The Effect of Maternal Employment and Child Care Choices on Children's Cognitive Development

Raquel Bernal*

Northwestern University, U.S.A.

April 2005

Abstract

This paper develops and estimates a dynamic model of employment and child care decisions of women after birth in order to evaluate the effects of maternal employment and daycare choices on children's cognitive ability. I use data from the NLSY to estimate the model. Results indicate that the effects of maternal employment and child care on children's ability are negative and rather sizeable. In fact, having a full-time working mother who uses child care during the first 5 years after the birth of her child is associated with a 10.4% reduction in child's ability test scores. Based on the estimates of the model, I assess the impact of policies related to parental leave, child care and other incentives to stay at home after birth on women's decisions and children's outcomes.

*I am grateful to Michael Keane for his guidance and support. I want to acknowledge helpful comments from Donghoon Lee, Raquel Fernandez, Chris Ferrall, Christopher Flinn, Chris Taber, Zvi Eckstein and participants of seminars at NYU, Queen's University, University Tor Vergata, Minnesota, UT Austin, Pittsburgh, Wisconsin, Brown, Rochester, Penn State, Yale, Carnegie Mellon, Western Ontario and Northwestern. I gratefully acknowledge financial support from the Colombian Central Bank and NYU.

1. Introduction

In this paper I study labor supply and child care decisions of women immediately after birth and assess how these decisions affect children's cognitive development. The effect of maternal time inputs on children's ability is theoretically ambiguous. First, we do not know whether maternal time is in fact effective in producing children's skills. Second, there is no a priori reason to believe that parental time is better in doing so than an alternative provider's time. The answers to these questions seem to be extremely important in terms of policy design. Apart from the relevance of the policy question, this topic is particularly interesting for economists. An extensive body of literature in labor economics has concluded that about 70% of the variance in wages corresponds to unobserved "individual effects" whose origin we do not really understand. This means that conditional on a set of observable characteristics including age, gender, education, race, parents' education, working experience, etc., there remains a significant proportion of the variation in wages that we have not been able to account for. A number of recent papers, such as Keane and Wolpin (1997) and Cameron and Heckman (1998) suggest that educational attainment, wages and other career outcomes are essentially driven by an initial skill endowment already determined by age 16. A better understanding of how this skill endowment is generated by inputs at early stages of life is critical for the understanding of labor market outcomes.

In this paper I use data from the National Longitudinal Survey of Youth (NLSY) to estimate a dynamic model of employment and child care choices of mothers after birth and evaluate the effects of mother's decisions on children's cognitive ability. A common limitation of previous studies that have used data from the NLSY to assess the impact of maternal employment on children's outcomes is that they have failed to fully control for potential biases that may arise from the following facts: (1) Women that work/use child care may be systematically different from women who do not work/do not use child care; (2) The child's cognitive ability itself may influence the mother's decisions of whether to work and/or place the child in day care.

Women are heterogeneous in their skill endowments, the constraints they face and their tastes. Likewise, children are heterogeneous in their cognitive ability endowments. Some of these characteristics might be unobserved by the researcher. Mothers' decisions of whether or not to work and whether or not to use child care will clearly depend on these unobserved heterogeneous characteristics of both mothers and children. To make this selection bias problem clear, I lay out a couple of examples. In the case of (1), for example, a woman with higher skills is more likely to have a child with high cognitive ability and also more likely to work. Then, a statistical analysis would attribute the effect of this woman's higher skills to employment, and the estimated effects of maternal employment on her child's cognitive outcomes would be upwardly biased. In the case of (2), mothers of low ability children may choose to compensate them by spending more time with them, in which case mothers are more likely to work if they have high ability children. Again, the estimated effect of maternal employment on child's cognitive outcomes would be upwardly biased. Clearly, these sample selection issues make evaluation of the effects of women's decisions on child outcomes very difficult. In this paper, I estimate a model of employment and child care choices jointly with a child cognitive ability production function. This type of estimation allows me to implement a selection correction in the sense that I can adjust for the fact that certain types of children are more likely to be put in child care and/or to have working mothers.

One can think of at least two different strategies that could be implemented to solve the selection problem. The first one is to run a regression of measures of the child's cognitive ability on several inputs, including mother's employment and child care use, by using instrumental variables for both of these inputs. However, it is quite difficult to come across valid instruments for both, work and child care choices. Researchers have used cross-sectional variation in prices, e.g., child care costs, and other location-specific characteristics as instrumental variables. Yet, there is little exogenous variation in such variables, which poses a question about their reliability as instruments. Location-specific variation in the ability endowments of its residents will plausibly be correlated with the demand for different market or publicly supplied services and products, e.g., child care. The second possibility is to run household fixed effects models in order to get rid of household specific unobserved components. However, it is plausible that mothers make time compensations for children depending on their ability type. In this case, using a household fixed effect model would not be appropriate, since maternal employment is correlated with the sibling-specific component of the cognitive ability endowment. Hence, I propose estimation of the structural model (of maternal employment and child care decisions) jointly with the cognitive ability production function as an alternative approach which, besides providing a plausible selection correction mechanism, serves the purpose of generating some interesting insights about women's dynamic behavior after birth and most importantly, of providing reliable estimates of preference and technological parameters which allows me to use the model to assess counterfactual policy experiments. This would not be possible using the reduced-form approaches mentioned above. In particular, I use the estimates of the model to evaluate the effects of policies related to parental leave, child care subsidies and other incentives for women to stay at home after birth on women's labor supply and child care choices and children outcomes.

The most important results of this paper are the following. First, the effect of maternal employment and child care on children's cognitive ability is negative and rather sizeable. In fact, having a full-time working mother who uses child care during one of the first five years after birth is associated with a 2% reduction in the child's test scores. Second, this effect is stronger for children with high ability endowments. In other words, there is a higher technological return to time spent with high ability children relative to low ability ones. Third, child care subsidies and a specific type of maternity leave policy are detrimental for children yet increase the mothers' expected lifetime utility, whereas a baby bonus received by the household after the birth of a child would have positive effects on both mothers' welfare and children's test scores.

The paper is organized as follows: In Section 2 I present a brief summary of the related literature. In Section 3 I describe the structure of the model. Section 4 discusses the solution and estimation methods as well as some issues related to identification. Section 5 describes the NLSY data on which I estimate the model and highlights the overall patterns in the data. Section 5 presents the estimates of the model, evaluates its ability to fit the data and discusses the importance of unobserved heterogeneity. Section 7 presents the results from several policy experiments. Section 8 concludes.

2. Related Literature

In this section, I briefly summarize the main results of the studies published to date examining the impact of early maternal employment and child care on child outcomes in the U.S. using data from the NLSY. For most part, these studies present simple correlations between inputs and child outcomes, or use a limited set of measures of family and child characteristics. In most cases, no correction for self-selection of children into child care arrangements or the group of working mothers was implemented. As will be clear by the end of this section, results are still far from conclusive.

A number of prior studies have found negative effects of maternal employment on child outcomes. Desai et al. (1989) used 503 four year olds from the NLSY in 1986. Results from multivariate regression analysis show a statistically significant adverse effect of maternal employment on children's intellectual ability but only for boys in higher income families. Baydar and Brooks-Gunn (1991) analyzed 572 white 3-4 year olds (in 1986) and concluded that the effects of maternal employment differ, depending on the timing of the employment. Mott (1991) analyzed 2,387 one to four year olds in 1986. He found maternal employment over 20 hours per week during the second quarter of the child's life to be negatively correlated with PPVT[1] scores. Harvey (1999) used children from 3 to 12 years old in 1986, 1988, 1990, 1992 and 1994. She found a negative effect of maternal work hours on PPVT and PIAT[2] scores for young children and a weaker or null effect at higher ages. Waldfogel et al. (2000) analyzed a sample of 1,872 children followed longitudinally from birth to age 7 or 8. They found small and persistent negative effects of maternal employment during the first year on children's cognitive outcomes for white children but not for African-American children. Han et al. (2001) used a cohort of 462 children followed from birth to age 7 or 8. They found that maternal employment in the first

[1] Peabody Picture Vocabulary Test. For details see Section 5.1.

[2] Peabody Individual Achievement Test. For details see Section 5.1.

year has a negative but small effect on cognitive outcomes for white children.

On the other hand, some prior studies have found positive effects of maternal employment on children's scores. Vandel and Ramanan (1992) studied 189 low income non-Hispanic second graders (in 1986) and found positive effects of maternal employment on PIAT and PPVT scores. Parcel and Menaghan (1994) used 768 3-6 year olds whose mothers were employed in 1986. They found positive effects from mothers' employment during the first year or first three years on PPVT scores.

And yet another group of studies finds inconclusive of insignificant results. Greenstein (1995) analyzed 2,040 4-6 year olds during 1986, 1988 and 1990. He found an insignificant relationship between maternal employment and PPVT scores. Moore and Driscoll (1997) analyzed 1,154 five to fourteen year olds (in 1992) whose mothers were on AFDC during 1986-1990. The authors found that maternal employment is associated with higher PIAT, although most of the effects are eliminated after controlling for maternal and household characteristics.

The main reason for the diversity of these results may well be the common limitation of failing to control for potential biases that may arise due to the endogeneity of employment and child care choices. A few recent studies have tried to overcome this issue by using a more extensive set of explanatory variables, using instrumental variables or estimating household fixed effects models. Blau and Grossberg (1992) used 874 3-4 year olds in 1986. To correct for potential heterogeneity bias, they estimated their basic equation using instrumental variables for maternal labor supply. However, the instruments, such as region of residence and maternal and child health, may be debatable since they may affect child outcomes directly. They concluded that maternal employment in the first year after birth is associated with lower PPVT scores, while the contrary is true for the second and third years of employment. Ruhm (2000) used a larger and more representative sample from the NLSY in an attempt to control for as many characteristics as possible. His results from multivariate regressions indicate that maternal labor supply during the first three years of the child's life is predicted to have a small negative effect on the verbal ability of 3 and 4 year olds and a significant negative effect on the reading and math achievement of 5 and 6 year olds. Finally, James-Burdumy (1998) used 2,119 three to four year olds (in 1986 and 1988) to estimate a household fixed-effect model using instrumental variables to control for unobserved heterogeneity in cognitive ability. She concluded that there is no effect of hours or weeks worked by the mother in years 1,2 or 3 on child test scores. However, it is plausible that mothers make time compensations for children depending on their ability type. In this case, using a household fixed effect model would not be appropriate, since maternal employment is correlated with the sibling specific part of the cognitive ability endowment.

3. The Model

At each period t after birth and before the child enters primary school, i.e. 5 years of age, the mother chooses whether to work full-time, work part-time or not work and whether to use child care or not. In the model the time periods correspond to 3 month intervals (quarters). Formally, let each alternative $j \in J = \{(h_t, I_t^c) : h_t = f_t + \frac{p_t}{2} \text{ and } I_t^c = 0, 1\}$. h_t corresponds to the employment choice while I_t^c refers to the child care decision. In particular, f_t is an indicator function that equals 1 if the woman works full-time in period t, p_t is an indicator function that equals 1 if the woman works part-time in period t and I_t^c is an indicator function that equals 1 if the woman used child care in period t. Finally, for simplicity, I define d_t^j to be an indicator function that equals 1 if alternative $j \in J$ is chosen at time t.

3.1. Utility Function

The current-period utility function for choosing alternative j is given by:

$$\begin{aligned} U_t^j = \; & \tfrac{1}{\alpha_1} c_t^{\alpha_1} + \alpha_2 h_t + \alpha_3 \left(\frac{A_t^{\lambda} - 1}{\lambda} \right) + \alpha_4 I_t^c + \alpha_5 h_t \left(1 - I_t^c\right) + \alpha_6 I_t^c (1 - I[\textstyle\sum_{\tau=1}^{t-1} I_\tau^c > 0]) \\ & + \alpha_7 I\left[t = 1\right] I_t^c + \alpha_8 I[t < 5] I_t^c + \varepsilon_t^j d_t^j, \qquad \text{for } j = 1, ..., 6 \end{aligned} \tag{3.1}$$

where consumption c_t is given by the budget constraint:

$$c_t = w_t.(500 h_t) + I^H - cc I_t^c$$

where w_t is the mother's real hourly wage, I^H denotes husband's average income during the period, and cc is the cost of child care services. Earned income is given by $w_t.(500h_t)$, because I define full-time work (for a quarter) as 500 hours and part-time work as 250 hours. This grouping of hours is necessary in order to keep the choice set discrete. Keane and Moffitt (1998) argue that this particular grouping is desirable given that hours are very concentrated at 20 and 40 per week, and because much of the variation away from those figures is likely to be measurement error.

The utility function (3.1) has the common CRRA form in consumption. The parameter α_2 is the disutility from working. The variable A_t is cognitive ability of the child. This is generated by a production function that is defined below. The mother gets utility from the child's cognitive ability according to the CRRA function with parameter λ. An estimated $\lambda < 1$ would imply that mothers get diminishing marginal utility from child ability, and will therefore have an incentive to engage in behaviors that compensate children with relatively low initial ability endowments.

The next set of terms in the utility function capture various aspects of the utility/disutility from child care use. This set of terms are necessary for the model to provide a good fit to the quantitative features of the NLSY data, in particular patterns of child care utilization. The parameter α_4 is a general non-pecuniary benefit/cost associated with the use of child care. The parameter α_5 is an extra disutility from working if child care is not available. The parameter α_6 is an extra cost of initiating child care if it has not been used before. This may capture the search time cost of finding a day care center, and/or the psychic cost of first time separation from the child. The parameter α_7 is an extra cost from using child care during the first quarter after birth ($t = 1$), and α_8 is an extra cost from using child care before the child is one year old ($t < 5$). Both of these parameters capture the fact that it is more difficult to find day care centers that will take infants, along with the fact that the psychic cost of separation from the child is greater when the child is very young.

Finally, ε^i is an alternative-specific random taste component. I allow these terms to be correlated across alternatives, to capture the fact that some alternatives are more similar than others. I assume that the random preference shocks $\varepsilon_t = \{\varepsilon_t^1, \varepsilon_t^2, \varepsilon_t^3, \varepsilon_t^4, \varepsilon_t^5, \varepsilon_t^6\}$ have a joint normal distribution $F(\varepsilon_t)$ and are serially uncorrelated.

3.2. Wage Formation

Besides the budget constraint, a woman faces two additional constraints that affect her employment and child care decisions. The first one, which is discussed in this section is her wage function and the second one is the child cognitive ability production function (which I turn to in the next section). It is useful to first define the woman's "initial wage", w_0, which is constructed based on her labor history prior to giving birth[3]. I assume that the initial wage is a function of a vector X_0 of observable characteristics that include education, age, age squared, race, and a "skill endowment" denoted by ξ_0[4]. This yields the following initial wage function:

$$\ln w_0 = X_0\theta + \xi_0 \tag{3.2}$$

After childbirth, the wage commanded by the woman if she returns to work is described by the following process:

$$w_t = w_0(1-\delta)^t \exp(\phi_1 E_t + \phi_2 f_{t-1} + \phi_3 p_{t-1} + \phi_4(E_t * ed) + \epsilon_t).$$

where δ is the depreciation rate of human capital, $E_t = \sum_{\tau=0}^{t-1} h_\tau$ is total work experience since birth, f_{t-1} and p_{t-1} indicate whether the woman worked full-time or part-time during the immediately preceding period, $E_t{*}ed$ is an interaction term of woman's experience and her education at birth[5], and ϵ_t is a measurement error with $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. This re-employment wage function captures the notion

[3] w_0 is measured as the average hourly wage of the 18-month period prior to birth. This is done in order to avoid dealing with measurement error in initial wages.

[4] X_0 includes education, age and age squared (which proxy for experience) because these are variables that augment initial skill endowment in a human capital earnings function (see Willis 1986, Heckman and Sedlacek 1985, Keane and Wolpin 1997). It is important to make a distinction between initial skill endowment and subsequent investments in order to get results that can be interpreted. In fact, the reason why AFQT scores are not included as part of X_0 is that these correspond to a mixture of the woman's initial ability endowment and subsequent educational inputs. If I introduce this variable in the wage equation, it would be both difficult to interpret that particular coefficient and less clear how to interpret several other coefficients, in particular the coefficient associated to education in the cognitive ability production function (γ_2).

[5] This interaction term significantly improves the fit of predicted wages by observed characteristics of women.

that, the longer a woman leaves the labor market after childbirth, the more her marketable skills will depreciate. However, that can be partially offset by accumulating working experience after birth.

3.3. Child's Cognitive Ability Production Function

Each mother derives utility from her child's cognitive ability, which she can observe. I assume that the child is born with a cognitive ability endowment A_0, which is correlated with some observable and unobservable variables according to the following equation:

$$\ln A_0 = \gamma_1 \xi_0 + \gamma_2 ed + \gamma_3 race + \gamma_4 \mu + \gamma_5 BW + \gamma_6 I[age<18] + \gamma_7 I[age>33] + \gamma_8 gender + \omega_\kappa \quad (3.3)$$

where ξ_0 is the mother's skill endowment (defined as noted in the previous section), *ed* is the mother's education, *race* is a dummy variable equal to 1 if the child is non-white, 0 otherwise , μ is the father's skill endowment[6], BW is child's birth weight, *I[age<18]* is a dummy variable that indicates if the mother is younger than 18, similarly for the dummy variable *I[age>33]* and *gender* is a dummy variable indicating if the child is a male. I include the age indicators because there is evidence that teenage mothers as well as older mothers[7] have less healthy children. Beyond that, I assume that age of the mother does not directly affect the child's skill endowment[8]. The term ω_κ represents unobserved heterogeneity in the child's cognitive ability type. I start by assuming a discrete distribution of types characterized by two types such that $\kappa = l, h$, (low or high).

An additional assumption of the model is that mothers know their child's cognitive ability endowment. Thus, mothers know ω_κ and $\ln A_0$. If this is the case, this can create a potential source of bias

[6] I assume that the labor income of father i is given by: $\ln I_{it}^H = \beta_o + \beta_1 age_{it} + \beta_2 age_{it}^2 + \mu_i + \varepsilon_{it}$ where age_{it} is father i's age in year t, μ_i is the father's skill endowment and ε_{it} is measurement error. Hence, the skill endowment is approximately given by:$\mu_i \approx \frac{1}{5}\sum_{t=1}^{5} \ln I_{it}^H - (\widehat{\beta}_o + \widehat{\beta}_1 age_{it} + \widehat{\beta}_2 age_{it}^2)$.The results of this regression are reported in Appendix 1.3.

[7] Given the fact that the NLSY is a sample of relatively young women, the upper threshold is set at 33, which would otherwise be too low to describe the set of older mothers.

[8] Different age effects were included in the specification of equation (3.3), however these never turned out to be significant.

in the estimates of the cognitive ability production function in the sense that mothers can engage in compensating behaviors by spending more type (and using less child care) with low endowment children[9]. While it is reasonable to assume that mothers know much more about the cognitive ability of their children than we do, I acknowledge that the the assumption that they have complete information may be less plausible. It could be possible to consider extensions such as incorporating learning in the model or allowing ω_κ to be a composite of two components, one of which is observed by the mother.

Finally, the cognitive ability production function maps the child's initial ability endowment A_0, along with subsequent inputs such as maternal time and alternative provider's time, into the child's (age adjusted) cognitive ability at time t, denoted A_t, according to:

$$\ln A_t = \ln A_0 + \gamma_9 E_t + \gamma_{10} C_t + \gamma_{11}(\ln A_0 * E_t) + \gamma_{12}(\ln A_0 * C_t) \tag{3.4}$$

where $E_t = \sum_{\tau=0}^{t-1} h_\tau$ and $C_t = \sum_{\tau=0}^{t-1} I_\tau^c$ denote the mother's total quarters of work experience and child care use, respectively, since birth. The coefficients γ_9 and γ_{10} capture the effects of cumulative maternal work and day care use on the child's cognitive outcome at time t. Additionally, interaction terms between the child's initial ability and mother's choices are included in order to allow the effect of women's choices to vary depending on the type of child[10]. The complete cognitive ability production function is obtained by substituting (3.3) into (3.4)[11].

[9] In this case, a sibling fixed effect estimator would not deal with the problem, because, if mothers can see the endowment differences across their children they may treat them differently.

[10] Alternative specifications of the production function were used to estimate the model. However, either the results presented in Section 6 did not vary significantly in terms of the main effects of interest or the model would underperform in terms of within-sample fit.

[11] Admittedly, the fact that specification (3.4) does not allow the timing of maternal inputs to matter is a strong assumption. The developmental psychology literature has long recognized that effects of maternal inputs are stronger during the first year after birth. This could be done, for example, by decomposing E_t and/or C_t into measures of employment and child care use when the child is in various different age ranges. Clearly, this would imply that the state space that the woman faces each period is not only characterized by cumulative work and child care decisions but also by these age-specific cumulative terms. Hence, this would considerably add to the computational burden of solving and estimating the model.

Note that (3.4) can be derived from a production function in which maternal time is a direct input. Suppose each woman has a stock of time available to spend with the child. Working or using day care reduce the amount of contact time. Then equation (3.4) can be obtained by substituting an equation for maternal time as a function of work and child care use into the production function.

Note that the econometrician does not observe actual cognitive ability of children, but instead has available a set of (age adjusted) cognitive ability test scores from which she has to infer the child's cognitive ability. Let S_t^A be the (age adjusted) test scores[12] observed in period t and let measurement error be specified as:

$$\ln S_t^A = \ln A_t + \eta_1 d_{1t} + \eta_2 d_{2t} + v_t \tag{3.5}$$

where d_{1t} and d_{2t} are cognitive ability test dummies[13] which capture the fact that the means on the different tests differ, and v_t is a measurement error with $v_t \sim N(0, \sigma_v^2)$.

3.4. Observed and Unobserved heterogeneity

I allow observed and unobserved heterogeneity in a number of dimensions. I have already noted that mothers and fathers are heterogeneous in their own skill type, given by ξ_0 and μ[14] respectively, and that children are heterogeneous in the endowment type, denoted by ω_κ in equation (3.3). Additionally, I allow mothers to be heterogeneous in their taste for work (α_2) and their taste for child care (α_4). Specifically, each preference parameter is described by the following equation:

$$\alpha_{i,k} = \alpha_{i1}\xi_0 + \alpha_{i2} ed + \alpha_{i3} race + \rho_i time + \overline{\alpha}_{i,k} \qquad \text{for } i = 2, 4 \text{ and } k = l, h$$

where *time* is a cohort trend[15] and $\overline{\alpha}_{i,k}$ is the unobserved component of tastes for work or day care. Initially, I assume that there are two different types in each case (low and high). This implies that

[12] I use the Peabody Picture Vocabulary Test and the Picture Individual Achievement Test (Math and Reading).

[13] $d_{1t} = 1$ if S_t^A is a PPVT score, 0 otherwise and $d_{2t} = 1$ if S_t^A is a PIAT (Math) score, 0 otherwise.

[14] In this framework each mother (and father) is her (his) own skill type given that the skill endowment is defined as the unexplained component of inital wages.

[15] I introduce the cohort term to allow for the possibility that changes in preference over time accounts for some of the increase in female labor participation rates over the sample period.

women will be classified into 8 types[16] by unobserved characteristics of both them and their children. $\overline{\alpha}_{i,k}$ are parameters to be estimated.

3.5. The Individual's Optimization Problem

There are two key sources of dynamics in this model that mothers consider when making choices: (1) they know how their decisions about working after childbirth will affect the depreciation of their market wages, i.e., the evolution of their human capital and (2) they know how their decisions about work and child care use will affect cognitive ability outcomes of their child.

In order to estimate the model, one first needs to solve the woman's optimization problem. I define S_t as the state at period t that arises as a result of the decisions made up to t. The problem is characterized by three state variables that evolve endogenously: work experience since birth (E_t), work decision during the immediately preceding period (h_{t-1}) and total child care used (C_t). Note that the state variables E_t and C_t are incremented in the obvious way at each age t based on the work and day care use decisions at $t-1$. Further I denote $S_t = \{E_t, h_{t-1}, C_t, \varepsilon_t\}$ as the state space at t and $\overline{S}_t = \{E_t, h_{t-1}, C_t\}$ as the deterministic part of the state space.

In addition, each individual woman has a set of individual specific state variables that stay fixed over time, or that are assumed to evolve exogenously. This include her skill endowment and her child's cognitive ability endowment, her race and education and her husband's average income. As a result of these variables, each woman faces her own unique optimization problem. However, I will focus only on the endogenously evolving state variables in describing the problem below.

I model mother's decisions from $t = 1$ (the first quarter after the child is born) until T=20. At $T+1$=21 the child reaches 5 years of age and goes to primary school. At any given quarter t after birth, the woman maximizes the expected present value of remaining utilities. The woman's optimization problem can be written in the following way. Let $V(S_t, t)$, the value function, denote the value a mother assigns to choosing alternative $j \in J$ at time t, given the individual's state S_t and discount

[16] All possible combinations of the unobserved types: child's ability (ω_κ), taste for work ($\overline{\alpha}_{2k}$) and for child care ($\overline{\alpha}_{4k}$).

factor β :

$$V(S_t,t)=\max_{d_t^j} E\left[\sum_{\tau=t}^{20}\beta^{\tau-t}\sum_{j=1}^{6}U_t^j d_t^j|S_t\right]. \tag{3.6}$$

The value function can be written as the maximum over alternative-specific value functions, each of which obeys the Bellman equation:

$$V(S_t,t)=\max_{d_t^j}\left\{V^j(S_t,t)\right\} \tag{3.7}$$

with $j\in J=\{(h_t,I_t^c):h_t=0,\frac{1}{2},1 \text{ and } I_t^c=0,1\}$. Each alternative-specific value function $V^j(S_t,t)$ is in turn given by:

$$V^j\left(S_t,t\right)=U_t^j(c_t,d_t^j,A_t)+\beta E\left[V(S_{t+1},t+1)|S_t,d_t^j=1\right] \quad \text{for } t\leq 20 \tag{3.8}$$

At the terminal period, $T+1=21$ quarters, when the child goes to primary school, the nature of the woman's decision problem changes fundamentally, so I do not model decisions beyond that point. Rather, I assume a terminal period value function that is a function of the values of the state variables at $T+1=21$:

$$V^j(S_{T+1})=U_{T+1}^j(c_{T+1},d_{T+1},A_{T+1})+\sum_{\tau=a}^{65}(\beta^4)^{\tau-a}\left(\frac{1}{\alpha_1}\widehat{c}^{\alpha_1}+\alpha_2\right)+\sum_{\tau=5}^{65}(\beta^4)^{\tau-5}\alpha_3\left(\frac{A_{T+1}^{\lambda}-1}{\lambda}\right)$$

where a is the age of the mother by the time her child goes to primary school. This equation indicates that the woman cares about the cognitive ability of her child and her own work experience since it will affect her future earning capacity. In particular, $\widehat{c_i}=E(c_i|w_{iT+1},E_{iT+1},d_{iT},I_{iT+1}^H,\xi_o)$ denotes predicted consumption, which is a function of the state variables at $T+1$ and accounts for the operativeness of the state vector at $T+1=21$ for future behavior of the woman[17]. Specifically:

$$\widehat{c_i}=[E(h_t)*\overline{w}_{iT+1}]+\nu I^H$$

where I^H is the husband's average income, $\overline{w}_{iT+1}$ is the predicted wage of individual i at period $T+1$ given the state variables at $T+1$ and the probability of employment status, $E(h_t)$, is given by a logit in various characteristics of the individual[18].

[17] Estimation results are not sensitive to the specification of the terminal value function.

[18] Estimations of this logit are reported in Appendix 1.3.

4. Solution and Estimation of the Model

4.1. Solution to the Individual's Decision Problem

The individual's optimization problem is solved recursively from the final period T. Consider an individual at period $T-1$ with particular values of the deterministic state space $\overline{S}_{T-1}$. Since she does not know in advance what the draws will be for the stochastic terms ε_T^j that affect the utility from each alternative ($U_T^j(S_T, T)$) then the value of being in a given state S_T at time $T = 20$ is uncertain from the perspective of $T-1$. Once the draws for ε_T^j are realized, she will choose the option with the highest value. Then, using E to denote the expectations operator, the individual must calculate the expected value of being in state S_T at time $T = 20$ from the perspective of time $T-1$. That is the expected maximum of all the time $T = 20$ alternative specific value functions given by (3.8):

$$\begin{aligned} E\max(S_T) &= E\max(V_T^1, V_T^2, V_T^3, V_T^4, V_T^5, V_T^6 | \overline{S}_{T-1}, d_{T-1}^j) \\ &= \int_{-\infty}^{+\infty} \max(V_T^1, V_T^2, V_T^3, V_T^4, V_T^5, V_T^6 | \overline{S}_{T-1}, d_{T-1}^j) dF(\varepsilon) \end{aligned} \tag{4.1}$$

for each $j = 1, ..., 6$. Even when ε_t's are stochastically independent, the Emax function is, in general, a multivariate integral. Furthermore, given that each choice $j = 1, ...6$, leads to a different state at time T, Emax must be calculated at each of these six time T state points for each element of $\overline{S}_{T-1}$.

Having calculated (4.1), the value functions at $T-1$, $V^j(S_{T-1}, T-1)$, are known up to the random draws of the ε_{T-1}'s. The individual receives a set of draws at $T-1$ and chooses the alternative with the highest value. From the perspective of $T-2$, however, this value is uncertain since the individual does not know the draws for ε_{T-1}'s in advance. Hence the individual must calculate in a similar fashion, the expected maximum of the alternative-specific value functions at $T-1$ from the perspective of $T-2$ for every feasible state point:

$$E\max(S_{T-1}) = E\left[\max_{d_{T-1}^j}[V^j(S_{T-1}, T-1)]\right] = E\left[\max_{d_{T-1}^j}(U_{T-1}^j + \beta E\max(S_T | d_{T-1}^j = 1))\right]$$

It is easy to see, that the functions $E\max(S_t)$ are determined recursively. We keep moving backwards, to compute the expected maximum of alternative-specific value functions at every t in a similar fashion.

This backward solution for the $E\max$ functions is repeated until $t = 0$, i.e., the moment of birth. Note that calculating the $E\max(S_t)$ function for any given value of the state space involves a six dimensional integration with respect to the ε_t vector. This calculation is performed by Monte Carlo integration, i.e., for each draw of the ε_t vector from the joint distribution, $\max_{d_t^j}[V^j(S_t, t)]$ is obtained and the sample mean is used as a numerical approximation of $E\max(S_t)$[19].

4.2. Estimation

Consider having data on a sample of individuals who are assumed to be solving the choice model previously described and for whom choices are observed in each of the periods $t = 1, ..., 20$ quarters. In addition, as usual, wages are assumed to be observed only in the periods in which market work is chosen. Then, having solved the dynamic optimization problem facing mothers, we can write the probability that a woman chooses alternative j at time t from her choice set J as:

$$\Pr(d_t^j = 1|\overline{S}_t) = \Pr\left(U_t^j + \beta \text{Emax}(S_{t+1}|d_t^j = 1) \geq U_t^k + \beta \text{Emax}(S_{t+1}|d_t^k = 1),\ \forall k \in J\right) \tag{4.2}$$

namely, the probability that the alternative j value function exceeds the other five. If choice j involves working then a wage will also be observed. Additionally, in some periods a child test score realization will be observed. The likelihood contribution of mother i in period t (age of the child in quarters) is the choice probability times the densities of the wage and the score (if these are observed) and will be given by:

$$L_i = \prod_{t=1}^{t=20} \left[\sum\nolimits_{j \in J} d_t^j \Pr(d_t^j = 1|w_t, \overline{S}_t)\right] .\phi(w_t|\overline{S}_t)^{(f_t + p_t)}.\ f(S_t^A|\overline{S}_t)^{I[S_t^A \ available]} \tag{4.3}$$

where $\phi(w_t|\overline{S}_t)$ is the density of the wage w_t conditional on the state space at t and $f(S_t^A|\overline{S}_t)$ is the density of a given test score S_t^A given the state at t which includes all prior periods' inputs into the cognitive ability production function. The likelihood function for the sample is the product of these probability statements over people.

[19] The algorithm uses 20 draws.

I have assumed that we have available a sample of individuals for whom choices $\{h_t, I_t^c\}$ are observed in each of the periods $t = 1, ..., 20$ quarters. Instead, the NLSY sample that I use contains individuals for whom employment choices are observed for the entire period ($t = 1$ to $t = 20$) while child care choices are observed only for the first three years after the mother gives birth ($t = 1$ to $t = 12$). If we do not observe a woman's child care choice in one period, then we do not fully observe her state space in subsequent periods, because it is not possible to know the value of the cumulative stock of child care use (C_t) with certainty. However it is possible to integrate over unobserved endogenous state variables when forming the likelihood function (see Keane and Wolpin 2001)[20]. Given that the number of possible histories increases significantly over time and the estimation can become burdensome, I use semester periods instead of quarters for the fourth and fifth years after the birth of the child. In order to do this it is only necessary to adjust the discount factor when needed.

Recall that we allow for K different types of individuals. In this case, independence over time is conditional on type. Let π_k be the proportion of the kth type in the population. Then, the procedure will consist of finding the parameter vector that maximizes the weighted average of type-specific likelihood contributions where the weights are the type proportions π_k and are parameters to be estimated (Heckman and Singer 1984). In other words, the likelihood function is a mixture of the type-specific likelihoods, $\sum_{k=1}^{K} \pi_k L_{ik}$, where L_{ik} is the likelihood of person i's observed choice sequence, and corresponding wage and score (if market work and test score are observed) if person i is of tastes/endowments type k.

Maximizing the sample likelihood with respect to the parameter vector would yield consistent and asymptotically normal estimates. Evaluation of the likelihood itself requires the calculation of six-

[20] For example, the probability of observing choice $\{f_{13}, p_{13}\}$ in $t = 13$ for every possible choice of I_{13}^c (which is not observed) will be given by:

$$\begin{aligned}\Pr(f_{13}, p_{13}|w_{13}, \overline{S}_{13}) = \; & \Pr(f_{13}, p_{13}, I_{13}^c = 0|w_{13}, \overline{S}_{13}). \Pr(I_{13}^c = 0|\overline{S}_{13}) \\ & + \Pr(f_{13}, p_{13}, I_{13}^c = 1|w_{13}, \overline{S}_{13}). \Pr(I_{13}^c = 1|\overline{S}_{13})\end{aligned}$$

where $\Pr(I_{13}^c = k|\overline{S}_{13}) = \frac{\Pr(f_{13}, p_{13}, I_{13}^c = k|w_{13}, \overline{S}_{13})}{\sum_{j=0}^{1} \Pr(f_{13}, p_{13}, I_{13}^c = j|w_{13}, \overline{S}_{13})}$, for $k = 0, 1$.

variate integrals[21]. In order to circumvent this problem, we use a GHK recursive probability simulator (Geweke, Keane and Runkle 1994) of the choice probabilities and form a simulated maximum likelihood estimator[22].

4.3. Identification Issues

In this section I discuss how the parameters that determine the effect of maternal work experience and child care use on child outcomes are identified. In the model, accumulated work experience and child care utilization will tend to be correlated with the child's initial cognitive ability level because women with higher skill endowment (which affects the child's cognitive ability outcomes) will tend to work more and use child care more, and they will also tend to have children with higher cognitive ability endowments. Thus, to identify the structural parameters of the cognitive ability production function we require factors that shift mothers' work and child care decisions but are unrelated to her skill endowment and the child's ability endowment. In other words, we need variables that enter the woman's decision rules for work and child care use, but that do not enter the child cognitive ability production function directly.

Note that exogenous variation in day care prices and maternal wages is hard to come by in the data, so a more subtle strategy is adopted here. The basic identification assumption relies on the permanent income hypothesis. The intuition is that short run movements in mothers' and husbands' wage rates enter the mother's working and child care use decision rules, but do not directly affect child's ability (in other words, they do not enter the cognitive ability production function). Only longed lived or permanent components of parents' income affect child outcomes. Two main channels can be thought to link permanent variations in the earnings capacity of parents and children's ability. The first one is that these factors determine resources available for child investments. The second one is that they may be directly correlated with children's skill endowments through genetic transmission

[21] The probability expressions (4.2) are six-variate integrals.

[22] The algorithm uses 25 draws.

or a better capacity to enhance the child's learning abilities. Both channels are incorporated in the model although no attempt is made to separate them.

In particular, the model implies that age of the mother at birth (and age squared), age of the father (and age squared) and mother's previous period employment decisions create short run variation in wages, and therefore in work and daycare decisions, which is not directly correlated with the child's cognitive ability. Consider a variable like mother's age, which generates systematic movement along the life-cycle wage path, holding permanent income constant (MaCurdy 1981). Movement along the age-wage profile is assumed to generate exogenous variation in the mother's working decisions. That is, if we take two otherwise identical mothers and one has a child at a point in the life cycle when her wage rate is higher, that mother will be more likely to work and use day care, ceteris paribus. In fact, the raw data (see Appendix 1.1) seems consistent with this story (if we condition on observables like education). The logic here is similar to that in Heckman and Walker (1990), who found evidence that shifts in the wage-age profile alter the life-cycle timing of work and births.

Age of the mother would be an implausible instrument if it affects the cognitive ability of the child directly[23]. However there is no strong evidence that this is true in either the medical or the developmental psychology literature, except perhaps in the case of teenage mothers. Even in this case, recent studies, including Lopez (2003) and Geronimus et al. (1994), show that differences in test scores of children of young mothers with respect to scores of children of older mothers seem to disappear once family background characteristics are controlled for. By comparing outcomes of children born to sisters who began childbearing at different ages, these authors conclude that mother's age does not have a causal effect on children's performance but rather that other characteristics of teenage mothers like lower levels of education, lower income levels and lower probability of being married account for these differences.

It is important to emphasize that the identification strategy assumes that mother's age (as well as

[23]This assumption is relaxed by including dummies for mother under 18 or over 33 in the production function. However, variation in mother's age between 18 and 33 is assumed to have no direct effect on children's ability.

father's age) is exogenous conditional on skill endowments, education and being older than 18 years (in the case of the mother). It is clear that higher skill people tend to have children later in the life cycle so this distinction is crucial.

An alternative way to estimate the effect of maternal employment and child care decisions on children's cognitive ability would be to estimate reduced form decision rules for both employment and child care choices together with a continuous outcome equation (test scores) and a wage equation. Appendix 1.2 outlines what the equations in this kind of setup would look like. From those, it is easier to see the exclusion restrictions that allow me to identify the key effects in my model. Variables that measure short-run variation in wages are assumed to enter employment and child care use decision rules while they do not directly affect child's outcomes. In particular, age of the mother at birth (and age squared), her employment decisions in the previous period, the child's age and the age of the father at birth (and his age squared) enter the work and child care probit but do not affect the test scores equation. It is interesting to note that in this case one would have to estimate 98 parameters as opposed to a total of 61 in the structural model.

It is worth reminding the reader that the identification assumption also implies that investments in children are made out of permanent income only. The identification assumptions take the life-cycle model seriously in the sense that they imply that transitory income of mothers during the first five years after the birth of the child is trivial for the household's permanent income level. This assumption is plausible given that the sample used consists of married women (and women cohabitating with a life-partner)[24]. This is to say that labor earnings of married women during five years of their entire life cycle net of child care costs amount for a considerably small portion of the household's permanent income. In fact, a typical woman in the sample works about 35 hours per week at an hourly wage of 6.7 (dollars of 1983). A simple calculation indicates that the average wife's labor earnings during the first five years after the birth of the child account for approximately 4.2% of total household permanent

[24] For details about the sample see Section 5.3.

income[25]. This percentage is reduced to 3.6% if the woman's labor income net of child care costs is used in the calculations of the net present value. This evidence suggests that it is plausible to assume that investments in children are made out of permanent income.

5. Data

The data are taken from the 1979 youth cohort of the National Longitudinal Surveys of Labor Market Experience (NLSY). The NLSY consists of 12,686 individuals, approximately half of them women, who were 14-21 years of age as of January 1, 1979. The sample consists of a core random sample and an oversample of blacks, hispanics, poor whites, and the military. Interviews were first conducted in 1979 and have been conducted annually to the present. On a regular basis, the NLSY79 has collected pre- and postnatal care information from the sample of women as they became mothers. Using data from the NLSY79 Workhistory File, it is possible to construct a detailed employment history for each mother in the sample for the period surrounding the birth of her child, i.e., up to four quarters before birth and each quarter interval since the child's birth for a period of five years.

In 1986 a separate survey of all children born to NLSY79 female respondents began. In addition to the data on the mother from the NLSY79, the child survey includes assessments of each child as well as additional demographic and development information collected from either the mother or the child. A battery of child cognitive, socioemotional, and physiological assessments as well as a variety of attitude, aspiration and psychological well-being questions have been administered biennially for children of appropriate age.

5.1. Child Assessments

I use as measures of the child's cognitive ability the scores on the Peabody Picture Vocabulary Test (PPVT) and the Peabody Individual Achievement Test Reading Recognition subtest (PIAT-R) and Mathematics subtest (PIAT-M). Both assessments are among the most widely used for preschool

[25] Calculated as the net present value of both husband's and wife's income until retirement at a 5% discount rate.

and early school-aged children. The PPVT is a vocabulary test for standard American English and provides a quick estimate of verbal ability and scholastic aptitude. The PIAT-M measures attainment in mathematics. Finally the PIAT-R measures word recognition and pronunciation ability.

The PPVT was administered to all children aged three and over. I examined the results for 3, 4 and 5 year olds. The PIAT-M and PIAT-R were administered to children 5 and over each survey year. I used data for children 5, 6 and 7 years old. The analysis is based on the "standard" cognitive assessment scores, which are transformations (on an age-specific basis) of the raw scores.

Table 1 displays average test scores in the NLSY by categories. The first three columns show average PPVT scores for children age 3, 4 and 5 in the PPVT. Columns four to six show standard scores in the PIAT Math subtest and the last three columns show PIAT reading scores for children 5, 6 and 7 years old. Simple t-tests indicate that, on average, children of older mothers (30 years or more), children of more educated mothers (more than 12 years of schooling) and children of white women in the NLSY perform better. With respect to maternal employment one can observe that children of women working (full-time or part-time) during the fourth quarter after birth obtained higher scores on the three tests at all ages. Children of mothers that returned within the first six months after birth to work also performed better. However, if we classify children by working status of their mother, i.e., full-time vs. part-time, there is no significant difference in test scores. Finally, children in the sample used to estimate the model perform better at all ages for reasons that will become clear.

5.2. Employment and Child care after Birth

Maternal employment is measured in the following way. Women reporting between 75 and 375 hours of work per quarter are assumed to be working part-time, women reporting more than 375 hours of work per quarter are assumed to be working full-time and women reporting less than 75 hours of work per quarter are assumed to be staying at home during the period. Women that report having used at least 10 hours per week of some kind of child care[26] are assumed to have used child care during the

[26] Relative or non-relative, day care center, nursery/preschool, regular school

corresponding period.

5.3. The Sample

The sample of women used consists of women with the following characteristics: (1) they worked at least once during the 18 months prior to the birth of the child (so I have a pre-childbirth wage measure available), (2) they did not have a second child before the first child was at least 5 years old and (3) they lived with their husband or co-resident male for the first five years after the birth of the child. The second condition is required to avoid modeling fertility decisions, and to avoid having to model mothers' time allocation among multiple preschool aged children. Presumably the amount of time that a mother can allocate to an individual child will differ, even conditional on her work and day care decisions, depending on how many children she has. Thus, the effect of day care and maternal employment on child outcomes may differ depending on the number of children[27]. The third condition is required to avoid having to deal with issues related to welfare participation that arise because single mothers are generally a low income group. It has been well documented that welfare participation affects single mother's labor supply decisions[28]. The final sample consists of 374 mothers and their children[29]. Admittedly, it is not clear whether the results that I report would generalize to other populations of interest, such as the sample of single mothers. However, given that by 2000 still 68% of children lived in two-parent households in the U.S., it is important to understand the behavior of this set of mothers and the effects of these women's choices on children's performance.

[27] It is worth noting that esentially all the "reduced form" work in this area has ignored this problem as well (i.e., they do not, in general, account for the fact that effects of maternal work and day care may differ depending on number of children).

[28] Both issues, fertility decisions and welfare participation, are undoubtedly very important to analyze when trying to understand mothers' employment and child care choices after birth. However, the computational burden implied by the model would be immensely complicated by the introduction of either of these.

[29] From the original 4,814 mothers, 2,922 worked during pregnancy. From these, only 1,444 had a husband or partner during the entire 5 year period after birth. From the 1,444, 499 did not have a second child before the fifth birthday of their first born. Finally, 125 observations have missing test scores data and have to be excluded for that reason.

Mothers in the sample were older than the average mother in the NLSY by less than a year and were also more educated by around one more year of schooling. About 28% of the sample was hispanic or black. Approximately 46% of women in the NLSY were working at some point during the first year after giving birth, while this proportion is equal to 78% in the sample. The hourly wage before birth was higher for women in the sample and equal to $6.75 (constant dollars of 1983). The average quarterly income of the spouse or partner was slightly higher in the sample ($4,463 vs. $4,307) but the difference is not statistically significant. Finally, women in the sample had on average 1.6 children, while women in the NLSY had 2.8 children on average.

Figure 1 displays employment and child care choices after birth of women in the sample. During the first quarter after birth, about 40% of mothers stayed at home and did not use child care, 33% returned to work (full-time or part-time) and used child care, 22.3% returned to work (full-time or part-time) and reported not having used child care and the remaining 5% stayed at home and used child care. By the end of the third year after birth, 40% of women were working full-time and used child care and 19% continued to stay at home and did not use child care.

6. Estimation Results

The model is estimated by maximizing the likelihood function as written in section 4.2. In order to do this I first solve the dynamic programming problem for each individual given a type and then write the probability expressions derived by comparing current utilities plus discounted future flows of utilities during the remainder of the period. Recall that the number of types (K) is fixed at 8. In assessing the model, I consider the reasonableness of the parameter values and the within-sample fit.

6.1. Inputs of the Model

Both the initial wage equation and the logit, for full-time and part-time probabilities in the terminal condition, are estimated before and used as inputs of the estimation algorithm. Given that all women in the sample were working at some point during pregnancy, initial wages are observed for everybody and it is straightforward to estimate this equation by OLS[30]. In addition, I estimate a logit using individual's characteristics to determine the probabilities of working full-time or part-time by the end of the 5-year period and until retirement[31]. The results are presented in Appendix 1.3. The coefficient on accumulated experience is positive and significant, which means that the probability of working (full-time or part-time) after the child goes to primary school increases with accumulated experience since birth. Similarly, the probabilities of working full-time and part-time increase with the age of the mother by the end of the period and average household income.

6.2. Parameter Estimates

Table 2 reports the estimates of the parameters in the utility function. These results indicate significant heterogeneity among mothers' types. On one hand, with respect to tastes for work, one of the types dislikes work 50% more than the other type, as can be observed from the two different values of α_2 (-7.80 and -12.47). Interestingly, mother's skill type, her education and her race (parameters α_{21}, α_{22} and α_{23}) are not significant determinants of the mother's taste for work. In order to have a clearer interpretation of some of these parameters, I express them in terms of consumption units. For example, for women with a high disutility from work, working full time during a given period reduces consumption by \$3,165.4 (while average consumption is equal to \$7,410), and for women with low disutility from work this reduction is equivalent to \$2,143.

Women are quite different in their tastes for child care as well. While one of the types derives

[30] Women with low education that are observed to be working are likely to have a high skill endowment. In this case, I expect the education coefficient to be biased downwards. However, this problem is not crucial because I am not modeling mothers' educational choices. The parameters that need to be estimated consistently are those in technological relationships (3.3) and (3.4), since women make decisions based on these constraints. Furthermore, my results are not sensitive to changes in the specification of the initial wage equation.

[31] I expect the coefficient on experience to be subject to heterogeneity bias. In other words, women who work more during pregnancy are probably the ones with higher taste for work and would work more after the end of the 5-year period regardless of what they did within this period.

disutility from child care (α_4=-.275), the other type derives a high utility from using child care in any given period (α_4=5.54). Again, mother's skill type, her education and race are insignificant in explaining differences in tastes for child care across women (parameters α_{41}, α_{42} and α_{43}). The disutility from using child care for women who actually dislike doing so is equivalent to \$186.7 while, the utility of using child care in the case of women that like child care is approximately \$1,777.

The cost to a parent of working without using child care is \$1,034. The cost of initiating child care (if never used before) is about \$1,151. The extra cost associated with having to use child care during the first quarter after birth is \$266 and the extra cost of using child care before the child is one year old is approximately \$378. Finally, the cost of child care per quarter is estimated to be \$161 (dollars of 1983) which corresponds approximately to \$291 in 2003. Although this amount may seem small, it is important to remember that this estimation averages over various types of child care which can have very different qualities and prices, including child care provided by relatives (which is in most cases free). Finally, note that the parameters on the cohort trend (parameters ρ_c and ρ_w) are not significant, which means that there are no significant differences in tastes for work or child care between women of younger and older generations.

Table 3 shows the estimates of the wage equation parameters. The experience effect on wages indicates that wages increase by 1% with each additional quarter of experience, which is in line with previous estimates implying that each additional year of experience increases wages by 4% (see Moffitt 1984 and Blau and Kahn 1997). The depreciation is equivalent to 0.4% per quarter, while having worked full-time during the preceding period increases wages by 1.9% and part-time by 1.5%.

The estimation results for the cognitive ability equation are displayed in Table 4. All inputs turn out to have the expected sign and most of them are statistically significant. Estimates of γ_1 to γ_5 have the expected signs. The positive coefficient on education implies that better educated mothers have a better technology for transferring human capital to their children. The fact that the age dummies come out insignificant suggests that mother's age is not directly correlated with children's ability. In this sense, it provides support to the identification strategy according to which the age of the mother creates short run variation in wages (affecting employment and child care decisions) but does not directly affect the child's cognitive ability. These results also indicate significant heterogeneity among children's ability types. High ability type children are 19% more able than low ability ones (4.79 vs. 4.60)[32].

In order to understand the net effect of experience or child care use on the child's cognitive ability, one has to take the derivative with respect to the relevant explanatory variable. For instance, the net effect of experience will be given by: $d\ln A_t/dE_t = 0.000913 - 0.000599\ln A_0$. Figure 2 plots this equation, i.e., the effect of mother's working experience on the child's cognitive ability, as a function of the child's ability endowment ($\ln A_0$). We are only interested in the relevant range of $\ln A_0$ which, given the estimated parameters, is between 4.60 and 4.81 in the sample (the region between the vertical lines). This means that the net effect of experience on the child's cognitive ability is between -0.184% and -0.197% per quarter. In fact, $d\ln A_t/dE_t$ evaluated at the average child's ability endowment ($\ln A_0$) is -0.192%. This implies that an additional year of mother's work experience is associated with a 0.8% reduction in the child's test scores (equivalent to 0.05 standard deviations). Further, note that given that $\gamma_{11} < 0$, the technological return on time spent with the child is higher in the case of high ability children.

The estimates imply that mother's provide a more stimulating environment than any alternative day care provider, and that this effect is stronger for higher ability children[33]. Note, however, that given the specification of the utility function, i.e., the CRRA functional form for child's cognitive ability, we are allowing for a compensation effect in the sense that parents may compensate low ability type children by devoting more time to them, depending on the curvature parameter λ. The net effect can only become clear by studying individuals' choices, which I do in the next section.

Similarly, we can calculate the net effect of child care use on the child's cognitive ability: $d\ln A_t/dC_t =$

[32] Increasing the number of types did not improve the fit of the model. Hence I keep the initial assumption of two discrete types: high and low.

[33] I also find that high ability children are in fact associated with high ability mothers. In this case, this result could also be interpreted as highly skilled mothers time inputs having stronger positive effect on children.

$-0.001780 - 0.000312 \ln A_0$. The effect of total child care use on a child's cognitive ability as a function of the child's ability endowment ($\ln A_0$) is plotted in Figure 3. As can be observed, the net effect of child care use on a child's cognitive ability in the relevant range of $\ln A_0$ is between -0.3214% and -0.3281% per quarter. The net effect evaluated at the average of $\ln A_0$ is equal to -0.3258%. This implies that an additional year of child care use is associated with a reduction of approximately 1.3% in the child's test scores (equivalent to 0.09 standard deviations). Again, given that $\gamma_{12} < 0$, there is a higher technological return to having high ability children spend less time at child care than in the case of low ability children. However, the magnitude of this technological difference is considerably smaller in this case than in the case of maternal employment.

In sum, the total effect of an additional quarter of maternal working experience and child care use on children's test scores is -0.51%[34]. This means that having a mother that works full-time and uses child care during one whole year (within the first five years after the birth of the child) is associated with a reduction in test scores of approximately 2%[35].

Table 5 shows the estimates of type proportions and the discount factor. Low (unobserved) ability endowment children correspond to approximately 35% of the sample and are approximately 19% less able than high ability types. Most of the mothers (86%) dislike having to use child care while 58% of them have a high distaste for work. Finally, the discount factor is estimated to be 0.99, which is reasonable since a period in the model is equal to three months.

I estimated an alternative version of the model in which age of the mother is excluded from the set of instruments an included in the cognitive ability production function[36]. The purpose of this

[34] The first column of the Table in Appendix 1.4 shows the results from running the cognitive ability equation as shown in section 2 by OLS using the same sample of women. Given that two interaction terms are included, i.e., employment and child care interacted with log(initial wage) and mother's education, one needs to take the derivate of log(scores) with respect to employment and child care in order to obtain the effect of mother's decisions on the child's cognitive ability. Evaluated at average log initial wages and average education level, this effect is approximately -0.09% per quarter.

[35] This is equivalent to a reduction of 0.14 standard deviations in cognitive ability test scores.

[36] That means that age of the father, mother's previous period employment decisions and the child's age are still assumed to affect the mother's employment and child care decision rules but not have a direct impact on the child's

estimation is to address a possible concern about age of the mother being directly correlated with the ability of the child. If a mother's readiness to be a parent was determined by her competency and maturity as measured by her age at childbirth, then it would be reasonable to assume that mother's age has a direct effect on children's outcomes. Results of this estimation indicate that an additional quarter of mother's full-time work experience is associated with a reduction of 0.181% (equivalent to 0.012 standard deviations)[37] in child's test scores if the effect is evaluated at the average child's ability endowment ($\ln A_0$). Likewise, an additional quarter of child care use is associated with a reduction of 0.303% in child's test scores (equivalent to 0.021 standard deviations)[38]. Note that both estimates are very similar to the ones obtained from the original model (which excludes mother's age from the ability production function). Interestingly, the parameter associated with mother's age in the ability production function is quantitatively very small (0.00096) and turns out to be insignificant (standard deviation is 0.00098). This provides further evidence in favor of the original assumption that mother's age does not have a causal effect on children's outcomes. Instead, characteristics associated with younger mothers, like lower income, lower education, higher likelihood of being from single-parent homes themselves, lower probability of finishing school and getting married (due to early childbearing) are more likely to be the reasons associated with poorer outcomes of children born to them.

6.3. Model Fit

Figure 4 depicts the fit of the model to the choice distributions in Figure 1, based on a simulation of 8,000 individuals. As can be observed, the model matches the data quite well, in particular, in the case of the most chosen alternatives, i.e., working full-time or part-time and using child care, and staying at home without child care. Table 6 shows the within-sample χ^2 goodness-of-fit test statistics[39]. The tests confirm the graphical results, with the fit being rejected in very few periods.

cognitive ability. This set of instruments still provides a source of identification of the parameters of interest.

[37] This effect was equal to -0.192% in the case in which mother's age was not included in the ability production function.

[38] This effect was found to be -0.325% in the original model.

[39] The χ^2 statistics have not been adjusted for the fact that the parameters of the model have been estimated.

Finally, predicted wages by mother's education and age, as well as predicted log average scores by age and by characteristics of the mother (figures not shown) fit the data quite closely.

6.4. Understanding Unobserved Heterogeneity

As has been emphasized, there is significant heterogeneity among individuals by unobserved characteristics. It would be interesting to try to describe these types even if the model is silent on how types are determined. As was mentioned in an earlier section, according to the parameter estimates, there is a higher technological return of spending time with higher ability children (since the parameter γ_{11} turned out to be negative) but women derive higher marginal utility from spending time with lower ability children (given that $\lambda < 1$). Since these two effects go in opposite directions, whether mothers engage in compensating behaviors such that they spend more time (and or less use child care) with low ability children is an empirical issue that we now turn to discuss.

Table 7 shows the proportion of mothers of low ability endowment children who work compared to the proportion of mothers of high ability endowment children who work. The right panel shows the same comparison in the case of child care use. One can observe that, on average, mothers of low ability children tend to work less and use less child care. For instance, during the first quarter after birth, 5 percentage points (8%) less women work and 5.29 percentage points less women use child care. The same is true for every period after birth. This pattern implies that mothers of low ability children compensate them by spending more time with them, in spite of the higher technological return of investing in high ability children. One can calculate experience and child care use (in number of quarters) of mothers of low ability children and mothers of high ability children by the end of the fifth year after birth. In the case of the former group, average working experience is 10.84 quarters and average day care use is 11.32 quarters. For the later group, average working experience is 11.51 quarters and average day care use is 11.94 quarters. Again, on average, mothers of low ability children accumulate less experience and use less day care than mothers of high ability children. Evaluated at the mean effect, this implies that the net average effect of experience and child care use (over the first five years after birth) on scores is -5.6% (0.36 standard deviations) in the case of mothers of low ability endowment children and -6.1% (0.43 standard deviations) in the case of mothers of high ability endowment children.

Columns (2) to (5) in the Table in Appendix 1.4 show the results of OLS estimations of the cognitive ability production function using simulated data based on the model and the estimated parameters. Column (2) shows the OLS estimation of the same equation presented in column (1)[40]. Interestingly, results from the estimation on simulated data seem to be very close to the estimation on actual data, except for a couple of exceptions like the estimated coefficient on education which seems considerably lower[41]. Once we calculate the total effect of employment and child care on child's cognitive development by taking the total derivative with respect to the corresponding variable, the net effect is approximately -0.15%[42]. Considering that the estimated effect from the structural model is -0.51% one can understand that the model is generating the same selection problem present in the data and it goes in the expected direction.

The third column shows the estimation of the same equation once we condition on both mothers' types and children's types (which are known in the simulation). One can observe that once we control for both types of heterogeneity, the estimated effect of experience and day care on test scores more than doubles from -0.15% to -0.31% per quarter. The intuition for this is that we are presumably controlling for both sources of selection bias once we introduce both types in the regression, i.e., high ability mothers are more likely to have high ability children and more likely to work, and mothers of low ability children tend to compensate them by spending more time with them. Since both sources generate an upward bias it is natural to expect that once we control for both sources of heterogeneity, the estimated effect is more negative. Columns (4) and (5) show the same regression, except mother's

[40] In this case, estimation was done using actual data.

[41] These regressions are based on a simulation of 8,000 individuals whereas the actual data contains 374 observations. The fact that the size of the simulated data is bigger can be the reason that some coefficients turn out to be significant while they were insignificant when estimated on actual data.

[42] When estimated on actual data, the effect of an additional quarter of employment and daycare use was -0.09%.

and children's types are included one at a time. Column (4) includes mother's skill types while excluding child type dummies. Results are very similar to the original equation in column (2) but the total effect of experience and day care increases from -0.15% to -0.2%. Finally, column (5) includes only children's ability types. In this case, the estimated effect increases to -0.27% per additional quarter of experience and child care use. This is interesting, in the sense that it suggests that children's unobserved heterogeneity accounts for a higher proportion of the bias in the original equation than mothers' heterogeneity.

Finally, if maternal employment and child care use are interacted with children's ability endowment types instead of mother's initial wage and education (regression not shown). The results of the estimation on simulated data match quite closely the results from the structural estimation shown in Table 4[43]. In this case, the estimated average effect of one additional quarter of experience and child care use on children's test scores is -0.49% (which was -0.51% in the structural model).

7. Policy Experiments

In this section we evaluate the effect of various policies on women's choice distributions and children's average test scores.

7.1. Child care Subsidy

The first experiment involves a 30% child care subsidy. In particular, the parameter cc is reduced from its estimated value of \$161.3 to \$112.9. As expected, the percentage of women choosing alternatives that include child care increases with respect to the baseline case. On average, there is an increase of 3 percentage points per period in the number of women that now choose to use child care (choice distribution not shown).

However, a priori it is not obvious what will happen with employment choices once the subsidy is

[43] This suggests that including unobserved child ability endowments nonlinearly might be a better way to describe the cognitive ability production function.

introduced. On the one hand, there is a substitution effect in the sense that the availability of cheaper child care might allow women to work more. On the other hand, there is an income effect given that the subsidy increases household income and hence might induce a reduction in the hours of work. The overall effect of the subsidy is to increase the percentage of women working in almost every period after birth by less that 1 percentage point which implies that the substitution effect dominates.

Table 8 displays the percentage difference in average log scores by ability type in the 30% child care subsidy case with respect to the baseline. The results indicate that the introduction of a subsidy is associated with a reduction in test scores for all ages and both types. Given the fact that child care has a negative effect on the child's cognitive ability, the incentive for mothers to move into child care alternatives is detrimental to children's scores even if it seems to increase parents' utility[44].

7.2. Maternity Leave Policy

In this section I analyze a maternity leave policy according to which there is no wage penalty for time out of the labor market after giving birth. In particular, I do this by setting the wage depreciation rate δ at 0. This means that if a woman did not work for a few periods after giving birth, her re-employment wage is drawn from the same wage distribution she had before giving birth[45].

Once the wage depreciation rate is set at 0 a higher percentage of women choose to work full-time relative to the baseline case, while moving out from both home and part-time alternatives (choice distribution not shown). In fact, in period 4 for example, the total percentage of women choosing to

[44] One can calculate the mother's mean expected present value of lifetime utility at t=1 and observe that it increases on average 0.1% once the child care subsidy is introduced.

[45] An alternative way to model the maternity leave policy would be to allow for the possibility of paid benefits during a given leave period. However, in this case the state space would have to be altered to account for the new state variable. In the U.S. the law requires only that employers hold the mother's job for 12 weeks after giving birth but does not mandate paid benefits during this time. Hence it seems reasonable to model this policy as one in which is not possible to discriminate against women depending on the time they spent out of the labor market after giving birth but paid benefits are not available.

work increases by 1.95 percentage points and the proportion of mothers choosing to use child care raises by 1.68 percentage points.

The intuition behind this result is as follows. Women are still getting zero wages during the periods in which they are away from the labor market after birth but the opportunity cost of staying at home has now increased relative to the baseline scenario. Foregone wages are higher during the current period and the discounted stream of future wages has increased as well. The expected gain derived from staying home with their children through their increased cognitive ability is not enough to compensate for the loss in foregone wages and, hence, women choose to work more.

As expected this has the effect of reducing average scores given that mothers are not only working more but also using more child care. Average scores are reduced by approximately 0.2% to 0.5%, depending on the test and age of the child. Mothers' mean expected present value of lifetime utility at time t=1 is increased by 0.66% with respect to the benchmark case. It is difficult to assess whether this type of policy is effective or not given the fact that while it increases women's lifetime utility, it decreases children's test scores which are, in turn, correlated with their future wages.

7.3. Baby Bonus

Finally, we assess the impact on women's decisions and children's average test scores of a $250 quarterly baby bonus after the birth of a child and until he or she is 5 years old[46]. In this case, less women choose alternatives that include work and more women decide to stay at home with their children (choice distributions are not shown).

Table 9 shows the proportion of women who choose each alternative in the baseline case as well as in the $250 baby bonus scenario in period 4. The last column shows the difference (in percentage

[46] To give a few examples, Australia just very recently implemented a baby bonus for a maximum of up to $2,500 per annum over five years. The minimum entitlement is $500 per year. In Singapore, the baby bonus amounts to $3,000 for the second child and $6,000 for the third child. Parents in Japan get a $70 allowance a month for the first two children until they enter pre-school.

points) between these. As can be observed, there is a reduction in the proportion of women working (full-time and part-time) of 1.79 percentage points and a reduction in the proportion of women using child care of almost 1 percentage point. The same pattern can be observed for almost every period after birth until the end of the fifth year. As a consequence of the change in input choices (maternal time and alternative provider's time), average scores increase for all tests and all ages. On average, test scores increase by 0.2% to 0.7% depending on the test and age of the child. At the same time, women's mean expected present value of lifetime utility in period 1 increases by approximately 1%.

8. Conclusions

In this paper I focus on the labor supply and child care decisions of women immediately following birth, in order to evaluate the effects of mothers' decisions on the well-being of their children. In particular, I am interested in assessing the impact of both employment and child care decisions on children's cognitive ability. Previous studies have provided evidence that test scores measured early in a person's life have significant effects on future educational and labor market outcomes[47]. It seems interesting to understand whether there are any parental inputs that can enhance cognitive ability of individuals during early stages of life. For this purpose I use data from the National Longitudinal Survey of Youth and, in particular, I look at the quarterly employment and child care histories of women after birth and until their child goes to primary school at age 5. I assess the impact of these histories on Peabody Picture Vocabulary Tests scores and Peabody Individual Achievement Test scores (Math and Reading Sections).

The key issue dealt with in the paper is the potential endogeneity problem that arises as a result of the existence of unobserved characteristics of both mothers and children. In fact, women are heterogeneous in both the constraints they face and their tastes. At the same time, children are heterogeneous in their cognitive endowments. As we would expect, mothers' decisions with respect

[47] Currie and Thomas (1999) show, for example, that men and women in the lowest quartile of the reading test score (PIAT in the NLSY) distribution have wages 20% lower at age 33 than those who scored in the highest quartile.

to working when children are young, and/or placing children in child care are influenced by these heterogeneous characteristics. Hence, children of working women or children of women who use child care will differ systematically from those whose mothers stay at home or do not use child care. Estimation of a structural model of women's employment and child care choices jointly with a cognitive ability production function (which includes mother's time and child care use as inputs) is suggested as a plausible way of implementing a selection correction. Additionally, the model provides interesting insights about women's dynamic behavior (in terms of employment and child care) after birth and allows us to explore the effects of counterfactual policy experiments on women's choices and children's outcomes[48].

Results suggest that the effects of maternal employment and child care during the first five years of life of the child are not negligible. In fact, an additional year of full-time work is associated with a reduction of about 0.8% in test scores while the impact of an additional year of child care use is a reduction of approximately 1.3% in test scores. This means that having a full-time working mother that uses child care during an entire year (within the first five years after the birth of the child) is associated with a 2% (0.14 standard deviations) reduction in ability test scores. Furthermore, I find that this effect is stronger for high ability children. In other words, there is a higher technological return to spending time with high ability children relative to time spent with low ability children.

The results of the policy experiments suggest that both child care subsidies and maternity leave entitlements can be detrimental for children, while increasing mothers' expected lifetime utility. On the one hand, child care subsidies provide incentives for women who derive high disutility from work and/or from child care to move into alternatives that include child care. This has a negative effect on children's average scores. On the other hand, by setting the wage depreciation at 0, in other words, eliminating the possibility of discriminating against women on the basis of the number of periods they were away from the labor market after giving birth, children are made worse off. The intuition behind

[48] Evaluation of counterfactual policy experiments would be unfeasible if a reduced-form estimation of the cognitive ability production function was performed instead.

this is as follows. Given the fact that women do not receive a wage (or a portion of it) while away from the labor market and foregone wages are higher relative to the case in which the depreciation rate is nonzero, women decide to work more under the new scenario. Not only is the current foregone wage higher but so is the expected stream of future wages. Hence, if the gain derived from increased child's cognitive ability is smaller than the opportunity cost of staying at home, women choose to work more with the expected detrimental effect that this has on average scores.

Finally, the effect of a \$250 quarterly baby bonus after the birth of a child and until he/she goes to primary school at age 5 has a positive impact on both mothers and children. On the one hand, mean expected present value of lifetime utility of mothers is increased by 1% and on the other hand average scores of children increase by about 0.5%. The raise in household's income provides an incentive for women to work less and stay at home with their children at the same time that it acts as a disincentive for child care use. Therefore, the net effect is to increase children's cognitive ability as well as mothers' utility.

In this paper, I have assumed that there is a homogeneous child care supply which includes several types of day care providers. An interesting extension of the model would include quality of child care as a choice variable of the mother. One might argue that the result according to which maternal employment has a significant and sizeable negative effect on children's cognitive ability is driven by the fact that most of the child care provided is of low quality. Clearly, introducing the quality of child care in the model might change the results in very interesting ways. A woman with higher wage might be able to purchase child care services of very high quality in which case it will not be so clear that her time investments will be as valuable.

Finally, one might think that different assumptions about what the mother knows about her child can possibly change the results. They way in which this would happen is not clear a priori. One could allow, for example, the child's unobserved ability endowment (ω_κ) to be a composite of two components one of which is observed by the mother. It is difficult to predict in which direction the results would change but it is plausible to think that this could describe parents' behavior better.

Appendix

1.1. Wages, experience and child care use of mothers by age

	Initial Wages (hourly wage dollars in 1983)			
Age of mother at birth	16-20	21-25	26-30	31-35
Education<=12 / Non-whites	4.03	4.50	4.59	8.78
Education<=12 / Whites	3.95	5.07	6.03	7.75
Education>12 / Non-whites	-	5.97	7.34	8.65
Education>12 / Whites	4.76	5.32	7.18	9.72

	Experience in number of quarters			
Age of mother at birth	16-20	21-25	26-30	31-35
Education<=12 / Non-whites	10.6	15.3	15.4	20.0
Education<=12 / Whites	11.7	14.0	13.8	14.8
Education>12 / Non-whites	-	16.3	14.8	14.8
Education>12 / Whites	14.0	15.0	15.9	16.0

	Child care use in number of quarters			
Age of mother at birth	16-20	21-25	26-30	31-35
Education<=12 / Non-whites	5.6	7.1	7.5	10.0
Education<=12 / Whites	5.5	6.1	6.1	7.2
Education>12 / Non-whites	-	7.5	8.0	7.9
Education>12 / Whites	5.6	8.1	8.4	8.4

NLSY women in the sample described in Section 5.3.

1.2. Reduced-form Model

An alternative way to estimate the effect of mother's employment and child care decisions on the child's cognitive ability is to estimate reduced form decision rules for work and day care together with a continuous outcome equation (test scores) and a wage equation. In this appendix, I describe briefly how the work probit, child care probit, outcome equation and wage equation would look like in such a setup. The work probit:

$$\begin{aligned} V_f^* = \ & \beta_0 + \beta_1 age + \beta_2 age^2 + \beta_3 educ + \beta_4 race + \beta_5 f_{t-1} + \beta_6 p_{t-1} + \beta_7 E_t + \beta_8 C_t + \\ & \beta_9 t + \beta_{10}\mu + \beta_{11} agef + \beta_{12} agef^2 + \beta_{13} BW + \beta_{14} gender + \beta_{15} I[age{<}18] \\ & + \beta_{16} I[age{>}33] + \beta_{17}\xi_0 + \beta_{18} I[C_t{=}0] + \beta_{19} I[t{=}1] + \beta_{20} I[t<5] + \varepsilon_f^* \end{aligned}$$

$$
\begin{aligned}
V_p^* = \;& \beta_{21} + \beta_{22} age + \beta_{23} age^2 + \beta_{24} educ + \beta_{25} race + \beta_{26} f_{t-1} + \beta_{27} p_{t-1} + \beta_{28} E_t + \\
& \beta_{29} C_t + \beta_{30} t + \beta_{31} \mu + \beta_{32} agef + \beta_{33} agef^2 + \beta_{34} BW + \beta_{35} gender + \beta_{36} I[age{<}18] \\
& + \beta_{37} I[age{>}33] + \beta_{38} \xi_0 + \beta_{39} I[C_t{=}0] + \beta_{40} I[t{=}1] + \beta_{41} I[t < 5] + \varepsilon_p^*
\end{aligned}
$$

where *age* is the mother's age at birth, educ is her education, f_{t-1} and p_{t-1} are the previous period employment decisions, E_t is accumulated experience, C_t is child care use, t is the age of the child, μ is father's skill endowment, *agef* is the father's age at birth, BW is the birth weight of the child, *gender* is the child's gender, *I[age<18]* indicates whether the mother is younger than 18, likewise *I[age>33]*, ξ_0 is the mother's skill endowment, $I[C_t = 0]$ indicates whether child care is used in t, $I[t = 1]$ is an indicator function that equals 1 if $t = 1$ and finally, $I[t < 5]$ is defined similarly. The child care probit:

$$
\begin{aligned}
V_c^* = \;& \beta_{42} + \beta_{43} age + \beta_{44} age^2 + \beta_{45} educ + \beta_{46} race + \beta_{47} f_{t-1} + \beta_{48} p_{t-1} + \beta_{49} E_t + \\
& \beta_{50} C_t + \beta_{51} t + \beta_{52} \mu + \beta_{53} agef + \beta_{54} agef^2 + \beta_{55} BW + \beta_{56} gender + \beta_{57} I[age{<}18] \\
& + \beta_{58} I[age{>}33] + \beta_{59} \xi_0 + \beta_{60} I[C_t{=}0] + \beta_{61} I[t{=}1] + \beta_{62} I[t < 5] + \varepsilon_c^*
\end{aligned}
$$

The continuous outcome equation would look like:

$$
\ln S_t = ab0 + \beta_{63} E_t + \beta_{64} C_t + \beta_{65}(ab0 * E_t) + \beta_{66}(ab0 * C_t) + \beta_{67} dPPVT + \beta_{68} dMATH + \varepsilon_s^*
$$

where

$$
ab0 = \beta_{69}\xi_0 + \beta_{70}\text{educ} + \beta_{71}\text{race} + \beta_{72} BW + \beta_{73}\mu + \beta_{74}\text{I[age<18]} + \beta_{75}\text{I[age>33]} + \beta_{76}\text{gender}
$$

and dPPVT and dMATH are test dummy variables.

And finally the wage equation is given by:

$$
\ln w_t = \beta_{77}\text{agem} + \beta_{78}\text{agem}^2 + \beta_{79}\text{educ} + \beta_{80}\text{race} + \beta_{81} t + \beta_{82} E_t + \beta_{83} f_{t-1} + \beta_{84} p_{t-1} + \beta_{85}(E_t * \text{educ}) + \varepsilon_w^*
$$

Assume $\{\varepsilon_f^*, \varepsilon_p^*, \varepsilon_c^*, \varepsilon_s^*, \varepsilon_w^*\}$ have a joint normal distribution $F(\varepsilon)$ and are serially uncorrelated[49]. Together with the parameters in the variance covariance matrix, there a total of 98 parameters in the reduced form model as opposed to 61 parameters in the structural model.

[49] We need two normalization for identification purposes. One in the child care probit and one in the work probit. This means that the total number of parameters in the variance covariance matrix is 13.

1.3. Inputs of the Model

Initial Wage Equation Estimation

	Log(Initial Wage)
Age of mother at birth	0.0623
	(0.03) *
Age of mother at birth^2	-0.0003
	(0.00)
Race of mother	-0.1340
	(0.06) *
Education of mother	0.0575
	(0.02) **
Constant	-0.4533
	(0.94)
No. of observations	374
R^2	0.24

Estimated by OLS

Father's Labor Income

	Log(Father's Wage$_t$)
Age of father in year *t*	0.1649
	(0.02) **
(Age of father in year *t*)^2	-0.0019
	(0.00) **
Constant	5.0330
	(0.33) **
No. of observations	1870
R^2	0.15

Estimated by OLS

Working Probabilities by the end of the 5-year period

	Full-time	Part-time
Age at end of period	2.087	1.641
	(0.53) **	(0.55) **
Age at end of period^2	-0.036	-0.030
	(0.01) **	(0.01) **
Education at end of period	0.038	0.012
	(0.02)	(0.01)
Race of mother	-1.077	-0.854
	(0.36) **	(0.37) *
Accumulated experience during the period	0.170	0.079
	(0.02) **	(0.02) **
Average Household Income	2.8E-05	1.6E-06
	(0.00) *	(0.00)
Estimation	Logit	
No. of observations	374	374
Pseudo-R^2	0.25	0.25

Accumulated experience is calculated as the sum of a dummy equal to 2 if the mother worked full-time during the period, 1 if she worked part-time and 0 otherwise.

1.4. OLS Estimations of the Cognitive Ability Production Function

Dep. Variable-> Log(Test Score)	Actual data	Simulated Data			
	(1)	(2)	(3)	(4)	(5)
Mother's age dummy	-0.0146	-0.0169	-0.0137	-0.0300	-0.0081
(=1 if mother's age<18)	(0.0241)	(0.0057) **	(0.0059) **	(0.0071) **	(0.0048)
Education of mother	0.0158	0.0040	0.0023	0.0070	0.0021
	(0.0051) **	(0.0008) **	(0.0010) **	(0.0012) **	(0.0008) **
Race of mother	-0.0738	-0.0325	-0.0100	-0.0009	-0.0086
	(0.0086) **	(0.0015) **	(0.0016) **	(0.0020)	(0.0013) **
Sex of child	0.0132	0.0032	0.0017	0.0024	0.0010
	(0.0076) *	(0.0015) *	(0.0014)	(0.0017)	(0.0011)
Mother's Skill Type			0.0078	0.0247	
			(0.0034) **	(0.0046) **	
Log(Hourly Wage Mother)	0.0406	0.0561			0.0053
	(0.0184) *	(0.0044) **			(0.0036)
Accumulated working experience of	0.0028	0.0015	-0.0043	-0.0146	-0.0020
the mother + total child care usage	(0.0019) *	(0.0004) **	(0.0011) **	(0.0014) **	(0.0004)
Father's Skill Type	0.0197	0.0301	0.0005	0.0052	0.0018
	(0.0056) **	(0.0018) **	(0.0015)	(0.0018) **	(0.0014)
Child's birth weight	0.0001	0.0001	0.0001	-0.0001	0.0000
	(0.0002)	(0.0000) *	(0.0000) *	(0.0000)	(0.0000)
Dummy PPVT Test	-0.0997	-0.1161	-0.1202	-0.1230	-0.1188
	(0.0096) **	(0.0017) **	(0.0018) **	(0.0022) **	(0.0015) **
Dummy MATH Test	-0.4680	-0.0615	-0.0622	-0.0622	-0.0611
	(0.0090) **	(0.0017) **	(0.0017) **	(0.0021) **	(0.0014) **
(Mother's Skill Type)*(Experience+Child care)			-0.0003	-0.0017	
			(0.0001) **	(0.0002) **	
Log(Mother's Wage)*(Experience+Child care)	-0.0009	-0.0010			-0.0002
	(0.0005) *	(0.0002) **			(0.0001)
(Mother's education)*(Experience + Child care)	-0.0002	-0.0001	-0.0001	-0.0002	-2.9E-05
	(0.0002)	(0.0000) *	(0.0000) *	(0.0000) **	(0.0000) **
Dummy Child Type Low			4.6359		4.5894
			(0.0286) **		(0.0141) **
Dummy Child Type High			4.8362		4.7828
			(0.0286) **		(0.0142) **
Constant	4.2475	4.4610		4.8392	
	(0.0756) **	(0.0170) **		(0.0391) **	
R^2	0.182	0.103	0.389	0.110	0.380
Effect of experience and child care on scores	**-0.09%**	**-0.15%**	**-0.31%**	**-0.20%**	**-0.27%**

Based on a simulation of 8000 individuals given the estimated parameters

(1) Estimated on Actual data, (2), (3), (4) and (5) estimated on simulated data (8000 individuals).

(2) Identical to column (1) using simulated data (3) Conditioning on both mothers' and children's types.

(4) Conditioning on mother's types only (5) Conditioning on children's types only

Table 1

AVERAGE ASSESSMENT SCORES FOR CHILDREN IN THE NLSY

	PPVT			***PIAT - Math***			***PIAT - Reading***		
Child's age in months	36-48	48-60	60-72	60-72	72-84	84-96	60-72	72-84	84-96
Total NLSY	88.2	84.7	90.2	98.5	99.4	99.8	107.0	102.6	103.2
No. of observations	1790	2691	2788	2913	2917	2818	2840	2873	2810
Mother's Age									
0-30 yrs	87.7	84.0	89.9	97.9	99.2	99.6	106.4	102.2	103.0
30 or more years	91.4	88.0	94.1	102.1	101.3	101.8	110.7	106.0	104.7
ttest (Ho:mean1~=mean2)	2.3 **	3.5 **	2.9 **	5.0 **	2.8 **	2.5 **	4.7 **	5.8 **	1.8
Mother's Education									
0-12 years	81.2	72.3	83.2	93.0	94.5	95.7	100.3	98.8	98.5
More than 12 years	90.8	89.2	92.4	100.6	101.6	101.4	109.6	104.3	105.1
ttest	10.0 **	16.2 **	11.1 **	12.7 **	15.0 **	12.2 **	15.0 **	13.6 **	13.0 **
Mother's Race									
White	95.7	95.8	98.3	102.2	103.3	103.0	109.2	104.0	105.2
Hispanic-Black	79.6	73.0	83.6	94.7	96.0	96.8	104.8	101.3	101.2
ttest	19.5 **	28.6 **	22.0 **	13.9 **	16.6 **	14.7 **	7.6 **	67.0 **	8.6 **
Presence of father									
Present first 3 yrs	91.9	89.4	95.5	100.5	101.2	101.1	108.4	103.5	104.4
Otherwise	81.6	76.3	86.3	95.0	96.7	98.1	104.6	101.1	101.6
ttest	11.9 **	14.2 **	12.7 **	9.7 **	9.7 **	6.8 **	6.2 **	6.3 **	6.0 **
Number of Siblings									
0-1	92.2	90.6	94.7	100.8	101.5	101.7	110.3	104.0	105.2
2+	84.7	79.8	87.2	96.6	97.8	98.4	104.3	101.5	101.7
ttest	8.4 **	12.3 **	10.4 **	7.5 **	8.0 **	7.8 **	10.5 **	6.2 **	7.4 **
Birth Order									
Firstborn	92.8	91.3	92.6	100.6	101.1	101.1	110.8	104.2	105.2
Otherwise	85.2	80.7	87.8	97.0	98.2	98.7	104.2	101.4	101.5
ttest	8.5 **	11.9 **	6.7 **	6.3 **	6.3 **	5.7 **	11.4 **	7.1 **	8.1 **
Mother's employment									
Working 4th qtr bef. birth	90.9	88.5	95.1	100.6	101.2	101.3	109.5	103.7	105.0
Otherwise	83.2	77.7	86.4	94.7	96.5	97.7	102.4	100.8	100.5
ttest	8.2 **	11.2 **	12.2 **	10.2 **	10.0 **	8.2 **	11.6 **	7.3 **	9.5 **
Mother returns to work									
0-6 months after birth	90.8	89.9	96.2	101.4	101.3	102.1	110.2	103.7	105.6
6+ months after birth	85.0	78.2	87.2	95.2	97.0	97.8	103.1	101.0	100.8
ttest	6.0 **	12.3 **	10.6 **	10.7 **	8.7 **	9.1 **	11.3 **	6.3 **	9.6 **
Obs	1522	2336	1973	2493	2507	2419	2421	2473	2416
Works full-time	90.6	88.7	95.3	101.5	101.5	101.8	110.2	104.1	105.8
Works part-time	91.6	90.7	95.5	101.3	101.4	102.0	110.2	103.6	105.6
ttest	0.7	1.6	0.2	0.2	0.1	0.3	0.0	0.8	0.3
Obs	739	1196	835	1222	1161	1093	1192	1145	1092
Mother belongs to the sample	96.1	97.9	102.3	104.1	103.9	103.2	113.9	105.5	106.4
Otherwise	87.7	83.8	89.7	98.1	99.1	99.6	106.5	102.4	103.0
ttest	4.3 **	8.9 **	7.3 **	5.4 **	5.3 **	3.6 **	6.7 **	3.6 **	3.2 **

NLSY and author's calculations.

Table 2

Estimation Results - Utility Function

Parameter	Estimate	Std. Errors
Consumption (α_1)	0.363337	(0.002078)
Mother's skill type on taste for work (α_{21})	0.029585	(0.227449)
Mother's education on taste for work (α_{22})	-0.007585	(0.034054)
Mother's race on taste for work (α_{23})	-0.034586	(0.253085)
Cohort trend in work preferences (ρ_w)	-0.001087	(0.002935)
Disutility from work ($\bar{\alpha}_2$) Type I	-12.475475	(1.841760)
Type II	-7.806053	(1.949133)
Utility from Child's ability (α_3)	0.138032	(1.381028)
Ability function (λ)	0.568200	(0.149655)
Mother's skill type in taste for child care (α_{41})	-0.010173	(0.124586)
Mother's education in taste for child care (α_{42})	0.002428	(0.037321)
Mother's race in taste for child care (α_{43})	0.051055	(0.174898)
Cohort trend in child care preferences (ρ_c)	0.003018	(0.002681)
Utility from childcare ($\bar{\alpha}_4$) Type I	-0.275665	(0.186625)
Type II	5.547423	(1.802276)
Cost of no childcare if working (α_5)	-3.708636	(0.061803)
Cost of initiating childcare (α_6)	-4.153411	(0.048892)
Extra cost of using childcare in qtr 1 (α_7)	-0.906821	(0.144037)
Extra cost of using childcare during 1st year (α_8)	-1.303910	(0.101233)
Childcare cost (*cc*)	161.376239	(0.023416)

Table 3

Estimation Results - Wage Equation

Parameter	Estimate	Std. Errors
Depreciation rate (δ)	-0.004263	(0.000518)
Experience (ϕ_1)	0.010082	(0.001067)
Previous full-time exp (ϕ_2)	0.019212	(0.003611)
Previous part-time exp (ϕ_3)	0.015191	(0.002371)
Experience*Education (ϕ_4)	0.000011	(0.000060)
Measurement error (σ_e^2)	0.437863	(0.009286)

Table 4

Estimation Results - Cognitive Ability Production Function

Parameter	Estimate	Std. Errors
Mother's skill type (γ_1)	0.001767	(0.000911)
Mother's education (γ_2)	0.001250	(0.000201)
Child's race (γ_3)	-0.004503	(0.000657)
Father's Skill (γ_4)	0.000060	(0.000069)
Child's birth weight (γ_5)	0.000081	(0.000127)
Mother's too young dummy (γ_6)	-0.007141	(0.024160)
Mother's too old dummy (γ_7)	0.007355	(0.029520)
Child's gender (γ_8)	0.000521	(0.000567)
Cognitive ability Type I	4.603163	(0.070898)
Type II	4.796719	(0.071798)
Experience (γ_9)	0.000913	(0.000286)
Childcare Usage (γ_{10})	-0.001780	(0.002453)
Experience*lnA_o (γ_{11})	-0.000599	(0.000207)
Childcare Usage*lnA_o (γ_{12})	-0.000312	(0.000122)

Table 5

Estimation Results - Estimated Type Proportions

Parameter	Estimate	Std. Errors
High disutility from work $\pi_{\alpha 2l}$	0.585030	(0.166318)
Low disutility from work $\pi_{\alpha 2h}$	0.414970	(….)
Low utility from child care $\pi_{\alpha 4l}$	0.860325	(0.178919)
High utility from child care $\pi_{\alpha 4h}$	0.139675	(….)
Low ability types $\pi_{\omega l}$	0.353814	(0.168029)
High ability types $\pi_{\omega h}$	0.646186	(….)
Discount factor	0.993960	(0.216958)

Table 6

Chi-squared Goodness-of-fit Tests of the Within-Sample Choice Distributions

	CHOICE								
Qtr.	Full-time & no child care	Part-time & no child care	Home & no child care	Full-time & child care	Part-time & child care	Home & child care		Row	
1	4.099	1.183	1.250	1.584	1.740	13.350	*	23.21	*
2	0.733	5.084	8.741	0.004	2.208	0.548		17.32	*
3	3.663	3.796	3.813	0.025	0.000	1.856		13.15	*
4	2.016	3.325	0.915	0.170	1.174	0.148		7.75	
5	0.370	0.015	2.499	0.267	2.832	0.588		6.57	
6	0.509	1.500	0.842	0.634	0.789	0.011		4.29	
7	2.410	1.065	0.136	0.211	0.644	0.768		5.23	
8	0.939	4.348	0.278	0.551	0.811	0.591		7.52	
9	0.988	3.810	2.775	0.070	0.371	0.186		8.20	
10	2.211	5.566	2.097	0.089	0.153	0.154		10.27	
11	5.931	0.696	1.350	0.149	0.625	1.194		9.95	
12	6.154	0.405	0.549	0.012	4.362	0.277		11.76	*

* Statistically significant at 0.05

Table 7

Mother's Compensation Effect

Choices by Child's Ability Endowment

(% of women choosing each alternative)

	Work			Child Care		
t	Low	High	Diff	Low	High	Diff
1	59.06	64.08	5.02	33.42	38.72	5.29
2	67.49	71.26	3.77	48.45	52.78	4.33
3	69.49	72.17	2.67	48.29	52.65	4.36
4	69.66	72.08	2.42	50.83	53.77	2.94
5	72.67	75.11	2.43	65.48	69.92	4.44
6	73.30	75.64	2.34	66.55	70.40	3.85
7	72.97	75.11	2.14	66.93	70.24	3.32
8	71.90	75.11	3.21	66.28	71.02	4.73
16	79.89	83.76	3.87	70.02	71.21	1.20

Based on a simulation of 8.000 individuals.

Table 8

Average Log Scores by Child's Type (30% Child Care Subsidy)

(% difference with respect to baseline)

	All Types	Low Type	High Type	Gap
PPVT1	-0.34	-0.20	-0.47	-0.27
PPVT2	-0.27	-0.22	-0.32	-0.10
PPVT3	-1.41	-0.23	-0.57	-0.34
MATH1	-0.45	-0.13	0.03	0.16
MATH2	-0.35	-0.55	-0.14	0.41
MATH3	-0.45	-0.64	-0.27	0.37
MATH4	-0.25	-0.05	-0.47	-0.42
READ1	-0.33	-0.22	-0.44	-0.22
READ2	-0.13	-0.02	-0.24	-0.22
READ3	-0.67	-0.76	-0.58	0.18
READ4	0.27	-0.13	0.67	0.80

Based on a simulation of 8000 individuals
PPVT1, PPVT2 and PPVT3 correspond to the PPVT score at age 3, 4 and 5 respectively
MATH1 to MATH4 correspond to the PIAT-MATH score at age 5 to 8 respectively
and similarly for READ1 to READ4

Table 9

Choices by Type in Period 4 ($250 Baby Bonus)

(% of people who choose each alternative)

	Baseline	Baby Bonus	Change (% points)
Full-time and no child care	10.39	9.87	-0.52
Part-time and no child care	11.48	11.58	0.09
Home and no child care	25.83	27.22	1.39
Full-time and child care	34.05	32.75	-1.30
Part-time and child care	14.95	14.88	-0.07
Home and child care	3.30	3.70	0.40
Work	**70.87**	**69.08**	**-1.79**
Child care	**52.30**	**51.34**	**-0.96**

Based on a simulation of 8000 individuals

Figure 1
Employment and Child Care Choices after Birth of Women in the Sample

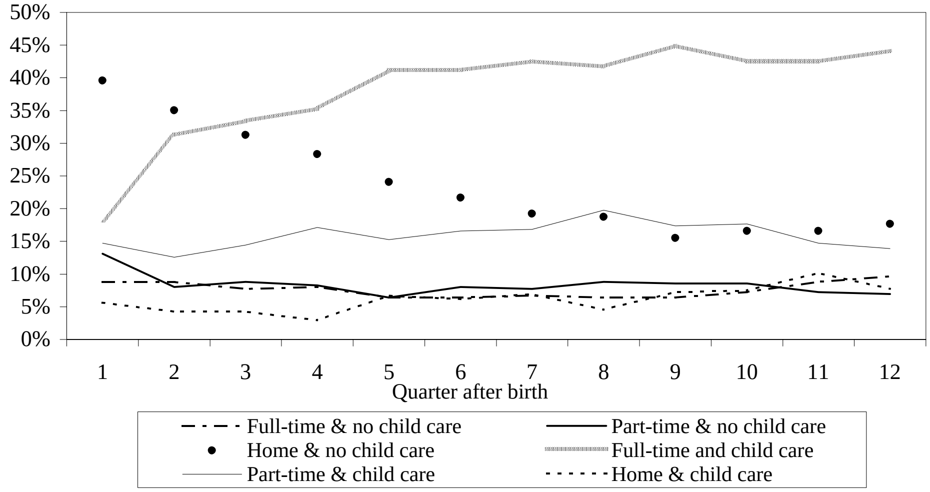


Figure 2
Effect of Working Experience on Cognitive Ability

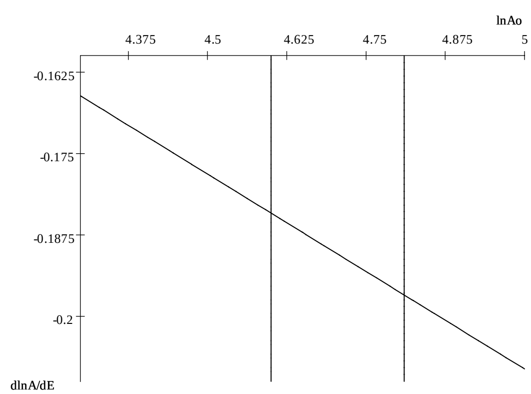


Figure 3
Effect of Child Care Use on Cognitive Ability

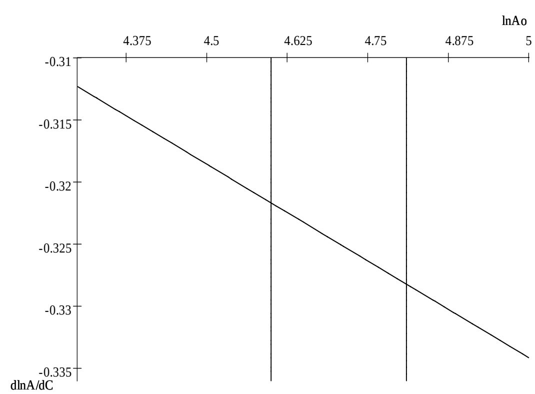
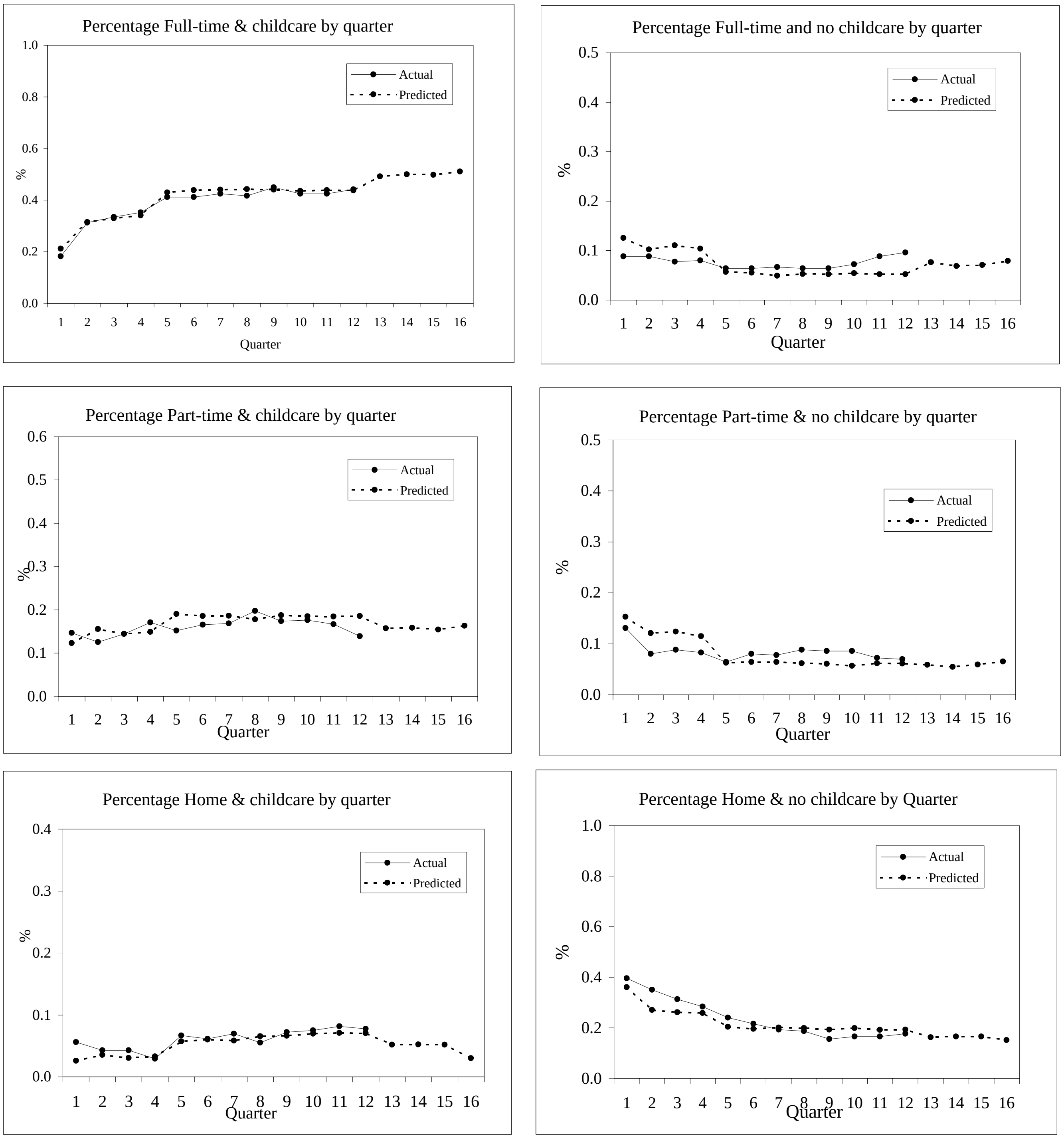


Figure 4
Model Fit to Choice Distributions



References

[1] Baydar, N. and J. Brooks-Gunn, "Effects of maternal employment and child-care arrangements on preschoolers' cognitive and behavioral outcomes: evidence from the children of the National Longitudinal Survey of Youth", *Developmental Psychology* 27 (1991), 932-945.

[2] Blau, F. and A. Grossberg, "Maternal Labor Supply and Children's Cognitive Development", *The Review of Economics and Statistics* 74 (1992), 474-481.

[3] Desai, S., P. Chase-Lansdale and M. Robert, 1989, "Mother or Market? Effects of maternal employment on the intellectual ability of 4-year old children", *Demography* 26 (1989), 541-561.

[4] Ermisch, J. and M. Francesconi, "The Effect of Parents' Employment on Children's Educational Attainment", IZA Discussion Paper No. 215 (2000), 1-35.

[5] Geronimus, A., Korenman, S., and M. Hillemeier, "Does Young Maternal Age Adversely Affect Child Development? Evidence from Cousing Comparisons in the United States", *Population and Development Review* 20 (1994), 585-609.

[6] Geweke, J., M. Keane and D. Runkle, "Alternative Computational Approaches to Inference in the Multinomial Probit Model", *The Review of Economics and Statistics* 76 (1994), 609-632.

[7] Greenstein, T., "Are the 'most disadvantaged' children truly disadvantaged by early maternal employment?", *Journal of Family Issues* 16 (Mar. 1995). 149-169.

[8] Harvey, E., "Short-term and long-term effects of early parental employment on children of the National Longitudinal Survey of Youth", *Developmental Psychology* 35 (1999), 445-459.

[9] Haveman, R. and B.Wolfe, 1995, "The determinants of children's attainments: a review of methods and findings", *Journal of Economic Literature* 33 (Dec. 1995), 1829-1878.

[10] Heckman, J. and G. Sedlacek, "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market", Journal of Political Economy 93 (1985), 1077-1125.

[11] Heckman, J. and B. Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data", *Econometrica* 52 (1984), 271-320.

[12] Heckman, J. and J. Walker, "The Relationship between Wages and Income and the Timins and Spacing of Births: Evidence from Swedish Longitudinal Data", *Econometrica* 52 (1990), 271-320.

[13] James-Burdumy, S., "The effect of maternal labor force participation on child educational attainment", mimeo, John Hopkins University, (October 1998).

[14] Keane, M. and R. Moffitt, "A Structural Model of Multiple Welfare Program Participation and Labor Supply", *International Economic Review* 39 (1998), 553-589.

[15] Keane, M. and K. Wolpin, "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence", *The Review of Economics and Statistics* 76 (1994), 648-672.

[16] Keane, M. and K. Wolpin, "The Career Decisions of Young Men", *Journal of Political Economy* 105 (1997), 473-522.

[17] Keane, M. and K. Wolpin, "The Effect of Parental Transfers and Borrowing Constraints on Education Attainment", *International Economic Review* 42 (2001), 1051-1103.

[18] Leibowitz, A., "Parental inputs and children's achievement", *Journal of Human Resources* 12 (1977), 242-251.

[19] Lopez, R., "Are Children of Young Mothers Disadvantaged Because of Their Mother's Age or Family Background?", *Child Development* 74 (2003), 465-474.

[20] MaCurdy, T., “An Empirical Model of Labor Supply in a Life-Cycle Setting”, *The Journal of Political Economy* 89 (1981), 1059-1085.

[21] Moore, K. and A. Driscoll, “Low-wage maternal employment and outcomes for children: a study”, *The Future of Children* 7 (1997), 122-127.

[22] Moffitt,R., “The Estimation of a Joint Wage-Hours Labor Supply Model”, *Journal of Labor Economics* 2 (October, 1984), 550-566.

[23] Mott, Frank, “Developmental Effects of Infant Care: The Mediating Role of Gender and Health”, *Journal of Social Issues* 47 (1991), 139-158.

[24] Norberg, K., “The effects of day care reconsidered”, NBER Working Paper No.6769 (1998), 1-49.

[25] Parcel, T. and E. Menaghan, “Early parental work, family social capital, and early childhood outcomes”, *American Journal of Sociology* 99 (1994), 972-1009.

[26] Ruhm, C., “Parental employment and child cognitive development”, NBER Working Paper No. 7666 (2000).

[27] Todd, P. and K. Wolpin, “Towards a Unified Approach for Modeling the Production Function for Cognitive Achievement”, *Economic Journal* (February 2003).

[28] Vandell, D. and J. Ramanan, “Effects of early and recent maternal employment on children from low-income families”, *Child Development* 63 (1992), 938-949.

[29] Waldfogel, J., W. Han and J. Brooks-Gunn, “Early Maternal Employment and Child Outcomes: A Longitudinal Analysis of Children from the NLSY”, mimeo, Columbia University (2000).

[30] Willis, R., “Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions”. *In Handbook of Labor Economics*, vol. 1, eds. Orley Ashenfelter and Richard Layard (Amerstardam: North-Holland, 1986).