MOTHER'S TIME ALLOCATION, CHILD CARE AND CHILD COGNITIVE DEVELOPMENT

Ylenia Brilli*†

April 2014

Preliminary - Do not quote without permission

Abstract. This paper analyzes the effects of maternal employment and non-parental child care on child cognitive development, taking into account the mother's time allocation between leisure and child care time. I estimate a behavioral model, where maternal labor supply, non-parental child care and time allocation decisions are the inputs in the child development process, using data from the Child Development Supplement and the Time Diary Section of PSID. I find that the elasticity of child's ability with respect to maternal time is higher than the one with respect to non-parental child care. Policies simulation suggests that only a policy improving non-parental child care productivity has positive effects on both child development and maternal labor supply.

JEL classification: D13, J13, J22.

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

*European University Institute and CHILD-Collegio Carlo Alberto. E-mail: ylenia.brilli@eui.eu

†I am indebted to Daniela Del Boca for her constant encouragement and advice. I am also grateful to Christopher Flinn, Cristian Bartolucci, Matthew Wiswall, Andrea Ichino, Marco Francesconi, Chiara Monfardini, Silvia Pasqua, Eva Mörk, James Heckman, Stefano Mosso, Frank Vella, Jérôme Adda, Marco Tonello, Brandon Restrepo, seminar participants at University of Turin, Collegio Carlo Alberto, EUI and participants at the XI Brucchi Luchino Labor Economics Workshop (Trento), the Workshop on Economics of Successful Children (Aarhus University), the XIII MOOD Workshop (EIEF, Rome), the II SSSI (University of Chicago) and the XXVIII EEA meeting (Gothenburg) for their useful comments and suggestions. This research was initially conducted at Collegio Carlo Alberto and when I was a visiting student at New York University: the hospitality of both institutions is gratefully acknowledged. I also acknowledge financial support from Collegio Carlo Alberto (project "Parental and Public Investments and Child Outcomes") and from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 320116 for the research project FamiliesAndSocieties. The paper previously circulated with the title "Mother or market care? A structural estimation of child care impacts on child development". Any errors 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 development is the time the child spends with the mother (Cunha et al. 2006). Maternal employment may have negative consequences on children's development, since it decreases maternal child care time (Cox et al. 1992); however, it may also matter as long as it boosts non-parental child care use (Almond and Currie 2011). Indeed, the increase in the maternal employment rate and the use of external forms of child care have raised concerns about the impacts that they may have on child development. The participation of mothers in the labor market has increased from around 40 percent in the 1970s to more than 60 percent at the end of the 1990s (U.S. Census Bureau 2000); this dramatic change has also been accompanied by a large increase in the number of children cared for by people outside of the family: in the same period, the percentage of 3-5 children enrolled in some forms of pre-primary educational programs increased from 7.9 to 51.7 percent for mothers in the labor force (Bianchi 2000). However, data from the American Time Use Survey shows that the amount of child care time does not vary crucially among employment status, while employed mothers spend a significantly lower amount of time in activities like socialization or doing sport, usually defined as *leisure* (U.S. Census Bureau 2013). Although there being several works assessing the effects of maternal employment on children's cognitive development, none has simultaneously taken into account maternal time allocation between leisure and child care time, as well as non-parental child care use.

This paper analyzes the effects of maternal employment and child care arrangements on children's cognitive development, distinguishing between maternal care and care provided by market services, and taking into account the additional choice that the mother makes concerning the trade-off between leisure and child care time. To understand the effects of maternal employment and non-parental child care use on children's development, I estimate a dynamic model of mothers' time allocation and labor supply, incorporating the child's human capital accumulation process. The model suggests that, after having controlled for maternal choices on labor supply and non-parental child care use, mothers may still differ in their time allocation between leisure time to spend alone and time to spend with the child, which can be productive for their child's cognitive development.

The contribution of this paper to the literature is twofold. First, I estimate a model incorporating three endogenous choices of mothers' time allocation, namely maternal labor supply, maternal child care time and non-parental child care use. This implies avoiding any assumptions about the relationship between maternal labor supply and maternal child care time. Existing studies assessing the effects of maternal employment and non-parental child care use (Bernal 2008; Bernal and Keane 2010) implicitly assume that mothers' time out of work is entirely spent by the mother with the child, being then productive for the child's cognitive development. However, employed mothers might allocate their time out of work in such a way as to give priority to the time spent with the child (Bianchi 2000; Hoffert and Sandberg 2001), so that there might not be a one-to-one relationship between time spent at work and child care time.

Second, this paper represents the first attempt to estimate the elasticity of a child's ability with respect to both maternal time and non-parental child care time in a cognitive ability production function framework. To the best of my knowledge, there are not studies that simultaneously evaluate the productivity of both inputs, taking into account the selection of mothers into work and child care use.

The proposed 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 time to use non-parental child care, 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 by 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 her observable characteristics, while mother's unobserved skills affect her taste for child's ability, mother's participation in 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 the child's skills and the parents' educational attainments.

The model is estimated using U.S. data from the Panel Study of Income Dynamics (PSID), linked to data from the Child Development Supplement (CDS) and the Time Diary (TD) Section. The CDS provides retrospective information on all child care arrangements used since birth and widely-recognized measures of child cognitive outcomes; the Time Diary (TD) component 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 retrieved using a Method of Simulated Moments estimator, minimizing the distance between several data statistics and their model counterparts.

The results show that more skilled 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 the elasticity of child's ability with respect to maternal time is higher (at any child's age) than that relating to non-parental child care. Thus, mother's time is always more productive than non-parental child care.

The estimated model is used to simulate the effects of several policies. Policies subsidizing non-parental child care have very little effect on child's development. Interestingly, only a policy increasing non-parental child care productivity to the level of mother's time productivity leads to an increase in both child's ability and mother's labor supply: this suggests that the quality of the service perceived by the mother plays a significant role not only for convincing the mother to use non-parental child care but also for boosting mother's participation in the labor market.

The rest of the paper is organized as follows. Section 2 presents the related literature and some stylized facts in non-parental child care use and maternal time allocation. Section 3 presents the model that is estimated, while Section 4 presents the econometric specification of the model and the empirical method used for the identification of parameters. 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 . Section 7 presents the results from the policy simulations and Section 8 concludes.

2. Background

The increase in female employment rate that has characterized all developed countries has raised concerns about the impact that maternal employment and non-parental child care may have on child development. This is one of the reasons why, in the last few decades, many studies have attempted to assess the effects of these choices.

Since the work of Becker and Tomes (1986), who first provided a framework for modeling 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's human capital and development. 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). Bernal and Keane (2011) report that one year of child care use decreases children's cognitive outcomes by 2.13 percent. Instead, in their evaluation of the early childhood program Head Start, Currie and Thomas (1995, 1999) find that children who attended the program get higher scores on 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. Recent studies focusing on Northern European countries (Datta Gupta and Simonsen 2011; Havnes and Mogstad 2011) show that the use of a high-quality formal child care service has beneficial effects on children's development.

The identification of the impacts of the effects of both maternal employment and non-parental child care on child development is hampered by several sources of endogeneity, mainly induced by the correlation of mothers' choices with mothers' and children's unobservables characteristics; moreover, the simultaneity of these choices makes it more difficult to take this selection into account. 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) and James-Burdumy (2005) 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 a mother's choices and a child's ability. While the first strategy is robust to the correlation of 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 makes it possible 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 the 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 (2014), who model household choices and investments in child quality from childbirth up to adolescence. They find that the productivity of mother's time investments declines over child's age, and that father's time becomes more productive as the child reaches middle childhood and adolescence. Differently from Del Boca et al. (2014), this paper does not model both parents' labor supply and time allocation decisions, focusing only on mothers' behavior and on the additional choice of using external forms of care; in other words, instead of considering fathers' time as a substitute for mothers' time with the child, the present study analyzes the role of external child care time as a substitute for maternal child care.

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 a 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) have evaluated the effects of maternal employment on a child's schooling, estimating the parameters of a conditional demand function for child's education; they find that one year more of a mother's full time employment reduces the probability that the child gets higher education by 11 percentage points.

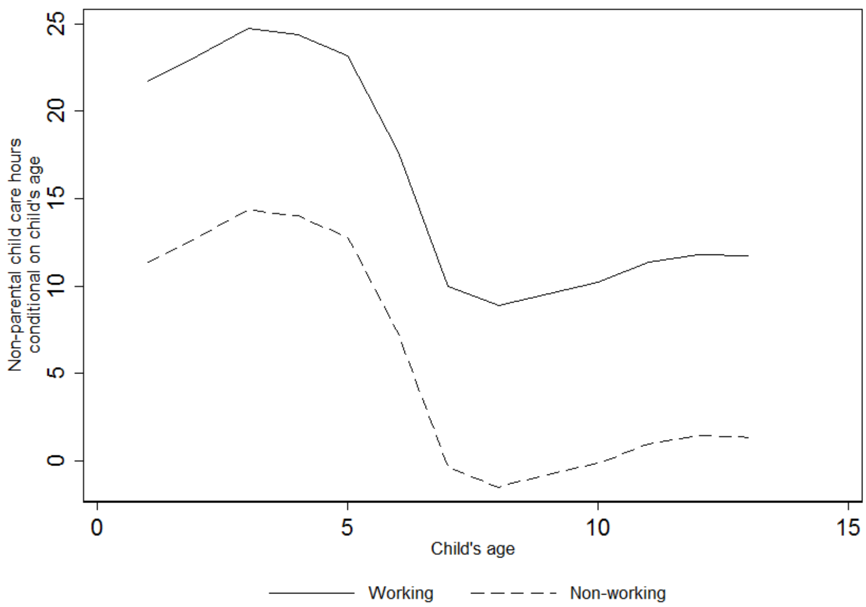
Bernal (2008) is the only study that evaluates the impact of both maternal employment and non-parental child care attendance on subsequent child outcomes using a structural approach. She finds that one year in external child care reduces the child's cognitive ability by 0.8 percent; however, the impact of maternal employment and external child care is more detrimental, since, together, they decrease a child's outcome by 1.8 percent.

Differently from all other studies on maternal employment and child development, this paper exploits the actual measure of maternal time to assess the effects of maternal employment and non-parental child care on children's development; this implies that the three choices on labor supply, maternal child care time and non-parental child care are treated as endogenous.

Notice 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 choose how to allocate their time between leisure and child care time. However, the actual investment made by the mother on the child through her contact time may differ according to how the mother allocates her time between leisure and child care time.

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. For instance, Leibowitz (1974, 1977) points out that more skilled 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, non-working mothers have also 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 years-old children enrolled in some forms of pre-primary educational programs increased from 4.8 to 44 percent for mothers not in the labor force. On the other hand, 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 (ATUS) 2005-2009, the amount of time spent

FIGURE 1
Non-parental child care time by mothers' employment status.



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

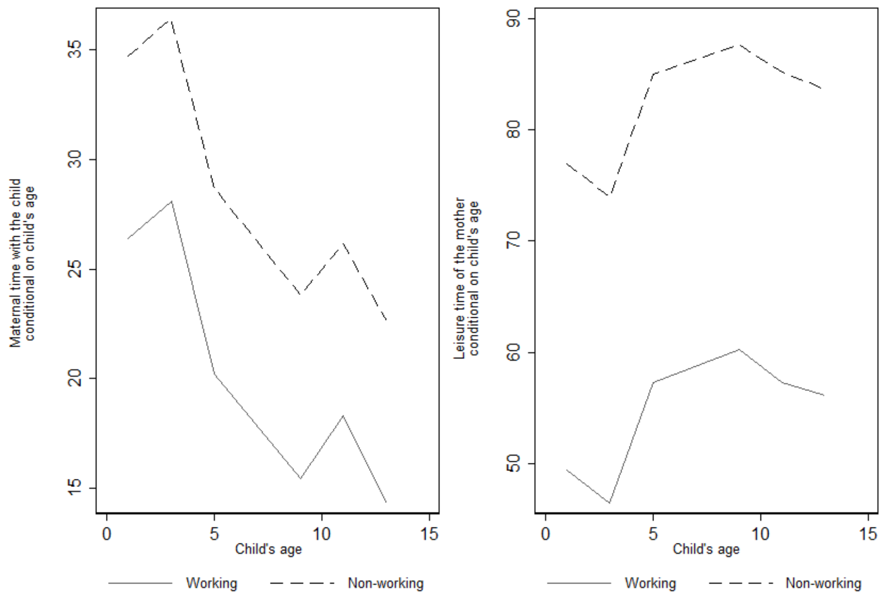
$$childcare_{it} = \eta_0 + \sum_{t=1}^{T} \eta_{1t} t_{it} + \eta_2 d_{it} + \epsilon_{it}$$

where $childcare_{it}$ represents (weekly) hours of non-parental child care in each year t, t_{it} are child's age fixed effects (with $t = 1, \ldots, 12$), d_{it} is a dummy variable equal to 1 if the mother of child i works in period t. $\eta_2 = 10.36$ 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 ($N = 3487$).

by mothers reading and playing with the child does not vary crucially across employment status (U.S. Census Bureau 2013). Guryan, Hurst, and Kearney (2008), exploiting ATUS data for 2003-2006, find that there is a striking positive education and income gradient in child care, while the gradient for leisure is negative: this means that more educated and wealthier mothers spend more time with their child even if working, preferring to renounce some leisure time.

Descriptive evidence from the data used in this paper supports the existence of these patterns. Figure 1 shows that non-working mothers also use a positive amount of non-parental 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 hours per week, the graph also confirms that child care is needed for its custodial purposes anytime the mother is working.

Figure 2
Maternal child care time and leisure by mothers' employment status.



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

$$\tau_{it} = \eta_0 + \sum_{t=1}^{T} \eta_{1t} t_{it} + \eta_2 d_{it} + \epsilon_{it}$$

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

$$l_{it} = \beta_0 + \sum_{t=1}^{T} \beta_{1t} t_{it} + \beta_2 d_{it} + \varepsilon_{it}$$

τ_{it} stems for (weekly) maternal time with the child and l_{it} 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_{it} are child's age fixed effects (with $t = 1, \ldots, 12$) and d_{it} is a dummy variable equal to 1 if the mother of child i works in period t. $\eta_2 = -8.32$ represents the difference in average maternal time (conditional on child's age) between working and non-working mothers. $\beta_2 = -27.49$ 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 ($N = 380$).

Figure 2 plots the fitted values of two regressions where the dependent variables are, respectively, maternal child care time and leisure time. The graph on the left (i.e., maternal child care time) 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 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 lower 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.

These patterns suggest that working and non-working mothers allocate their time out of work differently and that the choice of devoting time to the child instead of having leisure should be considered endogenous as the ones of labor supply and non-parental child care use.

3. The model

This section describes the theoretical model on which the estimation is based. Subsection 3.1 presents the basic structure, 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 and budget and time constraints. The functional form assumptions are based on the theoretical model developed in Del Boca et al. (2014).

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 non-parental child care to use. The choice variables are then: (i) h_t, representing hours of work; (ii) i_t, hours of non-parental 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 at each child's age t until the child reaches T years of age.[1]

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.[2] Finally, the simplification concerning the number of children allows

[1] $t = 1$ 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. $t = T = 13$ represents the terminal period of the model. It may be interpreted as the final period of middle childhood before the child enters adolescence.

[2] 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. 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.

to avoid modeling the fertility decisions of parents and to make additional assumptions on the different effects of investments on more siblings.[3]

The Mother's Utility Function

The 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.

The time constraint is defined as:

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

where TT is the mother's total time endowment.[4] 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.

[3]The sample selection may have an influence on the estimated parameters; this issue will be further discussed in Section 5.

[4]$TT = 112$ hours per week: it assumes 16 hours per day, excluding sleeping time, 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 the child's future ability, equation (4) shows that inputs operate with a lag. 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.

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).

Mothers' work is not explicitly included in the CAPF, because it may not have a direct impact on child development *per se*. Mothers' employment may indirectly affect child development through a change in mothers' 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, $\delta_{1t} \geq \delta_{2t}$.[5]

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 buy goods

[5]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 $\delta_{1t} \leq \delta_{2t}$.

effective for child development and (ii) this proportion is not affected by the mother's labor supply decisions.[6]

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 services with a potentially zero price in the market are also 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).

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.[7]

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

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

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

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

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

[6] The model implies that the additional labor income the mother gets from her labor supply is spent in non-parental child care.

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

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 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 noticed 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 absence from the labor market after childbirth. Although being a quite strong assumption for modeling mothers' labor supply, the exogenity of wage is necessary to estimate the model with continuous choice variables ad closed-form solutions. Notice that the use of continuous choice variables is needed to allow for three choices and to take into account the additional choice between leisure and time with the child. 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 of leaving the labor market after childbirth on wages, it is very likely to overestimate the proportion of mothers working and their labor supply on the extensive margin; this may lead to an overestimation of the amount of non-parental child care used and to an underestimation of maternal time with the child.

3.2. ***Terminal period value function and solution of the model.*** 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 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 the mother's subsequent 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 the child's future development during adolescence, from $T+1$ on.

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

$$V_{T+1} = \tilde{V}_{T+1} + \sum_{\kappa=0}^{+\infty} \beta^{\kappa} \alpha_3 lnA_{T+1} \tag{6}$$

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.[8] Equation (6) represents the terminal period value function 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.

[8]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.

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

$$i_t^c = \frac{\beta\delta_{2t}D_{t+1}}{p(\alpha_2 + \beta\delta_{2t}D_{t+1})}(w_t h_t + I_t) \tag{7}$$

$$\tau_t^c = \frac{\beta\delta_{1t}D_{t+1}}{(\alpha_1 + \beta\delta_{1t}D_{t+1})}(TT - h_t) \tag{8}$$

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 $T+1$ to period 1 is given by:[9]

$$\begin{aligned}
D_{T+1} &= \rho\alpha_3 \\
D_T &= \alpha_3 + \beta\delta_{4T}D_{T+1} \\
D_{T-1} &= \alpha_3 + \beta\delta_{4T-1}D_T \\
&\vdots \\
D_t &= \alpha_3 + \beta\delta_{4t}D_{t+1} \\
&\vdots \\
D_2 &= \alpha_3 + \beta\delta_{42}D_3 \\
D_1 &= \alpha_3 + \beta\delta_{41}D_2
\end{aligned} \tag{9}$$

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.[10] 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:

$$h_t^c = \frac{\alpha_2(TT - \tau_t^c)}{\alpha_1 + \alpha_2} - \frac{\alpha_1(I_t - p i_t^c)}{w_t(\alpha_1 + \alpha_2)} \tag{10}$$

[9]The same expressions can be derived computing $D_{t+1} = \frac{\partial V_{t+1}}{\partial A_{t+1}}$ instead of $D_{t+1} = \frac{\partial V_{t+1}}{\partial lnA_{t+1}}$ (See Appendix A, footnote 22). 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.

[10]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. 2014). This means that the model always predicts a positive amount of external child care, regardless of mother's working status or household income.

Substituting (7) and (8) in equation (10), the latent labor supply becomes:

$$h_t^* = \frac{TT(\alpha_2 + \beta\delta_{2t}D_{t+1})}{(\alpha_1 + \beta\delta_{1t}D_{t+1} + \alpha_2 + \beta\delta_{2t}D_{t+1})} - \frac{I_t(\alpha_1 + \beta\delta_{1t}D_{t+1})}{w_t(\alpha_1 + \beta\delta_{1t}D_{t+1} + \alpha_2 + \beta\delta_{2t}D_{t+1})} \tag{11}$$

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 (11), 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. Substituting (11) into (7) and (8) yields the unconditional demands for child care and time with the child.

Notice that mother's decision to work 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. Equation (7) shows 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.

4. Econometric strategy

In the following subsection, I present the empirical specification used to take the model to the data. Subsection 4.2 describes the econometric method used to estimate the model parameters and discusses identification issues. 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 $\gamma_1 = 0$, $\Gamma_2 = (\gamma_{21}, \gamma_{22})$ and $\Gamma_3 = (\gamma_{31}, \gamma_{32}, \gamma_{33})$ are vector of parameters representing the contribution of observable and unobservable characteristics to the corresponding preference parameter.[11] The functional forms for f_1, f_2, f_3 are specified in Appendix B.1. μ_0 represents mother's skills, whose distribution will be defined below.

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:

$$ln(w_t) = \mu_t + \epsilon_t \tag{12}$$

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:

$$\mu_t = \mu_0 + \mu_1 MotherEdu + \mu_2 MotherAge_t + \mu_3 MotherAge_t^2 + \mu_4 MotherRace \tag{13}$$

where μ_0 represents mother's skills. Equation (13) 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.

[11]A mother's education is expressed as years of education and ranges from 2 to 17; race is a dummy variable equal to 1 if the mother is white.

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:

$$I_t \overset{\text{iid}}{\sim} N(\mu_{inc}, \sigma^2_{inc}) \tag{14}$$

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

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:

$$\delta_i = exp(\xi_i t) \tag{15}$$

where $i = 1, 2, 3, 4$ and t represents the age of the child.[12]

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 a function of children's unobserved skills, parents' education and child's birth weight. Specifically:

$$A_1 = exp(\psi_{ck} + \eta_1 MotherEdu + \eta_2 FatherEdu + \eta_3 BirthWeight) \tag{16}$$

where ψ_{ck} represents child's skills. The child's skills, as well as mother's skills follow a discrete distribution (Heckman and Singer 1984), each with two points of support. So, for instance, child's skills are distributed as:

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

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.

Finally, the inclusion of mother's and father's education allows to capture a non-zero correlation between these observable characteristics and child's skills. It should be stressed

[12]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 may have lower influence.

that the coefficients in (16) capture merely correlations between observables and the child's cognitive ability endowment. As suggested by Bernal and Keane (2010), using as many observables as possible in the definition of (16) should also reduce the sensitivity of the results to the distributional assumptions on the unobserved heterogeneity term.

Recalling the value-added specification of the CAPF, defined in (4), the estimation 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 child's true cognitive ability is related to the measure of that given by the test scores. Existing studies using a structural approach 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 for the child's true 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. (2014) 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:

$$\pi_{score} = \frac{exp(A_t + v_t)}{1 + exp(A_t + v_t)} \tag{17}$$

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 (17) ensures this represents a value between zero and one. The test score measure is then defined as follows:

$$S_t = \pi_{score} * J_t \tag{18}$$

where J_t is the maximum number of items answered correctly at each child's age.[13]

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

[13]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 $t = 9, 10, 11, 12, 13$ $J = 57$.

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, mother's and child's unobservable characteristics are correlated and both enter the initial level of child's ability at birth, that the mother observes before making any decision.

4.2. ***Estimation method and identification.*** The model parameters are estimated using a Simulated Method of Moments estimator that minimizes the distance between several data statistics and their model counterparts. The data generating process implied by the model described in Sections 3 and 4.1 allows to simulate the same statistics for the individuals (mothers and children) in the sample over the child's life cycle. The full list of statistics used to construct the moment functions is reported in Table 1.

In order to recover the basic trends and levels of the real data, the statistics used to identify the model include the means of the mother's choice variables (hours of work, external child care time, maternal time with the child) and the outcome variable (test scores), both unconditional and conditional on child's and mother's age. They also include the average and standard deviation of wage and household income.

However, focusing on means is not sufficient to identifying all the parameters in the model (Adda, Dustmann, and Stievens 2012). More precisely, the estimation should shed light on the several trade-offs and mechanisms implied by the model: for instance, the trade-off that the mother faces between consumption and child's ability, implying that more hours of work determine higher consumption but also lower amount of time with the child, or the trade-off between mother's working opportunities and mother's preferences for child's ability, both functions of mother's unobserved skills.

To help identification of parameters in the mother's utility function, I use as moments the coefficients from OLS regressions of mother's education and race on mother's choices (hours of work, time with the child and non-parental child care). For the identification of parameters describing mother's unobserved heterogeneity, I use the variance of residuals from a regression of log wages on mother's age; moreover, the proportion of mothers with residuals lower than 1 standard deviation of wage can help in identifying the proportion of mothers in each skills level. The coefficients from a OLS regression of non-parental child

Table 1

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

Mother's choices
mean and std deviation mothers' hours of work, non-parental child care and mother's time with the child proportion of mothers working more than 20 hours per week
mean mothers' hours of work by child's age and conditional on mother's age mean non-parental child care time by child's age and conditional on mother's age mean mother's time with the child by child's age and conditional on mother's age
corr mother's wage and mother's hours of work corr mother's wage and non-parental child care time corr mother's wage and time with the child corr mother's hours of work and household income corr mother's hours of work and time with the child corr mother's hours of work and non-parental child care corr household income and mother's time with the child corr household income and non-parental child care corr non-parental child care and mother's time with the child
Test scores
mean test scores by child's age
Productivity parameters
corr mother's time with the child in 1997 (2002) and score in 2002 (2007) corr non-parental child care in t and score in $t+1$ corr mother's hours of work in t and score in $t+1$ corr household income in t and score in $t+1$
Outcomes transition probabilities (for children with 2 scores measures)
prop of children with score in range p_{97} in 1997 and p_{02} in 2002 prop of children with score in range p_{02} in 2002 and p_{07} in 2007
Wage equation and household income
mean and std deviation of mother's wage mean mother's wage conditional on mother's age regression of mother's log wage on education, age, age squared, race (coefficients) mean mother's wage by mother's education mean mother's wage by race mean mother's wage by mother's age mean, std deviation and median of household income
Mother's and child's unobservables
std deviation of the residuals from a mother's wage reg on age and hours of work prop of mothers with residuals lower than 1 std deviation wage
OLS regression of non-parental child care on wage residuals (coefficients) OLS regression of mother's time with the child on wage residuals (coefficients)
std deviation of residuals from a scores reg on age of the child prop children with residual lower 1 std deviation score
Preference parameters
OLS regression of mother's hours of work on education, race (coefficients) OLS regression of maternal time with the child on education, race (coefficients) OLS regression of non-parental child care on education, race (coefficients)
Child's initial ability
OLS regression of test scores on parents' education and child's birth weight (coefficients)

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. Mother's 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.

care time and maternal time with the child on mother's wage residuals and child's age may capture the contribution of mother's skills to mother's preference for child's ability.

The identification of parameters in the wage equation is reached through the coefficients of a linear regression of log wage on mother's education, age, age squared and race. Moreover, I use as moments the average mother's wage by mother's educational level, age and race. The parameters in the income process are identified using the average, standard deviation and median level of income.

The productivity parameters in the CAPF can be identified using the correlation between inputs (maternal time with the child, non-parental child care, hours of work and household income) and child's test scores over time. Moreover, the parameter δ_4 representing the productivity of child's ability in the previous period is recovered using transition probabilities from the first score measure available in the data (in 1997 or 2002) to the second score measure available in 2002 or 2007.

Finally, to identify the parameters in the initial level of child's ability, I use OLS regression coefficients of mother's and father's education and of child's birth weight on child's test scores, also controlling for child's age. The child's initial ability is also a function of children's unobserved skills. As described in Section 4.1, children's unobserved heterogeneity is assumed to have a mass point distribution. To identify the proportion of children in each type, I regress the test scores on child's age, compute the average residual for each child and then use the cross-Sectional variance of this average, which can be linked to the variance of ability in the sample, as a moment; as an additional moment, I use the proportion of children with this average residual lower than 1 standard deviation of the score.

The simulation of the data is obtained by taking $N * R$ 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 the error in the test score measure.[14] 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.

[14] $N = 430$ 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 decided 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 Simulated Method of Moments estimator is then:

$$\hat{\theta} = arg\ min\ \hat{g}(\theta)' W \hat{g}(\theta) \tag{19}$$

where

$$\hat{g}(\theta) = \hat{m} - \hat{M}(\theta) \tag{20}$$

$\hat{m}$ 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 provides further details on the estimation.

The estimation is done using the simplex algorithm, which is more robust to non-smooth objective function. Identification of the model parameters requires a unique solution for the minimization of the objective function defined by (19). In practice, identification depends on the uniqueness of the minimum and on the curvature around it. Due to the structure of available data, I expect that the identification of all parameters could be achieved, even if identification for parameters in the initial level of ability could be more tenuous since I can observe test scores starting only from age 4. To test for this, I estimate the model using different starting values and results do not differ from the ones presented in Section 6. 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 estimated parameters, with the only exception of the parameters in the initial level of ability, confirming that identification of these parameters is more problematic.

5. Data

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

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 was 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 under 18 years of age.

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 information to the primary caregiver on all non-parental child care arrangements used for the child since childbirth; a set of follow-up questions was 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.[15]

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 out by the child or by the child's primary caregiver about the child's activities over a specified 24-hour period.[16] Each participating child completed two time diaries: one for a weekday (Monday-Friday) and one for a weekend day (Sunday or Saturday). The TD additionally collects information on the social context of the 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

[15] 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.

[16] 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).

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.[17]

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.[18] The family information I gather includes each parent's hours of work, wage and non labor income in each period.[19]

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 430 observations.[20]

Before moving to the descriptive statistics, the importance of the sample selection should be stressed and it should be considered 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 contribute to the literature of child care and child development, because they are structurally obtained from a model based on economic theory, but the sample selection bias may come in different

[17]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 D.2.

[18]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.

[19]Between 1985 and 1997 PSID interviews were conducted annually but, 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.

[20]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 no information on parents' age, education and race, 85 have no test score measures in the period 1997-2007 and 4 have no information on weight at birth.

ways. The sample selection implies that all mothers' investments in child's ability are unrelated with the decision to marry or to cohabit and with fertility. However, if mothers in intact households have more marriage-oriented attitudes 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 time in their child. This may lead to overestimate the proportion of mothers not working or to overestimate mothers' preference for child's ability. Similarly, mothers with only one child may have higher preferences for child's ability and this may lead to overestimate mother's investments: this channel may imply either an overestimation of maternal time with the child or an overestimation of non-parental child care use. 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. Moreover, the fact of having only one child means that the mother has experienced only one work interruption due to childbirth, leading again to an overestimation of mother's attachment to the labor market. Even though it is difficult to derive a unique direction of the bias induced by sample selection, the arguments provided above suggest that it may oversample mothers who are more productive either in the labor market or at home with the child. This may turn to provide an upper bound of the proportion of mothers in the labor force or of the productivity of mother's time investments.[21]

5.1. ***Descriptive Statistics.*** 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. According to Table 2, mothers work, on average, 27 hours per week and use non-parental child care for almost 14 hours; moreover, they spend with their child, on average, 3 hours per day. Mother's wages are on average 14.25 US$, while household income represents, on average, around 800 US$ per week.

Table 3 provides some descriptive statistics on mother's work, child care and maternal time, by child's age. There are not large differences in mother's participation to the labor

[21]Table C.4 in Appendix C compares the characteristics of the subsample used for the analysis ($N = 430$) with the ones of the entire PSID-CDS sample ($N = 3243$).

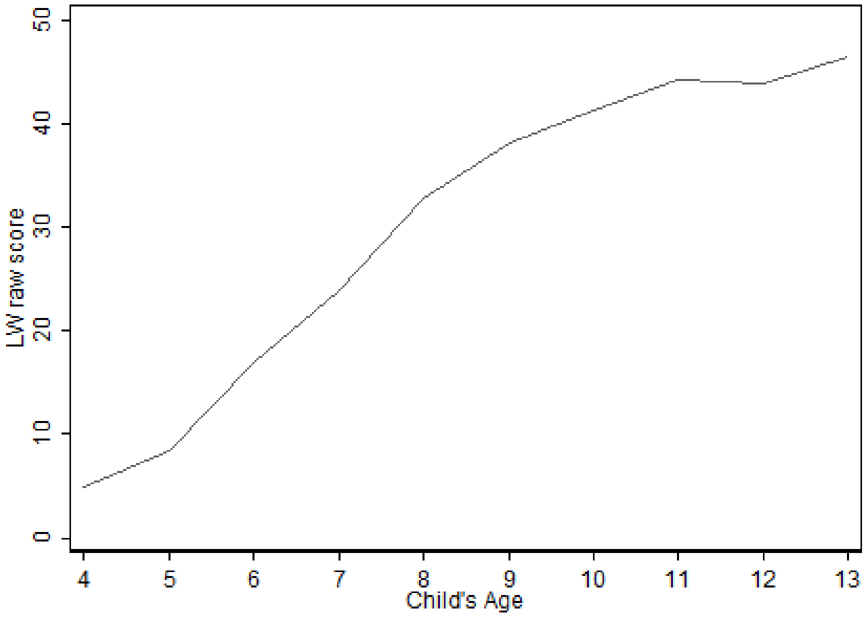
Table 2
Descriptive statistics on all variables for the entire period.

	Mean	SD	Min	Max
Child's LW raw score	35.08	14.50	1	57
Mother's hours of work	27.18	17.53	0	100
Proportion of working mothers	0.68	0.46	0	1
Non-parental child care hours	14.71	18.32	0	70
Mother's time with child	21.16	17.01	0.17	95.75
Child's gender: male	0.51	0.49	0	1
Child's birth weight	119.18	21.57	32	244
Mother's wage	14.26	10.19	5.01	133.93
Mother's age at child's birth	28.19	5.10	16	43
Mother's education	13.27	2.49	2	17
Mother's race: white	0.61	0.49	0	1
Father's education	13.30	2.47	1	17
Household income/10	79.25	66.59	0.01	1012.62

NOTE. Monetary variables deflated into 1997 US$. Child's birth weight is expressed in ounces (88 ounces = 2500 grams). The proportion of working mothers indicates the proportion of mothers working more than 20 hours per week. Household income includes father's labor income and household non labor income. Source: own elaboration from PSID-CDS data.

market across child's age. The number of hours worked by the mother ranges from 24 when the child is very young, to 29 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 10 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 3.4 hours per day, considering a working week of 5 days). When the child starts going to school, he spends out of home not only the time in external care but also a fixed amount of school time. If the child spends 6 hours per day at school, 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.

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

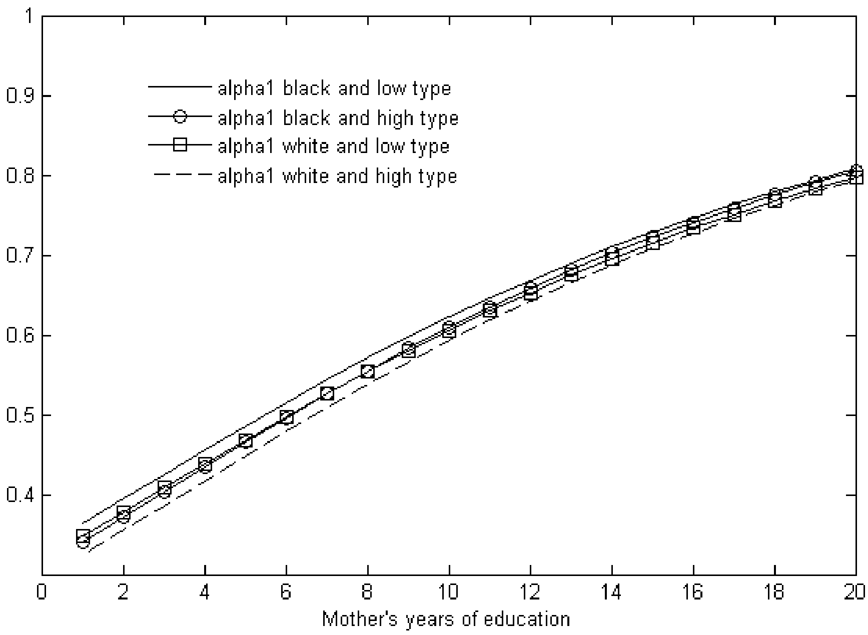
Child's Age	1-2	3-5	6-10	11-12
Mother's hours of work per week	24.53	26.20	28.07	29.44
	(17.65)	(17.48)	(17.22)	(17.25)
Non-parental child care hours	17.42	20.02	12.79	10.77
	(19.11)	(19.27)	(17.50)	(16.48)
Mother's time with the child	28.55	29.05	19.31	16.85
	(18.06)	(20.27)	(14.81)	(14.24)

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

6. Results

Table 4 shows the estimates of the parameters in the mother's utility function. The γs parameters represent the contribution of each observable and unobservable characteristic of the mother on her tastes for leisure, consumption and child's cognitive ability. Figures 4, 5 and 6 report the values taken by each preference parameter by subgroups and by mother's education. Figures 4 and 5 show that more educated mothers care more about leisure but less about consumption; moreover, there are positive differences in both leisure and consumption preferences by mother's skills level and by mother's race: white and low type mothers seem to care more about consumption, while high type mothers care less about leisure; the difference in consumption preference by mother's skills decreases for more educated women. Figure 6 shows the value of the preference parameter for child's ability. For any group, one more year of education implies a lower taste for child's ability. However, more skilled mothers care more about child's ability than the low skilled ones, while there are not significant differences induced by race. The estimates of the parameters in the utility function indicate that more educated mothers may work more to increase household consumption; however, for any given level of education, more skilled mothers care less about consumption and care more about child's ability than the low skilled mothers. Interpreting these patterns in terms of mother's behavior, it seems that more educated women work more and that the amount of time they spend with their child is lower than the corresponding one for the less educated. However, for any given level of education, mother's skills play a role in shaping mother's time allocation, since more skilled mothers may reallocate their time out of work in such a way to give priority to time with the child instead of leisure.

FIGURE 4
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.

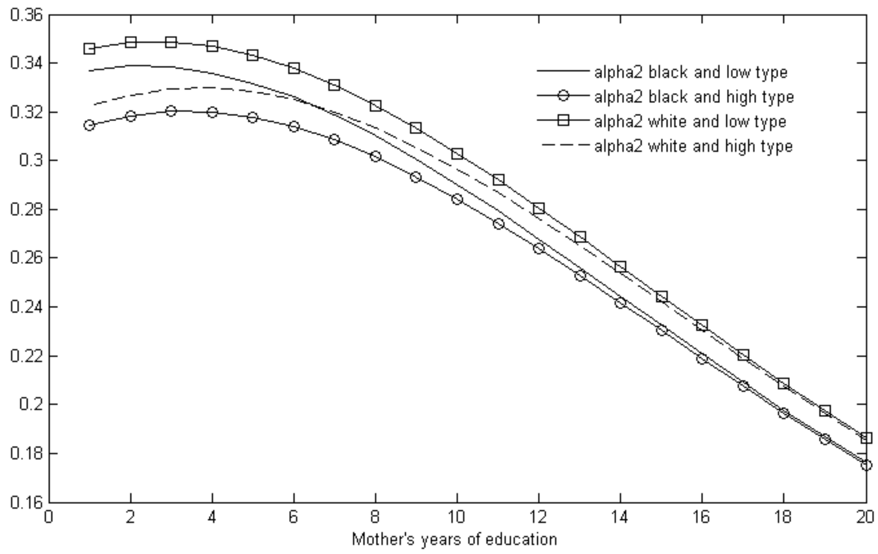
Table 4
Estimated parameters for mother's utility function.

Mother's Utility Function Parameters		
γ_{21}	contribution of mother's education to α_2	-0.0762
		(0.0046)
γ_{22}	contribution of mother's race to α_2	0.0699
		(0.0112)
γ_{31}	contribution of mother's education to α_3	-0.1957
		(0.0025)
γ_{32}	contribution of mother's race to α_3	0.0662
		(0.0118)
γ_{33}	contribution of mother's unobservable skills to α_3	0.2151
		(0.0142)
ρ	weight on future child's ability in the last period	24.0772
		(0.2297)
p	hourly price of child care	5.0902
		(0.0923)
Mother's Skills Distribution		
$\mu_{0\ high}$	skill level for high type	1.9184
		(0.0235)
$\mu_{0\ low}$	skill level for low type	1.4530
		(0.0950)
$\pi_{m\ high}$	proportion high skilled	0.5993
		(0.0160)
$\pi_{m\ low}$	proportion low skilled	0.4007
		(...)

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.

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 25 percent higher than the corresponding level for the low type. This implies a positive 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 40 percent. Recalling the estimated parameters for mother's preferences, the difference between high and low skilled mothers means that more skilled mothers have a higher probability of working induced by the higher wage offer that they receive, but they also care more about the ability of their child, increasing their time investments or the investments provided by non-parental child care. The choice between the two is strictly related to the values taken by the productivity parameters.

FIGURE 5
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.

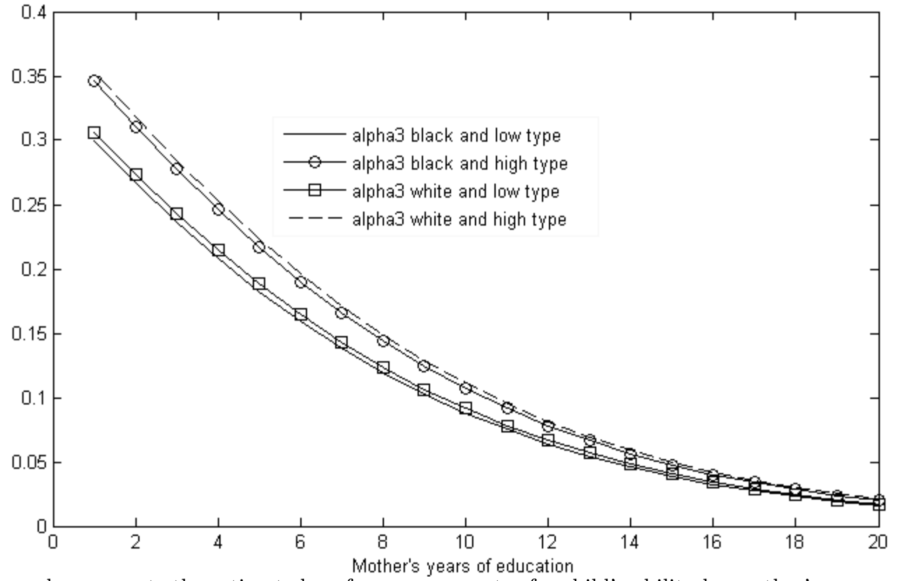
TABLE 5
Estimated parameters for the wage and income processes.

Wage Equation Parameters		
μ_1	coefficient of mother's education	0.0484
		(0.0011)
μ_2	coefficient of mother's age	0.0152
		(0.0009)
μ_3	coefficient of mother's age squared	−0.0003
		(0.0000)
μ_4	coefficient of mother's race	−0.0146
		(0.0011)
σ_ϵ	standard deviation transitory shock	0.3445
		(0.0031)
Household Income Process		
μ_{inc}	mean	34.3403
		(1.1065)
σ_{inc}	standard deviation	40.2330
		(3.7455)

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 4.8 percent with each additional year of education. This effect is in line with the one previously estimated by Del Boca et al.

Figure 6
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.

(2014), who use a similar specification for the mother's wage equation but do not control for mothers' unobserved heterogeneity, and also to the ones found by Bernal and Keane (2010) and Bernal (2008), who define an endogenous wage process and also control for mothers' unobservables.

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 7 and 8 show the time-varying elasticities as a function of child's age. Figure 7 reports the elasticities of child ability with respect to maternal time and non-parental child care time, while figure 8 reports the elasticities with respect to household income and the child's ability in the previous period. The first thing to notice is 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 7, the elasticity of child's cognitive ability with respect to non-parental child care is always lower than the elasticity with respect to maternal time with the child. The elasticity of child's ability with respect to maternal time ranges from 0.92

when the child is 1 year old to 0.5 when the child is 13 years of age. Instead the estimated elasticity with respect to non-parental child care ranges from 0.7 when the child is 1 year old to values close to 0 when the child grows up. The fact that the elasticity of non-parental child care strongly decreases over time may be due to two main reasons. On the 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, 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 the child's subsequent 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.

The relationship between the two elasticities shown in Figure 7 suggests that 1 percent increase in maternal time is always (for any child's age) more productive than one percent increase in external child care. Hence, if the mother works reducing the amount of time that she spends with the child, this may have negative consequences for child's ability and the effect is more detrimental during the child's first years of life. For instance, if the mother increases her labor supply, reducing her time with the child by 10 percent, when the child is 1 year old, this reduces child's ability in the following year by 9 percent. However, the model allows to identify all the mechanisms with which maternal employment affects child development, in particular the substitution between maternal time and non-parental child care that occurs when the mother works. In fact, working more, the mother not only reduces the amount of time shared with the child but also increases the use of non-parental child care.

In order to see whether non-parental child care is able to compensate for the reduction in child's ability induced by the decrease in maternal time, one should take into account both the estimated productivity parameters and the amount of both inputs that is used. Consider, for instance, an average mother spending with her 1-year-old child 28.55 hours per week and using non-parental child care for 17.42 hours per week. If this mother reduces her time with the child by 10 percent (corresponding to 2.8 hours) this reduces child's ability in the following period by 9 percent. If the same mother increases the use of external child care by the same amount of time (2.8 hours per week), this corresponds to an increase in child care use by 16 percent (with respect to the average value 17.42). The final effect of non-parental child care will be equal to a 9.6 percent increase in child's ability, since 1 percent increase in external child care time leads to an increase in ability by 0.6 percent. Thus, with the estimated productivity parameters the use of external child care can compensate for the negative effect on child's ability induced by the reduction in maternal time if non-parental child care is used for the same amount of time. However, if the difference between the estimated elasticities were higher and, at 1 year of age, $\delta_2 = 0.1$ instead of 0.6, the 16 percent increase in non-parental child care would have induced an increase in child's ability equal to 1.6 percent. In that case, non-parental child care would not be able to compensate for the reduction in child's ability. Eventually, even with the estimated values for the parameters δ_1 and δ_2, non-parental child care could not compensate for the reduction in ability induced by the lower maternal time if the mother were already using an higher amount of child care (e.g., higher than 17.42 hours per week). In fact, the decreasing marginal productivity of all inputs, including non-parental child care, implies that the additional hours are not productive enough to compensate for the reduction in child's ability induced by the mother's increase in labor supply.

Figure 8 shows the elasticity of child development with respect to 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).

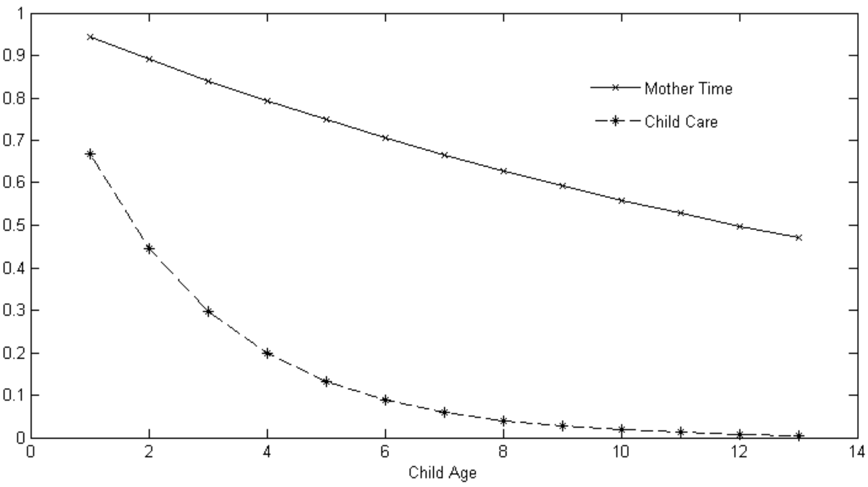
Finally, Table 6 reports the estimated parameters for the child's initial level of ability. High skilled children have a level of skills higher by 33 percent than the low skilled; the proportion of low skilled represents 49 percent of the sample. The coefficient for mother's education in the child's initial endowment is much higher than the coefficient for father's education and child's birth weight seems to explain more of the child's initial endowement. However, as pointed out in Bernal and Keane (2010), these coefficients just

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

CAPF Parameters		
ξ_1	slope productivity of maternal time	-0.0581 (0.0214)
ξ_2	slope productivity external child care	-0.4046 (0.0250)
ξ_3	slope productivity income	-0.2956 (0.0146)
ξ_4	slope productivity child's ability in previous period	-0.0470 (0.0017)
σ_v	standard deviation measurement error in test score	19.0927 (0.0894)
Child's Initial Ability Parameters		
$\psi_{0\ high}$	skill level for high type children	-55.5535 (0.6184)
$\psi_{0\ low}$	skill level for low type children	-82.7623 (0.9968)
$\pi_{c\ high}$	proportion high skilled children	0.5084 (0.0008)
$\pi_{c\ low}$	proportion low skilled children	0.4916 (...)
$\eta_{1\ MotherEdu}$	correlation child's endowment and mother's education	1.4949 (0.0870)
$\eta_{2\ FatherEdu}$	correlation child's endowment and father's education	0.5250 (0.0780)
$\eta_{3\ BirthWeight}$	correlation child's endowment and birth weight	1.5979 (0.0752)

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 7
Elasticity of child's ability with respect to mother's time with the child and non-parental child care.



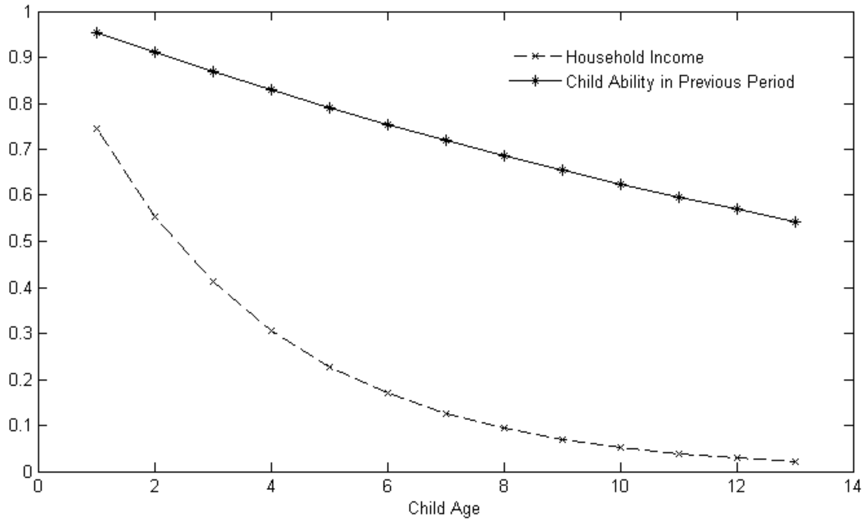
NOTE. This graph represents the productivity parameters for maternal time (τ_t) and non-parental child care (i_t) as a function of child's age $t = 1, 2, \ldots 13$. These parameters are defined as

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

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

represent correlations between the child's starting level of ability and parents' and child's observables characteristics and cannot be given a causal interpretation.

FIGURE 8
Elasticity of child's ability with respect to household income and child's ability in the previous period.



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

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

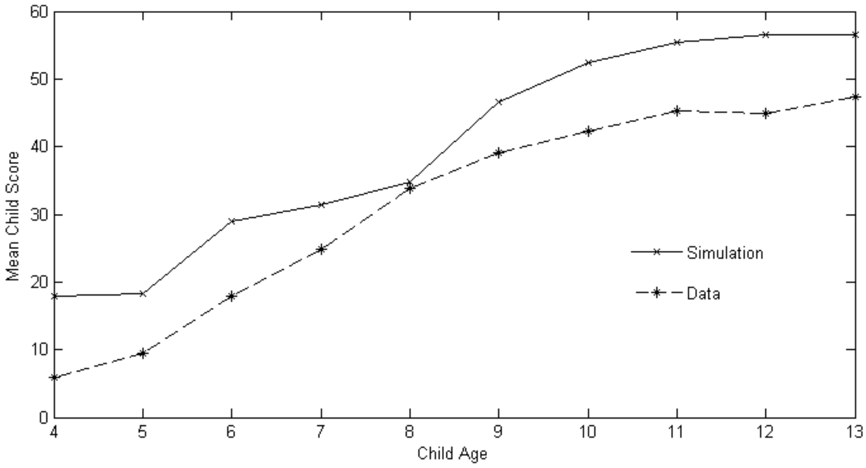
where $i = 3, 4$ and the estimated values for ξ_3 and ξ_4 are shown in Table 6.

These results are robust to alternative specifications that are presented in Appendix D. More precisely, I provide some sensitivity analysis testing the results on two dimensions: (i) the definition of the terminal period; (ii) the definition of maternal time. The first sensitivity check aims at focusing on a phase of the child's life when the inputs received from the mother seems more important than those from the father, according to the results provided by Del Boca et al. (2014). The second tests the sensitivity of the results to the different definition of maternal time with the child: more precisely, it shows the elasticity of child's development with respect to maternal time when maternal time is just active time with the mother and when maternal time also includes the father being around.

6.1. ***Goodness of fit of the model.*** Figure 9 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 child's subsequent 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

FIGURE 9
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.

income, observed in the actual and in the simulated data. The model predicts well the 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 and race.

Table 8 presents the average and standard deviations of the actual and simulated choice variables. The model predicts a slightly lower amount of working time and of non-parental child care with respect to the data. Finally, the model predicts well the average amount of time spent by the mother with the child.

7. POLICY SIMULATIONS

In this section, I use the estimated model to simulate the effects of policies (i) increasing the wage offers by 20 percent, (ii) subsidizing external child care, setting the price of the service at 1 US$ per hour, (iii) increasing the quality of non-parental child care, so that the productivity is the same as that of maternal time.

Panel (a) of Table 9 reports the percentage change of the variables induced by an increase in mother's wage by 20 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

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

	Actual Data	Simulated Data
All sample		
Mean mother's wage	14.2608	14.2934
Std mother's wage	10.1941	6.3130
Mean household income	39.6245	38.8525
Std household income	33.2903	33.5514
Wage by mother's education		
Some college education	16.3802	15.4950
High School	12.0260	13.4568
Wage by mother's race		
White	15.0270	14.5068
Black	12.9074	13.9572

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.

Table 8
Goodness of fit for mother's choices.

	Actual Data	Simulated Data
Mean Mother's hours of work	27.1812	24.9371
Std Mother's hours of work	17.5264	5.4826
Mean Non-parental child care hours	14.7107	9.6834
Std Non-parental child care hours	18.3211	14.1697
Mean Mother's time with the child	21.1600	21.4583
Std Mother's time with the child	17.0144	11.7665

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.

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 longer 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 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 mothers' participation in the labor market,

Table 9
Policy simulations.

(a) Increase in mother's wage

	Baseline	Percentage change
Child's ability in the last period	101.2550	1.0255
Hours of work	24.9371	1.8985
Maternal time with the child	21.4583	−0.5396
External child care time	9.6834	19.6097
Leisure	65.5501	−0.4626
Consumption	349.0612	19.5279
Utility	88.0353	5.0098

(b) Reduction in child care price

	Baseline	Percentage change
Child's ability in the last period	101.2550	17.5425
Hours of work	24.9371	0.2723
Maternal time with the child	21.4583	−0.0629
External child care time	9.6834	409.4884
Leisure	65.5501	0
Consumption	349.0612	0
Utility	88.0353	6.9867

(c) Increase in child care productivity

	Baseline	Percentage change
Child's ability in the last period	101.2550	21146.4896
Hours of work	24.9371	45.7298
Maternal time with the child	21.4583	−16.3999
External child care time	9.6834	407.8801
Leisure	65.5501	−11.9452
Consumption	349.0612	−12.3335
Utility	88.0353	34.5597

NOTE. This table reports percentage changes with respect to the baseline levels from (a) an increase in mother's wage by 20 percent, (b) a policy setting external child care price at 1 US$ per hour, and (c) a policy increasing non-parental child care productivity at the level of mother's time productivity. Child's ability in the last period is the value of the simulated child's ability at the end of period $t = 12$.

but can fail in having an effect on child's development although improving the economic conditions of the households.

Panel (b) in 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 5.09 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 are 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.

The final policy refers to the case where the non-parental child care option available to the parents is considered of high quality. More precisely, I simulate the mother's time allocation decisions and non-parental child care use in the case where the productivity of external child care is equal to the one of maternal time. Results are shown in panel (c) of Table 9. Notice that the price of child care does not change with its productivity, meaning that this policy may be consistent with backgrounds where the publicly and universally available child care is considered as of high quality, such as in Scandinavian countries. The literature on the determinants of external child care use points out that the perceived quality of the service plays a very important role in convincing parents to use it. However, it is not clear whether high-quality non-parental child care services can also change mother's behavior in terms of time allocation and labor supply. This exercise represents a very rough way of studying whether the availability of more productive non-parental child care arrangements may change mother's choices and time allocation. The results from this policy simulation confirm that the perceived quality of the service and the potential effects of non-parental child care play a very significant role in convincing the parents to choose it. In fact, the policy determines a strong increase in non-parental child care use and a resulting decrease in mother's time investments. More importantly, the higher quality of the service not only changes non-parental child care use but also determines a positive shift in mother's labor supply. The final effect on child's ability is positive and large. Hence, the implementation of a policy improving the productivity of external child care may have positive effects not only on external child care use and child development, but also on mother's labor supply.

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 existing studies, 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 mother's time with the child is more productive than non-parental child care at any age. Hence, maternal employment may have negative effect on child development, as long as the reduction in maternal time determines a decrease in child's ability that is not compensated for by the alternative forms of care available to the mother.

The policy simulations suggest that the policymaker 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 evaluate only "local" effects, they show that policies aimed at increasing participation of mothers in the labor market 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. Only a policy improving the perceived quality of non-parental child care (increasing its productivity up to the level of mother's time productivity) is found to be effective not only for increasing child's subsequent development but also for boosting mother's labor supply on the extensive margin.

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, these two choices may respond not only to different decision-making processes, but may also have diverse implications for child development. I leave this issue for future research.

References

Adda, J., C. Dustmann, and K. Stievens (2012). The career cost of children. EUI Working Paper ECO 2012/01.

Almond, D. and J. Currie (2011). Human capital development before age five. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*. Elsevier Science B.V.

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. 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.

Datta Gupta, N. and M. Simonsen (2011). Non-cognitive child outcomes and universal high quality child care. *Journal of Public Economics 94*(1-2), 30–43.

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 (2014). Household choices and child development. *Review of Economic Studies 81*(1), 137–185.

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*.

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

Havnes, T. and M. Mogstad (2011). No child left behind: Subsidized child care and children's long-run outcomes. *American Economic Journal: Economic Policy 3*(2), 97129.

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.

Heckman, J. J. and B. Singer (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica 52*(2), 271–320.

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.

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. 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.

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.

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/.

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:

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

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 (6).

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:

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

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

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

The corresponding derivatives[22] are given by the following expressions:

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

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

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

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

and the FOCs become:

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

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

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 $D_{T+1} = \frac{\partial V_{T+1}}{\partial ln A_{T+1}} = \rho\alpha_3$.

[22]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:

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

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).

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:

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

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:

$$i_{T-1}^c \Rightarrow \frac{\partial \bar{V}_{T-1}}{\partial i_{T-1}} + \frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial i_{T-1}} = 0 \tag{A.12}$$
$$\tau_{T-1}^c \Rightarrow \frac{\partial \bar{V}_{T-1}}{\partial \tau_{T-1}} + \frac{\partial V_T}{\partial lnA_T} \times \frac{\partial lnA_T}{\partial \tau_{T-1}} = 0 \tag{A.13}$$

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 (7) 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 (8) 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 (10) represents the optimal labor supply in each period as a function of i_t and τ_t; substituting (7) and (8), it yields the optimal labor supply choice for each period t, as defined by (11).

The unconditional demands for child care and time with the child are derived substituting the labor supply solution into equations (7) and (8) (corresponding to (A.25) and (A.31) in this Appendix).

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.[23]

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:

$$\alpha_1 = \frac{exp(\gamma_1)}{exp(\gamma_1) + exp(\gamma_{21} MotherEdu + \gamma_{21} MotherRace) + exp(\gamma_{31} MotherEdu + \gamma_{32} MotherRace + \gamma_{33} I(\mu_0 = \mu_{0high}))} \tag{B.1}$$

$$\alpha_2 = \frac{exp(\gamma_{21} MotherEdu + \gamma_{22} MotherRace)}{exp(\gamma_1) + exp(\gamma_{21} MotherEdu + \gamma_{22} MotherRace) + exp(\gamma_{31} MotherEdu + \gamma_{32} MotherRace + \gamma_{33} I(\mu_0 = \mu_{0high}))} \tag{B.2}$$

$$\alpha_3 = \frac{exp(\gamma_{31} MotherEdu + \gamma_{32} MotherRace + \gamma_{33} I(\mu_0 = \mu_{0high}))}{exp(\gamma_1) + exp(\gamma_{21} MotherEdu + \gamma_{22} MotherRace) + exp(\gamma_{31} MotherEdu + \gamma_{32} MotherRace + \gamma_{33} I(\mu_0 = \mu_{0high}))} \tag{B.3}$$

[23] See Mroz et al. (2010) for similar applications.

where 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.

Concerning the CAPF, the parameters in this case should be strictly positive. Hence, I implement the transformation defined by (15) 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 $\Gamma_2 = (\gamma_{21}, \gamma_{22})$, $\Gamma_3 = (\gamma_{31}, \gamma_{32}, \gamma_{33})$, $\Xi = (\xi_1, \xi_2, \xi_3, \xi_4)$, $k = (h, l)$, $\Upsilon = (\mu_1, \mu_2, \mu_3, \mu_4)$, $\Delta = (\eta_1, \eta_2, \eta_3)$ 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.[24]

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

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 the 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.[25] Then, 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

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

[25]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 was 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 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.

$\hat{m}$ 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 $N = 430$ 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,[26] 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 to compute also the standard errors. After having drawn B_{se} samples from the actual data,[27] I repeat the estimation of the parameters for each sample. This yields an empirical distribution of the parameters estimates, from which I can recover a bootstrap estimate of the variance, using the formula (Train 2009, pag. 201):

[26] $B = 200$.
[27] $B_{se} = 50$

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

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):

$$se_{\hat{\pi}_{lh}} = |g^{'}\left(\hat{z}_l\right)| se_{\hat{z}_l} \tag{B.9}$$

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

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 a 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 ($N = 430$) 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. 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 report only the results concerning the parameters for maternal time and external child care time in the CAPF.

D.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. Moreover, according to the results provided by Del Boca et al. (2014), mother's time is actually the most important input that the child may receive in the first 12 years of his life, while father's time becomes also important starting from this age on. In order to see whether the restrictions on father's labor supply imposed in the model play a role, I re-estimate it setting $T = 12$. This change yields a sample of 366 observations repeated for 12 periods.[28] Figure D.1 presents the results showing that mother's time productivity is slightly lower than the one estimated in the base analysis. This may confirm the point argued by Del Boca et al. (2014) saying that father's time also becomes important as the child grows up and enters adolescence. Hence, the estimated

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

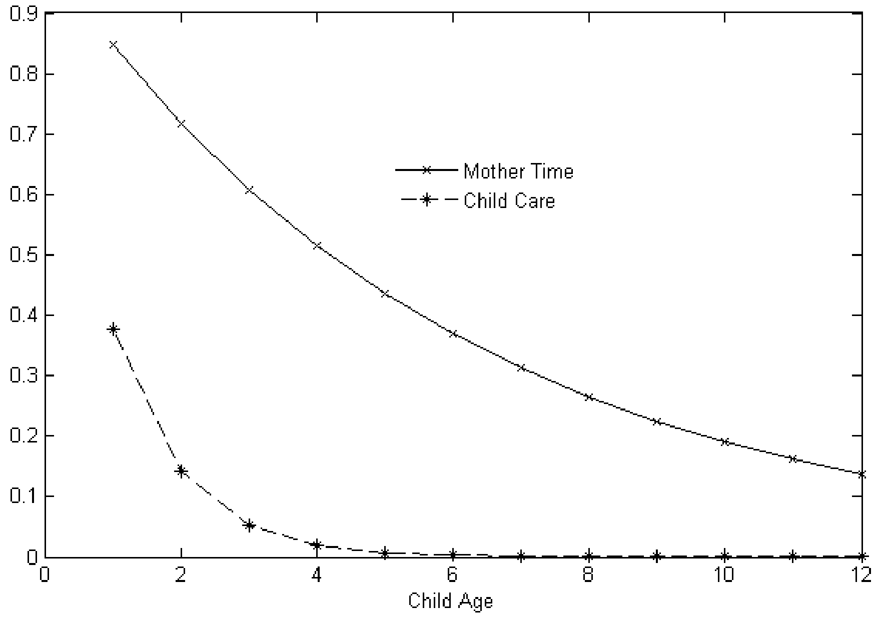
elasticity of child's development with respect to mother's time presented in Section 6 may provide an upper bound of the true value.

D.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 7 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 D.2: the elasticity of child's development with respect to maternal time is slightly lower than the one found in the main analysis. Moreover, the elasticity of child's development with respect to maternal time presents a sharper decrease over time than the one shown in figure 7. This may suggest that while active time with the mother is always productive for child's development, the passive time spells (when the mother is just around) also become important as the child ages.

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 the latter category also represents an input for child development, I repeat the estimation of the model defining the variable for maternal time adding also this category. The results shown in Figure D.3 provide a lower productivity than the one reported from the base analysis. This may suggest that the time spells when the father is around are somehow less productive for child development, pushing down the overall productivity of maternal time.

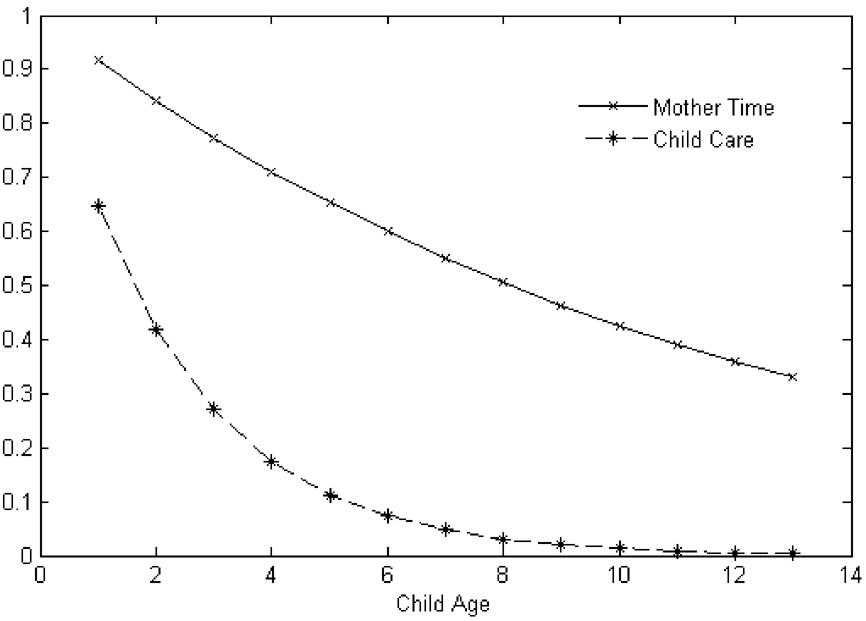
The results from all these additional analyses suggest that the true effect of maternal time on child's ability may range in between the one found in the main text and the ones reported in Figures D.2 and D.3. Thus, the estimated productivity in the base analysis is very likely to represent an upper bound of the true value.

FIGURE D.1
Elasticity of child's ability with respect to mother's time with the child and non-parental 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 $T = 12$.

FIGURE D.2
Elasticity of child's ability with respect to mother's time with the child and non-parental 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).

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*)

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

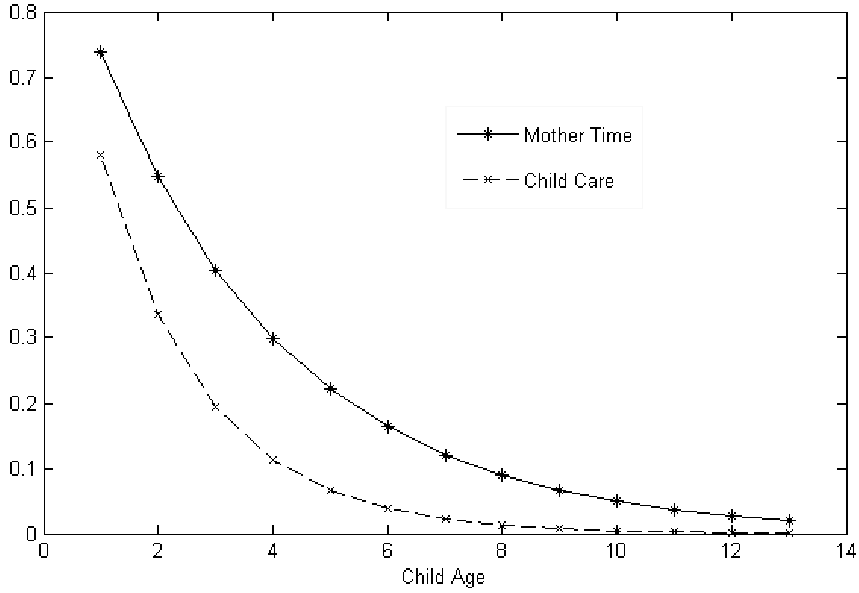
Table C.4
Mean characteristics of the sample with respect to PSID-CDS data.

	PSID-CDS	Sample	T-test
Mother's hours of work	23.59	27.18	−10.86***
	(0.14)	(0.29)	
Non-parental child care	12.21	14.71	−9.40***
	(0.10)	(0.25)	
Maternal time with the child	25.83	21.16	6.18***
	(0.32)	(0.68)	
Mother's wage before childbirth[a b]	10.98	11.24	−1.21
	(0.09)	(0.18)	
Mother's education	12.98	13.27	−7.84***
	(0.02)	(0.03)	
Mother's age at child's birth	26.98	28.19	−15.98***
	(0.04)	(0.07)	
Mother's race: white	0.61	0.61	−0.12
	(0.00)	(0.01)	
Child's gender: male	0.51	0.51	0.49
	(0.002)	(0.006)	
Child's birth weight	116.89	119.18	−7.38***
	(0.11)	(0.29)	
Father's hours of work	38.66	45.25	−30.28***
	(0.11)	(0.19)	
Father's education	12.66	13.29	−17.73***
	(0.01)	(0.03)	
Household non labor income[a]	16.86	12.87	2.44**
	(1.39)	(0.84)	
N	3243	430	

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

Figure D.3

Elasticity of child's ability with respect to mother's time with the child and non-parental child care if maternal time includes also time when the father is around.



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.