual policy simulations with all parameters re-estimated using the alternative technology specification (C-1). Comparing these results to the original estimates Table 8, we see only small differences in the counterfactual results. This indicates that the main counterfactual results are robust to the alternative technology specification.

Table C-1: Outcomes under Counterfactual Policy Simulations: Including Household Covariates)

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ Prog. Qual. (5)	Full Take Up (6)	$1/hour Co-pay (7)
Program Hours (τ_p)	0.00	18.52	0.00	0.00	1.40	45.00	6.56
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.98	87.50	87.50	86.10	42.50	80.94
Child cognitive skill, 36-months ($\ln h_1$)	86.70	94.01	88.61	88.69	88.23	107.53	90.94
Maternal-care hours (τ_m)	62.07	58.03	61.70	59.31	60.31	41.38	57.60
Nonmaternal-care hours (τ_n)	25.43	10.95	25.80	28.19	25.78	1.13	23.34
Earning hours (L)	16.74	15.62	7.69	21.76	18.00	18.66	18.51
Maternal-care quality ($\ln q_m$)	0.04	0.12	0.04	0.11	0.08	0.61	0.14
Nonmaternal-care quality (q_n)	3.00	3.20	4.70	3.86	3.32	4.05	3.40
Effort (e)	1.20	1.34	1.20	1.28	1.24	2.28	1.34
Policy cost, weekly $/child	0	336	336	84	10	831	117

Note: cells present averages of simulated endogenous variables in the treatment sample.

1. *Control*: the control state of the experiment, in which program care is not offered.
2. *Treatment*: the treatment state, free access to IHDP program care is offered.
3. *Cash Transfer*: cash transfer to each household equal to the IHDP's per capita cost. No care subsidy.
4. *Wage +25%*: boosts each mother's wage by 25%. No care subsidy.
5. ↓ *Program-Care Quality*: similar to IHDP's program-care offer but quality of program care reduced to equal the control-group's average nonmaternal-care quality, $3.14 hour.
6. *Full Take-up*: forces all households to take-up treatment program care 45 hours weekly and simulates choices they would make on all other margins.
7. *Co-pay*: like treatment care offer but program care requires a $1 per hour co-pay.

C.2 Alternative Skill-Production Technology: Using Instrumental Variables to Estimate the Skill-Production Technology

We next consider an alternative estimator for the skill-production technology (16). We use the specification above, including household covariates (C-1) and estimate the parameters of this specification using instrumental variables based on the random experimental treatment assignment, and the treatment assignment interacted with included household covariates.

Table C-3 reports estimates of the counterfactual policy simulations with all parameters re-estimated using the alternative IV estimator. Comparing these results to the original estimates Table 8, we see only small differences in the counterfactual results. This indicates that the main counterfactual results are robust to using the IV estimator.

Table C-2: Outcomes under Counterfactual Policy Simulations: Using Instrumental Variables to Estimate the Skill Production Technology)

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ Prog. Qual. (5)	Full Take Up (6)	$1/hour Co-pay (7)
Program Hours (τ_p)	0.00	18.52	0.00	0.00	1.58	45.00	6.76
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.98	87.50	87.50	85.92	42.50	80.74
Child cognitive skill, 36-months ($\ln h_1$)	86.70	94.01	87.35	87.45	86.96	108.46	90.03
Maternal-care hours (τ_m)	62.07	58.03	62.62	60.17	61.10	41.57	58.39
Nonmaternal-care hours (τ_n)	25.43	10.95	24.88	27.33	24.82	0.93	22.35
Earning hours (L)	16.74	15.62	6.85	20.66	16.98	18.48	17.53
Maternal-care quality ($\ln q_m$)	0.04	0.12	0.02	0.09	0.06	0.61	0.12
Nonmaternal-care quality (q_n)	3.00	3.20	4.44	3.54	3.05	3.65	3.11
Effort (e)	1.20	1.34	1.18	1.25	1.22	2.29	1.32
Policy cost, weekly $/child	0	336	336	84	10	831	117

Note: cells present averages of simulated endogenous variables in the treatment sample.

1. *Control*: the control state of the experiment, in which program care is not offered.
2. *Treatment*: the treatment state, free access to IHDP program care is offered.
3. *Cash Transfer*: cash transfer to each household equal to the IHDP's per capita cost. No care subsidy.
4. *Wage +25%*: boosts each mother's wage by 25%. No care subsidy.
5. ↓ *Program-Care Quality*: similar to IHDP's program-care offer but quality of program care reduced to equal the control-group's average nonmaternal-care quality, $3.14 hour.
6. *Full Take-up*: forces all households to take-up treatment program care 45 hours weekly and simulates choices they would make on all other margins.
7. *Co-pay*: like treatment care offer but program care requires a $1 per hour co-pay.

C.3 Alternative Maternal-Care Quality Function: Including Measurement Error in Maternal-Care Quality

We next consider an alternative to the specification of the maternal-care quality function (17). We assume that the measure of latent maternal-care is given by M_{mi}, where the measure is the learning and literacy items of the observer based HOME score, as described in the main text This measure relates to latent maternal-care quality $\ln q_{mi}$ through this measurement relationship:

$$M_{mi} = \ln q_{mi} + \omega_{mi} \tag{C-2}$$

ω_{mi} is a classical measurement error term which is distributed $\omega_{mi} \sim$ i.i.d.$N(0, \sigma^2)$. We assume the measurement error is independent of the latent maternal-care quality and all other observable variables. Latent maternal-care quality remains as specified above:

$$\ln q_{mi} = X_i' \delta_{q_m} + \ln e_i, \tag{C-3}$$

where measured maternal-care quality is then

$$M_{mi} = X_i' \delta_{q_m} + \ln e_i + \omega_{mi} \tag{C-4}$$

For this robustness exercise, we set the standard deviation of the measurement error σ such that 50 percent of the residual variation ($\widetilde{M} = M_{mi} - X_i' \hat{\delta}_{q_m}$) is "noise":

$$\frac{V(\omega_{mi})}{V(\ln e_i) + V(\omega_{mi})} = 0.5$$

Using this specification of the measurement process, we then re-estimate the effort level for each household, adjusting for measurement error using a simulation algorithm. For each household i, we draw a measurement error ω' from the given ω_{mi} distribution, and compute latent effort as

$$\ln e_i = M_{mi} - X_i' \hat{\delta}_{q_m} - \omega'.$$

This compares to the baseline specification where effort is computed without measurement error as

$$\ln e_i = M_{mi} - X_i' \hat{\delta}_{q_m}.$$

Using this measurement error process we recompute all preference parameters, and re-simulate the counterfactual policies using the new estimates, displayed in Table C-3. Comparing these results to the original estimates Table 8, we see only small differences. This indicates that the main counterfactual results are robust to the alternative maternal-quality measurement specification.

Table C-3: Outcomes under Counterfactual Policy Simulations: Including Measurement Error in Maternal-Care Quality)

	Control (1)	Treatment (2)	Cash Transfer (3)	Wage ↑ 25% (4)	↓ Prog. Qual. (5)	Full Take Up (6)	$1/hour Co-pay (7)
Program Hours (τ_p)	0.00	18.52	0.00	0.00	3.13	45.00	6.72
Non-Program Hours ($\tau_m + \tau_n$)	87.50	68.98	87.50	87.50	84.37	42.50	80.78
Child cognitive skill, 36-months ($\ln h_1$)	86.70	94.01	88.64	88.75	87.95	107.79	90.52
Maternal-care hours (τ_m)	62.07	58.03	62.75	60.18	60.43	41.58	58.59
Nonmaternal-care hours (τ_n)	25.43	10.95	24.75	27.32	23.94	0.92	22.19
Earning hours (L)	16.74	15.62	6.75	20.60	17.00	18.45	17.40
Maternal-care quality ($\ln q_m$)	0.04	0.12	0.01	0.09	0.07	0.61	0.11
Nonmaternal-care quality (q_n)	3.00	3.20	4.41	3.50	3.05	3.61	3.07
Effort (e)	1.20	1.34	1.44	1.54	1.53	2.87	1.62
Policy cost, weekly $/child	0	336	336	84	10	831	117

Note: cells present averages of simulated endogenous variables in the treatment sample.

1. *Control*: the control state of the experiment, in which program care is not offered.
2. *Treatment*: the treatment state, free access to IHDP program care is offered.
3. *Cash Transfer*: cash transfer to each household equal to the IHDP's per capita cost. No care subsidy.
4. *Wage +25%*: boosts each mother's wage by 25%. No care subsidy.
5. ↓ *Program-Care Quality*: similar to IHDP's program-care offer but quality of program care reduced to equal the control-group's average nonmaternal-care quality, $3.14 hour.
6. *Full Take-up*: forces all households to take-up treatment program care 45 hours weekly and simulates choices they would make on all other margins.
7. *Co-pay*: like treatment care offer but program care requires a $1 per hour co-pay.