MOTHER OR MARKET CARE? A STRUCTURAL ESTIMATION OF CHILD CARE IMPACTS ON CHILD DEVELOPMENT

Ylenia Brilli*†

April 2013
Do not quote without permission

Abstract. This paper analyzes the effects of maternal employment and external child care on child development, taking into account the additional choice the mother makes between leisure and time with the child. I estimate a behavioral model where labor supply, external child care and leisure time allocation are endogenously chosen by the mother and represent the inputs for the child cognitive development. The model is estimated using U.S. data from the Child Development Supplement and the Time Diary component of the Panel Study of Income Dynamics. Results show that a reduction in maternal time with the child induces a negative effect on child's ability that is compensated for by the use of external child care. This implies that maternal employment is not detrimental for child development. The estimated model is used to simulate the effects of policies aimed at boosting the economic conditions of the household or at subsidizing external child care: results show that an increase in household income has a positive effect on child development but negatively affects mothers' labor supply; instead, policies increasing mother's wage or subsidizing external child care may have small effects on child's development.

JEL classification: D13, J13, J22.

Keywords: non-parental child care, mother employment, mother time allocation, child development, structural estimation.

*Collegio Carlo Alberto, CHILD and Catholic University of Milan.
E-mail: ylenia.brilli@carloalberto.org

†I am indebted to Daniela Del Boca for her constant encouragement and advice. I am also grateful to Christopher Flinn, Cristian Bartolucci and Matthew Wiswall for their outstanding suggestions and useful discussions. I wish to thank Marco Francesconi, Chiara Monfardini, Silvia Pasqua, Erich Battistin, Ugo Colombino, Andrea Ichino, Daniele Checchi, Cheti Nicoletti, Nabanita Datta Gupta, Steven Rivkin, Lorenzo Cappellari, Claudio Lucifora, Alessandro Cigno, Marco Tonello, Toomas Hinnosaar, seminar participants at University of Turin, Catholic University of Milan, Collegio Carlo Alberto and participants at the 2012 Brucchi Luchino Labor Economics Workshop (Trento), the 2013 Workshop on Economics of Successful Children (Aarhus University), the 2013 ICEEE (University of Genoa) and the 2013 RES PhD Presentation Meeting (London) for their useful comments. This research has been mostly conducted at Collegio Carlo Alberto and when I was a visiting student at New York University: the hospitality of both institutions and financial support from Collegio Carlo Alberto are gratefully acknowledged. Any error should be attributed to the author.

1. Introduction

During the last decades, there has been a growing interest in studying the determinants of child cognitive achievement. Not only psychologists but also economists agree that one of the most valuable inputs for child's development is the time the child spends with the mother (Cunha et al. 2006). Indeed, the increase in maternal employment rate and the use of external forms of child care have raised concerns about the impacts that they may have on child's development. In the U.S., the participation rate of married women increased from around 40 percent in the 1970s to more than 60 percent at the end of the 1990s. Moreover, almost 65 percent of children aged 3-5 were enrolled in nursery schools before kindergarten during the 1990s (U.S. Census Bureau 2000). The majority of children of employed mothers regularly spend time in some forms of non-school non-parental care, so that the use of non-parental child care continues when the child reaches primary school age (Blau and Currie 2006).

The psychological literature argues that maternal employment and external child care may determine insecure mother-child attachments, which are formed in the first years of a child's life (Cox et al. 1992). The economic literature on these topics is large, but findings are mixed and very few studies adopt a credible identification strategy. Moreover, a research that evaluates the effects of maternal employment and non-parental child care taking into account the actual time spent by the mother with the child is lacking. This paper provides new evidence on the impact of maternal employment and external child care, taking also into account the additional choice of the mother on the amount of time to spend with the child.

The identification of the effects of both mother's employment and external child care choices on child's development is hampered by their correlation with both mother's and children's skills, as well as by their simultaneity. Most of the studies use Ordinary Least Square to evaluate the effect of maternal employment and external child care, but they are very likely to fail in taking into account these sources of endogeneity. Other identification strategies that have been adopted in the literature to account for these issues encompass Mother or Siblings fixed effects and Instrumental Variables estimators. While the former allows to get rid of time invariant unobserved heterogeneity of the mothers, the consistency of the latter relies on the assumption that the mothers react to the instruments without taking into account child's ability.

An alternative approach that has not yet been extensively used in this literature is structural estimation. This is the only strategy allowing to model different sources of endogeneity and, most importantly, to describe the decision making process of the mother for more than one endogenous choice. Moreover, it provides parameters estimates from theoretical model that can be used to simulate the effects of related policies.

Despite there being several studies in the child development literature using this approach, only Bernal (2008) applies this framework to the maternal employment and child care case. Moreover, a study that estimates these impacts taking into account the actual time spent by the mother with the child is lacking. In fact, both Bernal (2008) and all other studies using a reduced-form approach assume a specific relationship

between maternal employment time and the time the mother spends with the child (Keane 2010). More precisely, they estimate a cognitive ability production function with maternal employment as an input, arguing that the time the mother spends with the child can be proxied by the total amount of time available to the mother net of the time she spends at work. This definition of maternal time implies that the mother dedicates to the child all the remaining time out of work and that she does not care about having leisure. The most common finding of these studies is that maternal employment and external child care have a substantially negative effect on child's ability.

This assumption is very likely to fail if the mother decides how to allocate the time out of work between leisure and care of the child. Moreover, mothers' time allocation may substantially differ across employment status, since non-working mothers have more time out of work at disposal. For instance, Bianchi (2000) suggests that employed mothers allocate their time in such a way to give priority to the time they spend with the child. Hoffert and Sandberg (2001) show that there is not a one-to-one corresponding relationship between the time the mother spends out of work and the time she actually spends with the child.

In this paper, I estimate a behavioral model where the labor supply, non-parental child care and time allocation choices of the mother are endogenous. The model describes the mother's decisions to work, to use external child care and to spend time with the child starting from childbirth up to age 13. The model allows a direct estimation of the impact of maternal time on child's development, accounting for the fact that the mother not only chooses how many hours to work and how much external child care to use, but also how much time to devote to the child instead of having leisure. The mother's utility maximization problem is subject to the mother's time and budget constraints, as well as the child cognitive ability production function: the mother cares about consumption, leisure and the child's cognitive ability, while child's ability is specified with a value-added functional form and depends on the inputs received in the previous period. The empirical specification of the model introduces several sources of heterogeneity: the mother's preference parameters depend on mother's observable characteristics, while mother's unobserved skills affect the taste for child's ability, the participation to the labor market, the demand for non-parental child care and the choice between leisure and time with the child; finally, the child's initial endowment, i.e., the child's level of ability at birth, depends on both mother's and father's education, capturing a non-zero correlation between child's skills and parents' educational attainments.

The model is estimated using U.S. data from the Panel Study of Income Dynamics (PSID) and the Child Development Supplement (CDS) conducted in 1997, 2002 and 2007. The CDS provides retrospective information on all child care arrangements used since birth and widely-recognized measures of child's cognitive outcomes; the Time Diary (TD) section provides unique data on the amount of time the child spends with the mother. The main PSID surveys give detailed information on mother's work history and household income during the child's life cycle. The parameters of the model are

estimated using a Simulated Minimum Distance (SMD) estimator that minimizes the distance between several data statistics and their model counterparts.

This paper contributes to the literature evaluating the effects of maternal employment and external child care on child development in two ways. First of all, it provides estimates of the effect of maternal time and external child care relaxing the assumption that mother time out of work is a good proxy for maternal time with the child. In fact, differently from all existing studies on this topic, it takes into account the additional (endogenous) choice of the mother between leisure time to spend alone and time to spend with the child. Second, this paper represents the first attempt to estimate the elasticity of child's ability with respect to both maternal time and external child care time in a child cognitive production function framework. To the best of my knowledge, there are not studies that simultaneously evaluate the productivity of both inputs.[1]

The results show that more skilled and more educated mothers have higher preferences for child's ability. This implies that, even if they work more, they also make higher investments on their child's cognitive ability, either spending more time with the child or choosing more external child care or both. The estimated parameters in the child's cognitive ability production function show that, for an equal amount of maternal time and external child care time, the marginal productivity of maternal time is slightly lower than the one of external child care. Hence, if the mother works, a reduction in child's ability induced by a reduction in maternal time can be fully compensated for if the child spends the same amount of time in external child care.

In order to further understand this result, I re-estimate the model assuming leisure-minimizing preferences of the mother: the mother does not care about leisure and spends all the time out of work with the child. In this case, a reduction in maternal time due to mother's work induces a reduction in child's ability that is not compensated for by the use of external child care. In other words, the final effect of mother's employment is clearly negative. This result almost replicates what has been mostly found before in the literature. Hence, previous studies, defining maternal time as a residual from maternal working time, have overestimated maternal time productivity and the negative effect of maternal employment.

The estimated model is used to simulate the effects of policies improving the economic conditions of the household or decreasing the price of external child care. The results show that policies aimed at improving the economic conditions of the households have positive effects on child development while decreasing mother's labor supply; instead, policies aimed at boosting mothers' employment or at subsidizing external child care may have very little effects on child's development. Hence, the policy maker should take this into account when implementing such interventions.

The rest of the paper is organized as follows. Section 2 provides a background of the literature and presents some stylized facts in external child care use and maternal time allocation. Section 3 presents the model that is estimated: subsection 3.1 defines

[1]Recently Del Boca, Monfardini, and Nicoletti (2012) and Hsin (2009) have exploited PSID-CDS data to assess the effects of several time inputs on child development, one of them being maternal time. However, they do not consider external child care time as a substitute for maternal time with the child.

the basic structure, while subsection 3.2 discusses how the model is solved and presents the demand functions for all the choice variables. Section 4 presents the econometric specification of the model (subsection 4.1) and the empirical method used for the estimation (subsection 4.2). Section 5 describes the data and the sample used for the estimation, while section 6 presents the results and discusses the goodness of fit of the model (subsection 6.1). Section 7.2 presents the results from the counterfactual exercises and policy simulations and section 8 concludes.

2. Background

The increase in female employment rate that has characterized all developed countries has raised concerns for the impacts that maternal employment and external child care may have on child development. This is one of the reasons why, in the last decades, many studies try to assess the effects of these choices.

Starting from Becker and Tomes (1986), who first provide a framework for the implications of household decisions for children's subsequent utility and earnings, there has been a growing literature on the impacts of parental investments on children human capital and development. However, the studies on maternal employment and external child care present mixed findings. Several reduced-form studies find negative effects of maternal employment (Baydar and Brooks-Gunn 1991; Belsky and Eggebeen 1991; Desai, Chase-Lansdale, and Michael 1989; Ruhm 2004), while others find null effects (Chase-Lansdale et al. 2003; James-Burdumy 2005; Parcel and Menaghan 1994). Also studies on non-parental child care using reduced-form strategies provide ambiguous results. Bernal and Keane (2011) report that one year of child care use decreases children's cognitive outcomes, measured by the PIAT and PPVT scores, by 2.13 percent. Currie and Thomas (1995, 1999), instead, evaluate the impacts of the early childhood program Head Start and find that children who attended the program get higher scores at PIAT reading and Math test. Similarly, Magnuson, Ruhm, and Waldfogel (2007) find positive effects of having attended pre-kindergarten on academic achievement at kindergarten and primary school. Loeb et al. (2007) find that staying in center-based child care for more than 15 hours per week increases reading and Math score by almost 8 and 7 percent of a standard deviation.

The identification of the impacts of both maternal employment and external child care on child's development is hampered by three main sources of endogeneity. The first is due to the selection of mothers into employment and external child care use, induced by the correlation between mothers' choices and their unobservable skills. For instance, more skilled mothers may be more likely to work and to use external child care. If their skills are transmitted to their child in a way that the researcher cannot control for, this may overestimate the true effect of external child care. The second source of endogeneity depends on the correlation between mother's decisions and child's ability that the mother can (partially) observe, while the researcher does not. Finally, unobserved heterogeneity of both mothers and children is more difficult to take into account due to the simultaneity of these choices. While studies using OLS are very likely to fail in taking into account these sources of endogeneity, there are studies using

other techniques to handle these issues. Currie and Thomas (1995, 1999) use Mother fixed effects to control for time invariant unobserved heterogeneity of the mother, while Bernal and Keane (2011) use an Instrumental Variables estimator to take into account the correlation between mother's choices and child's ability. While the first strategy is robust to the correlation of the mother's decisions with mother's skills that do not vary over time, the second provides consistent estimates of the effects of interests only if it can be assumed that the mother reacts to the instruments without taking into account child's ability.

Structural estimation allows to account for the sources of endogeneity that may arise in this context, modeling the mother's decision making process for different choice variables. In this framework, each input is optimally chosen by the mother who maximizes her own utility function, with child's ability as an argument, and the child's ability production function is one of the constraints to this maximization problem. There are few studies using structural estimation in the child development literature. The model presented in this paper builds on Del Boca, Flinn, and Wiswall (2013), who estimate the impacts of parents' time inputs on children's development. They model household choices and investments in child quality from childbirth up to the last developmental period, when the child reaches adolescence, defining a dynamic discrete-time model where, in each period, both parents decide how many hours to work, how much time to spend with the child and the amount of expenditure for the child. In the definition of the choice variables, they distinguish between active and passive time of the mother, i.e., if the mother is directly involved in the child's activities or if she is just around without participating. They estimate that the elasticity of child ability with respect to maternal active time ranges from 0.25 when the child is two years old to 0.05 when he reaches 15 years of age; indeed, they find a very small elasticity for maternal passive time. Differently from Del Boca et al. (2013), this paper does not model both parents' labor supply and time allocation decisions, focusing only on mother's behavior and on the additional choice of using external forms of care when the mother is at work; in other words, instead of considering father's time as a substitute for mother's time with the child, the present study analyzes external child care time as a substitute for maternal child care time for the development of the child.

Mroz, Liu, and Van der Klaauw (2010) specify and estimate a behavioral model of household migration and maternal employment decisions in order to assess the effect of these choices on child's cognitive ability. They find that part-time employment of the mother reduces the child's score by 3 percent of a standard deviation while the mother's full-time status reduces the score by 5 percent of a standard deviation. Recently, Ermish and Francesconi (2012) evaluate the effects of maternal employment on child's schooling, estimating the parameters of a conditional demand function for child's education; they find that one year more of mother's full time employment reduces the probability that the child gets higher education by 11 percentage points.

Bernal (2008) is the only study that evaluates the impact of both maternal employment and external child care attendance on subsequent child outcomes using a structural approach.[2] She defines a single agent discrete-time multi-period model, where the mother decides among different combinations of work and child care use. The choice variables are discrete and the dichotomous variable for using child care is defined as being equal to one if the child ever used (in each period) any form of non-parental care. Bernal's main contribution is to consider the impact of the work and child care choices in the first five years of life of the child and to test whether the mother decides to work and to use external child care after having observed the child's initial ability endowment. She finds that one year in external child care reduces the child's cognitive ability by 0.8 percent; however, the impact of mother's employment and external child care is even more detrimental, as it decreases child's outcome by 1.8 percent.

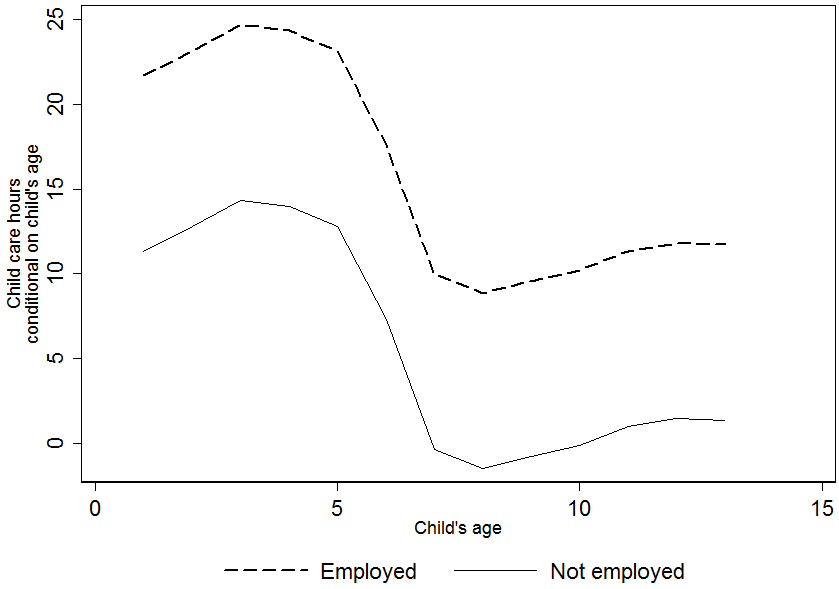
Differently from Bernal and from all other existing studies on maternal employment and child development, this paper exploits the actual measure of maternal time instead of using a proxy for that; this implies that all the three choices (i.e., labor supply, maternal child care time, external child care time) are treated as endogenous.

It should be noticed that the assumption concerning mothers' time allocation used in the previous studies may have implications for the effect that is actually estimated: in fact, arguing that maternal time with the child can be proxied by the amount of time the mother spends out of work rules out the possibility that mothers could choose how to allocate their time between leisure and time with the child; moreover, since employed mothers have less time at disposal, they necessarily spend a lower amount of time with their child. Indeed, even though data on mothers' and children's time use have become available only very recently, there have been some studies suggesting that mothers do not differ only in terms of participation decisions but also in terms of leisure time allocation. Leibowitz (1974, 1977) points out that more skilled and more educated mothers may also have a higher propensity to stay with their child, even if working. Recent studies on mothers' time use confirm this point, since they do not find significant differences across employment status in the amount of time mothers spend with their child (Bianchi 2000; Hoffert and Sandberg 2001).

The absence of significant differences in maternal time with the child between working and non-working mothers can be attributed to two main reasons. On the one hand, during recent years, also non-working mothers have started using external child care, so that children of non-working mothers may not be always available for maternal investments while attending external child care. For instance, Bianchi (2000) shows that from the end of the 1960s to the end of the 1990s, the percentage of 3-5 children enrolled in some forms of pre-primary educational programs increased from 7.9 to 51.7 for mothers in the labor force and from 4.8 to 44 percent for mothers not in the labor force. Blau and Currie (2006) report that this trend is confirmed for school-age children, who regularly spend time in some forms of after-school programs. On the other hand,

[2]Bernal and Keane (2010) propose a quasi-structural estimation of a model very similar to Bernal (2008). Their results suggest that one year of full time work and external child care use implies a reduction in child's ability of around 2 percent.

FIGURE 1
External child care time by mother's employment status.



NOTE. The vertical axis represents the fitted values of the following regression:

$$childcare_i = \eta_0 + \eta_1 t_i + \eta_2 d_i + \epsilon_i$$

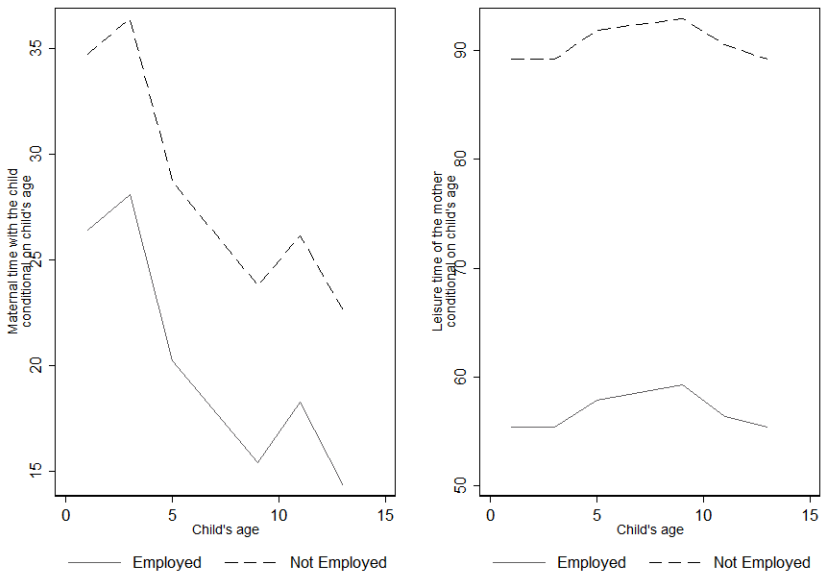
where $childcare_i$ represents (weekly) hours of external child care, t_i are child's age fixed effects, d_i is a dummy variable equal to 1 if the mother of child i works. $\eta_2 = 10.36$ represents the difference in average child care use (conditional on child's age) between working and non-working mothers. Source: own elaboration from PSID-CDS data ($N = 3510$).

working and non-working mothers may allocate their time out of work differently, so that the actual time that they spend with the child does not correspond to the time they spend out of work. According to data from the American Time Use Survey 2005-2009 (U.S. Census Bureau 2013), the amount of time spent by mothers reading and playing with the child does not vary crucially among employment status (4.76 for the working mothers vs 6.86 for the non-working ones); instead, employed mothers spend a significantly lower amount of time in activities like socialization or doing sport, i.e., activities usually defined as *leisure* (U.S. Census Bureau 2013).

Descriptive evidence from the PSID-CDS data used in this paper supports the existence of these patterns. Figure 1 shows that also non-working mothers use a positive amount of external child care for their child. This may happen if, for instance, they value the educational role of the service and choose it as an investment in their child's human capital. However, since the difference in average child care time between working and non-working mothers is equal to 10.36, the graph also confirms that child care is needed for its custodial purposes anytime the mother is working.

Figure 2 plots the fitted values of two regressions where the dependent variables are, respectively, maternal time with the child and leisure time. The graph on the left confirms that employed mothers allocate their time out of work in order to spend a positive amount of time with their child. Conversely, non-working mothers do not spend all their time with the child, but only around 30 hours per week when the child

FIGURE 2
Maternal time with the child and leisure by mother's employment status.



NOTE. The vertical axis in the graph on the left represents the fitted values of the following regression:

$$\tau_i = \eta_0 + \eta_1 t_i + \eta_2 d_i + \epsilon_i$$

while the vertical axis in the graph on the right represents the fitted values of the following regression:

$$l_i = \beta_0 + \beta_1 t_i + \beta_2 d_i + \varepsilon_i$$

τ stems for (weekly) maternal time with the child and l represents leisure time, computed as $l = TT - \tau - h$ where $TT = 112$ is the total time endowment and h represents weekly hours of work. t_i are child's age fixed effects and d_i is a dummy variable equal to 1 if the mother of child i works. $\eta_2 = -8.32$ represents the difference in average maternal time (conditional on child's age) between working and non-working mothers. $\beta_2 = -33.72$ represents the difference in average leisure time (conditional on child's age) between working and non-working mothers. Source: own elaboration from PSID-CDS data ($N = 624$).

is very young and around 25 when the child grows up. The graph on the right shows the fitted values of a regression on child's age fixed effects where the dependent variable is leisure time, computed as the difference between the total time endowment and the sum between working time and time with the child. Employed mothers spend a very low amount of time out of work in leisure, while the corresponding level for non-working mothers is considerably higher. Notice that while the difference in maternal time with the child between working and non-working mothers is equal to 8 hours per week, the difference in leisure is equal to 33 hours per week.

3. THE MODEL

This section describes the theoretical model on which the estimation is based. Subsection 3.1 presents the basic structure of the model, while subsection 3.2 derives the demand functions for all the choice variables.

3.1. ***Basic structure.*** The model follows a standard framework from Becker and Tomes (1986), where household preferences are described by a unitary utility function, with child's ability as an argument, and subject to a production function for

child's ability plus budget and time constraints.[3] The functional form assumptions are based on the theoretical model developed in Del Boca et al. (2013).

The model applies to intact households, where both the mother and the father are present. I consider only households with one child and I assume that the mother is the unique decision maker in the household concerning the work and external child care use decisions. This assumption implies that father's labor supply is exogenous with respect to child development[4] and that the father does not bargain with the mother concerning the external child care choice. However, the model allows the father to affect child development in two ways: first, the child's ability endowment depends also on father's education; secondly, father's labor income contributes to household earnings that are an input in the child cognitive production function and influence mother's choices concerning work, external child care and time with the child. Finally, the simplification concerning the number of children allows to avoid modeling the fertility decisions of parents and to make additional assumptions on the different effects of investments on more siblings.[5]

The model is dynamic and evolves in discrete time. In each period, the mother decides her own labor supply and time allocation, as well as the amount of external child care to use. The choice variables are then: (i) h_t, representing hours of work; (ii) i_t, hours of external child care and (iii) τ_t, the time the mother spends with the child. The timing is defined as follows: $t = 0$ represents the birth of the child and the mother makes all the decisions (in a relevant way for the child's development process) at each child's age t until the child reaches T years of age; $t = 1$ indicates the first 12 months of the child's life, $t = 2$ refers to the next 12 months of the child's life, and so on and so forth; $t = T$ represents the terminal period of the model. One may interpret this terminal period as the final of a specific developmental stage of the child, so that starting from this period both the mother's utility maximization problem and the child's cognitive production function change.[6]

[3]As pointed out by Blau (1999), the basic elements of any economic theory for the effect of an input on child development "are (i) a utility function that contains child outcomes as arguments; (ii) a production function for the child outcomes with inputs including the time of family members and purchased goods and services; (iii) budget and time constraints; and (iv) a specification of the information structure and the formation of expectations."

[4]Actually, this assumption mostly follows from the characteristics of the sample of intact households that I see in the data. In fact, all fathers in the sample work and the average working time does not change across child's age or across mother's participation decisions. Del Boca et al. (2013) allow also the father to decide his investments on child development, finding that fathers' time become more important than mother's time for child development starting from age 12. Since the present study considers mother's investments decisions up to age 13, the absence of father's time in the production function for child development would not affect the main results. Conversely, due to the importance of the alternative forms of care used for the child in this period of the child life, the model presented in this paper considers the external child care provider's time as a substitute for mother's own time.

[5]Of course, these assumptions on sample selection may determine biases in the parameters estimates; this issue will be treated in section 5.

[6]$T = 13$. It may be interpreted as the final period of middle childhood before the child enters adolescence. The definition of the terminal period theoretically implies that after T mother's investments do not have any impact on child development. According to table E.1, controlling for the choice variables, the child's age at which the cognitive test score is maximized is equal to 11.92. As a sensitivity analysis, I repeat the estimation of the model setting $T = 12$. Results are shown in appendix E.1.

The Mother's Utility Function

Mother's utility in each period is a function of her own leisure time (l_t), i.e, the time the mother spends alone without working, household consumption (c_t), including father's and child's consumption, and the child's cognitive ability (A_t). I assume a Cobb-Douglas form for preferences and I restrict the preferences parameters to be stable over time:

$$u(l_t, c_t, A_t) = \alpha_1 lnl_t + \alpha_2 lnc_t + \alpha_3 lnA_t \tag{1}$$

where $\sum_{j=1}^{3} \alpha_j = 1$ and $\alpha_j > 0$, $j = 1, 2, 3$.

The mother maximizes her utility subject to the budget and the time constraints. The budget constraint takes into account household consumption and the total income available in the family (from both parents' labor supply and non labor income) and is given by:

$$c_t = w_t h_t + I_t - pi_t \tag{2}$$

where w_t is mother's hourly wage; I_t represents household earnings (including father's labor income and household non labor income); i_t represents the number of hours that the mother uses non-parental child care and p is the hourly price of child care. The variable i_t includes any kind of non-parental child care arrangement. Finally, the mother does not make saving decisions, hence household income defined by I_t can be considered exogenous with respect to all mother's choices.[7]

The time constraint takes into account both the leisure time the mother spends alone and the time the mother devotes to the child:

$$TT = l_t + h_t + \tau_t \tag{3}$$

where TT is the mother's total time endowment,[8] h_t is the number of hours the mother works and τ_t is the number of hours the mother spends with the child. Notice that, in each period, the mother can choose to spend her leisure time alone (l_t) or to devote some time to the child (τ_t): hence, the model allows the mother to further choose between leisure and time with the child when she is not at work.

[7]The fact that the budget constraint is solved period by period rules out the possibility of taking into account parents' savings aimed at boosting the educational opportunities of the child in the future. However, the static nature of the budget constraint is needed in order to decrease the computational burden and the dimensionality of the state space; this simplification is quite common in dynamic models of labor supply and the specification of a model with saving decisions is beyond the scope of the paper.
[8]$TT = 112$ hours per week: it assumes 16 hours per day, that the mother should allocate between working, leisure and time with the child. All choice variables are defined on a weekly basis.

The Child's Cognitive Ability Production Function

The child's cognitive ability production function (hereafter CAPF) is defined using a value-added specification and taking a Cobb-Douglas form:

$$lnA_{t+1} = \delta_{1t} ln\tau_t + \delta_{2t} lni_t + \delta_{3t} lnI_t + \delta_{4t} lnA_t \tag{4}$$

where A_{t+1} is the outcome for a child at time $t+1$, τ_t and i_t are the inputs decided by the mother in each period t; I_t represents the income of the household, as already defined, and A_t is the level of child ability at period t. Since current ability influences future child's ability, equation (4) shows that inputs operate with a lag: development takes time. Moreover, the structure of equation (4) implies that when deciding the inputs on child development, the mother knows the productivity of each of them and the level of child's ability in the previous period.[9]

Despite posing some limitations on the substitution pattern across inputs due to the assumed functional form, the model allows the parameters in (4) to vary across child's ages in order to capture the fact that marginal productivity of inputs varies over the stages of child development (Cunha, Heckman, and Schennach 2010; Heckman 2007).[10]

Mother's work is not explicitly included in the CAPF, because it may not have a direct impact on child development *per se*. Mother's employment may indirectly affect child development through the change in the mother's time allocation, together with the use of non-parental child care. The child care input includes all contributions to child development due to the alternative care providers' time and may be more or less productive than mother's own time. This specification allows to test whether, in each period, maternal time is more productive than external child care time. If this is the case, then, for any period and for an equal amount of maternal time and child care time used, $\delta_{1t} \geq \delta_{2t}$.[11]

While the amount of non-parental child care can represent a measure of the services bought for the child, the household income in (4) proxies the expenditure in goods for the child (Todd and Wolpin 2003). The use of I_t as a proxy for the goods bought for the child relies on two assumptions: (i) a constant proportion of income is devoted to

[9]This implies that the estimated parameters in the CAPF are robust to the mother deciding her investments after having observed the child's level of ability. As pointed out by Keane (2010), "assumptions on what the mother knows are essential if the econometrician is to solve the mother's dynamic optimization problem". While assuming that the mother knows more about the child's ability production function than the econometrician does can be reasonable, it may be unrealistic to assume that the mother has complete information. However, the mother can learn about the productivity of each input and the child's ability after some realizations. This paper does not model this learning mechanism, simplifying to the mother's complete information case. See Fogli and Veldkamp (2011) or Fernandez (2007) for a model where the mother is uncertain about the effect of her employment on child's development and chooses her labor market participation according to the available information set.

[10]Notice also that, despite assuming a constant elasticity of child development in each period, the specification of the CAPF allows the marginal productivity of inputs to be decreasing with respect to each input level. This is consistent with a framework where the productivity of an input, e.g. external child care or maternal time, is higher for low levels and decreases as long as the amount of the input increases.

[11]For any period t, the marginal productivity of maternal time is given by $MP_{\tau_t} = \frac{\delta_{1t}}{\tau_t}$, while the marginal productivity of external child care is $MP_{i_t} = \frac{\delta_{2t}}{i_t}$. For $\tau_t = i_t$, $MP_{\tau_t} \geq MP_{i_t}$ if $\delta_{1t} \geq \delta_{2t}$; viceversa, $MP_{\tau_t} \leq MP_{i_t}$ if $\delta_{1t} \leq \delta_{2t}$.

buy goods effective for child development and (ii) this proportion is not affected by the mother's labor supply decisions.[12] Furthermore, income can have a direct impact *per se* as long as it captures the economic conditions where the household resides (Blau 1999; Levy and Duncan 2012).

Concerning the amount of external child care used by the child, the model does not distinguish between different kinds of service (for instance, formal vs. informal arrangements). Hence, it is assuming that all types of care have the same impact on child development and that the mother's decision making process for the two types of care is similar. The same homogeneity is then reflected in the price of external child care. The model predicts a strictly positive price of the service, regardless of its nature. This implies that also services with a potentially zero price in the market are characterized by a shadow price, representing, for instance, the limited availability of informal care or the value of the unpaid care provider's time in alternative activities (Blau and Currie 2006; Ribar 1992). However, the assumptions that different kinds of services have the same productivity is quite strong and also prevents from capturing differences in mothers' behaviors related to their use. Due to the importance of this topic, I leave for future research the estimation of a model where the mother not only chooses how many hours to work and how much time to spend with the child, but also how much time to use formal or informal child care arrangements.

Maximization Problem

In each period, the mother maximizes her expected life time utility, optimally choosing her labor supply, the child care input and the number of hours to devote to the child. In this decision-making process the mother takes into account the level of ability reached by the child in each period, the wage offer that she receives from the market and the level of income in the household. The child's cognitive ability represents an endogenous state variable, while the wage offer the mother receives in each period and household income are exogenous with respect to the maximization problem but differ for each mother in each period. The initial condition of the problem is given by the value of the state variables in the first period.[13]

The value function for the mother at period t is given by:

$$V_t(S_t) = max_{h_t, i_t, \tau_t} \;\; u(l_t, c_t, A_t) + \beta E_t V_{t+1}(S_{t+1}) \tag{5}$$

$$s.t. \;\; c_t = w_t h_t + I_t - pi_t$$

$$TT = l_t + h_t + \tau_t$$

$$lnA_{t+1} = \delta_{1t} ln\tau_t + \delta_{2t} lni_t + \delta_{3t} lnI_t + \delta_{4t} lnA_t$$

where $\beta \in [0, 1]$ and $S_t = \{A_t, w_t, I_t\}$ represents the vector of state variables. The timing of the model implies that after childbirth and during the first 12 months of child's life the mother observes the initial level of child's ability and the level of income

[12]The model implies that the additional labor income the mother gets from her labor supply is spent in external child care.

[13]The structure of the initial condition for child's ability as well as the draws from which the initial values of w_t and I_t are taken will be defined in subsection 4.1.

in the household and receives a wage offer; then she makes her decisions. Similarly, in the following periods, the mother chooses h_t, i_t and τ_t after having observed the corresponding level of A_t and I_t and after having received the wage offer from the labor market.

It should be noted that the maximization problem of the mother can be solved analytically only if the wage offer is exogenous with respect to the mother's past and current labor supply choices. This implies that the offer the mother receives in period t is not affected by her working decisions in $t-1$ and that it does not reflect any depreciation in mother's productivity due to the absence from the labor market after childbirth. This assumption can be quite restrictive in the framework of mothers' labor supply so that, when possible, existing discrete-choice models in this topic (for instance, Bernal (2008)) rule it out. However, in the present model, if the wage process were defined so to depend on previous labor supply choices, the model would become intractable and could not be estimated using continuous choice variables and closed form solutions. Notice that the use of continuous choice variables is necessary to allow for three choices and to take into account the additional choice between leisure and time with the child. Moreover, allowing the wage offer to depend on previous labor supply implies to put an extra-premium on wages due to the participation in the labor market in the previous period. With continuous choice variable, this implies that the mother will always choose $h_t > 0$.[14] For the same reasons, the wage offer cannot be defined as to depend on the current choice of labor supply, neither to vary according to whether the mother works part-time or full-time. Hence, the estimation of a model allowing for three endogenous choices defined as continuous variables comes at a cost that is given by the impossibility of taking into account the endogeneity of wage. This may have, of course, some implications on the estimated parameters. In fact, since the definition of the wage process does not take into account the potentially negative effect on wages of leaving the labor market after childbirth, it is very likely to overestimate the proportion of mothers working and their labor supply on the extensive margin; this also implies an overestimation of the amount of external child care used and an underestimation of maternal time with the child.

3.2. ***Terminal period value function and solution of the model.*** The problem defined by equation (5) can be re-written as:

$$Max_{\{h_t,i_t,\tau_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(l_t, c_t, A_t) \tag{6}$$

$$s.t.\ S_t = \{A_t, w_t\} \in \Omega_t$$

where Ω_t represents the state space in each period t. The solutions to this maximization problem involves three sequences of values.

The mother makes work, child care and time allocation decisions (that are relevant for the child development process described by equation (4)) in the first T years of the

[14] In a discrete-choice framework this does not happen because mother's labor supply is discretized in such a way that $h_t = 0$ whenever h_t is lower than a certain threshold.

child's life. After period T, both the mother's optimization problem and the child's ability production function change: the mother may continue to optimally choose labor supply and consumption, but she will not longer consider maternal and external child care choices.

The terminal level of child's cognitive ability is A_{T+1}, i.e., the level of ability reached in $T+1$, that will not be affected by subsequent mother's decisions. Thus, $A_t = A_{T+1}$ for any period $t = T+1, T+2, \ldots, \infty$. This level of ability may be interpreted as the starting point for future child's development during adolescence, from $T+1$ on.

The period $T+1$ maximization problem for an infinitely-lived household may be written as:

$$V_{T+1} = \tilde{V}_{T+1} + \sum_{\kappa=0}^{+\infty} \beta^{\kappa} \alpha_3 lnA_{T+1} \tag{7}$$

where

$$\tilde{V}_{T+1} = max_{h_{T+1}}\ \alpha_1 lnl_{T+1} + \alpha_2 lnc_{T+1} + \beta E_{T+1} \tilde{V}_{T+2}(l_{T+2}, c_{T+2})$$

and $\sum_{\kappa=0}^{+\infty} \beta^{\kappa} = \rho$ represents the value given by the mother to child's ability in the last developmental period.[15] Equation (7) represents the terminal period value function[16] and implies that the mother's maximization problem after period T does not depend on t and on the choices made in the previous period. Starting from period $T+1$, the mother decides only how much to work and, in each period, this choice affects only her current utility, without affecting the utility and decision-making process in the following periods.

The model is solved by backward induction and yields closed-form solutions for all the choice variables. The solution of the model involves the computation of the value function starting from the terminal period and the corresponding optimal solutions in each period. Following a two-stage process, I first derive the optimal solutions for external child care (i_t) and maternal time (τ_t), conditional on h_t, and then I compute the solutions for the mother's labor supply h_t. Analytical derivations of the results are in appendix A.

The demands for child care and time with the child, conditional on mother's labor supply, in each period, are given by:

$$i_t^c = \frac{\beta \delta_{2t} D_{t+1}}{p(\alpha_2 + \beta \delta_{2t} D_{t+1})}(w_t h_t + I_t) \tag{8}$$

$$\tau_t^c = \frac{\beta \delta_{1t} D_{t+1}}{(\alpha_1 + \beta \delta_{1t} D_{t+1})}(TT - h_t) \tag{9}$$

where $D_{t+1} = \frac{\partial V_{t+1}}{\partial lnA_{t+1}}$ represents the marginal utility the mother gets from child's future cognitive ability, in each period. The sequence of marginal utilities from period

[15] In the estimation, the discount factor is set at $\beta = 0.95$. In order to increase the flexibility of the model and to allow the discount factor of the mother to differ in the last period of investments with respect to the previous ones, the parameter ρ is estimated.

[16] The terminal period value function is similar to the one specified in Del Boca et al. (2013).

$T+1$ to period 1 is given by:[17]

$$\begin{aligned} D_{T+1} &= \rho\alpha_3 \\ D_T &= \alpha_3 + \beta\delta_{4T}D_{T+1} \\ D_{T-1} &= \alpha_3 + \beta\delta_{4T-1}D_T \\ &\vdots \\ D_t &= \alpha_3 + \beta\delta_{4t}D_{t+1} \\ &\vdots \\ D_2 &= \alpha_3 + \beta\delta_{42}D_3 \\ D_1 &= \alpha_3 + \beta\delta_{41}D_2 \end{aligned} \tag{10}$$

An implication of the Cobb-Douglas specification used in the mother's utility function and in the child cognitive ability production function is that any input should be strictly positive.[18] However, I do allow the possibility of corner solutions for the mother's labor supply decisions.

The mother's latent labor supply, conditional on i_t^c and τ_t^c, is given by:

$$h_t^c = \frac{\alpha_2(TT - \tau_t^c)}{\alpha_1 + \alpha_2} - \frac{\alpha_1(I_t - pi_t^c)}{w_t(\alpha_1 + \alpha_2)} \tag{11}$$

Substituting (8) and (9) in equation (11), the latent labor supply becomes:

$$h_t^* = \frac{TT(\alpha_2 + \beta\delta_{2t}D_{t+1})}{(\alpha_1 + \beta\delta_{1t}D_{t+1} + \alpha_2 + \beta\delta_{2t}D_{t+1})} - \frac{I_t(\alpha_1 + \beta\delta_{1t}D_{t+1})}{w_t(\alpha_1 + \beta\delta_{1t}D_{t+1} + \alpha_2 + \beta\delta_{2t}D_{t+1})} \tag{12}$$

The actual labor supply in each period is determined according to the following rule:

$$h_t = \begin{cases} h_t^* & \text{if} \quad h_t^* > 0 \\ 0 & \text{if} \quad h_t^* \leq 0 \end{cases}$$

According to equation (12), the mother's latent labor supply is negative or zero only if household income is strictly positive and sufficiently high. In general, there is always a negative income effect; instead, for an increase in wage a positive substitution effect prevails. The reservation wage of the mother, i.e. the wage offer for which the mother is indifferent between working and not working, is given by the following expression:[19]

$$w_t^* = \frac{I_t}{TT}\frac{(\alpha_1 + \beta\delta_{1t}D_{t+1})}{(\alpha_2 + \beta\delta_{2t}D_{t+1})} \tag{13}$$

Notice that the reservation wage of the mother is a function of both the preference parameters in the utility function and the productivity parameters in the child cognitive ability production function. The mother's reservation wage is higher if the mother cares

[17]The same expressions can be derived computing $D_{t+1} = \frac{\partial V_{t+1}}{\partial A_{t+1}}$ instead of $D_{t+1} = \frac{\partial V_{t+1}}{\partial ln A_{t+1}}$ (See appendix A, footnote 36). Notice that the marginal utility in $T+1$ is discounted for all the subsequent periods in which child's ability does not depend on mother's investments decisions.

[18]Concerning the child cognitive ability production function, if any factor is set at zero, the child ability is zero in all subsequent periods (since if $A_{t-1} = 0$, then for any t, $A_t = 0$) and the mother's utility will approach $-\infty$ as $A \to 0$, even if $\alpha_3 > 0$ (Del Boca et al. 2013). This means that the model always predicts a positive amount of external child care, regardless of mother's working status or household income.

[19]This expression is derived making $h_t^* = 0$ and solving for w_t.

more about leisure or if the elasticity of child development with respect to maternal time increases; instead, the reservation wage decreases if the mother cares more about consumption or if the productivity of external child care increases. Notice that mother's decision to enter the labor market also depends on the productivity of external child care, since if it increases with respect to maternal time, the mother may be more willing to substitute her time with the external child care provider's time.

Substituting (12) into (8) and (9) yields the unconditional demands for child care and time with the child:

$$i_t^* = \frac{\beta\delta_{2t}D_{t+1}}{p(\alpha_2+\beta\delta_{2t}D_{t+1})} I_t \left[1 - \frac{(\alpha_1+\beta\delta_{1t}D_{t+1})}{\alpha_1+\beta\delta_{1t}D_{t+1}+\alpha_2+\beta\delta_{2t}D_{t+1}}\right] + \\ + \left(\frac{w_t}{p}\right) \frac{TT\beta\delta_{2t}D_{t+1}}{\alpha_1+\beta\delta_{1t}D_{t+1}+\alpha_2+\beta\delta_{2t}D_{t+1}} \tag{14}$$

$$\tau_t^* = \left[TT - \frac{TT(\alpha_2+\beta\delta_{2t}D_{t+1})}{\alpha_1+\beta\delta_{1t}D_{t+1}+\alpha_2+\beta\delta_{2t}D_{t+1}} + \left(\frac{I_t}{w_t}\right) \frac{(\alpha_1+\beta\delta_{1t}D_{t+1})}{(\alpha_1+\beta\delta_{1t}D_{t+1}+\alpha_2+\beta\delta_{2t}D_{t+1})}\right] \times \\ \times \frac{\beta\delta_{1t}D_{t+1}}{(\alpha_1+\beta\delta_{1t}D_{t+1})} \tag{15}$$

Notice from equation (14) that demand for child care can be driven by necessity of custodial care, i.e., if the mother is working and needs someone looking after the child, or by valuing the educational role of the service. In fact, non-working mothers (for which $h_t = 0$) can demand of it if they value child's ability and they think child care can represent an input for child's development, as long as household income is strictly positive. An increase in household income determines an increase in both the demand of external child care and maternal time with the child that are defined as inputs for the production of child's ability. Conversely, if mother's wage increases, this shifts upward the demand for external child care (both because of the generated income effect and because of the increase in mother's labor supply), while it decreases maternal time with the child, which represents the opportunity cost of maternal time in the labor market. All these comparative statics will be relevant in the analysis when simulating the effects of policies, whose results are presented in section 7.2.

4. Econometric strategy

Structural estimation involves assumptions on how observed and unobserved heterogeneity enters the model described in the previous section. Section 4.1 presents the empirical specification used to take the model to the data, taking into account the available information at my disposal. Subsection 4.2 describes the econometric method used to estimate the model parameters. Further details on the empirical analysis performed to estimate the model are in appendix B.

4.1. ***Empirical specification.*** Unobserved and observed heterogeneity enters any stage of the decision-making process of the mother described in the previous section.

Consider first the utility function, where the parameters represent the tastes of the mother for leisure, consumption and child's ability. I allow observed and unobserved heterogeneity in preferences, defining these parameters as functions of some observed and unobserved characteristics. Specifically,

$$\alpha_1 = f_1(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$$

$$\alpha_2 = f_2(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$$

$$\alpha_3 = f_3(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$$

where $\gamma_1 = 0$, $\Gamma_2 = (\gamma_{2\ MotherEdu}, \gamma_{2\ MotherRace})$ and $\Gamma_3 = (\gamma_{3\ MotherEdu}, \gamma_{3\ MotherRace})$ are vector of parameters representing the contribution of each observable characteristic to the corresponding preference parameter. The functional forms for f_1, f_2, f_3 are specified in appendix B.1. μ_0 represents mother's skills, that are assumed to be distributed with a multinomial density and to take on two values representing two types of mothers: μ_{0high} represents the high-skilled type, while μ_{0low} is the low-skilled type. The values μ_{0high}, μ_{0low} and the probability that the mother belongs to each type (π_{mh} and $\pi_{ml} = 1 - \pi_{mh}$) are parameters to be estimated.[20] This specification allows the parameters in the mother's utility function to vary across subgroups in the sample[21]. Notice that observationally equivalent mothers can still have different preferences according to their skills.

As stated in section 3, in each period, the mother receives a wage offer and decides whether to enter in the labor market comparing the value of this offer with her reservation wage. The offer the mother receives is described by the following wage equation:

$$ln(w_t) = \mu_t + \epsilon_t \tag{16}$$

where

$$\epsilon_t \overset{\text{iid}}{\sim} N(0, \sigma_\epsilon^2)$$

is assumed to be uncorrelated over time and represents a transitory shock on wage that the mother can observe. The term μ_t is the mean of the log wage draws of the mother at time t and it is defined as follows:

$$\mu_t = \mu_0 + \mu_1 educ + \mu_2 age_t + \mu_3 age_t^2 + \mu_4 race \tag{17}$$

where μ_0 represents mother's skills, as already defined. Equation (17) states that the offer the mother receives from the market depends on her skills, her education and experience (captured by the age component and its square), but also on her race.

As for the wage process, also the income process is exogenous with respect to the mother's inputs decisions in each period. The evolution of the household income reflects the following structure:

$$I_t \overset{\text{iid}}{\sim} N(\mu_{inc}, \sigma_{inc}^2) \tag{18}$$

where μ_{inc}, σ_{inc} are parameters to be estimated.

[20] See appendix B.1 for the specification of the parameters in the utility function and for the mother's type proportions.

[21] Each group is defined by the combination of mother's years of education and race. Mother's years of education range from 2 to 17, while race is a dummy variable equal to 1 if the mother is white.

Concerning the child's cognitive ability production function, as stated in section 3.1, the parameters can vary across child's age. In order to respect the parameterization implied by the Cobb-Douglas functional form, the coefficients in equation (4) must be strictly positive; thus, they are defined as follows:

$$\delta_i = exp(\xi_i t) \tag{19}$$

where $i = 1, 2, 3, 4$ and t represents the age of the child.[22]

In order to estimate the model and to take into account the dynamic optimization problem faced by the mother, one needs to know the starting level of ability, i.e., the child's cognitive ability the mother observes in the first period before making her investments decisions. The initial ability endowment is assumed to be specified as follows:

$$A_1 = exp(\psi_{ck} + \eta_1 MotherEdu + \eta_2 FatherEdu) \tag{20}$$

where ψ_{ck} represents child's skills, that are distributed with a multinomial density:

$$f(\psi_{ck}) = P_k$$

with $P_k \geq 0$ and $\sum_k P_k = 1$. ψ_c can take on two values ($k = h, l$), representing high and low skilled children. As for the mother's types, the values ψ_{ch} and ψ_{cl} should be estimated, together with their corresponding probabilities π_{ck} $k = h, l$. The inclusion of mother's and father's education allows to capture a non-zero correlation between these observable characteristics and child's skills. Moreover, as suggested by Bernal and Keane (2010), using as much observables as possible in the definition of (20) should reduce the sensitivity of the results to the distributional assumptions on the unobserved heterogeneity term.[23] Recalling the specification of the CAPF, defined in (4), it provides consistent estimates of the productivity parameters for each input if the following conditions hold: (i) A_t is a sufficient statistics for the inputs history received by the child in the previous periods; (ii) the child's initial endowment A_1 (that the mother observes but the researcher does not) is only reflected in the level of ability in the subsequent period and does not affect child's ability in the future periods (Todd and Wolpin 2003).

Finally, it should be described how the true child's cognitive ability is related to the measure of that given by the test scores. Existing studies using a structural approach (Bernal 2008; Bernal and Keane 2010) define the test score measure as a continuous variable and identify a linear relationship between this variable and the child's cognitive ability, including a disturbance term. This notation interprets the test scores as a proxy

[22]Allowing the parameters to vary across child's age partially compensates for the lack of substitutability implied by the Cobb-Douglas functional form used to define the CAPF. Moreover, it allows to capture the (potentially) decreasing productivities of the inputs considered in (4): when the child reaches primary school age, other (unobserved) school inputs can contribute to his own cognitive development and family investments have lower influence.

[23]Due to the structure of the available data, the identification of more parameters in the child's initial endowment is hampered by the scarcity of test score observations for each child. In fact, I can observe at most 2 test score measures for each child and the test score measure is available only for children aged more than 4. In appendix E.4 I report the results of a sensitivity analysis where the child's initial endowment depends also on birth weight but all children are assumed to have the same level of skills.

for the true child's ability, but it does not take into account the fact that these measures represent just the number of questions answered correctly by the child. Following the approach suggested by Del Boca et al. (2013) and based on classical test theory (Novick 1966), I define the probability that the child answers correctly to each item as a function of the true child's ability:

$$\pi_{score} = \frac{exp(A_t + v_t)}{1 + exp(A_t + v_t)} \tag{21}$$

where $v_t \overset{\text{iid}}{\sim} N(0, \sigma_v^2)$ represents measurement error capturing the fact that test scores depict true child's ability with a noise. The structure of (21) ensures this represents a value between zero and one. The test score measure is then defined as follows:

$$S_t = \pi_{score} * J_t \tag{22}$$

where J_t is the maximum number of items answered correctly at each child's age.[24]

Summing up, the empirical specification of the model allows the mother's preference parameters to depend on mother's observable characteristics and unobserved ability, while mothers with higher skills receive, on average, higher wage offer, are more likely to work and to use more external child care. Moreover, similar mothers can receive different wage offers over time because of the transitory shock on wage that the mother can observe but the researcher does not. Finally, the definition of the child's initial endowment as a function of parents' education captures a non-zero correlation between parents' cognitive abilities and child's skills.

4.2. ***Estimation method.*** The model parameters are estimated using a Simulated Minimum Distance (SMD) estimator that minimizes the distance between a large number of data statistics and their model counterparts. The statistics used to construct the moment functions are summarized in table 1.

The simulation of the data generating process (DGP) implied by the model accounts for the selection of mother's participation in the labor market and the endogeneity issues arising for all the other choice variables. In other words, the DGP models the selection mechanisms underlying the work, external child care and time decisions. This point has practical consequences, in that it allows to recover non-randomly missing information, as mother's wage. In fact, when the information on mother's employment status is available in the data and the mother is not working, mother's wage is missing in an endogenous way. The simulation of the wage offer that each mother receives in every period allows to describe the participation decision as a function of the preference and productivity parameters.

Further, simulation is needed because the statistics and the moment functions recovered from the model are not in a tractable form. The minimum distance estimator involves the minimization of the distance between statistics provided by the data and statistics that are functions of structural parameters. For instance, define m as the

[24]The score measure used in the empirical analysis is the Letter Word test. To define the thresholds J_t I use the overall PSID-CDS data (3243 observations) and I identify the maximum number of items answered correctly at each age: in the age range 4-5 $J = 30$, in the age-range 6-8 $J = 50$ and finally, for $t = 9, 10, 11, 12, 13$ $J = 57$.

Table 1
Statistics of actual and simulated data used for the estimation of the model.

Inputs and outcome conditional on child's age	$N*T$
mean and std deviation of mother's hours of work	3518
mean and std deviation of child care hours	5598
mean and std deviation of maternal time with the child	624
proportion of mothers working	3518
average test score	615
Inputs statistics	$N*T$
mean and std deviation of mother's wage	2459
mean, std deviation and median of household income	3046
corr mother's wage and mother's hours of work	2459
corr mother's wage and household income	2148
corr mother's hours of work and household income	3044
corr mother's hours of work and time with the child	204
corr mother's hours of work and child care hours	3510
corr household income and time with the child	173
corr household income and child care hours	3039
Inputs and outcome correlation across time	$N*T$
corr maternal time with the child in 1997 and score in 2002	311
corr maternal time with the child in 2002 and score in 2007	111
corr child care hours in t and score in $t+1$	609
corr mother's hours of work in t and score in $t+1$	199
corr household income in t and score in $t+1$	168
Inputs conditional on parents' characteristics	$N*T$
mean mother's wage by mother's education, age and race	2459
mean mother's hours of work by mother's education, age and race	3518
mean maternal time with the child by mother's education, age and race	624
mean child care hours by mother's education, age and race	5598
Outcome conditional on parents' characteristics	$N*T$
average test score by parents' education	615
Outcomes transition probabilities (for children with 2 scores measures)	$N*T$
prop of children with score in range p_{97} in 1997 and p_{02} in 2002	1066
prop of children with score in range p_{02} in 2002 and p_{07} in 2007	1287

NOTE. These statistics are computed using PSID-CDS data on children aged 0-12 in 1997, with at least one test score measure and without siblings, and simulated data according to the model defined in section 3 and 4.1. Maternal time with the child is measured in 1997 and 2002; child's scores are measured in 1997, 2002 and 2007; from 1997 on, mother's hours of work, mother's wage and household income are measured every two years and these variables refer to the year before the survey (see section 5 and appendix C for a description of the data). Household income includes both father labor income and household non labor income. Child's age t ranges from 1 to 13. Mother's and father's education are classified as "college" (more than 12 years of education) and "high-school" (12 years of education); mother's race can be white or not white; mother's age is divided in two categories: more than 40 years old and younger than 30. Ranges p_y, with $y = 1997, 2002, 2007$ are defined according to the following ranges of the score distribution: $1st - 25th$ perc, $25th - 50th$ perc, $50th - 75th$ perc, $75th - 95th$ perc, higher than $95th$ perc. $N*T$ indicates the number of observations used for each statistic; for the all sample $N*T = 434*13 = 5642$.

data points and statistics and $M(\theta)$ as the functions of the parameters to be estimated. If the functions are not in an easily computable form, as in this case, they can be substituted with a simulator estimator $\hat{M}(\theta)$ for $M(\theta)$ (McFadden 1989; Pakes and Pollard 1989).

The simulation of the data is obtained by taking $N*R$[25] random draws from the initial distribution implied by the model, i.e., the child's and mother's skills distributions, and, for each period, from the wage and income distributions and from the distribution of

[25] $N = 434$ and $R = 5$. While R does not affect the consistency of the estimator, an higher number of simulation draws, with N fixed, can decrease the simulation noise and the variance, improving efficiency. However, I decide not to use more simulation draws, because the estimation is already time consuming. Using a laptop computer with Intel i7/1.5 GHz processor and Matlab Version 7.13, the estimating time is about 4 hours.

the error in the test score measure. The time invariant preference parameters are assigned to each mother, according to her observable characteristics and skills, while the productivity parameters are updated in each period. After having drawn the child's level of ability, the wage offer and the level of income in the first period, the optimal choices of the mother are obtained exploiting the optimal solutions derived in section 3.2. This process is repeated for every period, up to the final one T. The simulated data are used to compute the same statistics defined in table 1. Both actual and simulated statistics are used to construct the objective function to be minimized.

The Simulated Minimum Distance (SMD) estimator is then:

$$\hat{\theta} = arg\ min\ \hat{g}(\theta)' W \hat{g}(\theta) \tag{23}$$

where

$$\hat{g}(\theta) = \hat{m} - \hat{M}(\theta) \tag{24}$$

$\hat{m}$ is the vector of statistics defined from the actual data, while $\hat{M}(\theta)$ is the vector of simulated statistics according to the model. Given S number of moments, the weighting matrix is defined as:

$$W = \begin{pmatrix} \hat{V}[\hat{m}_1]^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{V}[\hat{m}_S]^{-1} \end{pmatrix}$$

where $\hat{V}[\hat{m}]$ is estimated with non-parametric bootstrap. Appendix B.2 provide further details on the estimation.

The SMD estimator consistently estimate the model parameters if the following conditions hold: (i) the estimated moments from the data are consistent estimates of the population moments; (ii) the model is identified, a necessary condition for which is that the number of moments is higher than the number of parameters to be estimated ($S \geq K$);[26] (iii) the model is correct: $m_0 = M(\theta_0)$, i.e., the population moments correspond to the simulated moments at the true parameters vector.

Identification of the model parameters requires a unique solution for the minimization of the objective function defined by (23). Due to the structure of available data, the identification of all parameters could be achieved, with the only exception of the ones in the definition of the starting level of ability. This is due to the scarcity of observations of tests scores measures in the data and on the fact that I can observe test scores only starting from age 4.[27] In practice, identification depends on the uniqueness of the minimum and on the curvature around it. To test for this, I estimate the model using different starting values and results do not differ from the ones presented in the following section. Moreover, I check that the objective function changes moving the values of the parameters. I find the value of the objective function to vary around the

[26] To estimate 26 parameters, I use 103 moments conditions.

[27] See the following section for further details on data availability. Notice that this issue arises for any structural analysis evaluating the effects of investments during early years and using test scores as outcomes, as long as it needs a specification for child's ability at birth that is not observed, such as in Bernal (2008) or Bernal and Keane (2010).

estimated parameters, with the only exception of the parameters in the initial level of ability, for which identification is more tenuous due to the reasons explained before.

5. Data

The model is estimated using data from the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS).

The PSID is a longitudinal study that begun in 1968 with a nationally representative sample of over 18,000 individuals living in 5,000 families in the United States. Starting from 1968, information about each family member is collected, but much greater detail is obtained about the head and the spouse. From 1997, the Child Development Supplement (CDS) gathers information on children aged 0-12 in PSID families through extensive interviews with their primary caregiver. The CDS has been replicated in 2002 and 2007 for children in this cohort who remain under 18.

For this analysis, I exploit the child cognitive ability measures and non-parental child care data provided in the Primary Caregiver Interview of the CDS, together with the time use details given in the Time Diary (TD) component of the CDS. The main PSID surveys are exploited to recover information on mother's work and household income.

The CDS supplement provides several measures of child cognitive skills, based on the Woodcock Johnson Achievement Test Revised (WJ-R) (Woodcock and Johnson 1989). The outcome measure considered in this study is the Letter Word (LW) test, which is applied to all children older than 4 and proves child's learning and reading skills (Hoffert et al. 1997). The raw LW score represents the sum of correct answers out of 57 items, ranging from 0 to 57. This measure is available in 1997, 2002 and 2007.

The CDS I (1997 wave) asks to the primary caregiver information on all child care arrangements used for the child since childbirth; a set of following-up questions is asked to the primary caregiver in the 2002 wave of the same supplement. Using both waves, I can recover the complete child care history for the children interviewed in 1997. The variable of interests is the number of hours the child uses non-parental child care at each age. This variable refers to any type of child care arrangement, either formal or informal, provided by people different from parents.[28]

In 1997 and 2002, the Child Development Supplement includes another instrument to assess the time use of children. The Time Diary (TD) is a unique feature of the CDS and consists in a chronological report filled by the child or by the child's primary caregiver about the child's activities over a specified 24-hour period.[29] Each participating child completed two time diaries: a weekday (Monday-Friday) and a weekend day (Sunday or Saturday). The TD additionally collects information on the social context of the

[28]The CDS questionnaire allows the primary caregiver to indicate more than one arrangement used at each child's age. If the primary caregiver used simultaneously more than one arrangement in a period, I define the child care variables exploiting the information on the arrangement used more hours per week. Notice that, in this case, the corresponding number of child care hours represents a lower bound of the true child care use.

[29]The primary caregiver completed the time diary for the very young children (e.g., younger than 3), while older children and adolescents were expected to complete the time diaries themselves (ISR 2010a,b).

activity by specifying with whom the child was doing the activity and who else was present but not engaged. The variable weekly time with the mother is constructed by multiplying the daily hours the child spends with the mother by 5 for the weekday and by 2 for the weekend day, and summing up the total hours in a week.[30]

I take information on mothers and fathers linking the CDS data to the main PSID surveys. Since children in 1997 have different ages, ranging from 0 to 13 years old, in order to identify the necessary information for all children in any period defined by the model, CDS data should be matched with family information from PSID surveys in the years 1985-2007.[31] The family information I gather includes each parent's hours of work, wage and non labor income in each period.[32]

All relevant variables are constructed for each child's age, defining age 1 as the first 12 months of child's life, age 2 as the next 12 months of the child's life, and so on. For the estimation of the model I consider all children without siblings interviewed in CDS I, living in intact households (where both mother and father are present), without missing data on personal and parents' demographic characteristics and with at least one test score measure. The final sample is composed by 434 observations.[33]

Before moving to the descriptive statistics, it should be stressed the importance of the sample selection and consider what biases might be introduced into the analysis by focusing on the subsample of children in intact households without siblings. The estimates derived from this sample are likely to be of general interest and can still contribute to the literature of child care and child development, because they are structurally obtained from a model based on economic theory. However, the sample selection bias may come in different ways and may also have an influence on the effects of the simulated policies. First of all, the sample selection implies that all mothers' investments

[30]More precisely, the TD distinguishes between contexts where the person with the child is directly involved in the activity ("active time") and others where the person is just around and not involved in the activity ("passive time"). The following time categories can be derived: (1) the child is with the mother, being the mother either involved in the activity or just around; (2) the child is with the mother, who is directly involved in the activity, but the father is around; (3) the child is with the father only; (4) the child is with the father and the mother is around; (5) the child is neither with the mother nor with the father. The analysis has been performed defining the variable weekly time with the mother using only category (1), so that all remaining time spells indicate that the child is not receiving investments from the mother. In order to see whether the results are sensible to this specification, I re-estimate the model using different definitions of maternal time. Results are reported in appendix E.2.

[31]For instance, to identify household information for all relevant periods for a child born in 1996 (1 year old in 1997) I need to use PSID surveys from 1997 to 2007; instead, if a child is born in 1986 (aged 11 years in 1997) I need to use PSID surveys from 1987 to 1999. Basically, all PSID surveys in the period 1985-2007 have been exploited. See appendix C, tables C.1 and C.2.

[32]Between 1985 and 1997 PSID interviews were conducted annually and, since then, interviews have been biennial. Note that all the variables that I use from the main PSID surveys concerning labor and non labor income of the household members refer to the year before the survey. All monetary variables are deflated into 1997 US$ using the Consumer Price Index (CPI) History for the U.S. See appendix C for further description of the data sources used for the analysis.

[33]Out of the 3,563 children interviewed in 1997, 314 do not have information on their parents, 2,069 have siblings and 602 live in households where one (both) parent(s) is (are) not present. Moreover, 52 children have not information on parents' age, education and race and 85 have not test score measures in the period 1997-2007.

in child's ability are unrelated with both the decision to marry or to cohabit and fertility, even though there has been evidence of their correlation (Francesconi 2002; Van Der Klaauw 1996). For instance, if mothers in intact households have more marriage-oriented attitudes[34] and unobservables determining their marriage/cohabitation decisions also influence their time allocation and fertility, they may be more likely to stay at home instead of working and to invest more on their unique child. This may lead to overestimate the proportion of mothers not working or to overestimate mothers' preferences for child's ability and their investments on the child. However, women in long-term relationship may also be more desirable in the labor market (since they may have good communication, conflict management skills, etc); if this is the case, this sample would be disproportionately represented by high productive mothers and may lead to overestimate the decision to work. Table C.4 in appendix C compares the characteristics of the subsample used for the analysis ($N = 434$) with the ones of the entire PSID-CDS sample ($N = 3243$) and indicates that mothers in this subsample are older, more educated, they work more, they use more external child care and they spend less time with their child. This may suggest that mothers in this sample can be more productive either in the labor market or at home with their child, leading to overestimate their labor supply or the productivity of their investments in child's ability. However, table C.4 also shows that mothers' wage before childbirth does not differ significantly across samples: as long as mothers' wage before childbirth can be considered a measure of their (initial) productivity at work, this seems to suggest that the selection is not oversampling mothers that are more productive in the labor market. Indeed, if they are more productive with their child, the sample selection may still end up overestimating the productivity of maternal time on child development. Eventually, if the skills of these 'more productive' mothers are transmitted to their child genetically, the subsample used for the analysis can be characterized by high-skilled children, leading to an overestimation of the level of child's ability. Hence, the assumptions implied by the sample selection should be kept in mind when presenting the estimated parameters and the results from the policy simulations.

5.1. ***Descriptive Statistics.*** This section provides some descriptive statistics on the sample used for the estimation.

Table 2 shows the average values of all the variables for the period considered in the model. In the sample, the average raw score is around 35 out of 57. Figure 3 shows the distribution of the average test score measure by child's age, while appendix D provides further descriptive statistics on the cognitive outcome measure. Mothers work, on average, 27 hours per week and use non-parental child care for almost 15 hours; moreover, they spend with their child, on average, 3 hours per day. Mother's wages are significantly lower than their male counterparts (on average 14.25 US$ versus 19.51 US$), and mothers work less than fathers. Household non labor income represents, on average, around 13 US$ per week.

[34] 82.7 percent of children in the sample lives in households where the mother and the father are married; the remaining part of the sample lives in households where parents cohabit.

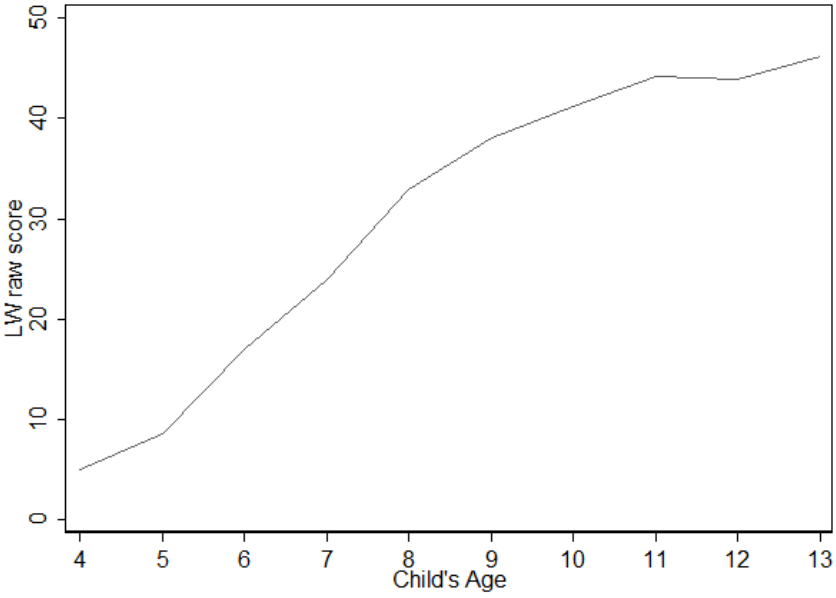
Table 2
Descriptive statistics on all variables for the entire period.

	Mean	SD	Min	Max
Child's LW raw score	35.10	(14.47)	0	57
Mother's hours of work	27.12	(17.55)	0	100
Proportion of working mothers	0.80	0.39	0	1
Non-parental child care hours	14.68	(18.32)	0	70
Mother's time with child	21.06	(17.01)	0.17	95.75
Child's gender: male	0.51	(0.50)	0	1
Mother's wage	14.25	(10.19)	5.01	133.93
Mother's age at child's birth	28.17	(5.10)	16	43
Mother's education	13.25	(2.50)	2	17
Mother's race: white	0.61	(0.49)	0	1
Father's hours of work	45.22	(10.95)	0.06	109.85
Father's wage	19.51	(12.96)	5.01	143.40
Father's age at child's birth	30.28	(6.11)	17	67
Father's education	13.27	(2.49)	1	17
Household non labor income	12.79	(49.10)	0	924.45

NOTE. Monetary variables deflated into 1997 US$. Source: own elaboration from PSID-CDS data.

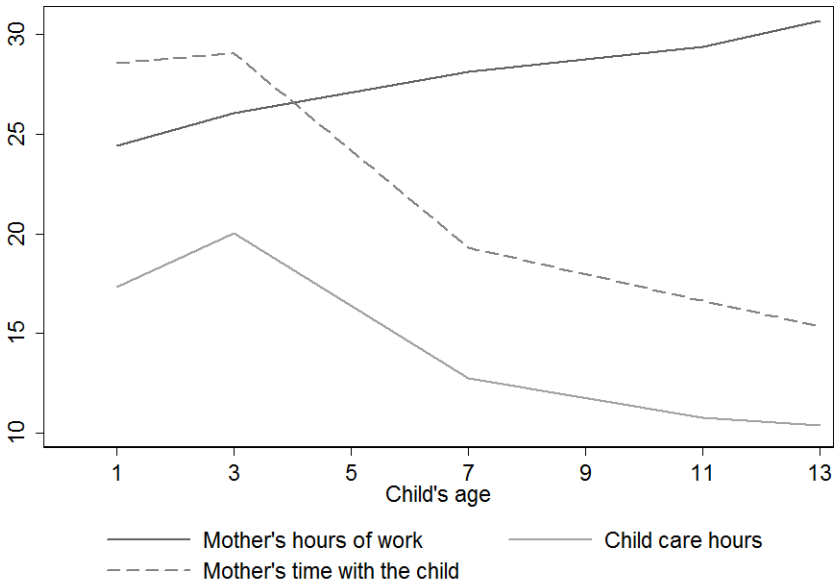
Table 3 provides some descriptive statistics on mother's work, child care and maternal time, by child's age. The temporal pattern of these variables is also reported in figure 4. There are not significant differences in mother's participation to the labor market across child's age. The number of hours worked by the mother ranges from 24 when the child is very young, to 30 when the child reaches 11 years of age; conversely, the average number of hours the child is cared for by someone other than his parents decreases as the child ages, ranging from 17 hours per week in the first years of life to 11 hours per week when he is 11 years old. Notice that the daily amount of time the mother spends at work when the child is younger than 6 almost corresponds to the time the child is cared for by someone else (4.8 hours per day vs 3.5 hours per day). When the child starts going to school, he does spend out of home not only the time in external care but also a fixed amount of school time. If the child spends at school 6 hours per day, he stays out of home almost 8 hours, while the mother works, on average, 5.6 hours per day. This difference shows that the amount of leisure time of the mother significantly increases when the child reaches school age. The average number of hours the child spends with the mother decreases as the child grows up: the mother spends with the child almost 4 to 5 hours per day when the child is younger than 5, while the time drops to 2 to 3 hours per day when the child reaches 6 years of age.

Figure 3
LW raw score by child's age.



NOTE. Source: own elaboration from PSID-CDS data.

Figure 4
Mother's hours of work, time with the child and non-parental child care time per week, by child's age.



NOTE. Source: own elaboration from PSID-CDS data.

Table 3
Descriptive statistics on maternal employment, non-parental child care and maternal time by child's age. Means and standard deviation in parentheses.

Child's Age	1-2	3-5	6-10	11-12
Mother's hours of work per week	24.40	26.03	28.10	29.34
	(17.67)	(17.55)	(17.21)	(17.29)
Proportion of working mothers	0.77	0.79	0.82	0.82
	(0.42)	(0.41)	(0.38)	(0.39)
Child care hours per week	17.39	20.02	12.76	10.76
	(19.11)	(19.28)	(17.49)	(16.49)
Mother's time with the child	28.55	29.05	19.31	16.64
	(18.06)	(20.27)	(14.80)	(14.21)

NOTE. Source: own elaboration from PSID-CDS data.

6. Results

This section presents the estimated parameters, while subsection 6.1 discusses the goodness of fit of the model.

Table 4 shows the estimates of the parameters in the mother's utility function. The γs parameters represent the contribution of each observable characteristic of the mother on mother's tastes for leisure, consumption and child's cognitive ability. Figures 5, 6 and 7 report the values taken by each preference parameter by subgroups and by mother's education. While the taste for leisure does not vary across mother's levels of education, more educated mothers care less about consumption. Low skilled mothers have higher preferences for leisure and consumption, while there are not differences in tastes induced by mother's race. Figure 7 shows the value of the preference parameter for child's ability. For any group, one more year of education implies an higher taste for child's ability, even though the marginal contribution of each year decreases as education increases. Moreover, more skilled mothers care more about their child's ability than the low skilled ones. These results seem in line with some recent findings on the effect of education on maternal time allocation, as provided by Guryan, Hurst, and Kearney (2008).

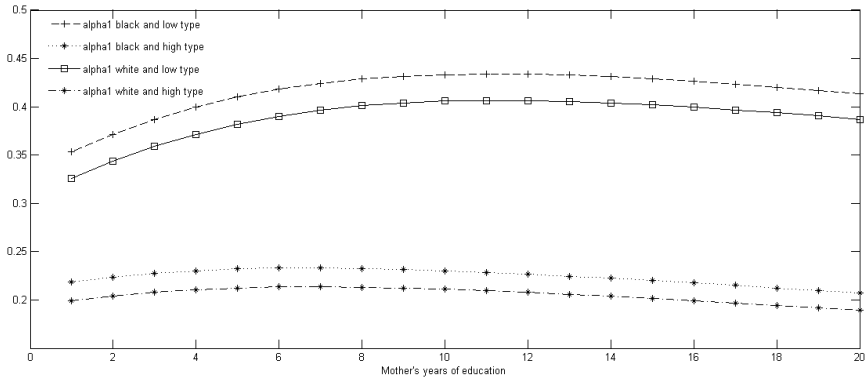
The parameter ρ indicates the value the mother poses on the child's level of ability reached in the last developmental period. The estimated value is roughly 34. To give an intuition to this number, consider the case of an infinitely lived household with a

Table 4
Estimated parameters for mother's utility function.

Mother's Utility Function Parameters		
$\gamma_{2\ MotherEducation}$	contribution of mother's education to α_2	-0.2094
		(0.0878)
$\gamma_{2\ MotherRace}$	contribution of mother's race to α_2	0.1369
		(0.1064)
$\gamma_{3\ MotherEducation}$	contribution of mother's education to α_3	0.0169
		(0.0425)
$\gamma_{3\ MotherRace}$	contribution of mother's race to α_3	0.1104
		(0.0514)
ρ	weight on future child's ability in the last period	34.1643
		(1.8554)
p	hourly price of child care	5.1502
		(0.6517)
Mother's Skills Distribution		
$\mu_{0\ high}$	skill level for high type	3.2011
		(0.8090)
$\mu_{0\ low}$	skill level for low type	1.2603
		(0.1833)
$\pi_{m\ high}$	proportion high skilled	0.3203
		(0.1006)
$\pi_{m\ low}$	proportion low skilled	0.6797
		(...)

NOTE. Standard errors are estimated with non-parametric bootstrap; standard errors for type proportions are computed using the delta method. See appendix B.3 for further details. Since type proportions should add to one, so that one of the type probabilities is obtained as a residual, I do not report standard errors in this case.

FIGURE 5
Preference parameters for leisure by subgroups and by level of education.

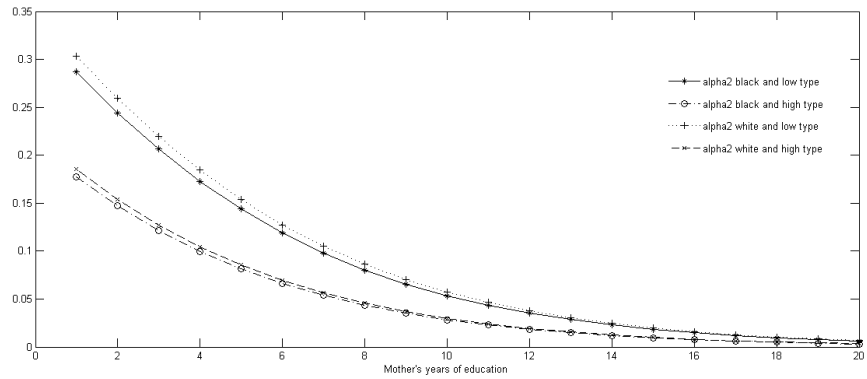


NOTE. This graph represents the estimated preference parameter for leisure by mother's years of education and for different subgroups, identified through mother's race and mother's skills level. The parameter is defined as $\alpha_1 = f_1(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$ where $\gamma_1 = 0$ and the estimated values for Γ_3, Γ_2 and μ_0 are shown in table 4. See appendix B.1 for further details.

discount factor in the last period that is equal to the factor in all previous periods. Since $\beta = 0.95$, the value of this parameter would be equal to $\rho = \sum_{k=0}^{\infty} \beta^k = \frac{1}{1-\beta} = 20$. Having found that the discount factor in the last period is higher than this value means that, in $T+1$, the mother poses additional weight on child's ability, because she may think it represents "an important initial condition for developmental processes that begin in the later teen years" (Del Boca et al. 2013).

The panel at the bottom of table 4 reports the parameters identifying the mother's skills distribution. The skills level of high type mothers is more than two times higher than the corresponding level for the low type. This implies a significant difference in the offer that mothers with different skills receive from the market and, as a consequence, in their employment decisions. The proportion of low skilled mothers in the sample is equal to 68 percent.

FIGURE 6
Preference parameters for consumption by subgroups and by level of education.



NOTE. This graph represents the estimated preference parameter for consumption by mother's years of education and for different subgroups, identified through mother's race and mother's skills level. The parameter is defined as $\alpha_2 = f_2(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$ where $\gamma_1 = 0$ and the estimated values for Γ_3, Γ_2 and μ_0 are shown in table 4. See appendix B.1 for further details.

Table 5
Estimated parameters for the wage and income processes.

Wage Equation Parameters		
μ_1	coefficient of mother's education	0.0788
		(0.0445)
μ_2	coefficient of mother's age	0.0018
		(0.0116)
μ_3	coefficient of mother's age squared	−0.0006
		(0.0008)
μ_4	coefficient of mother's race	−0.0156
		(0.0038)
σ_ϵ	standard deviation transitory shock	0.3031
		(0.0111)
Household Income Process		
μ_{inc}	mean	15.9606
		(0.3432)
σ_{inc}	standard deviation	12.8553
		(1.2765)

NOTE. Standard errors are estimated with non-parametric bootstrap. See appendix B.3 for further details.

Table 5 shows the results from the wage equation and the income process. All parameters in the wage equation have expected signs and reasonable magnitudes. The education effect on wages indicates that wage increases by 7.8 percent with each additional year of education. This effect is in line with the one found by Keane and Moffitt (1998) but slightly higher than the estimated effects in Del Boca et al. (2013) (4.8 percent), Bernal and Keane (2010) (4 percent) and Bernal (2008) (2 percent).

Table 6 presents the results of the parameters in the child's cognitive ability production function and the initial level of ability. The parameters shown in the first panel of this table represent the slope of each input productivity with respect to child's age. To simplify the presentation of the results, figures 8 and 9 show the time-varying elasticities as a function of child's age. Figure 8 reports the elasticities of child ability with respect to maternal time and non-parental child care time, while figure 9 reports the elasticities with respect to household income and the child's ability in the previous period. First of all, notice that the elasticity with respect to all inputs is higher during early years and decreases over time, as suggested by previous studies on human capital accumulation (Carneiro and Heckman 2003; Heckman 2008). According to figure 8, the elasticity of child's cognitive ability with respect to external child care is slightly higher than the one with respect to maternal time. The elasticity of child's ability with respect to maternal time ranges from 0.65 when the child is 1 year old to less than 0.05 when the child has 13 years of age. These estimates are higher than the values estimated by Del Boca et al. (2013) for the early years, but lower from age 5 on. For instance, Del Boca et al. (2013) find that the coefficient for maternal time when the child is 1 year old is equal to 0.25, while it is equal to 0.1 when the child is aged 8 years. At the same age, figure 8 reports a coefficient roughly equal to 0.05.

Table 6
Estimated parameters for the child's cognitive ability production function.

CAPF Parameters		
ξ_1	slope productivity of maternal time	−0.4388
		(0.1299)
ξ_2	slope productivity external child care	−0.3514
		(0.1169)
ξ_3	slope productivity income	−0.0566
		(0.0510)
ξ_4	slope productivity child's ability in previous period	−0.1026
		(0.1215)
σ_v	standard deviation measurement error in test score	15.2839
		(1.2061)
Child's Initial Ability Parameters		
$\psi_{0\ high}$	skill level for high type children	−52.3704
		(4.8011)
$\psi_{0\ low}$	skill level for low type children	−108.0380
		(13.7080)
$\pi_{c\ high}$	proportion high skilled children	0.0001
		(3.3533)
$\pi_{c\ low}$	proportion low skilled children	0.9999
		(...)
η_1	correlation child's endowment and mother's education	2.0618
		(1.2880)
η_2	correlation child's endowment and father's education	0.2255
		(1.1359)

NOTE. Standard errors are estimated with non-parametric bootstrap; standard errors for type proportions are computed using the delta method. See appendix B.3 for further details. Since type proportions should add to one, so that one of the type probabilities is obtained as a residual, I do not report standard errors in this case.

These estimates can be compared with existing studies evaluating the impacts of maternal employment and child care, using the information on mother's time out of work as a proxy for maternal time with the child. For instance, Bernal (2008) finds that one year of full-time work and external child care use reduces child's score by 1.8 percent; Bernal and Keane (2010, 2011) estimate a reduction in child's cognitive assessment due to one year of full-time work and external child care use of around 2 percent.

According to figure 8, if maternal time reduces by 10 percent when the child is 1 year old, because of mother's employment, the child's cognitive ability decreases by 6.5 percent. Assuming that the mother uses external child care for an equal amount of time, child's cognitive ability increases by 7.1 percent. This means that the reduction in child's ability due to the reduction in maternal time can be compensated by the use of external child care services, so that the overall effect of maternal employment turns out to be null or very small, but positive. This result is completely new in the literature and strongly confirms that using the time out of work as a proxy for maternal time with the child overestimates the real amount of time spent by the child with the mother, as well as its productivity. Section 7.1 further confirms this issue, presenting the results of a counterfactual exercise where the model has been estimated setting $\tau = TT - h$

and $\alpha_1 = 0$. In this scenario, the mother chooses zero leisure and maternal time with the child is proxied with the total time endowment net of working time.

Moreover, the specification of the model allows to identify the mechanisms with which maternal employment affects child development. In fact, despite having a negative effect due to the reduction in maternal time with the child, this impact can be compensated by child care facilities providing other inputs for child's cognitive development. Notice that maternal employment is not detrimental for child's development as long as external child care has at least the same productivity than maternal time.

Figure 8 shows that the coefficients of both external child care and maternal time strongly decrease over time. This pattern can be due to two main reasons. On one hand, it may depend on the fact that starting from compulsory school age, children start receiving other inputs that are unobserved in the data and not taken into account in the model. Hence, both maternal time and non-parental child care may play a weaker role. On the other hand, the steep fall in external child care productivity when the child starts going to kindergarten or primary school can be explained by the different purposes of external child care from the mother's point of view. In fact, the mother may choose a positive amount of child care if she works and needs someone looking after the child, but also if she thinks it can represent an input for subsequent child's development. The educational role of child care can be less important when the child starts going to school, because he is receiving other educational inputs from other institutions, so that from this age on the custodial role can be prevailing. As a consequence, child care productivity decreases even if the amount of time spent in external care remains constant.

Figure 9 shows the productivity of both household income and child's ability in the previous period. The result for household income seems in line with existing literature saying that economic conditions in early and middle childhood are more important for children's cognitive outcomes than those during adolescence (Duncan and Brooks-Gunn 1997; Duncan et al. 1998; Levy and Duncan 2012). It is also interesting to note that the productivity of the child's ability in the previous period is higher than the ones for maternal time and external child care. This can be explained by the fact that when the child ages he receives other inputs that are not included in the model (e.g., school) and that become more effective as he grows up.[35]

These results are robust to alternative specifications that are presented in appendix E. More precisely, I provide some sensitivity analysis testing the results on the following dimensions: (i) the definition of the terminal period; (ii) the definition of maternal time; (iii) the specification of the wage and income processes that does not allow any correlation between the two; (iv) the specification of the child's initial endowment.

6.1. ***Goodness of fit of the model.*** The reliability of the results and the credibility of the counterfactual exercise and policy simulations described below depend on how the model fits the data.

[35]Due to the assumption implied by the value-added functional form, if the level of child's ability in every period is a sufficient statistic for all the inputs received by the child in the previous period, the productivity of A_t should incorporate also the effects of inputs that I am not observing in the data.

Figure 10 shows the model fit for the child's score measure. Despite there being some differences between the actual and simulated data for the child's first years of life, the model predicts quite well the pattern of the score measure for subsequent child's ages.

Table 7 shows how the model performs in fitting the data concerning the wage and the income processes. Specifically, it shows the average and standard deviation of wage and income, observed in the actual and in the simulated data. The model behaves well for the prediction of average wage and income and there are not differences between the actual and simulated data concerning the standard deviation of income. Moreover, it reproduces the patterns in the data concerning the average wage by mother's education, while the average wage of black mothers is slightly overestimated.

Table 8 presents the average values for the choice variables. The model predicts an higher amount of working time and a slightly lower maternal time with the child with respect to the data. Finally, external child care time is narrowly underestimated. However, since these statistics are not used in the estimation of the model, the model seems to fit quite well the average values of mother's choices.

7. Counterfactual exercises

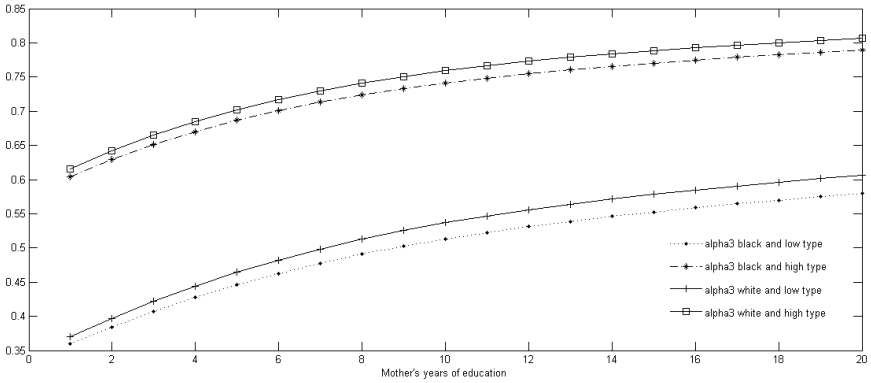
In this section, I use the estimated model to perform several counterfactual exercises. Subsection 7.1 presents the results from the estimation of the model where maternal time with the child is approximated from maternal working time, as it is usually done in the literature. Then, in subsection 7.2 I use the estimated model to simulate the effects of policies (i) increasing household income, by offering a 100 US$ lump-sum

Table 7
Goodness of fit for mother's wage and household income.

	Actual Data	**Simulated Data**
All sample		
Mean mother's wage	14.2510	14.7531
Sd mother's wage	10.1874	16.6083
Mean household income	17.6326	16.0272
Sd household income	12.8244	12.8255
Wage by mother's education		
Some college education	16.3756	16.6014
High School	12.0478	12.6266
Wage by mother's race		
White	15.0246	14.8084
Black	12.8967	14.6671

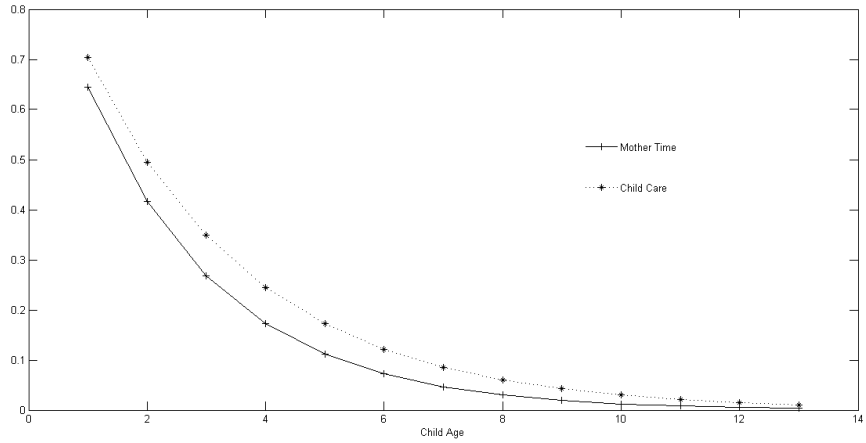
NOTE. Actual data represent PSID-CDS data on children aged 0-12 in 1997, with at least one test score measure and without siblings. See section 5 and appendix C for further details on the data. Simulated data represent the data obtained simulating the model described in section 3 and 4.1 and setting the parameters at the estimated values shown in tables 4, 5 and 6. Some college education stems for more than 12 years of education; high school education stems for 12 years of education.

Figure 7
Preference parameters for child's ability by subgroups and by level of education.



NOTE. This graph represents the estimated preference parameter for child's ability by mother's years of education and for different subgroups, identified through mother's race and mother's skills level. The parameter is defined as $\alpha_3 = f_3(MotherEdu, MotherRace, \gamma_1, \Gamma_2, \Gamma_3, \mu_0)$ where $\gamma_1 = 0$ and the estimated values for Γ_3, Γ_2 and μ_0 are shown in table 4. See appendix B.1 for further details.

Figure 8
Maternal time and non-parental child care productivity.



NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. These parameters are defined as

$$\delta_i = exp(\xi_i t)$$

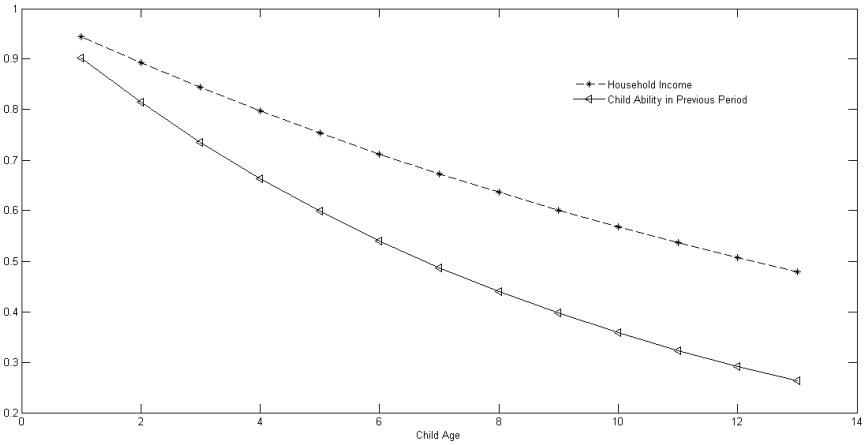
where $i = 1, 2$ and the estimated values for ξ_1 and ξ_2 are shown in table 6.

Table 8
Goodness of fit for mother's choices.

	Actual Data	**Simulated Data**
Hours of work	27.1166	30.8259
External child care hours	14.6839	10.2375
Maternal time with the child	21.0582	18.3568

NOTE. Actual data represent PSID-CDS data on children aged 0-12 in 1997, with at least one test score measure and without siblings. See section 5 and appendix C for further details on the data. Simulated data represent the data obtained simulating the model described in section 3 and 4.1 and setting the parameters at the estimated values shown in tables 4, 5 and 6.

FIGURE 9
Income and previous period child's ability productivity.

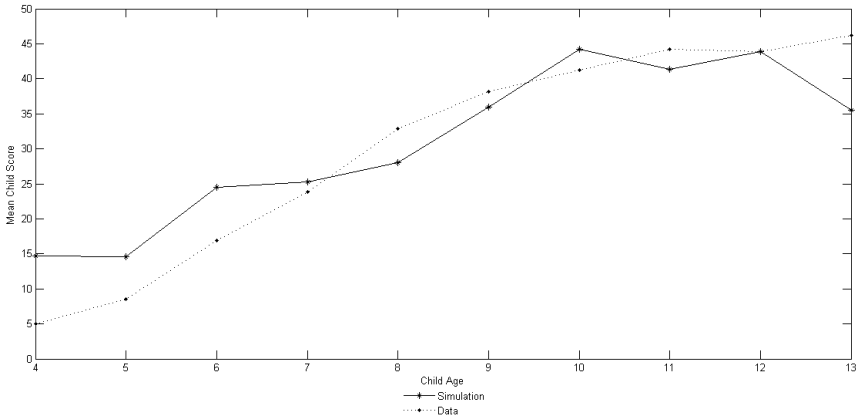


NOTE. This graph represents the productivity parameters for income (I_t) and child's ability (A_t) as a function of child's age $t = 1, 2, \ldots 13$. These parameters are defined as

$$\delta_i = exp(\xi_i t)$$

where $i = 3, 4$ and the estimated values for ξ_3 and ξ_4 are shown in table 6.

FIGURE 10
Goodness of fit for child's test score measure by child's age.



NOTE. Actual data represent PSID-CDS data on children aged 0-12 in 1997, with at least one test score measure and without siblings. See section 5 and appendix C for further details on the data. Simulated data represent the data obtained simulating the model described in section 3 and 4.1 and setting the parameters at the estimated values shown in tables 4, 5 and 6.

grant to households with children, (ii) increasing the wage offers by 50 percent, or (iii) subsidizing external child care, setting the price of the service at 1 US$ per hour.

7.1. ***Leisure-minimizing preferences.*** The main contribution of this paper is to estimate the effect of maternal employment and external child care taking into account the actual time spent by the mother with the child.

The literature has usually assumed that maternal time can be proxied by the mother's total time endowment net of working time. If this is the case, the mother does not care about having leisure and spends all the time out of work with the child. In order to see what are the implications of this assumption and to relate this model to the ones previously estimated in this literature, I re-estimate it setting the preference parameter

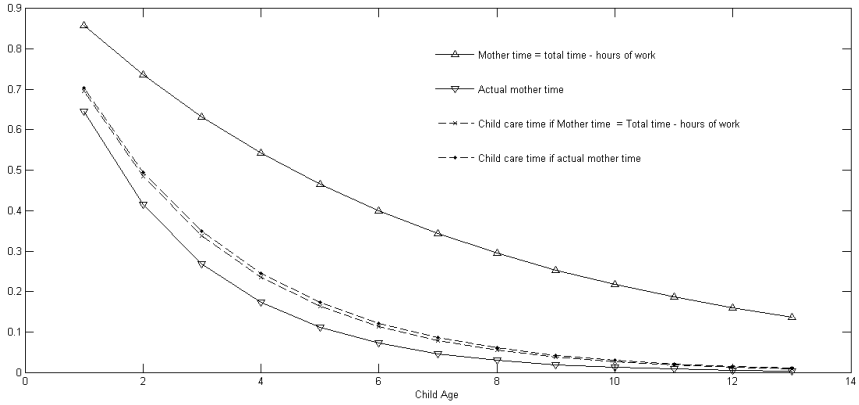
for leisure $\alpha_1 = 0$ and defining maternal time with the child as the difference between the total time endowment and the time spent at work, i.e. $\tau = TT - h$.

Figure 11 reports the elasticity of child's ability with respect to maternal time and external child care in this counterfactual scenario. For comparison purposes, the figure reports also the elasticities found in the main analysis and already presented in figure 8; they are labeled as "actual time". Notice that the parameters for external child care do not change across specifications, while the estimates for the elasticity of child's ability with respect to maternal time differ substantially. In this counterfactual scenario, the elasticity of child's ability with respect to maternal time ranges from 0.86 when the child is 1 year old to 0.1 when the child is aged 13. These values imply that if maternal time reduces by 10 percent when the child is 1 year old, the child's cognitive ability decreases by 8.6 percent. If the same amount of time is spent in external child care, child's cognitive ability increases by almost 7 percent. Thus, the reduction in child's ability due to the absence of the mother is only partially compensated for by the use of external child care. The net effect of mother's work during the child's first year of life is then a 2 percent decline in child's ability. It is interesting that this result is so similar with the effect of maternal employment that has been found before in the literature, in particular in studies using structural estimation. In fact, the model estimated in this paper is substantially different from, for instance, the one proposed by Bernal (2008), in terms of both functional form and identifying assumptions. The fact that, using the same specification for maternal time, the present model gives similar results is comforting. This suggests that the difference between the estimates found in this paper and the results found before in the literature is very likely to depend on the way of defining maternal time. If maternal time with the child is interpreted as all the time available to the mother net of working time, assuming that the mother does not have any leisure, this is very likely to overestimate maternal time productivity and the negative effect of maternal employment. Instead, if an actual measure of maternal time is used, allowing the mother to allocate the time out of work between leisure and care of the child, maternal employment is found not to be detrimental for child's development.

7.2. ***Policy simulations.*** In this section, I simulate the effects of some policies on mother's choices and behavior as well as on the level of child's ability reached in the last period. More precisely, I consider the following policies: (i) a lump-sum subsidy that increases household income by 100 US$ in each period; (ii) an increase in mother's wage by 50 percent, and (iii) a child care subsidy that set the price of child care at 1 US$ per hour. Results are shown in table 9.

The first panel of the table refers to a policy providing subsidies to households with children. More precisely, the policy implies that all households in any period receive an additional amount of 100 US$ that increases their (exogenous) income. One may think to this exercise as a simplification of subsidy policies targeted toward poor families and providing cash transfers not conditioned on specific behavior of the household members. As it has been shown in section 3.2, an increase in income determines a reduction in labor supply and an increase in the demand for all goods, i.e., consumption and child's

Figure 11
Productivity parameters for maternal time and child care obtained assuming a relationship between maternal time and working time. Comparison with the corresponding parameters found in the main analysis.



NOTE. 'Mother time = total time - hours of work' and 'child care time if mother time = total time - hours of work' represent the productivity parameters of maternal time and external child care, respectively, obtained from the estimation of the model setting $\alpha_1 = 0$ and $\tau = TT - h$, where α_1 is the preference parameter for leisure and $TT = 112$. 'Actual mother time' and 'Child care if actual mother time' represent the productivity parameters of maternal time and external child care, as already shown in figure 8.

ability. This effect is then reflected in the positive change in external child care use and maternal time with the child after the policy, since mothers adjust their investments in order to achieve higher level of child's ability. The final effect on child's ability is large, since the change in income determines a shift not only in the input related to the expenditure for the child (i.e., I_t), but also in external child care and maternal time. Notice that while this policy has a substantially positive effect on child's ability, it also influences the behavior of mothers inducing a strong reduction in their labor market participation. This may have consequences in terms of cost-benefit evaluation of the policy, since the current expenditure for providing 100 US$ to each household cannot be paid for by an increase in mothers' labor supply in the short-run, but only by higher capabilities of the child in the medium-run.

The second panel of table 9 reports the percentage change of the variables induced by an increase in mother's wage by 50 percent in any period. This may be due, for instance, to policies decreasing taxation on mothers' labor income or providing incentives for mothers' employment. The higher wage induced by the policy represents an higher opportunity cost of maternal time with the child, so that mothers work more and spend a lower amount of time with the child or in leisure. The demand for child care is positively affected by this policy through two different channels. First, since mothers are working more and need someone looking after the child for a larger amount of time, they also use more external child care. Second, they are also earning more, so that they can buy an higher amount of the service for their child. However, the policy has a very little effect on child's ability. This result stresses the importance of taking into account all the plausible channels with which the policy affects the outcomes of interests. In other words, a policy increasing mothers' labor income may be effective in increasing

Table 9
Policy simulations.

Increase in household income		
	Baseline	Percentage change
Test score measure in the last period	37.4241	50.8744
Child's ability in the last period	7.0119	431.3410
Hours of work	30.8259	−24.6236
Maternal time with the child	18.3568	4.4408
External child care time	10.2375	16.3742
Leisure	64.6404	9.2906
Consumption	482.5416	17.1592
Increase in mother's wage		
	Baseline	Percentage change
Test score measure in the last period	37.4241	0.1884
Child's ability in the last period	7.0119	1.0525
Hours of work	30.8259	2.3768
Maternal time with the child	18.3568	−0.5555
External child care time	10.2375	49.5170
Leisure	64.6404	−0.8584
Consumption	482.5416	49.5000
Reduction in child care price		
	Baseline	Percentage change
Test score measure in the last period	37.4241	0.8772
Child's ability in the last period	7.0119	4.5409
Hours of work	30.8259	7.1795
Maternal time with the child	18.3568	−0.0005
External child care time	10.2375	415.0235
Leisure	64.6404	−0.0045
Consumption	482.5416	0.0027

NOTE. This table reports percentage changes with respect to the baseline levels from (i) an increase in household income by 100 US$ per week, (ii) an increase in mother's wage by 50 percent, and (iii) a policy setting external child care price at 1 US$ per hour. Test score measure in the last period is the value of the simulated test score at the end of period $t = 12$. Child's ability in the last period is the value of the simulated child's ability at the end of period $t = 12$.

mothers' participation to the labor market, but can fail in having an effect on child's development although improving the economic conditions of the households.

The panel at the bottom of table 9 shows the percentage change of the variables after the implementation of a policy setting the price of child care at 1 US$ per hour. Similar policies have been implemented and evaluated during last years, especially in the U.S. and Canada. For instance, Baker, Gruber, and Milligan (2008) evaluate the effects of a policy setting the out-of-pocket price at 5$ per day in Quebec on maternal employment, child care use and child's outcomes. They find that the policy increases the use of the subsidized service and it also has a positive effect on maternal employment; they do not find any effect on the cognitive outcomes of children. The simulation of this policy has been done setting the hourly price of child care at 1 US$ instead of 5.15 US$, which is the estimated value shown in table 4. The results of this simulation are in line with the

ones found by Baker et al. (2008): the reduction in child care price, in fact, determines a large increase in the use of external child care but also an increase in mother's net wage. The substitution effect seems to prevail, since mothers' labor supply increases after the policy, while maternal time with the child and leisure seem to be unaffected. However, differently from Baker et al. (2008), the simulation of the policy yields a small increase in child's ability, induced by the increase in external child care use.

The results of these policy simulations may also depend on the sample selection issues exposed in section 5. The fact that both the increase in mothers' wage and the subsidization of child care have a small impact on mothers' labor supply seems to confirm that the subsample used to estimate the model is characterized by high-skilled mothers, already working in the labor market, for whom these policies do not play a relevant role. Similarly, if the sample is composed mainly by high-skilled mothers and children, the same policies have a small effect on child development, despite changing mothers' investments decisions.

8. Concluding remarks

This paper estimates a behavioral model where the labor supply, non-parental child care and time allocation choices of the mother are considered endogenous. In contrast to all existing studies in the literature, this paper takes into account the additional choice the mother makes concerning the time allocation between leisure and time with the child.

Maternal time and external child care serve as inputs in a child's development process that represents a constraint to the mother's utility maximization problem. The model is estimated using U.S. data from the Panel Study of Income Dynamics (PSID) and its Child Development Supplement (CDS) conducted in 1997, 2002 and 2007. The parameters of the model are estimated using a Simulated Minimum Distance estimator that minimizes the distance between several data statistics and their simulated counterparts.

The results suggest that maternal employment does not have a negative effect on child development, as long as the reduction in maternal time is compensated for by the alternative forms of care available to the mother. In fact, even though maternal employment reduces the amount of time the child spends with the mother, this negative effect on child's ability can be compensated for by a positive effect induced by non-parental child care attendance for the same amount of time.

Previous literature has neglected the additional choice of the mother between leisure time and maternal time with the child, defining the variable maternal time as the total time endowment available to the mother net of working time. It implies assuming that the mother does not care about leisure and that she spends all available time with the child. A counterfactual exercise that I perform using the estimated model shows that this assumption overestimates the amount of time spent by the mother with the child, as well as its productivity for child's development. In fact, if the model is estimated under this framework, maternal employment has a detrimental effect and the reduction in maternal time cannot be compensated for by the use of external child

care. This confirms that in order to estimate the effect of maternal employment and external child care on child's development it is important to take into account also the additional choice the mother makes concerning the time allocation out of work.

The policy simulations suggest that the policy maker should take into account all the potential effects and mechanisms with which the policies can affect the outcomes of interests. In fact, even though these simulations allow to only evaluate "local" effects, they show that policies aimed at increasing participation of mothers in the labor market, or at improving the economic conditions of poor households, may not necessarily have the same effect on child's development. Similarly, policies decreasing the cost of using external child care can induce an higher use of the service, but may have very small effects on either mothers' participation or child's ability.

Further research is needed in order to better understand the determinants of mothers' decisions concerning the usage of formal and informal child care arrangements and their effects on child subsequent development. In fact, the model presented in this paper does not distinguish between different kinds of child care and assumes that any type of care has the same productivity for child development. However, not only these two choices may respond to different decision making processes, but also they may have diverse implications for child development. Due to the importance of this topic, I leave for future research the specification of a model incorporating also the choices of using formal and informal child care.

References

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

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

Baydar, N. and J. Brooks-Gunn (1991). Effects of maternal employment and child care arrangements on preschoolers' cognitive and behavioral outcomes. *Developmental Psychology 27*, 932–945.

Becker, G. S. and N. Tomes (1986). Human capital and the rise and fall of families. *Journal of Labor Economics 4*, S1–S39.

Belsky, J. and D. Eggebeen (1991). Early and extensive maternal employment and young childrens socioemotional development: Children of the National Longitudinal Survey of Youth. *Journal of Marriage and the Family 53*, 10831110.

Bernal, R. (2008). The effect of maternal employment and child care on children's cognitive development. *International Economic Review 49*(4), 1173–1209.

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

Bernal, R. and M. P. Keane (2011). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics 29*(3), 459–512.

Bianchi, S. M. (2000). Maternal employment and time with children: dramatic change or surprising continuity? *Demography 37*(4), 401–414.

Blau, D. M. (1999). The effect of income on child development. *Review of Economics and Statistics 81*(2), 261–276.

Blau, D. M. and J. Currie (2006). Pre-school, day care, and after-school care: Who's minding the kids? In E. A. Hanushek and F. Welch (Eds.), *Handbook of The Economics of Education*, Volume 2.

Blundell, R., L. Pistaferri, and I. Preston (2008). Consumption inequality and partial insurance. *American Economic Review 98*(5), 1887 – 1921.

Cameron, A. C. and P. K. Trivedi (2005). *Microeconometrics. Methods and applications.* New York, USA: Cambridge University Press.

Carneiro, P. and J. J. Heckman (2003). Human capital policy. In J. J. Heckman, A. B. Krueger, and B. M. Friedman (Eds.), *Inequality in America: What Role for Human Capital Policies?*, pp. 77–239. Cambridge: MA: MIT Press.

Chase-Lansdale, P. L., R. A. Moffit, B. J. Lohman, A. J. Cherlin, R. L. Coley, L. D. Pittman, J. Roff, and E. Votruba-Drzal (2003). Mother's transitions from welfare to work and the well-being of preschoolers and adolescents. *Science 299*, 1548–1552.

Cox, M. J., M. T. Owen, V. K. Henderson, and N. A. Margand (1992). Prediction of infant-father and infant-mother attachment. *Developmental Psychology 28*(3), 474–483.

Cunha, F., J. J. Heckman, L. J. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. In E. A. Hanushek and F. Welch (Eds.),

Handbook of the Economics of Education, pp. 697–812. Amsterdam: North-Holland.

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

Currie, J. and D. Thomas (1995). Does Head Start make a difference? *American Economic Review 92*(3), 341–364.

Currie, J. and D. Thomas (1999). Does Head Start help Hispanic children? *Journal of Public Economics 74*(2), 235–262.

Davidson, R. and J. G. MacKinnon (2003). *Econometric theory and method*. New York: Oxford University Press.

Del Boca, D., C. J. Flinn, and M. Wiswall (2013). Household choices and child development. *Review of Economic Studies forthcoming*.

Del Boca, D., C. Monfardini, and C. Nicoletti (2012). Self investments of adolescents and their cognitive development. Collegio Carlo Alberto Notebooks N.265.

Desai, S., L. P. Chase-Lansdale, and R. T. Michael (1989). Mother or market? Effects of maternal employment on the intellectual ability of 4-year old children. *Demography 26*, 545–561.

Duncan, G. J. and J. Brooks-Gunn (1997). Income effects across the life span : integration and interpretation. In G. J. Duncan and J. Brooks-Gunn (Eds.), *Consequences of growing up poor*. New York: Russel Sage Foundation.

Duncan, G. J., W. J. Yeung, J. Brooks-Gunn, and J. R. Smith (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review 63*(3), 406–423.

Ermish, J. and M. Francesconi (2012). The effect of parental employment on child schooling. *Journal of Applied Econometrics forthcoming*.

Fernandez, R. (2007). Women, work and culture. *Journal of the European Economic Association 5*, 305–332.

Flinn, C. J. (2000). Modes of interaction between divorced parents. *International Economic Review 41*(3).

Fogli, A. and L. Veldkamp (2011). Nature or nurture? Learning and the geography of female labor force participation. *Econometrica 79*(4), 1103–1138.

Francesconi, M. (2002). A joint dynamic model of fertility and work of married women. *Journal of Labor Economics 20*(2), 336–380.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *Journal of Economic Perspectives 22*(3), 23–46.

Heckman, J. J. (2007). The economics, technology, and neuroscience of human capability formation. *PNAS 104*(33), 13250–13255. Proceedings of the National Academy of Sciences of the United States of America.

Heckman, J. J. (2008). Schools, skills, and synapses. *Economic Inquiries 46*(3), 289.

Hoffert, S. L., P. E. Davis-Kean, J. Davis, and J. Finkelstein (1997). *The Child Development Supplement to the Panel Study of Income Dynamics. The 1997 User Guide*. Ann Arbor, MI: Survey Research Center, Institute for Social Research, The University of Michigan.

Hoffert, S. L. and J. F. Sandberg (2001). How American children spend their time? *Journal of Marriage and the Family 63*, 295–308.

Hsin, A. (2009). Parent's time with children: Does time matter for children's cognitive achievement? *Social Indicators Research 93*(1), 123–126.

ISR (2010a). *The Panel Study of Income Dynamics Child Development Supplement. The User Guide Supplement for CDS-I.* Ann Arbor, MI: Survey Research Center, Institute for Social Research, The University of Michigan.

ISR (2010b). *The Panel Study of Income Dynamics Child Development Supplement. The User Guide Supplement for CDS-II.* Ann Arbor, MI: Survey Research Center, Institute for Social Research, The University of Michigan.

James-Burdumy, S. (2005). The effects of maternal labor force participation on child development. *Journal of Labor Economics 23*(1), 177–211.

Keane, M. P. (2010). Structural vs. atheoretic approach to econometrics. *Journal of Econometrics 156*, 3–20.

Keane, M. P. and R. A. Moffitt (1998). A structural model of multiple welfare program participation and labor supply. *International Economic Review 39*(3), 553–589.

Leibowitz, A. (1974). Home investments in children. In T. W. Schultz (Ed.), *Marriage, Family, Human Capital, and Fertility.* NBER. National Bureau of Economic Research.

Leibowitz, A. (1977). Parental inputs and children's achievement. *Journal of Human Resources 12*(2), 242–251.

Levy, D. and G. J. Duncan (2012). Using siblings samples to assess the effect of childhood family income on completed schooling. Northwestern University Working Paper.

Loeb, S., M. Bridges, D. Bassok, B. Fuller, and R. W. Rumberger (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review. Economics of Early Childhood Education Issue 26*, 52–66.

Magnuson, K. A., C. J. Ruhm, and J. Waldfogel (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review. Economics of Early Childhood Education Issue 26*, 33–51.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica 57*(5), 995–1026.

Mroz, T., H. Liu, and W. Van der Klaauw (2010). Maternal employment, migration and child development. *Journal of Econometrics 156*(1), 212–228.

Novick, M. R. (1966). The axioms and principal results of classical test theory. *Journal of Mathematical Psychology 3*, 1–18.

Pakes, A. and D. Pollard (1989). Simulation and the asymptotics of optimization estimators. *Econometrica 57*(5), 1027–1057.

Parcel, T. L. and E. G. Menaghan (1994). Early parental work, family social capital, and early childhood outcomes. *The American Journal of Sociology 99*(4), 972–1009.

Ribar, D. C. (1992). Child care and the labor supply of married women. Reduced form evidence. *Journal of Human Resources 27*(1), 134–165.

Ruhm, C. J. (2004). Parental employment and child cognitive development. *Journal of Human Resources XXXIX*, 155–192.

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

Train, K. (2009). *Discrete Choice Methods with Simulation* (2 ed.). Cambridge University Press.

U.S. Census Bureau (2000). *Statistical Abstract of the United States: 2000.*

U.S. Census Bureau (2013). American Time Use Survey 2005-2009. http://www.bls.gov/tus/.

Van Der Klaauw, W. (1996). Female labour supply and marital status decisions: A life-cycle model. *The Review of Economic Studies 63*(2), 199–235.

Woodcock, R. W. and M. E. B. Johnson (1989). *Tests of Achievement, Standard Battery [Form B].* Chicago, IL: Riverside Publishing.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge: MA: MIT Press.

Appendix A. Analytic solution of the model

In this appendix I derive analytically the closed-form solutions of the model, for all the choice variables.

The process of backward induction involves the solution of the optimization problem in each period, starting from the last one, T. Consider first the choice variables i_t and τ_t. The first step is to find the optimal child care and time input decisions at time T. The value function of the mother at period T can be written as:

$$V_T = max_{i_T,\tau_T}\ \alpha_1 ln(TT-h_T-\tau_T)+\alpha_2 ln(w_T h_T+I_T-pi_T)+\alpha_3 ln(A_T)+E_T\beta\{\tilde{V}_{T+1}+\rho\alpha_3 lnA_{T+1}\} \tag{A.1}$$

where the variables l_T and c_T have been already substituted using the time and budget constraints. Notice that the expectation operator in (A.1) is with respect to the terminal period value function, as defined in (7).

The optimal solutions for both i_T^c and τ_T^c at period T, conditional on h_T, are given by the solutions of the following first order conditions (FOCs):

$$\begin{aligned} i_T^c &\Rightarrow \frac{\partial V_T}{\partial i_T} = 0 \\ \tau_T^c &\Rightarrow \frac{\partial V_T}{\partial \tau_T} = 0 \end{aligned} \tag{A.2}$$

Due to the value-added specification of the child cognitive ability production function, as defined by (4), child ability in period $T+1$ is a function of the inputs received by the child at period T. Hence, (A.2) can be rearranged, using total differential, in the following way:

$$\begin{aligned} i_T^c &\Rightarrow \frac{\partial \bar{V}_T}{\partial i_T} + \frac{\partial V_{T+1}}{\partial lnA_{T+1}} \times \frac{\partial lnA_{T+1}}{\partial i_T} = 0 \\ \tau_T^c &\Rightarrow \frac{\partial \bar{V}_T}{\partial \tau_T} + \frac{\partial V_{T+1}}{\partial lnA_{T+1}} \times \frac{\partial lnA_{T+1}}{\partial \tau_T} = 0 \end{aligned} \tag{A.3}$$

where $\bar{V}_T$ is the current utility in period T:

$$\bar{V}_T = \alpha_1 ln(TT - h_T - \tau_T) + \alpha_2 ln(w_T h_T + I_T - pi_T) + \alpha_3 ln(A_T)$$

The corresponding derivatives[36] are given by the following expressions:

[36]The second term of the expressions defined in (A.3) is derived using the logarithm of A_{T+1} just for computational convenience. The results are the same computing $\frac{\partial V_{t+1}}{\partial A_{t+1}} \times \frac{\partial V_{t+1}}{\partial i_t}$ and $\frac{\partial V_{t+1}}{\partial A_{t+1}} \times \frac{\partial V_{t+1}}{\partial \tau_t}$, i.e. substituting the CAPF in exponential form:

$$A_{T+1} = \tau_T^{\delta_{1T}} i_T^{\delta_{2T}} I_T^{\delta_{3T}} A_T^{\delta_{4T}}$$

$$\frac{\partial \bar{V}_T}{\partial i_T} = \frac{-p\alpha_2}{w_T h_T + I_T - pi_T} \tag{A.4}$$

$$\frac{\partial \bar{V}_T}{\partial \tau_T} = \frac{-\alpha_1}{TT - h_T - \tau_T} \tag{A.5}$$

$$\frac{\partial V_{T+1}}{\partial lnA_{T+1}} \times \frac{\partial lnA_{T+1}}{\partial i_T} = (\beta\rho\alpha_3)\left(\frac{\delta_{2T}}{i_T}\right) \tag{A.6}$$

$$\frac{\partial V_{T+1}}{\partial lnA_{T+1}} \times \frac{\partial lnA_{T+1}}{\partial \tau_T} = (\beta\rho\alpha_3)\left(\frac{\delta_{1T}}{\tau_T}\right) \tag{A.7}$$

and the FOCs become:

$$i_T^c \Rightarrow \frac{-p\alpha_2}{w_T h_T + I_T - pi_T} + (\beta\rho\alpha_3)\left(\frac{\delta_{2T}}{i_T}\right) = 0 \tag{A.8}$$

$$\tau_T^c \Rightarrow \frac{-\alpha_1}{TT - h_T - \tau_T} + (\beta\rho\alpha_3)\left(\frac{\delta_{1T}}{\tau_T}\right) = 0$$

The solutions for both inputs at period T are given by:

$$i_T^c = \frac{\beta\delta_{2T}D_{T+1}}{p(\alpha_2 + \beta\delta_{2T}D_{T+1})}(w_T h_T + I_T) \tag{A.9}$$

$$\tau_T^c = \frac{\beta\delta_{1T}D_{T+1}}{\alpha_1 + \beta\delta_{1T}D_{T+1}}(TT - h_T) \tag{A.10}$$

where $D_{T+1} = \frac{\partial V_{T+1}}{\partial lnA_{T+1}} = \rho\alpha_3$.

These solutions can be substituted into the value function of the mother at period T, in order to get $V_T(i_T^c, \tau_T^c)$.

Consider now period $T-1$. The value function for this period is:

$$\begin{aligned} V_{T-1} = max_{i_{T-1},\tau_{T-1}}\ & \alpha_1 ln(TT - h_{T-1} - \tau_{T-1}) + \alpha_2 ln(w_{T-1}h_{T-1} + I_{T-1} - pi_{T-1}) + \alpha_3 ln(A_{T-1}) + \\ & + E_{T-1}\beta\{\alpha_1 ln(TT - h_T - \tau_T^C) + \alpha_2 ln(w_T h_T + I_T - pi_T^C) + \alpha_3 lnA_T + \\ & + \beta\{\tilde{V_{T+1}} + \rho\alpha_3[\delta_{1T} ln\tau_T^C + \delta_{2T} lni_T^C + \delta_{3T} lnI_T + \delta_{4T} lnA_T]\}\} \end{aligned} \tag{A.11}$$

The expectation in (A.11) is with respect to the value function at period T ($V_T(i_T^c, \tau_T^c)$) *and* the terminal period value function at period $T+1$.

Applying total differential, the solutions for both inputs in period $T-1$ are given by:

In this case, the second terms of the expressions in (A.3) become:

$$\frac{\partial V_{T+1}}{\partial A_{T+1}} \times \frac{\partial A_{T+1}}{\partial i_T} = \frac{\beta\rho\alpha_3}{\tau_T^{\delta_{1T}} i_T^{\delta_{2T}} I_T^{\delta_{3T}} A_T^{\delta_{4T}}}(\tau_T^{\delta_{1T}} i_T^{\delta_{2T}-1} I_T^{\delta_{3T}} A_T^{\delta_{4T}}) = \beta\rho\alpha_3\frac{\delta_{2T}}{i_T}$$

$$\frac{\partial V_{T+1}}{\partial A_{T+1}} \times \frac{\partial A_{T+1}}{\partial \tau_T} = \frac{\beta\rho\alpha_3}{\tau_T^{\delta_{1T}} i_T^{\delta_{2T}} I_T^{\delta_{3T}} A_T^{\delta_{4T}}}(\tau_T^{\delta_{1T}-1} i_T^{\delta_{2T}} I_T^{\delta_{3T}} A_T^{\delta_{4T}}) = \beta\rho\alpha_3\frac{\delta_{1T}}{\tau_T}$$

that are equivalent to (A.6) and (A.7).

$$i^c_{T-1} \Rightarrow \frac{\partial \bar{V}_{T-1}}{\partial i_{T-1}} + \frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial i_{T-1}} = 0 \tag{A.12}$$

$$\tau^c_{T-1} \Rightarrow \frac{\partial \bar{V}_{T-1}}{\partial \tau_{T-1}} + \frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial \tau_{T-1}} = 0 \tag{A.13}$$

where

$$\bar{V}_{T-1} = \alpha_1 ln(TT - h_{T-1} - \tau_{T-1}) + \alpha_2 ln(w_{T-1}h_{T-1} + I_{T-1} - pi_{T-1}) + \alpha_3 ln(A_{T-1})$$

and

$$\frac{\partial \bar{V}_{T-1}}{\partial i_{T-1}} = \frac{-p\alpha_2}{w_{T-1}h_{T-1} + I_{T-1} - pi_{T-1}} \tag{A.14}$$

$$\frac{\partial \bar{V}_{T-1}}{\partial \tau_{T-1}} = \frac{-\alpha_1}{TT - h_{T-1} - \tau_{T-1}} \tag{A.15}$$

$$\frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial i_{T-1}} = (\alpha_3 + \beta\alpha_3)\left(\frac{\delta_{2T-1}}{i_{T-1}}\right) \tag{A.16}$$

$$\frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial \tau_{T-1}} = (\alpha_3 + \beta\alpha_3)\left(\frac{\delta_{1T-1}}{\tau_{T-1}}\right) \tag{A.17}$$

Substituting (A.14),(A.15),(A.16) and (A.17) into (A.12) and (A.13) yields:

$$i^c_{T-1} \Rightarrow \frac{-p\alpha_2}{w_{T-1}h_{T-1} + I_{T-1} - pi_{T-1}} + (\alpha_3 + \beta\alpha_3)\left(\frac{\delta_{2T-1}}{i_{T-1}}\right) = 0 \tag{A.18}$$

$$\tau^c_{T-1} \Rightarrow \frac{-\alpha_1}{TT - h_{T-1} - \tau_{T-1}} + (\alpha_3 + \beta\alpha_3)\left(\frac{\delta_{1T-1}}{\tau_{T-1}}\right) = 0 \tag{A.19}$$

The solutions for both choice variables in period $T-1$, conditional on h_{T-1}, are then:

$$i^c_{T-1} = \frac{\beta\delta_{2T-1}D_T}{p(\alpha_2 + \beta\delta_{2T-1}D_T)}(w_{T-1}h_{T-1} + I_{T-1}) \tag{A.20}$$

$$\tau^c_{T-1} = \frac{\beta\delta_{1T-1}D_T}{\alpha_1 + \beta\delta_{1T-1}D_T}(TT - h_{T-1}) \tag{A.21}$$

where

$$D_T = \frac{\partial V_T}{\partial lnA_T} = \alpha_3 + \beta\delta_{4T}D_{T+1}$$

The solutions for period $T-1$, given by equations (A.20) and (A.21), can be substituted in (A.11) in order to get $V_{T-1}(i^c_{T-1}, \tau^c_{T-1})$. This expression can be used to write down the value function at period $T-2$. Using the same process described for periods T and $T-1$ and computing the corresponding derivatives yields the solutions for period $T-2$. The solutions for all the periods up to period $t=1$ can be retrieved similarly.

At the end, two sequences of optimal choices can be obtained. The sequence of optimal non-parental child care choices, conditional on mother's labor supply, is given by:

$$i_T^c = \frac{\beta\delta_{2T}D_{T+1}}{p(\alpha_2 + \beta\delta_{2T}D_{T+1})}(w_T h_T + I_T) \tag{A.22}$$

$$i_{T-1}^c = \frac{\beta\delta_{2T-1}D_T}{p(\alpha_2 + \beta\delta_{2T-1}D_T)}(w_{T-1} h_{T-1} + I_{T-1}) \tag{A.23}$$

$$i_{T-2}^c = \frac{\beta\delta_{2T-2}D_{T-1}}{p(\alpha_2 + \beta\delta_{2T-2}D_{T-1})}(w_{T-2} h_{T-2} + I_{T-2}) \tag{A.24}$$

$$\vdots$$

$$i_t^c = \frac{\beta\delta_{2t}D_{t+1}}{p(\alpha_2 + \beta\delta_{2t}D_{t+1})}(w_t h_t + I_t) \tag{A.25}$$

$$\vdots$$

$$i_2^c = \frac{\beta\delta_{22}D_3}{p(\alpha_2 + \beta\delta_{22}D_3)}(w_2 h_2 + I_2) \tag{A.26}$$

$$i_1^c = \frac{\beta\delta_{21}D_2}{p(\alpha_2 + \beta\delta_{21}D_2)}(w_1 h_1 + I_1) \tag{A.27}$$

Equation (A.25) is equal to (8) in the main text. Instead, the sequence of optimal choices for time with the child, conditional on mother's labor supply, is given by:

$$\tau_T^c = \frac{\beta\delta_{1T}D_{T+1}}{(\alpha_1 + \beta\delta_{1T}D_{T+1})}(TT - h_T) \tag{A.28}$$

$$\tau_{T-1}^c = \frac{\beta\delta_{1T-1}D_T}{(\alpha_1 + \beta\delta_{1T-1}D_T)}(TT - h_{T-1}) \tag{A.29}$$

$$\tau_{T-2}^c = \frac{\beta\delta_{1T-2}D_{T-1}}{(\alpha_1 + \beta\delta_{1T-2}D_{T-1})}(TT - h_{T-2}) \tag{A.30}$$

$$\vdots$$

$$\tau_t^c = \frac{\beta\delta_{1t}D_{t+1}}{(\alpha_1 + \beta\delta_{1t}D_{t+1})}(TT - h_t) \tag{A.31}$$

$$\vdots$$

$$\tau_2^c = \frac{\beta\delta_{12}D_3}{(\alpha_1 + \beta\delta_{12}D_3)}(TT - h_2) \tag{A.32}$$

$$\tau_1^c = \frac{\beta\delta_{11}D_2}{(\alpha_1 + \beta\delta_{11}D_2)}(TT - h_1) \tag{A.33}$$

Equation (A.31) is equal to equation (9) in the text. The sequence of values for D_{t+1} is defined in the main text.

Once having found the solutions for both the child care and the time allocation decisions, the solutions for the labor supply can be computed using the same backward procedure. Equation (11) represents the optimal labor supply in each period as a

function of i_t and τ_t; substituting (8) and (9), it yields the optimal labor supply choice for each period t, as defined by (12).

The unconditional demands for child care and time with the child are derived substituting the labor supply solution into equations (8) and (9) (corresponding to (A.25) and (A.31) in this appendix). The final expressions for them are reported in equations (14) and (15).

Appendix B. Empirical analysis and estimation

This appendix provides additional details on the empirical analysis performed to estimate the model.

B.1. ***Empirical specification.*** As stated in section 4.1, the definition of the model parameters should ensure that they respect the requirements imposed by the functional form restrictions. In order to respect these requirements without posing additional constraints to the estimation algorithm, I use a suitable transformation of the original parameters for any coefficient on which the model imposes restrictions due to functional form or empirical specification assumptions.[37]

Concerning the parameters in the mother's utility function, they should be positive and sum to one. In order to respect these requirements, I define them as multinomial probabilities and equal to the following expressions:

$$\alpha_1 = \frac{exp(\gamma_1)}{exp(\gamma_1) + exp(X'\gamma_2) + exp(X'\gamma_3 + I(\mu_0 = \mu_{0high}))} \tag{B.1}$$

$$\alpha_2 = \frac{exp(X'\gamma_2)}{exp(\gamma_1) + exp(X'\gamma_2) + exp(X'\gamma_3 + I(\mu_0 = \mu_{0high}))} \tag{B.2}$$

$$\alpha_3 = \frac{exp(X'\gamma_3 + I(\mu_0 = \mu_{0high}))}{exp(\gamma_1) + exp(X'\gamma_2) + exp(X'\gamma_3 + I(\mu_0 = \mu_{0high}))} \tag{B.3}$$

where $X = [MotherEducation, MotherRace]$, γ_1 is normalized to being 0 and μ_0 represents mother's skills. The distribution of mother's skills is explained in the text. A similar transformation has been implemented for parameters representing probabilities, i.e. type proportions of high and low skilled mothers and children. More precisely, the proportion of high skilled mothers is defined as

$$\pi_{mh} = exp(z_m)/(1 + exp(z_m)) \tag{B.4}$$

while the proportion of high skilled children is

$$\pi_{ch} = exp(z_c)/(1 + exp(z_c)) \tag{B.5}$$

The proportions of low skilled mothers and children are, respectively, $\pi_{ml} = (1 - \pi_{mh})$ and $\pi_{cl} = (1 - \pi_{ch})$. The parameters z_m and z_c are actually estimated and are used to recover the type proportions of high skilled mothers and children.

[37] See Flinn (2000) and Mroz et al. (2010) for similar applications.

Concerning the CAPF, the parameters in this case should be strictly positive. Hence, I implement the transformation defined by (19) that exploits the properties of the exponential function.

The vector of parameters to be estimated is the following:

$$\Theta = \{\Gamma_2, \Gamma_3, \rho, \Xi, \sigma_v, \mu_{0k}, z_{mh}, \Upsilon, \sigma_\epsilon, \psi_{ck}, z_{ch}, \Delta, p, \theta_{inc}\} \tag{B.6}$$

where $\Gamma_2 = (\gamma_{2\ MotherEdu}, \gamma_{2\ MotherRace})$, $\Gamma_3 = (\gamma_{3\ MotherEdu}, \gamma_{3\ MotherRace})$, $\Xi = (\xi_1, \xi_2, \xi_3, \xi_4)$, $k = (h, l)$, $\Upsilon = (\mu_1, \mu_2, \mu_3, \mu_4)$, $\Delta = (\eta_1, \eta_2)$ and $\theta_{inc} = (\mu_{inc}, \sigma_{inc})$.

The parameter p represents the hourly price of child care. It is estimated because the actual distribution of that measure in the data has a large mass toward zero, also for children actually using the service. This may be due to the usage of informal child care, that can have a zero market price. Using the direct measure available in the data yields an infinite demand for external child care for those using an arrangement with a zero price, regardless of mother's labor income and household earnings.

B.2. ***Estimation.*** The estimation has been done in two-stages, after having set the discount factor $\beta = 0.95$: the parameters of the income process have been estimated in the first stage, while all remaining parameters have been estimated in the second stage.[38]

After having computed the statistics defined in table 1 for the actual data, I proceed with the first-stage estimation of the income parameters. This involves the simulation of the income process, after having drawn from a standard normal distribution $N * R$ times, for every period. This distribution is actually a function of the two parameters that should be estimated, i.e., μ_{inc} and σ_{inc}. The statistics used to estimate these parameters are the average, standard deviation and median income for all the periods. I compute these points for both the actual and the simulated income processes. The SMD estimator for this first stage minimizes an objective function where each moment condition is the distance between the income data moments and their simulated counterparts. Each moment condition is weighted using the inverse of the corresponding statistics in the data. The vector of first-stage estimated parameters is then: $\hat{\theta}_{inc} = (\hat{\mu}_{inc}, \hat{\sigma}_{inc})$.

The second-stage involves the estimation of all remaining parameters using the same estimator. First of all, I simulate the data according to the DGP implied by the model, taking $N * R * T$ draws for wage, error in test score measure and income and $N * R$ draws for child's and mother's skills. Following Keane and Moffitt (1998), I re-draw the errors to simulate the income distribution using the parameters estimated in first stage. In each period, the values for mother's labor supply, non-parental child care and maternal time are derived using the optimal solutions implied by the model.[39] Then,

[38]Results do not change estimating all parameters in only one stage. However, the estimation in two-stages is less time consuming.

[39]To test numerically the accuracy of the solutions given by the theoretical model, I also perform a grid search, assuming that the mother's decision to work were actually discrete. In other words, I compute the value of the demands for child care and time with the child, as well as the mother's inter temporal utility, for different levels of mother's labor supply (with the number of hours of work ranging from 0

after having simulated the data for all the periods, I compute the statistics defined in table 1 from the simulated data.

The estimator used in this second-stage minimizes an objective function where each moment condition is the distance between the data statistics and the simulated counterparts, as summarized by table 1:

$$\hat{\theta} = arg\ min\ \hat{g}(\theta)' W \hat{g}(\theta)$$

where

$$\hat{g}(\theta) = \hat{m} - \hat{M}(\theta)$$

$\hat{m}$ is the vector of statistics defined from the actual data, while $\hat{M}(\theta)$ is the vector of simulated statistics according to the model that are functions of the structural parameters to be estimated. W is a positive definite diagonal weighting matrix. According to Cameron and Trivedi (2005, pag. 203), the most efficient minimum distance estimator uses a weighting matrix whose elements are estimates of the inverse of the covariance matrix of the vector $\hat{m}$; this is the so-called optimal minimum distance (OMD) estimator. Since Altonji and Segal (1996) provide evidence of small sample biases in the OMD estimator, I use the diagonally weighted minimum distance estimator proposed by Blundell, Pistaferri, and Preston (2008). Given S number of moments, the weighting matrix is then defined as:

$$W = \begin{pmatrix} \hat{V}[\hat{m}_1]^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \hat{V}[\hat{m}_S]^{-1} \end{pmatrix}$$

where $\hat{V}[\hat{m}]$ is estimated with non-parametric bootstrap and according to the formula (Davidson and MacKinnon 2003, p. 208):

$$\hat{V}\left[\hat{m}\right] = \left[\frac{1}{B}\right] \sum_{b=1}^{B} \left(\hat{m}_b^* - \bar{m}^*\right) \left(\hat{m}_b^* - \bar{m}^*\right)' \quad \text{(B.7)}$$

Non-parametric bootstrap (with replacement) has been implemented according to Wooldridge (2002, p. 379): I used a random number generator to obtain N integers, where $N = 434$ represents the sample size of the actual data, and these integers index the observations drawn from the actual distribution of data. Repeating this process B times,[40] it yields B bootstrap samples on which the statistics defined in table 1 can be computed: $\hat{m}_b^*$ represents a statistic computed for the sample b, while $\bar{m}^*$ is the average of the statistics across the B samples.

B.3. ***Standard errors.*** Non-parametric bootstrap with replacement has been used also to compute the standard errors. After having drawn B_{se} samples from the actual data,[41] I repeat the estimation of the parameters for each sample. This yields an

up to the total time endowment) and I define as optimal choices those that provide the highest utility. The solutions do not differ from the ones provided by the theoretical model.

[40] $B = 200$.

[41] $B_{se} = 50$

empirical distribution of the parameters estimates, from which I can recover a bootstrap estimate of the variance, using the formula (Train 2009, pag. 201):

$$\hat{V}\left[\hat{\theta}\right] = \left[\frac{1}{B}\right]\sum_{b=1}^{B}\left(\hat{\theta}_b^* - \bar{\theta}^*\right)\left(\hat{\theta}_b^* - \bar{\theta}^*\right)' \quad \text{(B.8)}$$

Taking the square root of (B.8) yields the bootstrap estimate of the standard errors $se_{\hat{\theta}}$.

The standard errors for the type proportion parameters π_{mh}, π_{ch} are computed applying the delta method to the non-linear functions (B.4) and (B.5). Defining $g(z_l) = exp(z_l)/(1 + exp(z_l))$ as the function to be approximated, respectively for mothers ($l = m$) and children ($l = c$), the standard errors of the parameters π_{mh} and π_{ch} are given by (Davidson and MacKinnon 2003, chapter 5.6):

$$se_{\hat{\pi}_{lh}} = |g'(\hat{z}_l)|\, se_{\hat{z}_l} \quad \text{(B.9)}$$

where $l = m, c$ and $g'(\hat{z}_l) = \frac{\partial g(\hat{z}_l)}{\partial \hat{z}_l}$.

Appendix C. PSID-CDS data

This appendix provides further details on the data used to estimate the model.

The overall dataset is composed by different supplements of the Panel Study of Income Dynamics (PSID) gathered in the period 1985-2007. Table C.1 summarizes the main information on availability and sources of data. Notice the difference in the availability of information between data taken from the main PSID surveys or related to the external child care information, and the other variables taken from the CDS supplements of the PSID. PSID surveys and the retrospective nature of questions on child care use allow to cover all the periods considered in the model. Instead, the information on maternal time and child's cognitive outcomes are available only at the year of the CDS survey, i.e., 1997, 2002 or 2007.

The merging procedure between PSID and CDS data is done exploiting information on the relationship of each CDS child with respect to the head of the household and the primary caregiver. The final sample is composed by all children aged 0-12 in 1997 without siblings and with both parents living in the household, without missing information on child and parents characteristics and with at least one test score measure. As summarized in table C.2, birth cohorts of children in this sample range from 1984 to 1996, while the terminal period of the model ($T = 13$) corresponds to 1997 for those born in 1984 and to 2009 for those born in 1996.

Table C.3 summarizes the available data for a child born in 1996. This table stresses the existence of long time-gap of missing data due to the structure of the surveys and the timing of the interviews. In fact, while the child care information is available for all periods, data on maternal time and child's cognitive outcomes are available only in the years of the CDS supplement, i.e., 1997, 2002 and 2007.

Table C.4 shows the average characteristics of the sample used for the estimation ($N = 434$) and the total sample of children in CDS, for whom it has been possible to derive information on their parents (3243 observations). This comparison sample includes both families with only one child and families with more children. Mothers in the sample used for the analysis spend less time with their child, work more and use a slightly higher amount of external child care; moreover, they are older and more educated than the mothers in the PSID-CDS data. However, they do not differ in terms of wage at childbirth and race.

Appendix D. Additional descriptive statistics

See table D.1.

Appendix E. Sensitivity analysis

The results presented in section 6 are robust to several sensitivity analyses that will be described below. For the sake of brevity, I only report the results concerning the parameters for maternal time and external child care time in the CAPF.

E.1. ***Terminal period.*** The model has been estimated setting $T = 13$. However, looking at figure 3, it seems that the child's ability measure increases up to age 12 and starting from this point it becomes flatter. This issue is further confirmed by table E.1, showing that the age at which the child's outcome is maximized is around 12. Then, I re-estimate the model setting $T = 12$. This change yields a sample of 368 observations repeated for 12 periods.[42] Figure E.1 presents the results, that are very close to the ones presented in section 6.

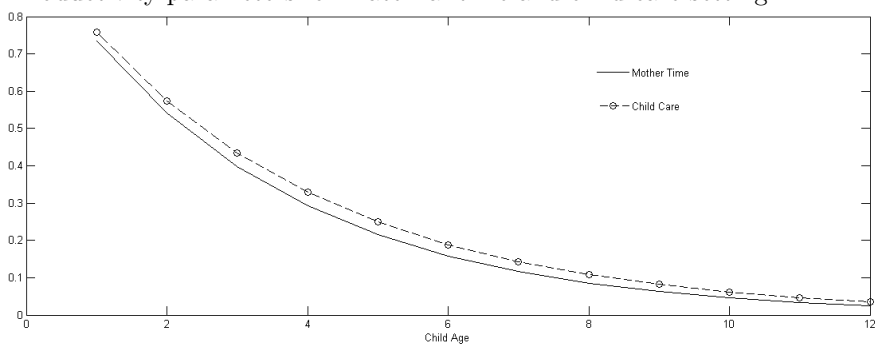
E.2. ***Definition of maternal time.*** The variable weekly time with the mother has been defined considering the time spells in which only the mother was present, either being directly involved in child's activities or being just around and not participating. I test the robustness of this choice on two dimensions. First, the category of time when the mother is not actively involved with the child may include housework activities, that may not represent an investment in child's human capital. If this is the case, the estimated coefficient reported in figure 8 overestimates the true effect of maternal time. I can test for this issue defining the variable maternal time in such a way that only activities when the mother is directly participating are included. Results are reported in figure E.2: the elasticity of child's development with respect to maternal time is lower than the one found in the main analysis. Second, the definition of the variable in the main analysis does not consider as maternal time the time spells when the mother is involved in child's activities but also the father is present. In order to test whether also the latter category represents an input for child development (meaning that the results shown in figure 8 underestimate the true effect), I repeat the estimation of the model defining the variable for maternal time adding also this category. The results shown in figure E.3 do not differ from the ones presented in section 6. The results from these additional analyses suggest that the true effect of maternal time on child's ability

[42]The reduction in sample size is due to observations that have only one test score measure at age 13.

Table C.1
Information on availability and sources of data.

Set of Variables	Source	Survey Years	Additional Info
Non-parental child care	CDS	1997-2002	Retrospective questions on all arrangements used since birth and questions on arrangements used at the time of the survey
Child cognitive outcomes	CDS	1997-2002-2007	Only for children older than 3
Child demographic characteristics	CDS	1997-2002	Time-invariant (except *age*)
Maternal time with the child	CDS-TD	1997-2002	Available only for the year of the survey
Parents' hours of work	PSID	1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1999, 2001, 2003, 2005, 2007	Referred to the year before the survey
Parents' wages	PSID	1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1999, 2001, 2003, 2005, 2007	Referred to the year before the survey
Parents' non labor income	PSID	1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1999, 2001, 2003, 2005, 2007	Referred to the year before the survey
Parents' demographic characteristics	PSID	1997	Time-invariant (except *age*)

Figure E.1
Productivity parameters for maternal time and child care setting $T = 12$.



NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 12$. The final period of the model is $T = 12$.

may range in between the one found in the main text and the one reported in figure E.2.

E.3. ***Correlation between household income and mother's wage offers.*** The specification of the mother's wage offer in (16) and household income in (18) prevents these two components to be correlated. However, it is very likely that mothers in wealthier families receive different wage offers from the market; moreover, a correlation between mother's wage and father's labor income (included in household income) may indicate assortative mating between the two. An easy way to allow these components to be correlated is to include household income as a determinant for the average wage draw the mother receives in each period. Hence, the wage offer described in (16) becomes:

$$ln(w_t) = \mu_0 + \mu_1 educ + \mu_2 age_t + \mu_3 age_t^2 + \mu_4 race + \mu_5 I_t + \epsilon_t \tag{E.1}$$

Figure E.4 presents the estimated values of the coefficients for maternal time and external child care and they do not differ from the ones presented in section 6.

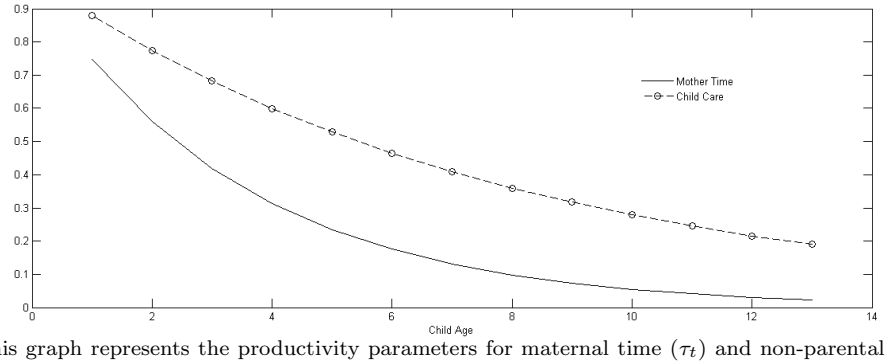
E.4. ***Specification for child's initial endowment.*** As pointed out in section 4.1, the identification of parameters in (20) is hampered by the paucity of test score observations in the available data. This fact is further confirmed by the high standard errors of the estimated parameters, as reported in the panel at the bottom of table 6. Hence, I re-estimate the model using a specification of the child's initial endowment depending on more observable characteristics and on less parameters describing the distribution of child's unobserved skills. Equation (20) becomes:

$$A_1 = exp(\eta_0 + \eta_1 MotherEdu + \eta_2 FatherEdu + \eta_3 BirthWeight) \tag{E.2}$$

where η_0 is a constant and η_3 is an additional parameters to be estimated. Notice that this specification assumes that all children have the same level of unobserved skills. To perform this replication, I keep only children without missing data in birth weight, ending up with 340 observations. Moreover, in order to allow the identification of the new parameter η_3, I add to the existing moment conditions the average child's score conditional on the child's birth weight.[43] Results reported in figure E.5 do not differ from the ones presented in section 6, even though the productivity coefficient for maternal time is slightly lower.

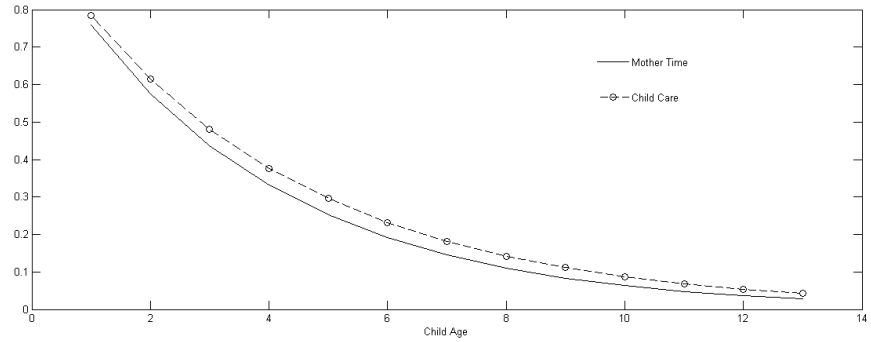
[43]More precisely, I add the following statistics: average score if child's birth weight is higher than 90 ounces; average score if child's birth weight is lower than 90 ounces (90 ounces correspond to almost 2552 grams). Hence, the number of moment conditions turns to be equal to 105.

Figure E.2
Productivity parameters for maternal time and child care if maternal time is only active time with the mother.



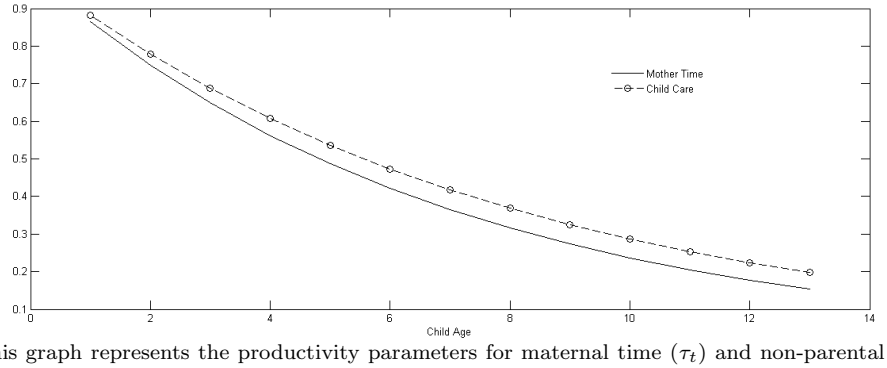
NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. τ includes all activities where the mother is actively participating with the child (active time) and excludes the ones where is present but not engaged (passive time).

Figure E.3
Productivity parameters for maternal time and child care if maternal time includes also time when the father is not involved in child's activities.



NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. τ includes all time spells when the mother is with the child and also those when the mother is present and the father is around but not involved in child's activities.

Figure E.4
Productivity parameters for maternal time and child care allowing household income to affect wages.



NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. The wage offer is defined as:

$$ln(w_t) = \mu_0 + \mu_1 educ + \mu_2 age_t + \mu_3 age_t^2 + \mu_4 race + \mu_5 I_t + \epsilon_t$$

where I_t represents household income.

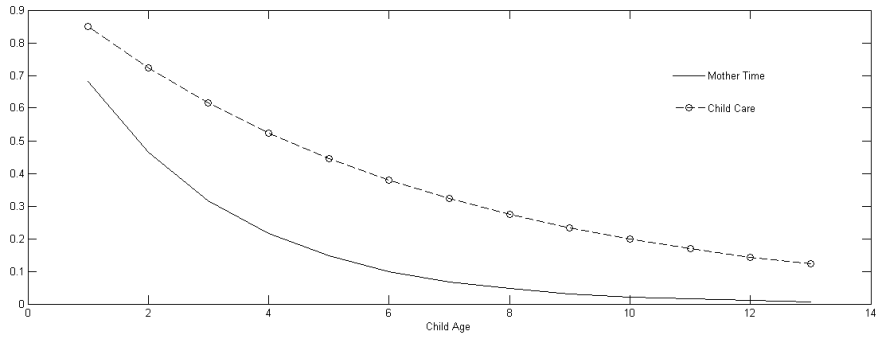
Table C.2
Cohorts of children in the final sample.

Year of Birth	Child's Age					
$t = 0$	$t = 1$	$t = 2$	$t = 3$	$\cdots$	$t = 12 = T - 1$	$t = 13 = T$
1984	1985	1986	1987	$\cdots$	1996	1997
1985	1986	1987	1988	$\cdots$	1997	1998
1986	1987	1988	1989	$\cdots$	1998	1999
1987	1988	1989	1990	$\cdots$	1999	2000
1988	1989	1990	1991	$\cdots$	2000	2001
1989	1990	1991	1992	$\cdots$	2001	2002
1990	1991	1992	1993	$\cdots$	2002	2003
1991	1992	1993	1994	$\cdots$	2003	2004
1992	1993	1994	1995	$\cdots$	2004	2005
1993	1994	1995	1996	$\cdots$	2005	2006
1994	1995	1996	1997	$\cdots$	2006	2007
1995	1996	1997	1998	$\cdots$	2007	2008
1996	1997	1998	1999	$\cdots$	2008	2009

Table C.3
Available data for a child born in 1996.

	Child's age (t)													Source	Survey Year
	1	2	3	4	5	6	7	8	9	10	11	12	13		
Non-parental child care	X	X	X	X	X	X	X	X	X	X	X	X	X	CDS	1997, 2002
Child cognitive outcomes						X					X			CDS	2002, 2007
Child demographic charact.	X					X					X			CDS	1997, 2002, 2007
Maternal time with the child	X					X								TD	1997, 2002
Parents' hours of work		X		X		X		X		X				PSID	1999, 2001, 2003, 2005, 2007
Parents' wages		X		X		X		X		X				PSID	1999, 2001, 2003, 2005, 2007
Parents' non labor income		X		X		X		X		X				PSID	1999, 2001, 2003, 2005, 2007
Parents' demographic charact.	X		X		X		X		X					PSID	1997, 1999, 2001, 2003, 2005, 2007

Figure E.5
Productivity parameters for maternal time and child care with a different specification for child's initial endowment.



NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. The child's initial endowment is defined as:

$$A_1 = exp(\eta_0 + \eta_1 MotherEdu + \eta_2 FatherEdu + \eta_3 BirthWeight)$$

where η_0 is a constant and η_3 is an additional parameter to be estimated. The estimation of this version of the model has been done adding the following statistics to the moment conditions: average score if birth weight is higher than 90 ounces; average score if birth weight is lower than 90 ounces.

Table C.4
Mean characteristics of the sample with respect to PSID-CDS data.

	PSID-CDS	Sample	T-test
Mother's hours of work	23.59	27.12	−10.69***
	(0.14)	(0.29)	
Non-parental child care	12.21	14.68	−9.34***
	(0.10)	(0.24)	
Maternal time with the child	25.83	21.06	6.34***
	(0.32)	(0.68)	
Mother's wage before childbirth[a b]	10.98	11.24	−1.25
	(0.09)	(0.18)	
Mother's education	12.98	13.25	−7.23***
	(0.02)	(0.03)	
Mother's age at child's birth	26.98	28.17	−15.77***
	(0.04)	(0.07)	
Mother's race: white	0.61	0.61	0.35
	(0.00)	(0.01)	
Father's hours of work	38.66	45.22	−30.22***
	(0.11)	(0.19)	
Father's education	12.66	13.27	−16.98***
	(0.01)	(0.03)	
Household non labor income[a]	16.86	12.79	2.50**
	(1.39)	(0.84)	
N	3243	434	

[a] Monetary variables deflated into 1997 US$.
[b] Mother's wage before childbirth refers to the year before the child was born.

Table D.1
Descriptive statistics on LW raw test scores, by child's age and by subgroups.

Child's age	4-5	6-10	11-13
All Sample ($N = 434$)	6.70	30.05	44.84
	(4.13)	(11.68)	(6.23)
Female	6.88	30.69	45.23
	(3.97)	(11.18)	(5.65)
Male	6.53	29.40	44.47
	(4.31)	(12.18)	(6.76)
Non-White	7.26	30.16	43.10
	(4.10)	(11.19)	(6.42)
White	6.32	29.98	45.93
	(4.15)	(12.05)	(5.86)
Mother Education:	7.24	31.86	43.39
years of schooling> 12	(4.44)	(11.62)	(5.60)
Mother Education:	5.62	27.71	43.24
years of schooling= 12	(3.44)	(11.19)	(6.52)

Table E.1
OLS estimates of LW raw score on child's age.

Dep. Var. LW raw score		
Child's Age	18.195***	18.163***
	(1.883)	(1.829)
Child's Age Squared	−0.764***	−0.762***
	(0.097)	(0.095)
Controls		
Maternal Time	X	X
Child Care Time	X	X
Household Income	X	X
Mother Education		X
Father Education		X
N	195	195
Child's age maximizing the outcome		
	11.90	11.92

NOTE. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered at child level. The age maximizing outcome (last row) is computed setting $\frac{\partial LW score}{\partial t} = 0$.