

Life-cycle and intergenerational effects of child care reforms

MARC K. CHAN

Department of Economics, Faculty of Business and Economics, University of Melbourne

KAI LIU

Faculty of Economics, University of Cambridge, Department of Economics,
Norwegian School of Economics, and IZA

We investigate the importance of various mechanisms by which child care policies can affect life-cycle patterns of employment and fertility among women, as well as long-run cognitive outcomes among children. A dynamic structural model of employment, fertility, and child care use is estimated using Norwegian administrative data. The estimation exploits a large-scale child care reform, which provided generous cash transfers to mothers who did not use formal child care facilities. We find that the reform generates sizable changes in employment and fertility decisions, especially among low-education women. We then use the mothers' unobserved heterogeneity in the structural model as a control function to examine the effects of mothers' behavior on long-run cognitive outcomes of children. The reform leads to lower reading scores among children, primarily as a result of mothers shifting to inferior forms of care. In counterfactual simulations, we compare the effects of an alternative child care subsidy, an expanded maternity leave program, and a tax deduction for mothers with children.

KEYWORDS. Female labor supply, child care reform, fertility, cognitive development of children, discrete choice dynamic programming.

JEL CLASSIFICATION. C35, J13, J22, H31, I28.

1. INTRODUCTION

The dramatic increase in labor force participation among women in the past few decades was accompanied by fundamental changes in how families raise their young children. Stay-at-home mothers became increasingly rare, and the use of nonmaternal child care became increasingly common. In 1950, only 12 percent of married women with children under six were working in the United States; by 2000, this number had increased to over 60 percent (Blau and Currie (2006)). The labor force participation rate

Marc K. Chan: marc.chan@unimelb.edu.au

Kai Liu: kai.liu@econ.cam.ac.uk

We are especially grateful to Kjell Salvanes for his support. We also thank the editor, three anonymous referees, Nina Drange, Chris Ferrall, Drew Griffen, Katrine Løken, Robert Moffitt, David Ribar, Erik Sørensen, Shintaro Yamaguchi, and the participants of various seminars and conferences for helpful comments and fruitful discussions. All remaining errors are ours.

among married and cohabiting women with children has also risen substantially in European countries.¹

In recent years, there has been an increasing advocacy for more government intervention in child care at early ages. In the State of the Union address in 2013, President Obama proposed to make “high-quality preschool available to every single child in America.” In many other economies, child care reforms aiming to provide affordable child care have been or are being implemented (Bennett and Tayler (2006)). However, child care policies are costly to the government due to their generosity and broad coverage of the population. Changes to child care policies have also led to widespread concerns about the consequences for the well-being of children.

In this paper, we use a structural life-cycle model and a large-scale child care reform in Norway to understand the effect of child care policies on life-cycle decisions among women and long-run cognitive outcomes among children. There are a number of key questions being addressed: (i) What are the implications of child care policies for labor supply, child care use, and fertility decisions of women over the life cycle? (ii) What are the effects on long-run cognitive development of children, and which underlying mechanisms are important? (iii) What are the effects of child care policies compared to tax policies and maternity leave programs?

Answering these questions has proven to be quite difficult, as there remains a large knowledge gap in the area. Existing research on the incentive effects of child care policies is largely confined to static analysis with an emphasis on maternal employment.² These studies do not address important issues such as human capital formation and fertility decisions. Although there exists a large literature on dynamic labor supply and fertility decisions (e.g., Moffitt (1984), Hotz and Miller (1988), Francesconi (2002), Gayle and Miller (2006), Adda, Dustmann, and Stevens (2017)), child care decisions are often overlooked. Traditionally, analysis of the effect of child care policies on children’s cognitive outcomes has been limited by the lack of significant policy changes and detailed data.³ Only recently much progress has been made by using specific policy reforms to indirectly evaluate the mechanisms that determine cognitive development (e.g., Bernal and Keane (2011), Dahl and Lochner (2012)).⁴

¹For instance, in Norway, the labor force participation rate among married and cohabiting women with children under 16 years of age increased from 17 percent in 1970 to over 80 percent in 2001 (Ljones (1979), Statistics Norway (2001)).

²For instance, see Blau and Robins (1988), Connelly (1992), Michalopoulos, Robins, and Garfinkel (1992), Ribar (1992, 1995), and Averett, Peters, and Waldman (1997).

³As a result, the literature has found inconclusive evidence that maternal employment can worsen child outcomes (Blau and Currie (2006), Blau (1999), Gregg, Washbrook, Propper, and Burgess (2005)). The literature is also relatively silent on the roles of the underlying mechanisms. On the one hand, maternal employment crowds out parental time with children. On the other hand, it increases family income, which affects child development directly or via the use of nonmaternal child care.

⁴The above studies use welfare reform and the expansion of the earned income tax credit, respectively. Also see, for example, Baker and Milligan (2010), Dustmann and Schönberg (2012), Carneiro, Løken, and Salvanes (2015) on evidence from maternity leave reforms, and Baker, Gruber, and Milligan (2008) and Havnes and Mogstad (2011) on evidence from child care reforms. These studies do not directly estimate the cognitive ability production function of children.

In this paper, we construct a discrete choice dynamic programming model to determine the importance of various mechanisms by which child care policies can affect life-cycle patterns of employment and fertility among women, as well as cognitive outcomes among children. In the model, the women's fertility decisions are formulated jointly with labor supply and child care use decisions; both skill endowments and preferences are subject to heterogeneity. The budget constraint includes income tax, the deduction schedule of child care expenses, and major family transfer programs including maternity leave, child subsidy, and the "cash-for-care" program discussed below.

Our structural model is estimated using administrative data from Norway between 1993 and 2005. We exploit a large-scale child care reform in the period of study as a source of identification to the model.⁵ In 1998, Norway implemented the cash-for-care reform (*kontantstøtte*), which provided cash to families with young children who *did not* use formal child care facilities. The reform provided a means for mothers to substitute cash benefits for formal care, which was heavily subsidized by the government. Payment of benefits did not involve employment restrictions, and mothers could freely decide how to use the benefits. The reform resulted in a large exogenous change in the relative price of child care facilities. The child age restrictions on program eligibility creates variations in exposure to benefits across child cohorts. In addition, the implementation of the reform generates notable differences in life-cycle exposure to benefits across women cohorts. By exploiting the above features and large-scale administrative data, we are able to select several key cohorts that can maximize the variation in the degree of exposure to the reform.

Combining with administrative data on national test scores beyond age 10, we study the implications of various policies for children's long-run cognitive outcomes. There is an extensive literature suggesting that the production of cognitive ability is determined by early inputs, and in the estimation of the production technology, it is important to correct for endogeneity bias resulting from unobserved child-specific endowment effects (Todd and Wolpin (2003, 2007), Cunha, Heckman, and Schennach (2010)). Our cognitive ability production function accounts for the main features that are considered important by this literature. We address the potential source of endogeneity bias by a control function approach, where the estimated structural model is used to predict a mother's unobserved characteristics, in particular, skill endowment (and hence the child endowment), conditional on her observed behavior. Although this greatly simplifies the model, there are also limitations. We discuss the properties of our approach relative to the full structural approach, which jointly estimates the production function with a behavioral model.

Closely related to our paper is a recent important paper by Bernal (2008), who estimates a dynamic model of maternal employment and child care decisions using data

⁵In recent years, progress has been made in synergizing the methodological approaches undertaken by reduced-form and structural studies (e.g., Todd and Wolpin (2006), Attanasio, Meghir, and Santiago (2011), Ferrall (2012) on policy experiments; Kaboski and Townsend (2011), Chan (2013) and Blundell, Dias, Meghir, and Shaw (2016) on large-scale reforms).

from the National Longitudinal Survey of Youth (NLSY79).⁶ Her model focuses on maternal decisions within the first 5 years after child birth.⁷ We attempt to build upon Bernal (2008) along the following dimensions.⁸ First, we incorporate fertility decisions, so that employment, child care use, and pregnancy depend on both the number and the age of children, which are endogenous in the model. Second, we extend the model and data to a life-cycle framework, so as to analyze women's decisions both prior to and after child birth. Finally, by formally utilizing the child care reform in estimation, the identification strategy is potentially more robust than using preexisting variations of women's behavior in the data to identify the model.

We find that child care policies have important implications for the life-cycle decisions of women. By heavily subsidizing mothers, an early exposure to the cash-for-care program can generate sizable changes in employment and fertility decisions, especially among low-education women. If the program begins implementation at age 19, then by age 30, it will have reduced the employment rate of these women by 2.4 percentage points. In addition, although the program substantially reduces formal care use among mothers by 6.8 percentage points, the *overall* rate of formal care drops only slightly due to increased fertility. The total fertility effect is an increase of 0.23 children (+12.8 percent) per low-education woman.

Among the three test subjects—reading, mathematics, and English—we find that the form of child care only affects reading scores at a later age. We consider three categories of care: (i) formal care, (ii) maternal care, and (iii) nonmaternal informal care.⁹ Nonmaternal informal care leads to worse reading skills than formal care, and the gap is particularly large among low-education mothers. For example, among low-education mothers, an additional year of nonmaternal informal care (as opposed to formal care) will reduce the test score by 6.2 percent (0.14 standard deviation (s.d.)). Maternal care is inferior to formal care among low-education mothers, but it has similar technological returns to formal care among high-education mothers. We also find that the skill endowment

⁶Del Boca, Flinn, and Wiswall (2014), Gayle, Golan, and Soyatas (2014), and Griffen (forthcoming) are important recent contributions. In Del Boca, Flinn, and Wiswall (2014), households make labor supply decisions and decide on the allocation of parental time and pecuniary investments in child quality production. They emphasize the importance of parental time inputs on child development, and do not focus on other dimensions such as child care use, fertility, and unobserved heterogeneity. Focusing on explaining racial differences in the intergenerational transmission of human capital, Gayle, Golan, and Soyatas (2014) estimate a dynastic model of parental time and monetary inputs in early childhood with endogenous fertility, home hours, labor supply, marriage, and divorce. Their results suggest significant returns to parental time investment in children mainly through improved education outcomes. Griffen (forthcoming) builds on Bernal (2008) by considering heterogeneity in price and quality of child care programs. In addition, mothers can choose to enroll in Head Start, a federally funded preschool program for poor children. The above features are very important in the United States. By contrast, the Norwegian system is highly homogeneous (see Section 2).

⁷The sample consists of mothers for the first 5 years after the birth of the child and who do not have an additional child during that period.

⁸Using cognitive scores from the preschool period, Bernal (2008) focuses on how early cognitive development can affect mothers' decisions. By contrast, we focus on how early postnatal intervention can affect the child's cognitive outcomes at a much later age. See the estimation section for more details.

⁹Maternal care corresponds to "not employed and not using formal care"; nonmaternal informal care corresponds to "employed and not using formal care." For more details, see Section 3.

of high-education mothers is significantly associated with the child's score: in all subjects, a 0.6 s.d. increase in maternal skill endowment (evaluated at 0 years of experience) is associated with a 0.05 s.d. increase in the child's score. Taking into account various mechanisms, we find that the cash-for-care program reduces the reading score by 1.14 percent (0.03 s.d.) among children of low-education mothers, primarily due to mothers shifting away from formal care. The effect is small among children of high-education mothers.

We examine several counterfactual policies as alternative options, including a partial cash-for-care program in which workers are ineligible for benefits, an expansion of maternity leave, and tax deductions for the presence of children. Although the partial program generates a larger work disincentive than the full program, it has a smaller effect on fertility and is far less expensive to implement. Moreover, the partial program has a small effect on children's reading scores, as mothers move into maternal care. The maternity leave expansion and tax deduction tend to have a more balanced impact in the population and a minimal effect on children's reading scores.

The paper is organized as follows. Section 2 provides an institutional background of the reform. Section 3 describes the structural model. Section 4 presents the data, sample construction, and summary statistics. Section 5 discusses identification and the estimation strategy. Section 6 presents estimation results and conducts counterfactual policy analysis. Section 7 concludes. Additional results are provided in the Appendix, which is available in a supplementary file on the journal website, <http://qeconomics.org/supp/617/supplement.pdf>.

2. INSTITUTIONAL BACKGROUND

Norway offers generous support to families with children. During the period of our study, in the first year since the birth of a child, parents were entitled to 42 weeks of parental leave with full compensation or, alternatively, 52 weeks with 80-percent wage compensation.¹⁰ After the leave expires and until the child comes of school age, family welfare policy focuses on the provision of subsidized child care. When the child becomes 1 year old, families have the option to use child care centers at a price that is heavily subsidized relative to the cost.¹¹ Our definition of formal child care is subsidized child care, which includes both public and private child care centers. These centers are regulated by the same law: they follow the same national curriculum, have the same price schedule, and are equally subsidized (Drange and Rege (2013)). The costs of a day care center are shared between the state, the municipality, and the parents. In 1998, the monthly parental payment was approximately 3500 Norwegian kroner (NOK; equal to

¹⁰In the period of study, 4 weeks out of the 42 weeks of paid parental leave were reserved exclusively for the father (paternity quota). Apart from the exclusive quotas, parents could share the remaining periods of parental leave between them as they desired, with the restriction that mothers and fathers could not both take leave at the same time. The majority of fathers (close to three quarters) take exactly four weeks of the quota (Cools, Fiva, and Kirkebøen (2015)).

¹¹Subsidized child care saw its largest expansion in the late 1970s supported by increased funding from the federal government. From a total coverage rate of less than 10 percent for 3–6-year-old children in 1975, coverage had gone up to over 40 percent by 1985 (Havnes and Mogstad (2011)).

approximately 470 U.S. dollars at the 1998 exchange rate). The Day Care Act (*Barnehageloven*) stipulates national standards regulating both public and private care centers. There are national requirements concerning size, child-to-staff ratio, staff qualifications, playground facilities, and total area within the center. The curriculum is centrally determined, with a strong focus on learning through social relationships both with other children and with adults in the day care centers.¹²

The cash-for-care reform (*kontantstøtte*) was introduced in 1998. According to the legislation, there were three main purposes of this reform: give more freedom of choice to parents regarding the form of care, provide parents with more time to be with their children, and redistribute to families that do not benefit from publicly subsidized day care (*Kontantstøtteloven*, Section 1, 1998). From August 1, 1998, the cash-for-care benefit was available to children aged 13–24 months, and from January 1, 1999, it was expanded to apply to children aged 13–36 months. All parents with children in this age group who do not use publicly subsidized day care are entitled to the subsidy. To receive the full subsidy, the child must not attend a publicly subsidized day care center.¹³ In addition, there is no obligation for parents who claim the benefit to stay at home and care for the children themselves. The subsidy is a flat, tax-free payment, paid out monthly from the month after the child is 1 year old (from month 13), until the month the child is 3 years old (36 months). The subsidy was set to 3000 NOK per month in 1998.¹⁴ The subsidy was approximately equal to the state subsidy for a place in a day care center.

3. ECONOMIC MODEL

The decisions of the adult individual are described as follows. In each decision period t (year), individual i chooses her level of labor supply, which involves no work ($h_{it}^p = h_{it}^f = 0$), part-time work ($h_{it}^p = 1, h_{it}^f = 0$), or full-time work ($h_{it}^p = 0, h_{it}^f = 1$). The employment indicator is denoted by $h_{it} \equiv h_{it}^p + h_{it}^f$. Individuals with fewer than three children can decide whether to become pregnant ($p_{it} \in \{0, 1\}$). In addition, if the individual's first or second child is between ages 1 and 3, she faces the decision of whether to put that child in a formal child care facility ($c_{it} \in \{0, 1\}$).

The choice process is further simplified according to the underlying data structure and policy environment. If the individual has a child of age 0 (i.e., first year following birth), she cannot become employed or pregnant for that period. For a mother whose first and second child are *both* between ages 1 and 3, both children will be in child care if the mother uses child care at all. Therefore, the total number of feasible choices can be

¹²Apart from formal child care that is strictly regulated and publicly subsidized, families may also choose informal care, that is, care provided by nannies or close family such as grandparents. Informal care is ineligible for public subsidy and is not subject to public regulations.

¹³Parents of children who attend publicly subsidized day care on a part-time basis may receive a share of the full benefit depending on weekly attendance. In the data, more than 80 percent of the benefit recipients use day care for less than 10 hours per week.

¹⁴The subsidy was reduced to 2263 NOK per month in 1999, before being adjusted upward to 3000 NOK per month in 2000 and then to 3657 NOK per month in August 2003.

1, 2, 3, 6, or 12 depending on the state variables.¹⁵ The key state variables that determine the choice set are the number of children ($n_{it} = 0, 1, 2, 3$) and the ages of the first and second child ($a_{1it}, a_{2it} = 0, 1, 2, 3, 4$).¹⁶ The latter state variables are crucial because the entitlement of cash-for-care benefits and maternity leave depends on children's age. The individual can have up to three children, and it is assumed that there are no child care choices related to the third child.¹⁷

Utility function

The deterministic part of the utility function takes the form

$$\begin{aligned} \bar{u}_{ijt} = & y_{it} + \alpha_h^p h_{it}^p + \alpha_h^f h_{it}^f + \alpha_c c_{it} + \alpha_p p_{it} + \alpha_{hc} h_{it} c_{it} + \alpha_{hp} h_{it} p_{it} \\ & + (\alpha_{hy} h_{it} + \alpha_{hy}^f h_{it}^f + \alpha_{cy} c_{it} + \alpha_{py} p_{it}) y_{it} + (\alpha_{hc}^f c_{it} + \alpha_{hp}^f p_{it}) h_{it}^f \\ & + (\alpha_{hn1} n_{it} + \alpha_{hn2} \mathbf{1}_{\{1 \leq a_{1it} \leq 3\} \cup \{1 \leq a_{2it} \leq 3\}}) h_{it}^f + \alpha_{hn3} \mathbf{1}_{\{n_{it} > 1\}} h_{it}^p \\ & + (\alpha_{cn1} (n_{it} - 1) + \alpha_{cn2} \mathbf{1}_{\{a_{1it} \geq 3\}}) \mathbf{1}_{\{n_{it} \geq 1\}} c_{it} + (\alpha_{pn1} + \alpha_{pn2} \mathbf{1}_{\{1 \leq a_{1it} \leq 2\}}) \mathbf{1}_{\{n_{it} = 1\}} p_{it} \\ & + \alpha_{hh} h_{i,t-1} h_{it} + \mu_{hj} h_{it} + \mu_{pj} p_{it}. \end{aligned}$$

We assume that the individual consumes all her income each period, that is, income is equivalent to consumption. Her utility thus depends on her income (y_{it}), which is determined by a budget constraint that is discussed in detail below. She faces direct utilities of employment (α_h^p, α_h^f), child care use (α_c), and pregnancy (α_p). We expect α_h^p and α_h^f to be negative because they reflect the opportunity cost of leisure. In addition, workers can face different utilities of child care use and pregnancy than nonworkers (α_{hc}, α_{hp}).

The utility parameters can have the following behavioral interpretation related to the form of child care. The choices represent three types of child care: maternal care ($h = 0, c = 0$); nonmaternal informal care ($h = 1, c = 0$); formal care ($c = 1$). Formal care is directly observed in the data. Maternal care and nonmaternal informal care are not directly observed, but can be inferred from the data by combining women's labor supply decisions and formal child care choices. According to this formulation, the parameter α_c

¹⁵For instance, if the individual's first child is aged between 1 and 3, and the second child is aged 0, then she can neither be employed nor pregnant, so she only faces a binary decision of whether to put the first child in child care.

¹⁶Child's age 4 is an absorbing state; for instance, for child 1, $a_{1it} = 0$ if $n_{it} = 0$ or $n_{it} p_{i,t-1} = 1$, else $a_{1it} = \max\{a_{1i,t-1} + 1, 4\}$.

¹⁷We avoid modeling small categories because they provide limited identifying information regarding the underlying parameter(s) and do not justify the extra computational and modeling burden. The data related to the characterization of the model are described as follows. In the last period of the estimation sample (combining low- and high-education samples), only 5 percent of women have three children, and 2.5 percent have a third child aged 1 or above. Among all women with more than one child, 5 percent have births in consecutive years (first- and second-order births), and 80 percent have a birth gap between 2 and 4 years. Among all observations where child care choice is relevant, 82 percent involve child care choice for one child only. The estimated model is thus most useful for generating predictions that are related to the larger categories in the estimation sample. This does not preclude extending the model; for instance, it is possible to model up to four children in simulation exercises *after* estimation.

can be broadly interpreted as the utility of formal care relative to maternal care; $\alpha_c + \alpha_{hc}$ can be broadly interpreted as the utility of formal care relative to nonmaternal informal care.

In the model, the marginal utility of income can differ by labor supply, child care, and pregnancy status (α_{hy} , α_{hy}^f , α_{cy} , α_{py}). The disutility of full-time work can differ by child care and pregnancy status (α_{hc}^f , α_{hp}^f). The model also allows the utilities to differ by certain state variables. In particular, employment, child care use, and pregnancy depend on both the number and the age of children, with corresponding utilities captured by parameters α_{hn1} , α_{hn2} , α_{hn3} , α_{cn1} , α_{cn2} , α_{pn1} , and α_{pn2} . The motivation is both methodological and empirical. Empirically, women's choice patterns differ nontrivially by the number and age of children, so the empirical model allows for some flexibility along these dimensions.¹⁸ This is important given that the model is estimated by the method of maximum likelihood and the estimation procedure is based on matching the choice probabilities conditional on every state. Methodologically, the specification allows mothers to adjust their employment, child care, and fertility behavior according to the number and age of children.¹⁹ Such adjustments may implicitly reflect the mother's preference over the number and well-being of children; for instance, she may stay at home nurturing a young child and she may even adjust her pregnancy pattern depending on the child's age. However, we emphasize that these adjustments can also be consistent with the hypothesis that mothers *do not* care about their children, for example, the mother may simply enjoy the presence of her child by staying at home.

An alternative approach is to allow for a specific dimension of children's well-being, such as cognitive ability, to enter into the adult's utility function directly. This approach is tractable when there is one child (e.g., Bernal (2008)), but the problem becomes substantially more complicated when fertility is endogenous. The model will require explicit assumptions regarding the woman's information set on each child's cognitive ability. For the model to be tractable, the extra elements may need to enter in a highly stylized way. More importantly, there are practical limitations due to data availability. We only observe a child's test score after age 10, which is well beyond the age when mothers make child care decisions. Because early measures of cognitive ability are unavailable, it is extremely hard to investigate dynamic aspects of child development, that is, how cognitive ability feeds back into mother's behavior. In addition, due to the sample window (see Section 4), the majority of test scores of the second child and almost all test scores for subsequent children in our sample are missing. These limitations prevent us from adopting a more structural approach.

¹⁸Certain parameters, which may appear more arbitrary than the others, are chosen partly due to strong empirical evidence along certain dimensions. We conducted a number of sensitivity analyses with alternative specifications (e.g., including $h_{it}^p n_{it}$), and found qualitatively similar results.

¹⁹Although the utility function does not contain direct terms on the number of children (e.g., n_{it} and n_{it}^2), these parameters can be estimated under the current parametric assumptions, with qualitatively similar results. Preference for children is primarily identified from observed pregnancy patterns in the data. We normalize the preference for no child to zero, so that the preference for the first child is reflected in parameter α_p . The parameter also reflects the direct utility cost of pregnancy. Preference for subsequent children is reflected in parameters α_{pn1} and α_{pn2} .

One main drawback of our approach is that we are unable to disentangle how much of the woman’s behavioral response to the child care reform can be attributed to preferences for the cognitive ability of her children. For example, given the monetary incentives of cash subsidies from the Norwegian reform, we expect mothers to reduce formal care and they may decrease their labor supply. However, if mothers also care about children’s development and if informal care is inferior to formal care, they may choose a smaller change in behavior. Because our estimation sample covers cohorts of women who are exposed to the reform to varying degrees, we are able to measure the *overall* behavioral response from the data. However, we are unable to disentangle the relative roles of both types of incentives. Despite this drawback, we will explain in Section 5.3 how information from the structural model can be combined with the reform to estimate a cognitive ability production function that is “outside” of the structural model.

The parameter α_{hh} captures the degree of state dependence in work preference. The panel feature of the data allows for the modeling of unobserved heterogeneity, which is characterized by several “types” of individuals that differ in unobserved permanent characteristics (e.g., Heckman and Singer (1984)). The model allows for five types of individuals. Individual i knows her own type, which is denoted by $j \in \{1, 2, 3, 4, 5\}$. For this “type- j ” individual, she has type-specific utilities of employment and pregnancy, which are denoted by μ_{hj} and μ_{pj} , respectively. The types are only identified relative to an excluded category. The type-specific utilities of a type-1 individual are normalized to zero, that is, $\mu_{h1} = 0$ and $\mu_{p1} = 0$. The properties of the unobserved types will be discussed further in later sections.

Consider an individual who faces a given choice set. The utility of alternative k , where k is an index representation of the choices, is the sum of the “deterministic” choice-specific utility $\bar{u}_{ijt}(k)$ and a choice-specific preference shock ε_{cikt} :

$$u_{ijk_t} = \bar{u}_{ijt}(k) + \varepsilon_{cikt}.$$

The vector of choice-specific shocks is denoted by ε_{cit} , and is assumed to follow an independent and identically distributed (i.i.d.) extreme value distribution with means at Euler’s constant and standard deviations at $(\pi/\sqrt{6})\sigma_c$, where $\pi/\sqrt{6} \approx 1.2825$ is a normalization constant.

Budget constraint and wage equation

The individual’s income is determined by the budget constraint

$$\begin{aligned} y_{it} = & w_{it}(h_{it} + h_{it}^f) - T(w_{it}(h_{it} + h_{it}^f), D_c(c_{it}, n_{it})) + B_n n_{it} \\ & - (P_{c1} + \mathbf{1}_{\{1 \leq a_{1it} \leq 3\}} \mathbf{1}_{\{1 \leq a_{2it} \leq 3\}} P_{c2}) c_{it} \\ & + B_c Z_{it} (\mathbf{1}_{\{1 \leq a_{1it} \leq 3\}} + \mathbf{1}_{\{1 \leq a_{2it} \leq 3\}}) (1 - c_{it}) - \mathbf{x}'_{cit} \boldsymbol{\beta}_c c_{it} \\ & + (b_m \min\{w_{it}(h_{i,t-1} + h_{i,t-1}^f), \bar{B}_m\} \\ & - T(b_m \min\{w_{it}(h_{i,t-1} + h_{i,t-1}^f), \bar{B}_m\}, 0)) p_{i,t-1} h_{i,t-1} + x_{yit}. \end{aligned}$$

Gross earnings is the product of the wage rate (w_{it}) and work hours ($h_{it} + h_{it}^f$). A full-time worker is assumed to work twice as many hours as a part-time worker. The individual pays income and payroll taxes, which are determined by a piecewise linear tax function $T(\cdot)$ defined in the data section. Income tax is a function of gross earnings as well as the amount of deduction due to expenses related to child care ($D_c(\cdot)$). The individual also receives a child subsidy that pays B_n per period per child.

If the individual uses a formal child care facility ($c_{it} = 1$), she pays P_{c1} for the first child and a discounted price P_{c2} for the second child. If she does not use formal child care ($c_{it} = 0$), she may receive a cash-for-care benefit B_c for each child who is within the eligible age. However, she can only receive this benefit if the cash-for-care program is present ($Z_{it} = 1$ if the policy is in place; $Z_{it} = 0$ otherwise). We include supply-side factors that influence behavior exclusively through the relative price of formal/informal care: (i) the coverage rate of formal child care facilities at the individual's municipality of residence, and (ii) whether the adult individual lives close to her parents (=1 if they are in the same municipality; =0 otherwise). These factors enter into the budget constraint via \mathbf{x}_{cit} .²⁰

The individual is entitled to a maternity leave benefit if she was both pregnant ($p_{i,t-1} = 1$) and employed ($h_{i,t-1} = 1$) last period. The benefit amount is calculated on the basis of earnings last period, which is proxied by $w_{it}(h_{i,t-1} + h_{i,t-1}^f)$.²¹ Thus, the benefit is larger if the individual worked full time instead of part time during pregnancy. The benefit is taxable and it is equal to a proportional adjustment b_m of earnings, up to an earnings cap of \bar{B}_m .

The individual may receive an income from her partner (x_{yit}). If she cohabits with her partner, we set x_{yit} to be her partner's post-tax earnings during the period.²² Otherwise, we set $x_{yit} = 0$. In estimation and counterfactual policy simulations, we assume that the partner's wage is determined outside the model and is invariant to changes in the policy environment. The literature typically finds very small labor supply elasticity with respect to child care cost among men. To the extent that men's labor supply is inelastic especially at the extensive margin, this assumption is less restrictive than it otherwise would be.²³ One drawback is that fathers may change their active or passive time with children (Del Boca, Flinn, and Wiswall (2014)) even though they do not adjust their

²⁰Both factors capture differences in the supply of formal/informal child care across regions. To allow for broad changes over time, a linear function of calendar year is also included. In estimation and simulation, its value is normalized to zero in 1999 and is bounded by the observed years in the sample.

²¹For computational reasons, the model does not include lagged wage ($w_{i,t-1}$) as a state variable. Lagged pregnancy status $p_{i,t-1}$ is defined from other state variables as follows: its value equals to 1 if $n_{it} = 1$ and $a_{1it} = 0$ or if $n_{it} = 2$ and $a_{2it} = 0$.

²²Following Eckstein and Wolpin (1989), Francesconi (2002), and Bernal (2008), we assume that the woman and her cohabiting partner pool their resources together. Under the assumption that father's labor supply is exogenous, our model can be interpreted as a neoclassical household model, where the "household" (which is equivalent to the female individual in our paper) is making choices of maternal employment, child care, and fertility. Ideally, we would have liked to allow only a fraction of the partner's income to enter the budget constraint, but that fraction is likely to be endogenous and determined by a model of strategic interactions within the family.

²³For instance, Del Boca, Flinn, and Wiswall (2014) find that fathers' labor supply is invariant over a wide range of children's ages (between 3 and 15). Over 95 percent of fathers work, except for one-child families

labor supply. It will be ideal if this feature could be incorporated into our model, but it is not possible due to lack of data. Nevertheless, [Del Boca, Flinn, and Wiswall \(2014\)](#) find that the amount of time input by fathers is roughly half of the time input by mothers over a wide range of children's age. The implications for child development are complex because they find that in one-child families, mothers' active time input is substantially more productive (in generating the child's cognitive ability) than fathers' active time input at preschool age, but the evidence is more mixed at a later stage and for two-child families.²⁴

The log wage equation is given as

$$\ln w_{ijt} = \beta_{w0} + \beta_{we1}\mathcal{E}_{it} + \beta_{we2}\mathcal{E}_{it}^2 + \beta_{w1}x_{wit} + \mu_{wj} + \varepsilon_{wit}, \quad (1)$$

$$\mathcal{E}_{it} = \mathcal{E}_{it}^f + \beta_{wp}\mathcal{E}_{it}^p. \quad (2)$$

The wage depends on the individual's "full-time equivalent" work experience (\mathcal{E}_{it}), which is a weighted function of her cumulative periods of part-time work ($\mathcal{E}_{it}^p = \sum_{s=0}^{t-1} h_{is}^p$; $\mathcal{E}_{i0}^p = 0$) and cumulative periods of full-time work ($\mathcal{E}_{it}^f = \sum_{s=0}^{t-1} h_{is}^f$; $\mathcal{E}_{i0}^f = 0$). The parameter β_{wp} determines the weight of part-time experience versus full-time experience. The wage also depends on the unemployment rate in the individual's municipality of residence (x_{wit}), which enters into the model through the wage equation exclusively. A type- j individual has a type-specific skill endowment μ_{wj} , which represents an unobserved permanent component in the log wage equation. For a type-1 individual, we normalize μ_{w1} to zero, that is, $\mu_{w1} = 0$. Therefore, for $j = 2, 3, 4, 5$, μ_{wj} can be interpreted as the relative wage difference between a type- j individual and a type-1 individual, all else being equal. Note that for a type- j individual, her "gross" skill endowment when evaluated at 0 years of experience is given by the term $\beta_{w0} + \mu_{wj}$. The log wage is subject to a normally distributed shock ε_{wit} , which has standard deviation σ_w and is serially uncorrelated and independent of the preference shocks.

Unobserved heterogeneity

The modeling of unobserved heterogeneity serves the following purposes. First, individuals who consistently pursue different choices may differ substantially in unobserved permanent characteristics, and failure to control for this source of difference may result in biased estimates of policy effects.²⁵ Second, the observed wage can be endogenous

with a child at 3 years of age (93.7 percent). In addition, the average weekly working hours range between 43 and 47 hours depending on the child's age. By contrast, mothers' labor supply varies widely by children's age: from 65 to 89 percent, and from 23 to 39 hours, respectively. More relevant in our context, by exploiting the Norwegian cash-for-care reform as a natural experiment, [Bettinger, Hægeland, and Rege \(2014\)](#) and [Drange \(2015\)](#) find that the reform had no significant effect on fathers' labor force participation.

²⁴They find that, in general, the productivity of time inputs declines substantially with child's age. In addition, they find that the active time spent by parents in one-child families is more productive (in their scale, as high as around 0.16 for fathers and 0.24 for mothers), but the productivity of other types of parental time input in one-child and two-child families are generally much lower (lower than 0.1).

²⁵For instance, in [Chan's \(2013\)](#) empirical analysis, he finds that structural models with unobserved heterogeneity tend to generate lower behavioral elasticity measures than models without such features.

due to self-selection into work based on unobserved permanent characteristics (beyond the fact that only workers' wages are observed). Because the extensive margin of female labor supply is important, the panel of observed wages needs to be corrected for self-selection before it can be used to recover the unobserved skill endowment ("ability") in the labor market. Third, the structural model will be used to predict a woman's unobserved skill endowment in the labor market conditional on her observed behavior. This information will enter into the cognitive ability production function as a control function, based on the premise that a mother's skill endowment is correlated with her child's unobserved endowment effect.

For a type- j individual, her overall type-specific characteristics are given by the tuple $(\mu_{wj}, \mu_{hj}, \mu_{pj})$. Due to the role of unobserved heterogeneity in the estimation of the cognitive ability production function, we adopt a relatively flexible specification with five unobserved types.²⁶ In addition, for each individual, the probability of the unobserved types may be associated with her partner's characteristics (education and "permanent income"; see Section 5.2 for details). There are two objectives of this approach. First, the distribution of unobserved heterogeneity can be captured more precisely using the partner's information. Second, it relaxes the implicit assumption that the partner's income (x_{yit} in the budget constraint) is uncorrelated with the woman's unobserved heterogeneity. For example, a high-skill individual (i.e., high μ_{wj}) may be more likely to have a high-skill partner. If that is the case, x_{yit} will be positively correlated with the woman's unobserved skills (i.e., women who receive high x_{yit} will be systematically different from those who receive low x_{yit}). Our model explicitly accounts for this potential correlation, albeit in a restrictive way. Nevertheless, we emphasize that our approach merely captures the association between the woman's unobserved heterogeneity and her partner's characteristics. The association does not necessarily have a causal interpretation.²⁷ Section 5.2 contains further details of our approach.²⁸

²⁶Estimating multiple unobserved types can be demanding due to computational burden and requirements on data. This is feasible in our analysis due to a relatively large sample size and the fact that the cognitive ability production function is not jointly estimated with the structural model. In Bernal (2008), a joint estimation procedure is performed with two unobserved skill endowment types and a sample of 529 mothers. Her sample includes women who live with their husband or co-resident male for the first 5 years after the birth of the child and who do not have an additional child for 5 years after the birth of that child. The sample contains quarterly data on employment for up to 5 years after child birth and child care (an indicator variable including formal or informal care) up to 3 years after child birth. Del Boca, Flinn, and Wiswall's (2014) sample consists of 105 one-child households and 132 two-child households, which are separately used for estimation. Although they perform a joint estimation approach, due to sample size and the short panel that they consider (i.e., two waves of letter-word score data and three waves of parents' data), they do not model unobserved heterogeneity in their analysis.

²⁷The full approach, which we do not pursue, requires estimating a matching model that jointly determines the matching outcome between the woman and her partner. Note that our approach will not be justified if the policy environments considered in the model can affect the matching outcomes substantially.

²⁸The existing literature on female labor supply takes a variety of approaches. Earlier papers such as Eckstein and Wolpin (1989), Van der Klaauw (1996), Francesconi (2002), and Bernal (2008) construct the partner's "expected wage" and put it in the woman's budget constraint only. For example, in Francesconi (2002), the partner's expected wage is constructed as a linear combination of the woman's demographic characteristics. This approach reduces statistical power because it removes most of the sample variation in

Intertemporal optimization problem

At age t_0 , individual i (who belongs to type j) maximizes her expected present discounted value of utility from the current period to the end of the time horizon T ,

$$\max E_{t_0} \sum_{t=t_0}^T \delta^{t-t_0} u_{ijk_{it}t},$$

where δ denotes the discount factor and k_{it} is the individual's choice at period t .²⁹ Generically, the intertemporal optimization problem can be written in recursive form as

$$V_{ijt}(\mathbf{S}_{it}, \boldsymbol{\varepsilon}_{it}) \equiv \max_{k \in C_{it}} [u_{ijk_{it}t} + \delta E_t V_{ij,t+1}(\mathbf{S}_{ik,t+1}, \boldsymbol{\varepsilon}_{i,t+1})], \tag{3}$$

where $V_{ijt}(\cdot)$ is the value function with two sets of state variables \mathbf{S}_{it} and $\boldsymbol{\varepsilon}_{it}$, and C_{it} denotes the index representation of the choice set. The deterministic part of the state space $\mathbf{S}_{it} = (h_{i,t-1}, h_{i,t-1}^f, n_{it}, a_{i1t}, a_{i2t}, \mathcal{E}_{it})$ is carried around explicitly as an argument in the expected value function, and may evolve according to a law of motion.³⁰ The error space $\boldsymbol{\varepsilon}_{it} = (\boldsymbol{\varepsilon}_{cit}, \boldsymbol{\varepsilon}_{wit})$ contains preference and wage shocks that are integrated out in each period of the backward recursion procedure. In the recursion procedure, the length of the time horizon is assumed to be 18 years. In the terminal period, the value function is assumed to be³¹

$$V_{iT} = \psi_{n1} n_{iT} + \psi_{n2} n_{iT}^2 + \psi_{e1} \mathcal{E}_{iT} + \psi_{e2} \mathcal{E}_{iT}^2. \tag{4}$$

The status of the cash-for-care program affects not only the static budget constraint (which determines contemporaneous utility $u_{ijk_{it}t}$), but also the value function. Two sets of value functions are computed. The first set of value functions is computed from the

the partner's wage. Bernal (2008) constructs the expected wage from the partner's average income during the sample period. This measure is likely to be correlated with the woman's unobserved skills.

²⁹The discount factor is set at 0.9 per annum.

³⁰Because \mathcal{E}_{it} is a weighted function of \mathcal{E}_{it}^f and \mathcal{E}_{it}^p , no extra dimensions of the state space are needed. The value function is obtained by interpolation for non-integer values of \mathcal{E}_{it} . In addition, when the individual solves the dynamic programming problem, she perceives the variables outside \mathbf{S}_{it} and $\boldsymbol{\varepsilon}_{it}$ (except for Z_{it}) to remain unchanged over time. In particular, if she cohabits with her partner, she will perceive $x_{yit} = \bar{x}_{yi}$ in all future periods, where \bar{x}_{yi} denotes the partner's "average post-tax earnings." Otherwise, she will perceive $x_{yit} = 0$ in all future periods. We compute \bar{x}_{yi} as the average post-tax earnings of the partner during the calendar years that coincide with the sample periods in the woman's panel. Thus, the measure can include earnings prior to the event of cohabitation. The value function is obtained by interpolation with several grid points of x_y .

³¹The intercept of the terminal value function is normalized to 0 because it is not identified. One reason is that the intercept term affects the terminal values of all states by an equal amount, so it does not influence choices prior to T (choices are affected by differences in terminal values across states). Individuals with no children and no work experience have 0 terminal value; while this can be interpreted as a 0 relative value of the baseline state, the intercept can be normalized to other values without loss of generality. See also Keane and Wolpin (2001) and Fang and Silverman (2009), who adopt a similar specification for the terminal value function (polynomial function of state variables and no intercept) in discrete choice dynamic programming models.

dynamic programming problem *without* the cash-for-care program for all t . Denote this value function by $V(\mathbf{S}, \boldsymbol{\varepsilon}, z = 0)$. The second set of value functions is computed *with* the program for all t . Denote this by $V(\mathbf{S}, \boldsymbol{\varepsilon}, z = 1)$. Throughout the time horizon, the individual faces two different policy environments: one before the cash-for-care reform ($t < \tau_i$ and $Z_{it} = 0$) and the other after the reform ($t \geq \tau_i$ and $Z_{it} = 1$). When $t < \tau_i - 1$ or $t \geq \tau_i$, the individual perceives that the current policy status Z_{it} will remain unchanged in the future. The optimal choice is

$$k_{it}^* \equiv \operatorname{argmax}_{k \in \mathcal{C}_{it}} [u_{ijk_t}(Z_{it}) + \delta E_t V_{ij,t+1}(\mathbf{S}_{ik,t+1}, \boldsymbol{\varepsilon}_{i,t+1}, z = Z_{it})].$$

When $t = \tau_i - 1$ (one period prior to the reform), there is no cash-for-care program in the current period, but the individual perceives that it will be in place starting from the next period. The optimal choice is

$$k_{i,\tau_i-1}^* \equiv \operatorname{argmax}_{k \in \mathcal{C}_{i,\tau_i-1}} [u_{ijk,\tau_i-1}(Z_{i,\tau_i-1} = 0) + \delta E_t V_{ij\tau_i}(\mathbf{S}_{ik\tau_i}, \boldsymbol{\varepsilon}_{i\tau_i}, z = 1)].$$

4. DATA

4.1 Construction of samples

Our data are based on several administrative registers from Statistics Norway covering the entire resident population of Norway. The sample construction process requires linking these registers.³² Below we describe the data and the construction of the final sample used in estimation.

Women's panel Our life-cycle model of women's behavior requires panel data where we observe labor supply, child care choices, and fertility. For each woman in each year, we have information on her hours of work per week and labor earnings.³³ The earnings data are collected from tax records.³⁴ We also observe the individual's country of birth, education level, and municipality of residence. We include only women who were born in Norway. Education is defined at the level measured in 2009 to capture the highest completed level of education. We include only women whose maximum years of schooling is between 10 and 16 years. Municipality of residence is used to define whether the woman lives in the same municipality as her parents and the local coverage rate of formal child

³²Each register contains unique and consistent individual identifiers that allow us to match observations of the same individual across different registers. For example, the Central Population Register, spanning from 1967 to 2010 and updated annually by the local population registries, provides information that allows us to link parents to their children. Given the additional information we have on each child's date of birth, we are able to construct fertility histories in each family.

³³Hours of work per week is reported by the employer each year in one of the following categories: less than 20 hours, 20–29 hours, and 30 hours or more. Earnings is the sum of pre-tax labor income (from wages and self-employment) and work-related cash transfers (unemployment benefits and short-term sickness benefits). Labor earnings are deflator-adjusted prior to estimation.

³⁴Unlike those data from tax records available in most other countries, there is minimal attrition in the Norwegian income data due to the lack of a need to ask permission from individuals to access their tax records. Also, our earnings data pertain to all adult individuals, not only to jobs covered by social security.

care facilities.³⁵ We also use variables related to spousal education and income, and the municipality of residence of the woman's mother.³⁶

Child care choices among mothers are identified from two sources. After the cash-for-care reform in 1998, the cash-for-care register provides detailed benefit receipts for each child in every month. We can construct measures of child-specific child care usage from this register. For the years prior to the reform, child care attendance can be inferred using information on tax deductions for child care expenses from tax records of the parents (Black, Devereux, Løken, and Salvanes (2014), Drange, Havnes, and Sandsør (2016)). Parents are allowed to deduct up to 25,000 NOK (\approx 4310 USD) from taxes in one calendar year for the first child for formal child care. We have access to this data from 1993 to 2005. For years when both sources are available, we define that the mother was using formal child care if she claimed less than 6 months of cash-for-care benefits (therefore using formal child care for at least half of the year) and the child care tax deduction for the household was over 10,000 NOK. Otherwise, the mother is defined as using formal child care when the child care tax deduction exceeded 10,000 NOK.

We construct two separate panels of women, one containing women with high school education and the other including women with college education. Structural estimation will be conducted on each sample separately, thereby allowing each parameter in the model to differ by education group. In the empirical analysis, each panel contains the first 9 years of data since the beginning of the woman's decision-making process, which is defined as the expected age of school graduation. For low-education women, the first decision period starts from age 19. For high-education women, the first decision period starts from age 23.

The earliest and latest calendar years on which the panel data are based are 1993 and 2005, respectively. During this period, there were no other significant changes in income tax schedules and work-family related policies.³⁷ In addition, each panel differs by cohort composition so as to maximize the variation in exposure to the cash-for-care reform throughout the life cycle. In the Appendix, Table A.1 lists the calendar years covered by each cohort in each panel. For low-education women, we select those who were born between 1974 and 1978. As a result, the 1978 cohort will have the longest exposure to the cash-for-reform (since age 20, or period 2); the 1974 cohort will have the shortest exposure to the reform (since age 24, or period 6). For high-education women, we select those who were born between 1970 and 1974. In this sample, the 1974 cohort will have

³⁵The municipality-level (smallest administrative unit in Norway) coverage rate is defined as the number of children aged between 1 and 3 who are in formal care, divided by the population within the same age range.

³⁶For each woman, we link her partner's annual income and highest level of education over the sample period. See the subsection on construction of fertility histories for the household relation variable that is used for linkage.

³⁷Prior to the period we study, there were large expansions in paid parental leave between 1986 and 1993. In 1986, Norwegian parents were granted 18 weeks of paid parental leave. In subsequent years, leave rights were gradually extended to 35 weeks in 1992 and to 42 weeks in 1993 (or 52 weeks with 80-percent pay). From 1993 to 2005, there were no changes to the parental leave policy. Fully-compensated parental leave was extended from 42 weeks to 44 weeks in 2006, and then to 46 weeks in 2009 (or 56 weeks of 80-percent compensation).

the longest exposure to the cash-for-reform (since age 24, or period 2); the 1970 cohort will have the shortest exposure to the reform (since age 28, or period 6). For both panels, the earliest age in which the individual is exposed to the reform ranges from age 20 (1978 cohort) to age 28 (1970 cohort).

Fertility histories and sample selection Construction of fertility histories of women requires linking children with their mothers. From the Central Population Register, we first select all children born between 1991 and 2010. For each child, we have information on the date of birth as well as information from the Birth Register such as gender, birth weight, and household relation at the time of birth. The household relation variable is used to link children with mothers and to select mothers with “stable” household compositions (discussed further below).³⁸

Information on the date of birth of each child is used to construct fertility histories. We exclude women who never had any birth by 2010, the last accessible year in the children’s data. Their fertility decisions may be constrained by other types of factors not captured by the model.³⁹ To avoid any left-censoring of fertility due to the way we construct the panel, mothers who had given birth prior to the age corresponding to the first period in the model are also excluded. To keep the estimation tractable, we keep mothers with no more than three children. To focus on stable households and minimize impacts from changes to household compositions due to divorce and remarriage, we impose two additional selection restrictions. First, we exclude single mothers because there were other welfare programs targeted to single mothers (Mogstad and Pronzato (2012)). Therefore, women in our sample are either cohabiting or married at the time of every child birth.⁴⁰ Second, given the way the father enters into the model, we drop mothers who ever gave birth to a child by 2010 whose biological father is different from the father of her previous child. Hence, for women with more than one child in 2010, all children will have the same father living in the household at the time of every child birth.⁴¹ We compare

³⁸One important advantage of the household relation variable in the Birth Register (especially when compared with data in the United States) is that it clearly distinguishes between cohabiting and married couples at the time of child birth.

³⁹In 2010, the high-education women in the analysis sample were between ages 36 and 40, and the low-education women were between ages 32 and 36. Note that we still have women who are non-mothers during the sample period: they are women whose first birth arrived by 2010 and after the last period in the sample.

⁴⁰Single mothers can be identified at the time of birth from the birth register. During the period we study, close to 50 percent of children were born when parents were married. Around 40 percent of children were born when parents were cohabiting. The remaining 10 percent consists of children born by single mothers or in other types of households.

⁴¹The “stable family” selection criterion is applied until 2010, which is beyond the sample period used for estimation. The sample overrepresents women who have exactly one child as measured *in 2010*, as they are never subject to the selection criterion (i.e., the child has no siblings). In the low-education sample, the proportion of one-child families (as measured in 2010) is 25.9 percent; in the high-education sample, the proportion is 18.3 percent. To gauge the sample selection issue, we try to “rebalance” the sample by imposing a co-residence restriction. Among two- and three-child families (as measured in 2010), the share of partners who are not co-residing in the first 4 years after first child’s birth is less than 1 percent. Among one-child families (as measured in 2010), the corresponding share is 11 percent. Therefore, the degree of overrepresentation is $0.259 \times 0.11 \approx 2.8\%$ in the low-education sample and $0.183\% \times 0.11 \approx 2.0\%$ in the high-education sample. Note that around 40 percent of one-child families (as measured in 2010) do *not* have a child during the sample period for estimation.

characteristics of households excluded from the analysis as a result of family instability and premarital births. Women in our final sample tend to be more educated, work more, and have children at a later age (see Appendix Table A2). The differences are small, with the exception of a noticeably lower employment rate when single mothers are added to the sample.

To ease computational burden in estimation, we use a 10-percent random sample of the matched women–children data. It consists of a balanced panel of 2330 women in the low-education sample and 3465 women in the high-education sample.⁴²

The test score data From 2007 to 2010, we have registry data on national standardized test scores for all children at the beginning of grades 5 and 8 (at ages 10 and 13, respectively). These tests are mandatory, and the test scores are used to provide feedbacks on teaching and inform policy makers on municipalities facing special challenges (Deborah, Lorna, William, and Claire (2011)). The tests include three subjects: (i) general reading in Norwegian, (ii) mathematics, and (iii) English. The tests have been standardized and administered in a consistent manner since 2007. Because the tests are standardized and graded externally, the test scores can be compared across cohorts and schools.

Because the data are available between 2007 and 2010, we can only observe the test scores of those children who were born between 1994 and 2000.⁴³ Children who were born after 2000 would still be lower than grade 5 by 2010, so their scores are unavailable. In the construction of the test score data, we focus on the first or second child of the women in the women's panel, and select those children who were born between 1994 and 2000. Children who were born between 2001 and 2006 still appear in the fertility histories of women's panel, but they are excluded from the test score data.⁴⁴ The final data set contains 1087 children who have a low-education mother, and 2043 children

⁴²There are five cohorts of women in each sample. In the Population Register, the size of each women cohort is around 30,000. First, by keeping women with 10–16 years of schooling, 30 percent of the sample is excluded: around 20 percent of the women has less than 10 years of schooling and 10 percent has more than 16 years of schooling. After this selection criterion, we have a preliminary sample of around 21,000 women in each cohort. From this sample, we exclude women who were born outside Norway (3.8 percent of the preliminary sample; same definition below), who out-migrated within the first 10 years since labor market entry (4.4%), who were non-mothers by the end of 2010 (13.6 percent), who were mothers with more than 3 children (3.9 percent), who were mothers who gave birth before the first period (3.8 percent), who were single mothers (9.3 percent), and who were mothers of children with different biological fathers (4.1 percent). Relative to the *population*, the overall sample selection rate is around 40 percent. This results in a total of around 6000 women (low education + high education) in the final sample ($30,000 \times 5 \text{ cohorts} \times 40\% \times 10\% = 6000$).

⁴³The time frame of the data implies that we observe grade 8 scores of children who were born between 1994 and 1997, and grade 5 scores of children who were born between 1997 and 2000. In estimation, grade 5 scores are used whenever available; otherwise (normalized) grade 8 scores are used.

⁴⁴A more ideal data set should contain children who were born between 1994 and 2006. The pregnancy decisions of their mothers would then occur between 1993 and 2005, which coincides with the calendar years of the women's panel (see Appendix Table A1). The number of children who were born between 1994 and 2000 represents 46 percent of the number of children who were born between 1994 and 2006. This explains the relatively small sample size of the final test score data. Within these child cohorts, 95 percent have nonmissing scores. Note that our children cohorts (1994–2000) coincide with early periods of the mother's life cycle. This implies that we will have a disproportionately high fraction of first-born children in the test score data.

who have a high-education mother. We will further discuss how we use the test score data in Section 5.3.

Policy environment The income tax schedule ($T(\cdot)$) is constructed from tax tables.⁴⁵ The unit of taxation in Norway is an individual. Standard deductions include basic allowance (equal to 20 percent of gross earnings with a ceiling at 32,600 NOK) and personal allowance of 25,000 NOK. The maximum annual tax deduction for child care expenses ($D_c(\cdot)$) was 25,000 NOK for families with one child, and 30,000 NOK for families with more than one child. The base marginal tax rate, imposed on income after the above deductions, is 28 percent. Marginal tax rates of 37.5 percent and 51 percent are levied on individuals with net income greater than 248,000 NOK and 272,000 NOK, respectively. On top of the income tax, there is also a mandatory national insurance contribution, which constitutes 7.8 percent of gross earnings.

Data on other policy parameters are briefly described as follows. The child subsidy (B_n), which is not taxable, is 11,640 NOK per year for each child under 18 years of age. The annual subsidized cost of formal child care for the first child (P_{c1}) is fixed at 28,325 NOK; when two children are in formal child care, the total cost ($P_{c1} + P_{c2}$) is 45,202 NOK (Rauan (2013)).⁴⁶ To align the cash-for-care benefit with the decision period in the model (calendar year), we assume that the child is eligible for three periods from age 1 to 3, with a per-period benefit (B_c) of 24,000 NOK.⁴⁷ The earnings cap for the calculation of maternity leave benefit (\bar{B}_m) is 240,000 NOK, and the rate of coverage is 80 percent ($b_m = 0.8$).

4.2 Sample characteristics

Table 1 reports the choices made at each age by low-education and high-education women, respectively. Among low-education women, the employment rate is 8.9 percent at age 19 (year 1 of the model). By age 23 (year 5), the employment and pregnancy rates have increased to 59.3 percent and 14.3 percent, respectively. The overall formal care usage rate remains low at 2.8 percent. Employment starts to fall steadily beyond age 23, as the fraction of mothers continues to increase. At age 27 (year 9), the employment and pregnancy rates are 55.1 and 18.0 percent, respectively. The overall formal care usage rate increases to 11.5 percent, and the average number of children per individual is 0.98.

⁴⁵We thank Erik Sørensen for providing us with formulas to calculate income tax from annual labor earnings. The tax schedule is constructed using the schedule at 1998. There is no major tax reform in the period we study. Since the last tax reform in 1992, there have been some minor changes and most of the tax payers are unaffected by these changes (Aarbu and Thoresen (2001)).

⁴⁶The price schedule is nationally regulated. Most of the price variations are due to the number of children (i.e., price discount for an additional child if the first child is in formal care). There are remaining regional price variations, which are driven by prices paid by extremely low-income households who are eligible for additional subsidy on the cost of child care (Black et al. (2014)). The average household income in our sample far exceeds the income cutoffs eligible for these additional subsidies; hence we do not consider such price variations in the model.

⁴⁷The cash-for-care benefit eligibility is based on the age of the child, but the decision period is a calendar year. In the data, we observe many individuals who receive cash-for-care benefits for three consecutive calendar years. This is primarily due to children born in the middle of the calendar year. After the alignment, the total amount of benefits that the child is eligible for in the model is $3 \times 24,000 = 72,000$ NOK, which is the same as 24 months of benefits of 3000 NOK.

TABLE 1. Observed choices by age.

Age	Work (%)			Among Mothers: ^a			Non-Maternal Informal Care (%)	Pregnant (%)	Number of Children
	PT or FT	PT	FT	Overall Formal Care (%)	Formal Care (%)	Maternal Care (%)			
<i>Low-education women:</i>									
19	8.9	3.9	5.1	0.0	n/a	n/a	n/a	3.7	0.02
20	30.6	12.2	18.4	0.1	5.1	87.2	7.7	6.9	0.05
21	44.4	15.7	28.7	0.4	7.3	75.0	17.7	8.9	0.12
22	53.2	18.2	35.0	1.1	9.0	65.3	25.6	10.2	0.22
23	59.3	18.5	40.8	2.8	14.7	56.9	28.3	14.3	0.32
24	57.6	16.4	41.2	3.8	15.7	55.1	29.2	13.9	0.47
25	58.7	16.7	42.0	5.3	16.9	50.9	32.2	17.1	0.62
26	56.3	17.1	39.2	7.8	21.6	44.9	33.5	17.4	0.80
27	55.1	17.3	37.9	11.5	27.2	42.6	30.2	18.0	0.98
28	53.2	16.7	36.5	10.3	23.4	42.5	34.0	- ^b	1.17
<i>High-education women:</i>									
23	34.3	11.7	22.6	0.0	n/a	n/a	n/a	5.8	0.04
24	44.0	12.9	31.1	0.8	19.5	56.4	24.1	8.6	0.10
25	54.5	12.3	42.2	2.4	25.4	42.0	32.6	11.3	0.19
26	62.8	13.2	49.5	5.7	33.6	36.1	30.4	14.4	0.30
27	65.9	12.6	53.4	8.7	35.5	32.0	32.5	17.9	0.46
28	64.8	13.2	51.6	12.8	39.6	28.1	32.2	18.3	0.64
29	64.2	13.8	50.4	15.4	37.9	26.9	35.2	19.4	0.84
30	63.5	14.4	49.1	19.3	41.4	24.1	34.5	18.0	1.04
31	63.9	14.5	49.4	25.3	50.4	19.0	30.7	18.7	1.23
32	62.9	15.6	47.2	20.3	41.4	21.4	37.3	- ^b	1.42

^aMothers with at least one child aged between 1 and 3.

^bNot computed from analysis sample.

Among high-education women, the basic choice patterns are qualitatively similar. Nevertheless, they exhibit more intensive employment, fertility behavior, and formal care use. The employment rate is 34.3 percent at age 23 (year 1 of the model). By age 27 (year 5), the employment and pregnancy rates have increased to 65.9 percent and 17.9 percent, respectively. The overall formal care usage rate is 8.7 percent. At age 31 (year 9), the employment rate and pregnancy rates are 63.9 percent and 18.7 percent, respectively. The overall formal care usage rate increases to 25.3 percent, and the average number of children per individual is 1.23.⁴⁸

⁴⁸It is potentially interesting to compare the sample with U.S. women, where a similar class of models has been applied. For instance, consider Keane and Wolpin (2010), who base their estimation on women from NLSY79. In their sample, teenage pregnancy is common; by age 20, white women already have an average of 0.28 children. However, given that the pregnancy rate is generally lower (usually between 4 and 8 percent), the first age at which white women have had one child on average is 27, which is similar to low-education women in Norway. Another interesting difference is the high teenage employment rate. The most rapid increase in the employment rate among white women occurs between ages 16 and 18 (from 30 percent to 63 percent), and the employment rate increases steadily afterward, to 72.6 percent at ages 30–33.

Table 1 also reports the forms of care chosen by mothers with at least one child aged between 1 and 3. Among low-education mothers, maternal care is consistently the most popular form of care, which is followed by nonmaternal informal care and formal care. Among high-education mothers, the same pattern is observed at an early age; however, starting from age 27, formal care has become the most popular form of care, and maternal care has become the least popular form of care.

Table 2 reports the summary statistics of other selected variables that are used in the analysis. High-education workers earn an average of 218,000 NOK per year, which is 23 percent higher than the average earnings of low-education workers (177,000 NOK). A high-education woman is more likely to have a partner with higher levels of education and income. They are less likely to have a parent who lives in the same municipality.

TABLE 2. Summary statistics of selected variables.

Variable	Low-Education Women		High-Education Women	
	Mean	Std. Dev.	Mean	Std. Dev.
<i>Panel Data of Adult Women:</i>				
Yearly gross earnings (workers only, 100,000 NOK)	1.77	0.48	2.18	0.67
Local child care coverage rate (%)	48.50	11.67	50.32	11.54
Grandparent lives close by (%)	65.67	47.48	49.41	49.99
Local unemployment rate (%)	2.45	0.57	2.46	0.59
Partner's education (=1 if more than high school) (%)	11.16	31.49	41.21	49.22
Partner's "permanent income" (100,000 NOK/year) ^a	2.19	0.82	2.66	1.03
Number of adult individuals	2330		3465	
<i>Children Data for Test Score Regressions:</i>				
Ln (Child's reading test score)	2.86	0.45	3.04	0.40
Ln (Child's mathematics test score)	3.17	0.42	3.36	0.37
Ln (Child's English test score)	3.00	0.40	3.14	0.36
Child's gender (= 1 if male) (%)	51.83	49.98	50.19	50.01
Child's birth weight (kilograms)	3.54	0.59	3.56	0.58
Mother under age 21 at birth (%)	9.35	29.10	0.00	0.00
Sibling born within 4 years of first child's birth (%)	74.49	43.60	81.80	38.59
First child in the family (%)	77.33	41.88	74.07	43.83
Child's school quality quintile:				
1st, lowest (%)	18.44	–	11.66	–
2nd (%)	24.31	–	20.14	–
3rd (%)	24.68	–	24.45	–
4th (%)	20.73	–	24.01	–
5th, highest (%)	11.83	–	19.75	–
Number of children ^b	1087		2043	

^a Defined as the average gross earnings of the partner during the calendar years that coincide with the sample periods in the woman's panel.

^b Children who were born between 1994 and 2000 and have a nonmissing reading score.

Unfortunately, the child care choices are not directly comparable due to differences in data definitions: the child care variable in NLSY79 is an indicator for whether nonmaternal care (formal or informal) is used for at least 10 hours per week during the last month (Bernal (2008)). In Bernal's sample of NLSY79 women, the child care (formal or informal) usage rate is around 60 percent by the end of the third year after child birth.

The local child care coverage rate and unemployment rate are similar between low- and high-education women.

The average difference in log reading scores between the children of low- and high-education women is 0.18 (i.e., $3.04 - 2.86 = 0.18$), which is less than 0.5 standard deviations.⁴⁹ Similar differences (measured in s.d.) are found in log math and English scores. Children of high-education women are more likely to attend a school with higher quality. They have similar birth weight to children of low-education women. In our sample, the overall fraction of first-born children is around 75 percent.

5. IDENTIFICATION AND ESTIMATION

5.1 *Policy reform and the structural model*

The cash-for-care reform resulted in a large exogenous increase in the relative price of formal child care facilities. In Section 4, we discussed the selection of women cohorts who are different in the degree of life-cycle exposure to the reform. In the discussion below, we focus on the child age restrictions on cash-for-care eligibility, which also create variations in exposure to benefits across child cohorts. As an illustration, we focus on the variations implied by the *model*. In the model, the cash-for-care program is restricted to children between ages 1 and 3. Children who were born in 1994 or earlier were never exposed to benefits because they were aged 4 or older when the program started in 1998. Children who were born in 1995 turned age 3 in 1998, which implies that they were exposed to one period of benefit eligibility. By a similar argument, children who are born in 1996 were exposed to two periods of benefit eligibility, whereas children who were born in 1997 or later were exposed to three periods of benefit eligibility.

Thus, there is variation in the exposure to benefit eligibility among mothers whose children were born in 1998 or earlier. This variation results in different incentives for employment and child care use. In addition, the reform was *unanticipated* by all these mothers. This is because the reform was announced around 6 months before implementation, and the pregnancy decision was made one period prior to child birth. By contrast, for mothers whose children were born in 1999, the reform was *anticipated* when the pregnancy decision was made. Even though they face the same benefit exposure as mothers whose children were born in 1998, there is a notable difference in fertility incentives due to the anticipation. These sources of variations can help identify the effect of the reform on employment, child care use, and fertility.

Apart from the above policy features, the identification argument of the structural model is rather standard and has been discussed in the literature. For instance, because wages are observed for workers only, the distributional assumptions on the choice and wage shocks are important for identification, as in a similar class of models. Nevertheless, exclusion restrictions are included for robustness: these include the local unemployment rate in the wage equation, and the local child care coverage rate and proximity

⁴⁹Although the summary statistics are not directly comparable, Griffen (forthcoming) finds that the average difference in children's cognitive skills (measured at preschool ages) is 0.39 s.d. between mothers with a high school diploma or less and mothers with at least some college education. We would like to thank Drew Griffen for the discussion of this point.

to grandparents in the budget constraint (via affecting child care). Note that the income tax schedule creates piecewise linearity in the budget constraint; in addition, the policy reform affects the budget constraint and decisions in various ways as described above.

5.2 Likelihood function of the dynamic programming model

The model is estimated by the method of maximum likelihood. To compute the likelihood function, the unobserved types have to be taken into account and integrated out. The probability that the individual belongs to type j takes the multinomial logit form

$$\Pr(j|\mathbf{x}_{ai}) = \begin{cases} \frac{\exp(\mathbf{x}_{ai}\boldsymbol{\beta}_{aj})}{1 + \sum_{l=2}^5 \exp(\mathbf{x}_{ai}\boldsymbol{\beta}_{al})}, & \text{if } j = 2, 3, 4, 5, \\ 1 - \sum_{j=2}^5 \Pr(j|\mathbf{x}_{ai}), & \text{if } j = 1. \end{cases} \quad (5)$$

The vector of covariates \mathbf{x}_{ai} includes a unit constant as well as the partner's "permanent" income and education. The *permanent income* is defined as the average gross earnings of the partner during the calendar years that coincide with the sample periods in the woman's panel. Thus, it may include earnings prior to the event of cohabitation. The partner's education is defined as the highest completed years of education during the same horizon. Both measures reflect the partner's "ability," and they may be associated with the probability that the woman belongs to a certain ability type.

The model parameters reside in the utility function, wage equation, type probability function, and terminal value function. There are 58 parameters in total. The type-specific parameters are $(\mu_{wj}, \mu_{hj}, \mu_{pj}, \boldsymbol{\beta}_{aj})$, where $j \in \{2, 3, 4, 5\}$. To facilitate the interpretation of estimation results, we rank the types in descending order according to the unobserved ability in the wage equation, so $\mu_{w5} < \mu_{w4} < \mu_{w3} < \mu_{w2} < \mu_{w1} = 0$ (i.e., type-1 individuals have the highest ability). Although the types are discrete, the tuple $(\mu_{wj}, \mu_{hj}, \mu_{pj})$ differs across types so that the unobserved taste for work, taste for pregnancy, and ability can be correlated.

The likelihood function is constructed as follows. Consider individual i in period t . Conditional on her wage w_{it} (observed for workers), state variables \mathbf{S}_{it} , and type j , the alternative-specific value to alternative k that is *exclusive* of the preference shock can be defined as

$$\begin{aligned} \bar{V}_{ikt}(w_{it}, \mathbf{S}_{it}, j) &\equiv \bar{u}_{ijt}(k; w_{it}, \mathbf{S}_{it}) \\ &+ \delta E_t V_{ij,t+1}(\mathbf{S}_{ik,t+1}, \boldsymbol{\varepsilon}_{i,t+1}). \end{aligned} \quad (6)$$

Due to the distributional assumption for the choice shock, the choice probability is

$$P_{ikt}(w_{it}, \mathbf{S}_{it}, j) \equiv \frac{\exp(\bar{V}_{ikt}(w_{it}, \mathbf{S}_{it}, j)/\sigma_c)}{\sum_{l \in C_{it}} \exp(\bar{V}_{ilt}(w_{it}, \mathbf{S}_{it}, j)/\sigma_c)}. \quad (7)$$

Given observed choice k_{it} , which includes information about work status h_{it} , the likelihood contribution is

$$L_{ik_{it}}(j) = \begin{cases} P_{ik_{it}}(w_{it}, \mathbf{S}_{it}, j)f(w_{it}|\mathbf{S}_{it}, j), & \text{if } h_{it} = 1, \\ \int P_{ik_{it}}(w, \mathbf{S}_{it}, j)f(w|\mathbf{S}_{it}, j) dw, & \text{if } h_{it} = 0, \end{cases} \quad (8)$$

where $f(\cdot)$ is the probability density function of the wage equation.⁵⁰ In a panel with N individuals and \mathcal{T} periods, the log likelihood function is

$$LL = \sum_{i=1}^N \ln \sum_{j=1}^5 \Pr(j|\mathbf{x}_{ai}) \prod_{t=1}^{\mathcal{T}} L_{ik_{it}}(j). \quad (9)$$

5.3 Cognitive ability production function

5.3.1 Accounting for the endogeneity of inputs Our objective is to obtain unbiased estimates of the effect of early maternal inputs on the child’s cognitive ability, as measured by his/her test score beyond age 10. Before discussing the full empirical specification, the key sources of potential bias and identification are first highlighted. For illustration purposes, consider the stylized model

$$s_{iT} = \gamma_0 + \mathbf{x}'_{si} \boldsymbol{\gamma}_1 + \mathbf{Y}'_{it} \boldsymbol{\gamma}_2 + \nu_i + \varepsilon_{siT}, \quad (10)$$

where s_{iT} denotes child i ’s cognitive ability at child’s age T , \mathbf{x}_{si} is a vector of child-specific covariates, \mathbf{Y}_{it} denotes a vector of inputs by the child’s mother at child’s age t , ν_i represents the child’s unobserved endowment effect, and ε_{siT} is a pure transitory shock. We consider the case where the measurement period of cognitive ability (T) is much later than the period of maternal inputs (t).

Suppose the child endowment ν_i is positively correlated with the mother’s unobserved ability as measured by her skill endowment in the labor market (μ_{wi}). Then maternal inputs (\mathbf{Y}_{it}) that are related to labor supply will become correlated with ν_i , resulting in ordinary least squares (OLS) estimation bias. For example, mothers with high unobserved ability tend to have high child endowment and they also tend to work more because of higher wages.

To address this issue we use a control function approach, whereby the structural model is used to predict a mother’s unobserved ability conditional on her observed behavior. More specifically, we control for the child endowment effect ν_i using a prediction

⁵⁰For nonworkers, the integral is computed using a Gauss–Hermite quadrature.

of the mother's unobserved ability as a covariate:

$$\begin{aligned}
 E(\nu_i | \mathbf{Y}_{it}, \mathbf{x}_{si}) &= \delta_w E(\mu_{wi} | \mathbf{Y}_{it}, \mathbf{x}_{si}) \\
 &\approx \delta_w \sum_{j=1}^5 \mu_{wj} \Pr(j | \mathbf{Y}_{it}, \mathbf{x}_{si}) \\
 &= \delta_w \sum_{j=1}^5 \mu_{wj} \frac{\Pr(\mathbf{Y}_{it} | j, \mathbf{x}_{si}) \Pr(j | \mathbf{x}_{si})}{\sum_{l=1}^5 \Pr(\mathbf{Y}_{it} | l, \mathbf{x}_{si}) \Pr(l | \mathbf{x}_{si})} \equiv \delta_w Q_{wi}.
 \end{aligned} \tag{11}$$

The parameter δ_w , which is expected to have positive sign, captures the association between the mother's unobserved ability and the child endowment effect. In the structural model, there are five discrete unobserved types (i.e., $j = 1, \dots, 5$), so the conditional mean is obtained by a weighted average Q_{wi} . The weighted average can be computed, as the type-specific ability μ_{wj} , conditional probability $\Pr(\mathbf{Y}_{it} | j, \mathbf{x}_{si})$ given type j , and type probability $\Pr(j | \mathbf{x}_{si})$ can all be retrieved directly from the structural model. Note that Q_{wi} can differ substantially across mothers with different observed behavior because of selection into work, pregnancy, and child care choices.⁵¹

In our analysis, the measurement of cognitive outcomes occurs at a much later age than maternal inputs. The emphasis on long-run effects of maternal inputs implies that potential feedback mechanisms play a limited role in the analysis.⁵² Although this greatly simplifies the model, \mathbf{Y}_{it} may still be correlated with ν_i even after we control for the mother's ability (Q_{wi}). An illustrative scenario is described as follows. Suppose the mother knows her child's endowment ν_i when she makes decisions at child's age t . Suppose she also knows the technology in equation (10) and her terminal value function is concave in her child's cognitive ability s_{iT} . Then, all else being equal, a mother facing low ν_i may prefer to care more for the child by working less (assuming maternal care provides the highest technological return), so that s_{iT} can become higher in the terminal period. By contrast, a mother facing high ν_i may prefer to stay at work because of diminishing marginal (terminal) value of s_{iT} . This will result in a positive correlation between the mother's labor supply and her child's endowment, even after we control for the mother's ability.

We acknowledge that the above scenario cannot be thoroughly investigated without a full structural model. Nevertheless, it may be possible to mitigate the estimation bias,

⁵¹Attanasio, Cattan, Fitzsimons, Meghir, and Rubio-Codina (2015) estimate production functions for cognitive and socio-emotional skills as a function of maternal skills and child's past skills, as well as maternal and time investments that are treated as endogenous. They also use a two-stage control function approach, where parental investments are determined by reduced-form equations meant to approximate the decision rules of parental investments. The control function in our paper is predicted by the estimated dynamic structural model of women's behavior, which allows us to conduct policy simulations on the intergenerational effects of child care reforms.

⁵²This stands in contrast to Bernal (2008), where cognitive scores at an early age and maternal inputs from the same periods are used. In her model, the feedback mechanism (e.g., cognitive outcomes affect maternal inputs) is the primary object of interest.

using an additional control function from our partial structural model. We construct an extra covariate, Q_{hi} , which is a prediction of the mother’s unobserved heterogeneity in work preference (μ_{hi}) in the structural model

$$E(v_i|\mathbf{Y}_{it}, \mathbf{x}_{si}) = \delta_w E(\mu_{wi}|\mathbf{Y}_{it}, \mathbf{x}_{si}) + \delta_h E(\mu_{hi}|\mathbf{Y}_{it}, \mathbf{x}_{si}) \approx \delta_w Q_{wi} + \delta_h Q_{hi}, \tag{12}$$

where $Q_{hi} \equiv \sum_{j=1}^5 \mu_{hj} \frac{\Pr(\mathbf{Y}_{it}|j, \mathbf{x}_{si}) \Pr(j|\mathbf{x}_{si})}{\sum_{l=1}^5 \Pr(\mathbf{Y}_{it}|l, \mathbf{x}_{si}) \Pr(l|\mathbf{x}_{si})}$. A mother with high child endowment may consistently work more, which is recorded as having a high Q_{hi} in our structural model. Therefore, v_i may be positively associated with Q_{hi} , which suggests a positive δ_h . We emphasize that δ_h *cannot* be interpreted as a causal effect and it is not a primary object of interest. The purpose of including $\delta_h Q_{hi}$ is to serve as a control function, such that we can obtain an unbiased estimate of the effect of \mathbf{Y}_{it} on s_{iT} . That is, after controlling for Q_{wi} and Q_{hi} , \mathbf{Y}_{it} becomes uncorrelated with v_i .

To summarize, our approach uses the structural model to extract information about the mother’s unobserved heterogeneity, which is then used as a control function to mitigate the estimation bias of the cognitive ability production function. A main shortcoming of this hybrid approach is that we lose a unifying theory that explains children’s development process. For example, the full structural approach involves jointly estimating the production function with the mother’s decision problem.⁵³ The mother’s well-being is a function of the child’s cognitive ability (either through concurrent utility or a terminal value). Depending on the technological and utility returns, mothers will have incentives to invest in their children. Of course, the full process can only be formally specified under an explicit set of assumptions, for example, the mother knows the production function. Under the hybrid approach, this process is kept implicit in the structural model. In particular, we can no longer estimate the “deep” utility parameter of interest—the importance of the child’s cognitive ability to the mother’s utility. The estimation of the production function is mainly treated as an econometric issue.

5.3.2 Empirical specification Let s_i be the test score of child i . We first consider the empirical specification

$$\ln(s_i) = \bar{s}_i + \gamma_h^e \sum_{a=1}^3 h_{ia} + \gamma_c^e \sum_{a=1}^3 c_{ia} + \gamma_y^e \ln \sum_{a=0}^3 y_{ia} + \varepsilon_{si}, \tag{13}$$

where \bar{s}_i represents the time-invariant characteristics of the child:

$$\bar{s}_i = \gamma_0^e + \mathbf{x}'_{si} \boldsymbol{\gamma}_1^e + \delta_w^e Q_{wi} + \delta_h^e Q_{hi}. \tag{14}$$

Estimation is conducted separately for children with low-education mothers and children with high-education mothers. Hence, all the parameters differ by the mother’s education group (noted by superscript e).⁵⁴

⁵³In Section 3, we discuss the limitations that prevent us from pursuing the full structural approach.

⁵⁴The effects of maternal inputs on cognitive ability can differ by the mother’s education level. For instance, mothers with high education may be more effective in caring for the child on their own or utilizing financial resources for their child.

The covariates in equation (13) include total years of maternal employment (h_{ia}) between child's ages 1 and 3, total years of formal child care use (c_{ia}) between child's ages 1 and 3, and the logarithm of total parents' income (y_{ia}) between child's ages 0 and 3. Parental income is regarded as a proxy for investment expenditure related to children, under the assumption that households spend a fixed fraction of income on child-related expenditure. For detailed discussions of a similar specification, see Bernal (2008). The time-invariant term \bar{s}_i contains demographic characteristics such as father's education, child's gender and birth weight, proximity of grandparents, and birth order and sibling status. It also contains the conditional means of the mother's unobserved ability Q_{wi} and work preference Q_{hi} .

A shortcoming of this specification is that it is not directly related to the form of child care. As discussed in Section 3, we consider three forms of care: (i) maternal care ($h = 0$ and $c = 0$), (ii) nonmaternal informal care ($h = 1$ and $c = 0$), and (iii) formal care ($c = 1$). Based on this categorization, we have the preferred empirical specification

$$\begin{aligned} \ln(s_i) = & \bar{s}_i + \gamma_{hc1}^e \sum_{a=1}^3 h_{ia}(1 - c_{ia}) + \gamma_{hc2}^e \sum_{a=1}^3 (1 - h_{ia})(1 - c_{ia}) \\ & + \gamma_y^e \ln \sum_{a=0}^3 y_{ia} + \varepsilon_{si}. \end{aligned} \quad (15)$$

Under this specification, formal care is considered as a benchmark (or excluded) category. Thus, γ_{hc1}^e represents the effect of *nonmaternal informal care* relative to formal care; γ_{hc2}^e represents the effect of *maternal care* relative to formal care.

We also consider additional specifications. For example, equation (15) assumes that maternal employment (h) does not matter when the child is in formal care ($c = 1$). This restriction can be relaxed by adding a covariate that interacts maternal employment with formal care use ($\sum_{a=1}^3 h_{ia}c_{ia}$). Then the benchmark category becomes "nonemployment with formal care use."⁵⁵ For completeness, equation (14) can also include the conditional mean of unobserved heterogeneity in pregnancy preference. The results are further discussed in Section 6.2.

One concern with our specification is that it ignores other inputs in the production function that may become important after early childhood. In particular, school inputs can be an important determinant of cognitive ability. They may also be correlated with early age inputs and result in omitted variable bias. For instance, if additional formal care also induces the mother to send her child to good schools, our estimate will be interpreted as the total impact on cognitive ability, which includes the direct marginal effect of formal care and the indirect effect through increased school inputs. See Todd and Wolpin (2003) and Liu, Mroz, and Adair (2009) for an in-depth discussion of the interpretation of the measured impact of observed early inputs when there could be other decisions made in later time periods that are not observed.⁵⁶

⁵⁵Among low-education mothers with at least one child aged 1–3, 6.5 percent are nonemployed and use formal care simultaneously. Among high-education mothers, the proportion is 13.2 percent.

⁵⁶Liu, Mroz, and Adair (2009) show that the underlying mechanisms can be quite complex because of complementarities of inputs. Although they acknowledge the importance of estimating the marginal ef-

Although there is growing evidence that early life environments are critical for skill formation (e.g., Heckman and Mosso (2014)), our model is silent on the relationship between early age inputs and school inputs. We can gauge the importance of school inputs by controlling for the quality of the school of the child in the production function. We divide all schools in Norway into five quintiles based on each school's average test score of its students.⁵⁷ For each quintile except for the lowest one, a dummy variable is constructed to serve as a proxy for school quality in that quintile. We then link each child in the sample with the dummy variables using the school identifier. The inclusion of indicator variables for school quality would control for differential school inputs that were previously left out and potentially correlated with early maternal inputs. We discuss our findings in detail in Section 6.2. In short, the effects of early age inputs are qualitatively similar even after controlling for school quality in our production function.

Due to the nature of the education system, the issue of differential school input may be less severe in Norway than other countries such as the United States. Primary and lower secondary education in Norway are founded on the principle of equity and adapted education for all pupils in a school system based on the same national curriculum. The national test is taken during the 9-year compulsory schooling period, in which the schools do not track students. Existing research suggests that within-school variations are more important than between-school variations in explaining student performance.⁵⁸

For each child, there are three different subject scores: reading, mathematics, and English. Estimation is carried out separately for each type of score and each education group of the mother, using the same set of covariates. Estimation is carried out in two stages. First, the conditional means of unobserved heterogeneity corresponding to each child's mother are computed using the structural model.⁵⁹ Then the model is estimated

fect, they also argue that the evaluation of some policies might better be carried out without conditioning on many inputs. In some cases, the total effect of an observed input can be more policy-relevant than its marginal effect. For instance, for policies aimed at increasing access to prenatal care to improve birth weight, the ceteris paribus effect of a prenatal care visit is likely trivial because it is not a direct health input (e.g., better nutrition or more exercise). Prenatal care visit represents an "intermediating" input, which induces changes in various direct health inputs. By estimating a health production function that depends on prenatal care visits only, we will obtain a "contaminated total" effect. However, this contaminated effect likely captures a policy-relevant interpretation of the impact of receiving prenatal care.

⁵⁷The data are constructed directly from the population register. For a given subject (reading, mathematics, English), the average test score of a school is computed from the scores of all students who took the subject's national test at the beginning of grade 5 (or 8) between 2007 and 2010. The final average test score of a school is obtained by averaging the scores across the subjects.

⁵⁸For instance, Hægeland, Raaum, and Salvanes (2005) assess the effect of family background, school resources, and teacher qualifications, during the 3-year lower secondary school, on pupil achievement at age 16 in Norway. They find that between-school variations altogether explain no more than 10 percent of the variations in student performance at age 16.

⁵⁹In the empirical model, the conditional type probabilities ($\Pr(j|Y, \mathbf{x})$) are conditional on all information in the panel data that are used for structural model estimation. This includes the mother's choices, wages (observed if employed), covariates, and cohort. This yields the best summary of the mother's unobserved type given the data available. The conditional probabilities are computed by applying Bayes rule on the likelihood contribution in equation (8) (i.e., $\prod_{t=1}^T L_{ik_{it}}(j)$) and type probabilities in equation (5).

using OLS, using the mother's actual inputs as well as the conditional means as covariates. Our baseline sample consists of up to the first two children of the mother.

6. ESTIMATION RESULTS

6.1 *Structural model estimates*

Tables 3 and 4 report estimates from the structural model for low-education women and high-education women, respectively. In both samples, the utility costs of employment and pregnancy tend to be larger than that of formal care use.⁶⁰ There is a positive interactive utility of being employed (full-time) and using formal care, or being employed while pregnant. The utility costs vary by family size; for instance, with an additional child, full-time work becomes less attractive. The standard deviation of the choice shock (σ_c) is smaller among low-education women.⁶¹

In both samples, the return to a year of full-time equivalent work experience is around 10 percent among individuals with no experience. However, the weight of part-time relative to full-time experience is larger among low-education women (0.772) than high-education women (0.678). The standard deviation of the wage shock (σ_w) is slightly smaller among low-education women.

In each sample, the unobserved types are sorted in descending order by the type-specific skill endowment. Note that for a type- j individual, her gross skill endowment when evaluated at 0 years of experience is given by $\beta_{w0} + \mu_{wj}$ (see Section 3). The unobserved types are different between both samples because the models are estimated separately. To avoid confusion, we use L_j to denote a type- j low-education woman and H_j to denote a type- j high-education woman. For instance, a type- $L1$ woman has a log wage intercept of 5.48 (β_{w0}), while a type- $L5$ woman has a log wage intercept of $5.48 - 0.64 = 4.84$. The log wage intercepts for type- $H1$ and type- $H5$ women are 5.844 and $5.844 - 0.91 = 4.934$, respectively. Hence, in both samples, there is substantial unobserved heterogeneity in the wage equation, especially when compared to the standard deviation of the wage shock ($\sigma_w \approx 0.20$).⁶² There is also noticeable unobserved heterogeneity in work and pregnancy preferences.

The overall distribution of unobserved heterogeneity depends on the type probabilities. One useful metric is unconditional type proportions. For low-education women,

⁶⁰Note that the relative sizes of utility costs depend on the level of income. For instance, at zero income, the utility cost of formal care (α_c) is -211.52 for low-education women. At an income of 200,000 NOK, the utility cost becomes $\alpha_c + 200\alpha_{cy} = -211.52 + 200 \times 0.29 = -153.52$.

⁶¹The coefficients in the budget constraint indicate that formal care use is positively associated with a higher coverage rate in the local area, and it is negatively associated with the proximity of grandparents. The terminal value function is not precisely estimated; for low-education women, there is some evidence that it is associated with the number of children.

⁶²The unobserved types are important in explaining the variance of earnings. For each individual, we first simulate her present discounted value (PDV) of gross earnings between year 1 and year 12 of the model. We then decompose the cross-sectional variance into a part explained by unobserved types and another part explained by idiosyncratic shocks. Among both low-education and high-education women, unobserved types explain approximately half of the total variance in PDV gross earnings. In a similar class of model using data from the United States (NLSY79), Keane and Wolpin (2010) find that unobserved types explain 61 percent of the variance in full-time wage offer of white females at age 30.

TABLE 3. Structural model estimates, low-education women.^a

		Utility function:							
<i>Basic parameters:</i>				<i>Other interactions:</i>					
Part-time work (α_h^p)	-426.82	(21.02)	α_{hy}	-0.02	(0.02)	α_{hn1}	-37.76	(4.96)	
Full-time work (α_h^f)	-503.22	(20.31)	α_{hy}^f	0.16	(0.02)	α_{hn2}	-25.29	(6.52)	
Child care use (α_c)	-211.52	(20.18)	α_{cy}	0.29	(0.04)	α_{hn3}	44.04	(5.76)	
Pregnancy (α_p)	-466.64	(45.30)	α_{py}	0.63	(0.06)	α_{cn1}	36.06	(6.71)	
Work \times Child care use (α_{hc})	5.35	(9.82)	α_{hc}^f	44.65	(10.96)	α_{cn2}	104.17	(10.65)	
Work \times Pregnancy (α_{hp})	160.78	(31.87)	α_{hp}^f	-118.85	(9.48)	α_{pn1}	24.42	(25.46)	
Work \times Lagged work (α_{hh})	174.58	(13.35)				α_{pn2}	15.26	(8.46)	
Std. dev. choice shock (σ_c)	103.54	(7.21)							
Wage equation:		Budget constraint:				Terminal value function:			
Intercept	5.480	(0.012)	β_{c1} (coverage)	-132.10	(16.75)	ψ_{n1}	3.23	(1.52)	
FT-equivalent experience	0.099	(0.002)	β_{c2} (grandparent)	9.01	(4.76)	ψ_{n2}	-0.88	(0.40)	
FT-equivalent experience sq.	-0.012	(0.001)	β_{c3}	-7.11	(0.91)	ψ_{e1}	-0.78	(0.86)	
Weight of PT to FT experience	0.772	(0.035)				ψ_{e2}	0.05	(0.08)	
Local unemployment rate	-0.001	(0.043)							
Std. dev. wage shock (σ_w)	0.181	(0.001)							
Type-Specific parameters:	Type L2		Type L3		Type L4		Type L5		
<i>Utility function:</i>									
Work (μ_h)	40.06	(10.63)	113.24	(11.05)	54.60	(10.94)	107.42	(12.62)	
Pregnancy (μ_p)	-43.32	(14.67)	70.86	(15.58)	12.06	(14.81)	123.66	(19.91)	
<i>Wage equation:</i>									
Skill endowment (μ_w)	-0.25	(0.01)	-0.42	(0.01)	-0.54	(0.01)	-0.64	(0.02)	
<i>Type probabilities (MNL):</i>									
Intercept	2.24	(0.24)	2.42	(0.24)	2.79	(0.25)	1.50	(0.35)	
Partner's education	-0.53	(0.46)	-0.81	(0.44)	-0.95	(0.44)	-0.41	(0.49)	
Partner's "permanent income"	0.09	(0.21)	0.03	(0.20)	-0.12	(0.20)	0.06	(0.23)	

^aNumber of individuals = 2330, number of observations = 18,640, log-likelihood = -20,536.08. Standard errors are given in parentheses.

TABLE 4. Structural model estimates, high-education women.^a

			Utility function:			Other interactions:				
<i>Basic parameters:</i>										
Part-time work (α_{ht}^p)	-590.36	(30.18)	α_{hy}	0.05	(0.02)	α_{hn1}	-28.35	(4.49)		
Full-time work (α_{ht}^f)	-618.58	(26.89)	α_{hy}^f	0.24	(0.03)	α_{hn2}	-62.88	(6.89)		
Child care use (α_c)	-231.31	(19.98)	α_{cy}	0.47	(0.04)	α_{hn3}	74.30	(7.75)		
Pregnancy (α_p)	-775.49	(63.91)	α_{py}	0.84	(0.07)	α_{cn1}	52.85	(6.44)		
Work \times Child care use (α_{hc})	-4.74	(8.49)	α_{hc}^f	39.07	(8.88)	α_{cn2}	144.40	(12.30)		
Work \times Pregnancy (α_{hp})	298.57	(42.58)	α_{hp}^f	-137.11	(11.51)	α_{pn1}	197.17	(40.18)		
Work \times Lagged work (α_{hh})	265.84	(19.46)				α_{pn2}	54.20	(10.95)		
Std. dev. choice shock (σ_c)	164.67	(11.03)								
Wage equation:			Budget constraint:			Terminal value function:				
Intercept	5.844	(0.013)	β_{c1} (coverage)	-142.10	(13.35)	ψ_{n1}	-0.39	(0.92)		
FT-equivalent experience	0.103	(0.002)	β_{c2} (grandparent)	12.27	(3.60)	ψ_{n2}	0.25	(0.34)		
ft-equivalent experience sq.	-0.012	(0.001)	β_{c3}	-3.39	(0.60)	ψ_{e1}	0.24	(1.29)		
Weight of PT to FT experience	0.678	(0.028)				ψ_{e2}	-0.05	(0.12)		
Local unemployment rate	0.139	(0.033)								
Std. dev. wage shock (σ_w)	0.211	(0.001)								
Type-Specific parameters:			Type H2		Type H3		Type H4		Type H5	
<i>Utility function:</i>										
Work (μ_h)	39.27	(19.84)	32.49	(16.63)	123.82	(17.74)	11.64	(19.37)		
Pregnancy (μ_p)	33.73	(21.69)	1.11	(12.55)	232.00	(23.68)	33.32	(15.53)		
<i>Wage equation:</i>										
Skill endowment (μ_w)	-0.30	(0.01)	-0.56	(0.01)	-0.74	(0.01)	-0.91	(0.02)		
<i>Type probabilities (MNL):</i>										
Intercept	1.71	(0.23)	3.61	(0.21)	2.93	(0.22)	2.65	(0.22)		
Partner's education	0.18	(0.39)	-0.09	(0.35)	-0.53	(0.35)	-0.84	(0.36)		
Partner's "permanent income"	0.05	(0.17)	-0.42	(0.15)	-0.35	(0.16)	-0.35	(0.16)		

^aNumber of individuals = 3465, number of observations = 27,720, log-likelihood = -33,555.32. Standard errors are given in parentheses.

the type proportions (L1–L5) are 2.9, 26.1, 11.3, 36.5, and 23.2 percent, respectively; for high-education women, the type proportions (H1–H5) are 1.4, 7.9, 24.4, 47.2, and 19.1 percent, respectively.⁶³ In both samples, partner's education is positively associated with the woman's skill endowment type.

6.1.1 Model fit and within-sample effects of the reform Figure 1(a) and (b) shows model fit by age of the women. The left figure shows the employment (part-time and full-time) and pregnancy rates. The right figure shows the overall formal care usage rate, and the formal care usage rate among mothers with at least one child aged between 1 and 3.⁶⁴ The simulations capture the essential data patterns including the peak of employment, and rising pregnancy and formal care usage rates. Although the model predicts part-time work closely, it tends to overpredict full-time work especially at younger ages. Among low-education women, the model also tends to underpredict the pregnancy rate.⁶⁵

Figures 2 and 3 compare the simulated outcomes with actual data by calendar year. This allows us to gauge the data movements before and after the cash-for-care reform, and connect the fit of the model to possible changes in outcomes from the policy.⁶⁶ We also simulate a counterfactual scenario with no cash-for-care reform. The within-sample effects of the reform are given by the difference between both simulations.

Among low-education women (Figure 2), the employment rate (and full-time work) increases quickly from 1994 to 1998 as the cohorts become older; then it remains at a steady level afterward. The pregnancy rate trends upward and it has a small jump in 1999. Among mothers with young children, there are noticeable data movements in child care between 1999 and 2000: formal care drops while the other two forms of care increase. Overall, the changes in outcomes before and after the reform are far from dramatic. Although the model does not perfectly fit the data, it can follow the general data patterns before and after the reform. We find that the within-sample effects of the reform are quite modest, and they tend to pick out subtle changes in outcomes mentioned above.

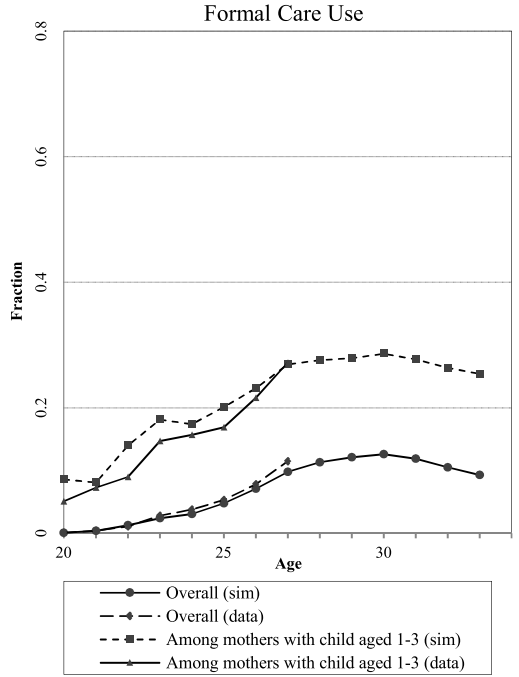
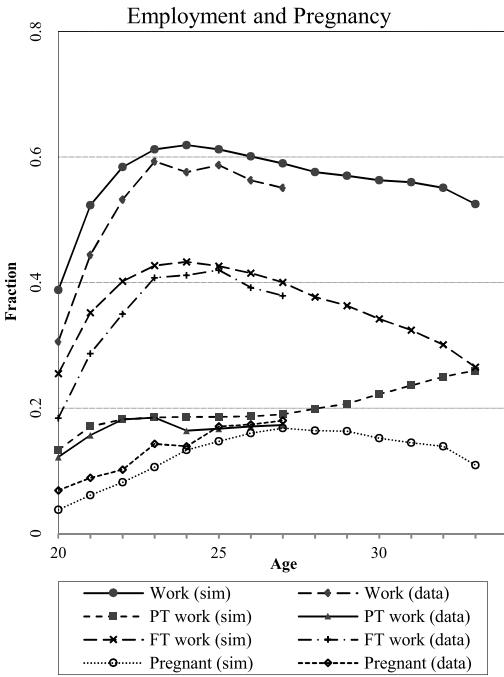
We now turn to high-education women (Figure 3). The employment rate increases until 2000; in fact, there was a *jump* in full-time work between 1998 and 1999. The pregnancy rate has a smooth upward trend without a jump in 1999. Nevertheless, there are still noticeable data movements in child care among mothers shortly after the reform.

⁶³Based on the type-specific log wage intercepts and type proportions, the average log wage intercepts for low-education and high-education women are 5.021 and 5.161, respectively. Hence, the average difference in log wage intercept by education is 0.14 (when evaluated at 0 years of experience).

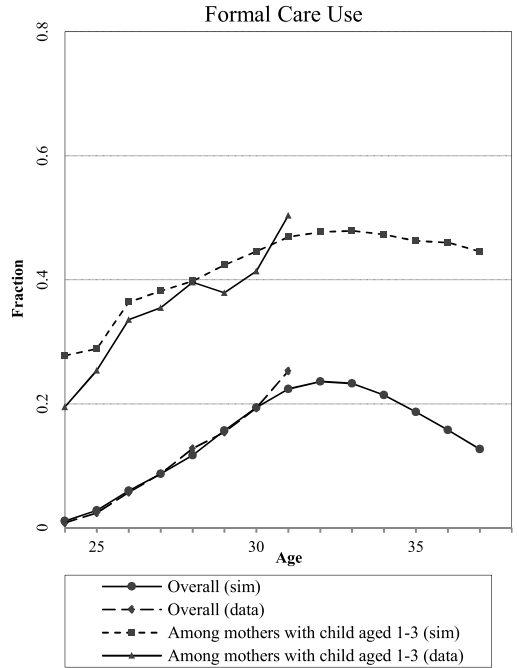
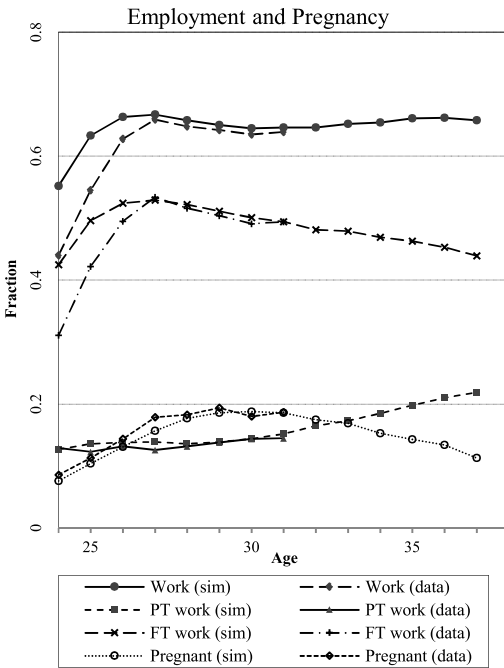
⁶⁴Twenty simulations per individual are conducted, and simulated outcomes are reported until year 15 of the model.

⁶⁵In Appendix Table A.3, we simulate the PDV of gross earnings and net government benefits (i.e., benefits minus tax) between year 1 and year 12 of the model. For each individual, we compute her effective tax rate by dividing the negative of her PDV net government benefits by her PDV gross earnings. Among both groups of women, the unweighted average effective tax rate (across women) is close to 0. This is due to generous benefits paid to women with multiple children.

⁶⁶Note that unlike Figure 1, the number of observations can differ by calendar year. See Appendix Table A1 for the relationship between calendar year and the number of cohorts.

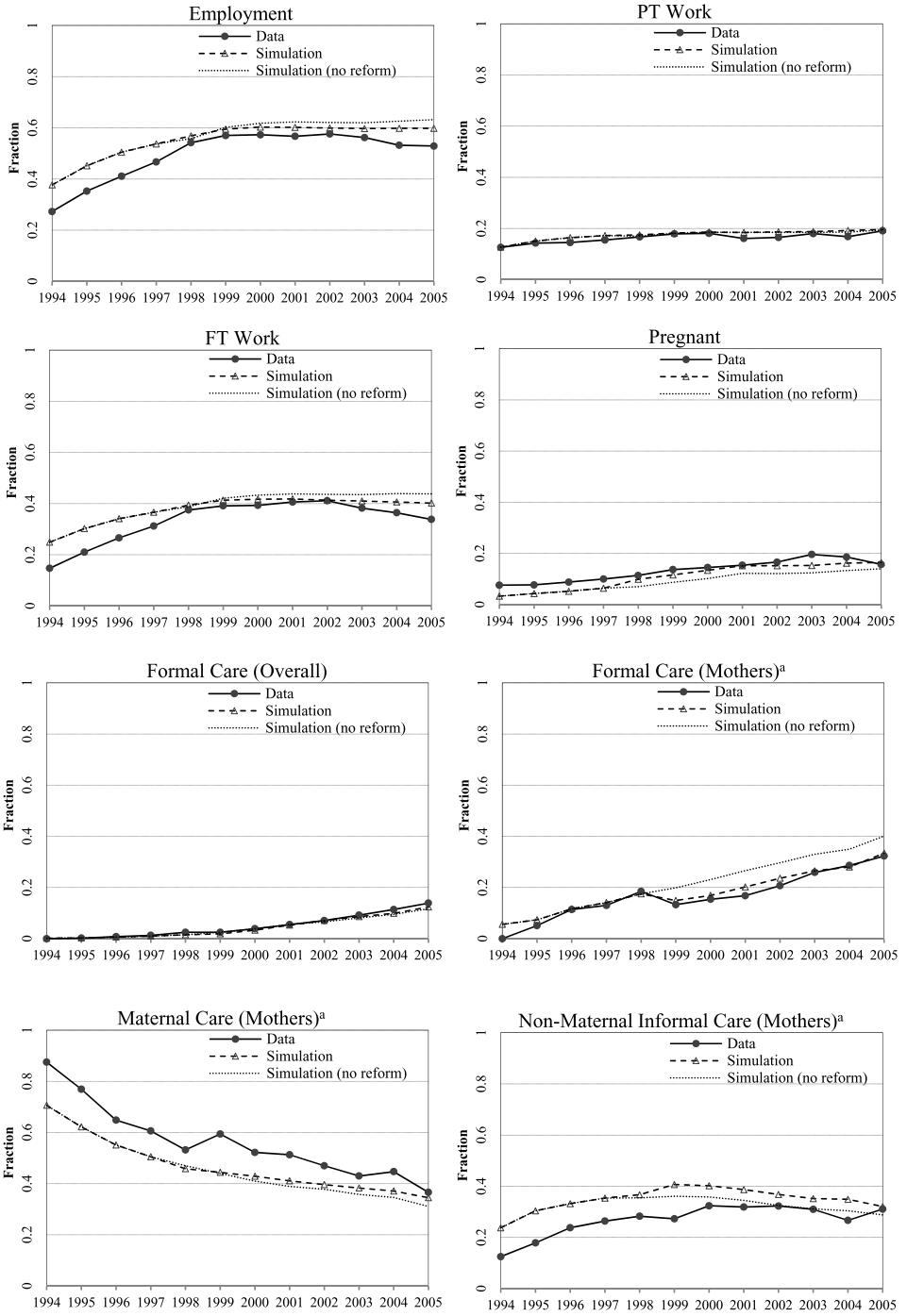


(a)



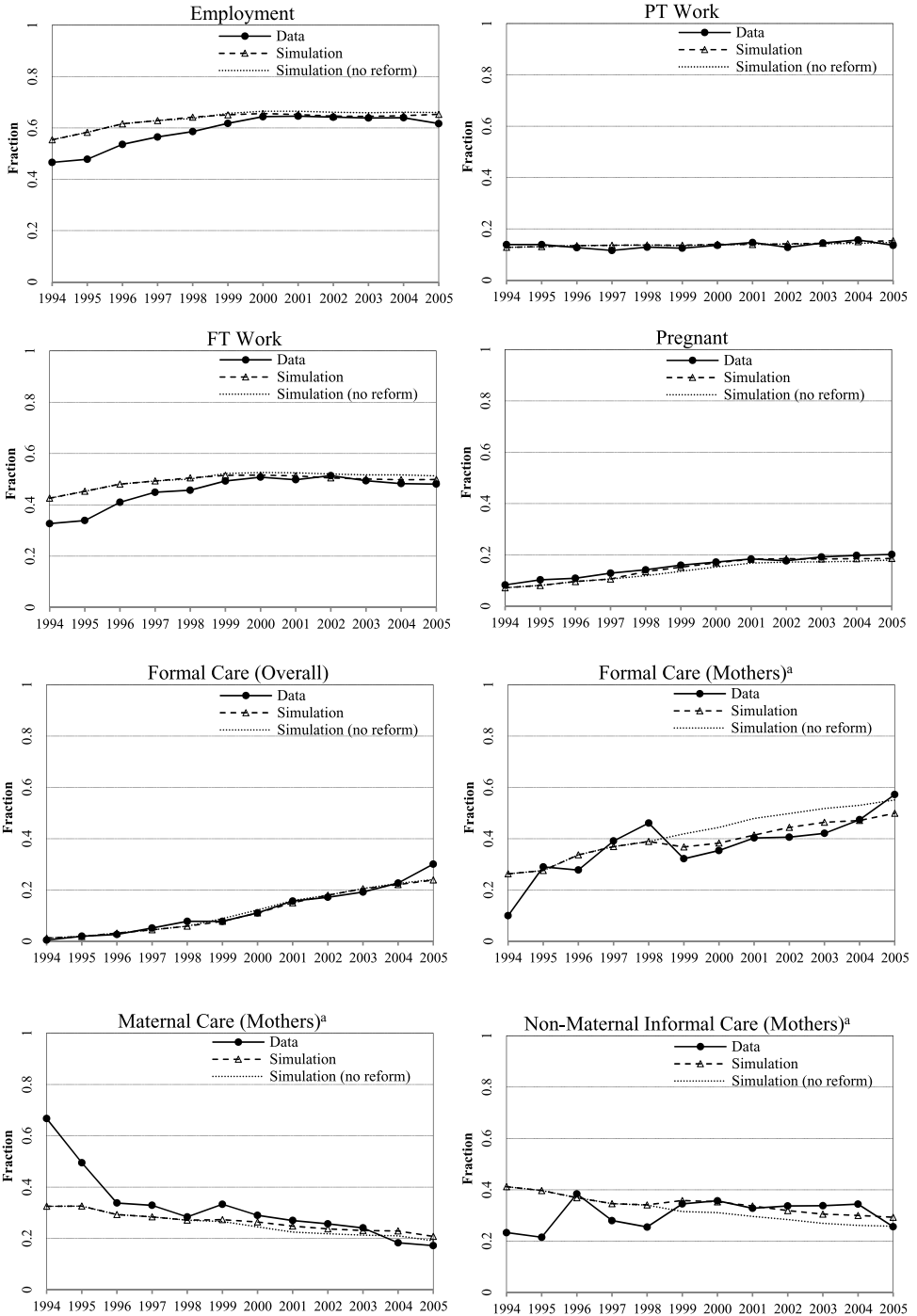
(b)

FIGURE 1. (a) Model fit by age (low-education women). (b) Model fit by age (high-education women).



a: Mothers with at least one child aged between 1 and 3.

FIGURE 2. Model fit by calendar year (low-education women).



a: Mothers with at least one child aged between 1 and 3.

FIGURE 3. Model fit by calendar year (high-education women).

Consistent with the general data patterns, the reform predicts small within-sample effects on employment and pregnancy, but the effects on child care among mothers remain sizable.

6.2 Test score estimates

Table 5 and 6 report estimates from reading score regressions among children of low-education and high-education mothers, respectively.⁶⁷ Columns 1 and 2 are based on the empirical specification in equation (13). Column 1 does not have a control function. After adding a control function (column 2), the coefficient on maternal employment becomes more negative. Among both groups of mothers, an additional year of maternal employment will reduce the reading score by 2.5 percent or 0.06 standard deviations. This is similar to Bernal's (2008) estimate of 0.07 standard deviations.⁶⁸

As discussed earlier, a shortcoming of the above specification is that it is not directly related to the form of care. To differentiate the impacts of the three forms of care, we turn to columns 3 and 4, which are based on our preferred empirical specification in equation (15). For low-education mothers with a control function (column 4 in Table 5), the coefficients on nonmaternal informal care and maternal care are both negative and statistically significant. An additional year of nonmaternal informal care (as opposed to formal care) will reduce the test score by 6.2 percent (0.14 s.d.), whereas an additional year of maternal care (as opposed to formal care) will reduce the test score by 3.8 percent (0.08 s.d.).⁶⁹ Among high-education mothers (column 4 in Table 6), the corresponding coefficients on nonmaternal informal care and maternal care are -2.5 percent (-0.06 s.d.) and $+0.2$ percent ($+0.01$ s.d.), respectively.

Overall, nonmaternal informal care leads to worse reading skills than formal care, and the gap is particularly large among low-education mothers.⁷⁰ For maternal care, the situation is more complicated. Maternal care is less effective than formal care among low-education mothers, but it has similar returns to formal care among high-education mothers. Our results highlight the role of quality of child care in determining children's

⁶⁷As discussed earlier, the estimation sample includes up to the first two children of the mother. There are 1087 children with low-education mothers, and 2043 children with high-education mothers. Both postnatal maternal employment and child care use are constructed directly from the mother's choices in the data. Postnatal income is constructed from the parents' gross earnings in the data plus government benefits minus income tax, which are both computed from the model's budget constraint.

⁶⁸The coefficient on formal care is positive and it is larger among low-education mothers. Although Bernal (2008) finds that child care use reduces children's test scores, the results are not directly comparable, as her definition of child care includes *informal care*. We will discuss this finding further below. In addition, after adding the control function, the income coefficient becomes smaller and it is not significantly different from 0.

⁶⁹Although some of these estimates are large, they do not necessarily translate into large policy effects because the policies need to have a strong impact on mother's work and child care use patterns simultaneously. See Section 6.3.2 for more discussions.

⁷⁰Few of the cases in nonmaternal informal care can be explained by care by the father. When the child is between ages 1 and 3, the father's employment rate (defined as employed if pre-tax earnings exceed 100,000 NOK) among children who are subject to nonmaternal informal care is 93.8 percent; among children who are not subject to such care, the father's employment rate is 92.1 percent.

TABLE 5. Reading score regression: children of low-education mothers.^a

Dependent Variable: Ln(Reading Score)	(1)	(2)	(3)	(4)
Years of maternal employment (age 1–3)	–0.010 (0.014)	–0.025 (0.017)		
Years of formal care use (age 1–3)	0.044 (0.018)	0.046 (0.019)		
Years of “non-maternal informal care” (age 1–3)			–0.054 (0.020)	–0.062 (0.021)
Years of “maternal care” (age 1–3)			–0.045 (0.018)	–0.038 (0.018)
Ln(mother + father total income, age 0–3)	0.005 (0.039)	–0.006 (0.040)	–0.001 (0.038)	–0.011 (0.039)
<i>Control function:</i>				
Skill endowment (log wage) (Q_w)		0.144 (0.151)		0.117 (0.151)
Work preference (in NOK 100,000) (Q_h)		0.139 (0.083)		0.126 (0.083)
<i>Socioeconomic covariates:</i>				
High education father	0.061 (0.027)	0.058 (0.027)	0.060 (0.027)	0.057 (0.027)
Child’s gender (male = 1)	–0.084 (0.026)	–0.086 (0.026)	–0.083 (0.026)	–0.085 (0.026)
Ln(birth weight)	0.027 (0.069)	0.025 (0.070)	0.028 (0.069)	0.026 (0.069)
Grandparent lives close	–0.004 (0.027)	–0.005 (0.027)	–0.004 (0.027)	–0.005 (0.027)
Mother under age 21 at birth	–0.122 (0.050)	–0.133 (0.050)	–0.121 (0.050)	–0.132 (0.050)
Young sibling born within 4 years of first child’s birth	–0.014 (0.032)	–0.015 (0.032)	–0.010 (0.032)	–0.009 (0.032)
Not the first child in the family	–0.011 (0.033)	–0.002 (0.033)	–0.010 (0.033)	–0.002 (0.033)
F-test of work preference and skill (p -value)		0.245		0.315
R-squared	0.125	0.127	0.127	0.129

^aThere are 1087 children in the regression. Standard errors are given in parentheses. Grade scores are normalized.

cognitive development. Relative to formal care, other forms of care arranged by low-education mothers are presumably of lower quality.

The coefficient on the skill endowment control function (Q_{wi}) is positive in both samples. However, it is larger and only statistically significant among high-education mothers (column 4 in Table 5 and 6). For each 10-percent increase in a high-education mother’s skill endowment (evaluated at 0 years of experience), her child’s reading score is higher by 1.87 percent. This translates roughly into a 0.6 s.d. increase in maternal skill endowment being associated with a 0.05 s.d. increase in the child’s reading score.⁷¹ Father’s education is also more positively associated with children’s reading score among

⁷¹Appendix Figure A.1 plots the distribution of predicted skill endowment conditional on observed behavior (Q_{wi}) for low-education and high-education women. In both samples, there is sizable variation in

TABLE 6. Reading score regression: children of high-education mothers.^a

Dependent Variable: Ln(Reading Score)	(1)	(2)	(3)	(4)
Years of maternal employment (age 1–3)	–0.018 (0.009)	–0.025 (0.010)		
Years of formal care use (age 1–3)	0.021 (0.008)	0.017 (0.008)		
Years of “non-maternal informal care” (age 1–3)			–0.027 (0.010)	–0.025 (0.010)
Years of “maternal care” (age 1–3)			–0.007 (0.010)	0.002 (0.010)
Ln(mother + father total income, age 0–3)	0.043 (0.026)	0.027 (0.026)	0.038 (0.025)	0.022 (0.026)
<i>Control function:</i>				
Skill endowment (log wage) (Q_w)		0.185 (0.062)		0.187 (0.062)
Work preference (in NOK 100,000) (Q_h)		0.014 (0.026)		0.012 (0.026)
<i>Socioeconomic covariates:</i>				
High education father	0.116 (0.022)	0.113 (0.022)	0.117 (0.022)	0.114 (0.022)
Child’s gender (male = 1)	–0.064 (0.016)	–0.065 (0.016)	–0.064 (0.016)	–0.064 (0.016)
Ln(birth weight)	0.065 (0.045)	0.073 (0.045)	0.066 (0.045)	0.074 (0.045)
Grandparent lives close	–0.008 (0.016)	–0.004 (0.016)	–0.008 (0.016)	–0.004 (0.016)
Young sibling born within 4 years of first child’s birth	0.046 (0.022)	0.050 (0.022)	0.049 (0.022)	0.054 (0.022)
Not the first child in the family	–0.047 (0.020)	–0.042 (0.020)	–0.049 (0.020)	–0.045 (0.020)
F-test of work preference and skill (p -value)		0.011		0.011
R-squared	0.191	0.194	0.190	0.194

^aThere are 2043 children in the regression. Standard errors are given in parentheses. Grade scores are normalized.

high-education mothers (the coefficient is 0.114 versus 0.057 among low-education mothers). The coefficient on the work preference control function (Q_{hi}) is larger among low-education mothers, but it is statistically insignificant.

We perform a battery of robustness checks based on discussions of additional specifications in Section 5.3.2: (i) relax the restriction that maternal employment does not matter when the child is in formal care; (ii) include a pregnancy preference control function instead of a work preference control function; (iii) control for school quality using dummy variables constructed from school quintiles; (iv) include first-born chil-

predicted skill endowment. High-education women tend to have a larger spread in predicted skill endowment than low-education women (standard deviation is 0.17 versus 0.13).

dren only. The full results are reported in Appendix Tables A4 and A5.⁷² The coefficients on nonmaternal informal care and maternal care are robust to the above specifications. In particular, school quality is an important predictor of reading scores (p -value from a joint F -test is lower than 0.01), but they have a minor impact on the size and statistical significance of both coefficients. Given that school quality represents factors including (and beyond) school inputs, this suggests that missing school-age inputs is unlikely to create a substantial bias in our estimates. We also find that school quality tends to play a more important role among children of low-education mothers.⁷³

Relative to reading score, the results from math and English score regressions are dramatically different. The full results are reported in Appendix Tables A.6 and A.7. Among both groups of mothers, the coefficients on nonmaternal informal care and maternal care are close to 0 and they are mostly statistically insignificant. This suggests that the form of care does *not* have a significant impact on math and English scores.⁷⁴ However, maternal skill endowment is significantly associated with math and English scores among high-education mothers. The degree of association is similar to reading score—a 0.6 s.d. increase in maternal skill endowment being associated with an approximate 0.05 s.d. increase in either score. School quality is also an important predictor of both scores, especially among children of low-education mothers.

The differential impacts of child care method on reading and math scores are broadly consistent with findings from the literature that emphasize the importance of language environment in early childhood in shaping reading abilities later in life. Reading scores measure skills related to language ability, and the environment at early ages is shown to be an important factor in shaping these skills.⁷⁵ Interestingly, the results in reading do not seem to carry over to the acquisition of a foreign language (in this case, English). Given that English skills are primarily developed during school age, it may not be at all surprising to see a smaller effect of the child care method on English scores.

⁷²In column 1, we find that when the child is in formal care, maternal employment has a small negative effect but the difference is not statistically significant. In column 2, the conditional means of unobserved heterogeneity for work and pregnancy preferences are highly correlated. To avoid problems of multicollinearity, only one of them is included in the equation. Among low-education mothers, the results are similar, with the exception that the coefficient on the skill endowment control function becomes close to 0. Among high-education mothers, there are almost no changes in results.

⁷³Among low-education mothers, the difference in log score between children belonging to the highest and lowest quintiles of school quality is 0.219 (0.49 s.d.), all else being equal. Among high-education mothers, the difference is 0.155 (0.39 s.d.). In addition, after including school quality, the coefficient on the skill endowment control function becomes close to 0 among low-education mothers. The coefficient remains similar and statistically significant among high-education mothers.

⁷⁴An exception is the effect of nonmaternal informal care on English scores among children of high-education mothers: the coefficient is -0.019 (significant at the 10-percent level) when the control function and school quality are present.

⁷⁵Studies have found that children from disadvantaged environments are exposed to a substantially less rich vocabulary than children from more advantaged families (Hart and Risley (1995), Fernald, Marchman, and Weisleder (2013)). For instance, at age 3, children from professional families speak 50 percent more words than children from working-class families and more than twice as many compared to children from welfare families.

6.3 Counterfactual analysis

6.3.1 *Effects on adult individuals* Table 7 compares simulation results from several policy scenarios for both low-education and high-education women in year 12 of the model. Simulation results for year 6 are available in Appendix Table A.8. The baseline scenario (column 1) assumes that individuals are never subject to the cash-for-care program, but other aspects of the policy environment remain the same as in the data. In subsequent columns, the following policies are implemented since the *first* period of the model: a full cash-for-care program (column 2); a partial cash-for-care program where workers are ineligible (column 3); a 20-percent increase in the coverage rate of maternity leave, keeping the cap unchanged ($1.2 \times b_m$; column 4); an income tax deduction of

TABLE 7. Effects of counterfactual policies, year 12.

Variable	Difference From Baseline				
	Baseline (no Cash- for-Care) (1)	Full Cash-for- Care (2)	Partial Cash-for- Care (3)	Expand Maternity Leave (4)	Tax Deduction for Children (5)
<i>Low-education women:</i>					
Work (%)	58.7	-2.4	-3.6	-0.6	+1.7
Part-time work (%)	21.0	+1.6	-0.4	+0.4	+2.0
Full-time work (%)	37.7	-4.0	-3.2	-1.1	-0.3
Overall formal care (%)	13.0	-0.7	-0.5	+0.9	+1.7
Formal care among mothers (%) ^b	35.4	-6.8	-4.0	-0.0	+1.2
Maternal care among mothers (%) ^b	35.0	+2.5	+7.5	-0.3	-3.3
Non-maternal informal care among mothers (%) ^b	29.6	+4.3	-3.4	+0.3	+2.0
Pregnancy (%)	13.7	+1.2	+0.7	+0.6	+0.3
Has a child aged 0 to 3 (%)	47.2	+7.0	+3.4	+3.0	+3.5
Number of children	1.06	+0.25	+0.11	+0.10	+0.15
PDV earnings (1000 NOK) ^a	689.20	-16.23	-20.41	-3.64	-0.91
PDV net government benefits (1000 NOK) ^a	-95.58	+56.73	+30.52	+23.72	+23.63
PDV utility ^a	1007.20	+25.10	+12.00	+12.30	+9.80
<i>High-education women:</i>					
Work (%)	66.1	-0.6	-1.1	-0.1	+1.8
Part-time work (%)	18.2	+0.4	-0.3	+0.3	+0.5
Full-time work (%)	47.9	-1.0	-0.8	-0.4	+1.3
Overall formal care (%)	22.6	-1.7	-0.5	+0.5	+0.9
Formal care among mothers (%) ^b	52.8	-5.6	-2.1	-0.0	+1.1
Maternal care among mothers (%) ^b	21.6	+1.6	+3.0	-0.2	-1.5
Non-maternal informal care among mothers (%) ^b	25.6	+4.0	-0.9	+0.3	+0.4
Pregnancy (%)	15.1	+0.0	+0.0	+0.1	-0.3
Has a child aged 0 to 3 (%)	53.0	+1.8	+0.7	+1.1	+0.7
Number of children	1.55	+0.12	+0.04	+0.07	+0.14
PDV earnings (1000 NOK) ^a	963.37	-12.26	-14.10	-5.56	-5.18
PDV net government benefits (1000 NOK) ^a	-135.20	+45.13	+20.20	+29.30	+32.20
PDV utility ^a	1789.70	+30.80	+10.10	+20.60	+26.20

^aFrom year 1 to year 12.^bMothers with at least one child aged between 1 and 3.

20,000 NOK for each child, up to two children (column 5).⁷⁶ Individuals are aware that the policy is in place for all periods.

We first consider low-education women at year 12. In the baseline scenario, low-education women have a lower employment rate than high-education women (58.7 versus 66.1 percent). They have a lower overall rate of formal care use (13.0 versus 22.6 percent), lower pregnancy rate (13.7 versus 15.1 percent), and fewer children (1.06 versus 1.55). They have a lower present discounted value (PDV) of gross earnings from year 1 to year 12 (689,200 versus 963,370 NOK).

Column 2 reports the effects of the full cash-for-care program. The policy reduces the employment rate by 2.4 percentage points. This is due to a reduction in full-time work (−4.0 percentage points), which is partially offset by an increase in part-time work.⁷⁷ Interestingly, overall formal care use reduces only slightly (−0.7 percentage points) due to two offsetting forces: (a) a reduction in formal care use among mothers with young children (−6.8 percentage points) and (b) an increase in the fraction of women with young children (+7.0 percentage points). Among mothers with young children, the policy increases nonmaternal informal care by a larger amount than maternal care.

Column 3 reports the effects of the partial cash-for-care program that covers non-workers only. Under this program, mothers can only collect benefits if they pursue maternal care ($h = 0, c = 0$). The main purpose of this policy adjustment is to discourage mothers from pursuing nonmaternal informal care ($h = 1, c = 0$), which has a negative effect on children's cognitive scores. However, because workers are ineligible for benefits, the reform generates a sizable work disincentive. The employment rate reduces by 3.6 percentage points due to reductions in both part-time and full-time work. Among mothers with young children, maternal care becomes more popular (+7.5 percentage points) while the other two forms of care become less popular. The partial program has a much smaller fertility effect than the full program and it is also less expensive.

Column 4 reports the effects of expanded maternity leave. Although it involves a similar increase in government expenditure to the partial cash-for-care program, the behavioral effects are relatively small. Employment reduces slightly due to a reduction in full-time work. There is an increase in the overall formal care rate due to increased fertility. Among mothers with young children, there are almost no changes in the relative popularity of the three forms of care.

Column 5 reports the effects of the child-based tax deduction, which involves a similar increase in government expenditure to the expanded maternity leave. The deduction resembles an earnings-dependent tax credit for children (e.g., see the United States).

⁷⁶The tax deduction is directly applied to the mother's gross earnings for the purpose of income tax calculations. It is independent of child care use.

⁷⁷Previous empirical studies using the difference-in-difference approach document that the cash-for-care allowance decreased eligible mothers' full-time employment by 4–5 percentage points and decreased their labor force participation by 2–3 percentage points (Schöne (2004), Drange (2015)). These are short-term effects identified by comparing mothers with children of the same age born in different time periods. In contrast, our model allows us to evaluate both short-term and long-term effects and to distinguish between responses at different points of a woman's life cycle.

The policy increases employment, especially part-time work (+1.7 and +2.0 percentage points, respectively). The overall formal care rate increases due to an increase in formal care among mothers *and* an increase in fertility. Among mothers with young children, maternal care reduces while the other two forms of care become more popular.

We now discuss the effects on high-education women at year 12. Relative to low-education women, the policy effects are typically much smaller. For instance, the full cash-for-care program (column 2) reduces the employment rate only slightly and the fertility effect is smaller. However, among mothers with young children, there is still a sizable reduction in the formal care usage rate (−5.6 percentage points). The effects of the partial cash-for-care program (column 3) are also small. Due to higher earnings, the maternity leave expansion and tax deduction tend to be more generous toward high-education women. Nevertheless, the effects of the expanded maternity leave are small. By contrast, the deduction generates similar effects on employment and fertility relative to low-education women.

The fertility effects discussed above may not necessarily reflect life-cycle effects because women may adjust their timing of births. To investigate this issue, we look at the average number of children in the last year of the model, which represents a “completed” fertility profile (see Appendix Table A.9). In the baseline scenario, the means are 1.80 and 2.30 among low-education and high-education women, respectively. In the full cash-for-care scenario, the completed fertility effects are +0.23 children (+12.8 percent) and +0.11 children (+4.8 percent) among low-education and high-education women, respectively.⁷⁸ In addition, the average mother’s age at birth of the first child becomes younger by 0.99 and 0.32 years, respectively. The larger fertility effect of the cash-for-care program among low-education women is consistent with the fact that many of these women do not use formal care and can potentially benefit more from the program. In addition, the cash-for-care benefits constitute a higher fraction of their income because their earnings are typically lower.

6.3.2 Effects on children Table 8 reports the effects of the above counterfactual policies on children’s reading scores. To compute the effects on children, policy effects on mothers are first simulated, which are then used to compute the predicted change in reading scores based on estimation results in specification column 4 of Tables 5 (low-education women) and 6 (high-education women). The table reports changes in reading scores among *preexisting* children, who are defined as children who are born in *both* the baseline and the counterfactual policy scenarios. Because the policies increase fertility, the table also reports the reading scores among *new* children, who are defined as children who are born in the counterfactual policy scenario only. The scores among new children can be systematically different from preexisting children due to differences in their mothers’ postnatal choices as well as unobserved ability. Results are separately reported for the first child of low-education mothers and high-education mothers who are born by the twelfth year of the model.

⁷⁸Although the results are not directly comparable, the fertility effects are roughly in line with the sizable increase in the total fertility rate (TFR) in Norway during the 2000s (see Appendix Figure A.3). The TFR dropped from 1.86 in 1993 to 1.75 in 2002, then increased to 1.95 in 2010. Note that the TFR may change slowly because it is constructed from age-specific fertility rates of all fecund women cohorts.

TABLE 8. Decomposition of policy effect on reading scores

Variable	Full Cash-for- Care (1)	Partial Cash-for- Care (2)	Expand Maternity Leave (3)	Tax Deduction for Children (4)
Children of Low-Education Mothers:				
<i>Preexisting Children (born in both baseline and counterfactual, first child):</i>				
Effect on reading test score	-1.14%	-0.15%	-0.01%	-0.03%
Due to:				
Years of non-maternal informal care (age 1–3)	-0.82%	+0.72%	-0.04%	-0.27%
Years of maternal care (age 1–3)	-0.25%	-0.85%	+0.08%	+0.27%
Ln(total parents' income, age 0–3)	-0.08%	-0.02%	-0.04%	-0.03%
<i>New Children (born in counterfactual only, first child):</i>				
Fraction of new children to preexisting children	16.68%	8.85%	6.61%	9.14%
Reading test score relative to preexisting children (in counterfactual scenario):	-0.20%	-0.50%	-0.19%	-0.00%
Due to:				
Difference in mother's post-natal choices and income	+0.06%	+0.08%	-0.02%	-0.16%
Difference in mother's unobserved type	-0.26%	-0.58%	-0.17%	+0.16%
Children of High-Education Mothers:				
<i>Preexisting Children (born in both baseline and counterfactual, first child):</i>				
Effect on reading test score	-0.19%	+0.15%	+0.07%	+0.00%
Due to:				
Years of non-maternal informal care (age 1–3)	-0.29%	+0.11%	-0.00%	-0.01%
Years of maternal care (age 1–3)	+0.01%	+0.02%	-0.00%	-0.00%
Ln(total parents' income, age 0–3)	+0.09%	+0.03%	+0.07%	+0.02%
<i>New Children (born in counterfactual only, first child):</i>				
Fraction of new children to preexisting children	5.56%	2.32%	3.06%	5.42%
Reading test score relative to preexisting children (in counterfactual scenario):	-1.88%	-2.57%	-1.58%	-1.08%
Due to:				
Difference in mother's post-natal choices and income	-0.67%	-0.70%	-0.71%	-0.70%
Difference in mother's unobserved type	-1.21%	-1.87%	-0.87%	-0.38%

We first consider preexisting children of low-education mothers. The full cash-for-care program (column 1) reduces the average reading score by 1.14 percent, or 0.03 standard deviations. This is due to mothers shifting away from formal care: the increase in nonmaternal informal care contributes to a reduction of 0.82 percent, and the increase in maternal care contributes to a reduction of 0.25 percent. By contrast, the partial cash-for-care program (column 2) barely changes the average reading score (-0.15 percent). This is due to mothers moving into maternal care, which has a more positive effect on reading score than nonmaternal informal care, but a more negative effect on reading score than formal care. Both the expanded maternity leave (column 3) and tax deduction (column 4) have little effect on the reading score.

In all policy scenarios, the new children of low-education mothers have a marginally lower average reading score than the preexisting children. Under full and partial cash-

for-care scenarios, this is primarily due to the increased fertility of low-ability mothers, who are less attached to the labor market but can potentially benefit more from the program.

We now consider children of high-education mothers. The policies barely affect the average reading score of preexisting children. This is because the three forms of care do not have substantially different effects on reading score. However, in all policy scenarios, new children tend to have moderately lower reading scores than preexisting children (ranging between 1.08 and 2.57 percent). This is mostly attributed to the increased fertility of low-ability mothers, although differences in mothers' postnatal income and choices also play a role. Nevertheless, new children constitute a very small proportion of children of high-education mothers.

By combining the outcomes from adults and children, it is possible to assess, albeit in a highly speculative way, the desirability of the cash-for-care program. A primitive analysis is provided as follows. Let $W(v, s; \theta) \equiv \theta \ln v + (1 - \theta) \ln s$ be a "social welfare function," where v is the adults' outcome, s is the children's outcome, and θ is the weight on adults. Denote outcomes in the baseline scenario by (v_0, s_0) and denote outcomes in the counterfactual policy scenario by (v_1, s_1) . For simplicity, suppose $v_1 > v_0$ and $s_1 \leq s_0$, so that there is a trade-off in the well-being of adults and children. Then the threshold weight θ^* such that $W(v_1, s_1; \theta^*) = W(v_0, s_0; \theta^*)$ is given by $\theta^* = -\frac{\Delta \ln s}{\Delta \ln v - \Delta \ln s} \geq 0$, where $\Delta \ln v \equiv \ln v_1 - \ln v_0$ and $\Delta \ln s \equiv \ln s_1 - \ln s_0$. When $\theta > \theta^*$, we have $W(v_1, s_1; \theta^*) > W(v_0, s_0; \theta^*)$; when $\theta < \theta^*$, we have $W(v_1, s_1; \theta^*) < W(v_0, s_0; \theta^*)$.

The threshold weight θ^* can be interpreted as the smallest adult weight in the social welfare function (θ) such that the counterfactual policy scenario is "rationalized," that is, it is at least as good as the baseline scenario. As an illustration, consider the full cash-for-care program (column 2 in Table 7 and column 1 in Table 8). Among low-education women, the policy increases PDV utility by 2.49 percent and it reduces the average test score of preexisting children by 1.14 percent. If we use these outcomes in the social welfare function, we will have $\theta_{\text{lowed}}^* = -\frac{-0.0114}{0.0249 - (-0.0114)} = 0.314$.⁷⁹ Among high-education women, the policy increases PDV utility by 1.72 percent and it reduces the average test score of preexisting children by 0.19 percent. Then $\theta_{\text{highed}}^* = -\frac{-0.0019}{0.0172 - (-0.0019)} = 0.099$. Therefore, if the adult weight in the social welfare function (θ) is at least 0.314, the full cash-for-care program will be rationalized among both groups of women.

7. CONCLUSION

In this paper, we used a hybrid approach to analyze the effect of child care policies on life-cycle decisions among women and long-run cognitive outcomes among children. We first estimated a structural model in which women's fertility decisions were formulated jointly with labor supply and child care use decisions. The sample period covered a large-scale child care reform in Norway, which resulted in a large exogenous change in the relative price of formal care facilities. Combining with administrative data on national test scores, we further examined the effects of maternal inputs on long-run cognitive outcomes of children. We used the structural model to extract information about

⁷⁹Alternatively, we can use PDV income (i.e., sum of earnings and net government benefits) as a more tangible measure of adults' outcome.

the mother's unobserved heterogeneity, which was then used as a control function to mitigate the estimation bias of the cognitive ability production function.

We found that an early life-cycle exposure of the child care reform would result in sizable changes in employment and fertility decisions, especially among low-education women. Mothers reduced formal care use and some of them became employed at the same time; this resulted in an increase of maternal care and especially nonmaternal informal care. These behavioral changes had significant intergenerational consequences, especially on children's reading ability as measured beyond age 10. Among low-education mothers, we found that an additional year of nonmaternal informal care and maternal care between ages 1 and 3 (as opposed to formal care) could reduce the child's average reading score by 6.2 percent (0.14 s.d.) and 3.8 percent (0.08 s.d.), respectively. Although these estimates were large, the intergenerational effect depended on the size of policy impact on mother's work and child care use patterns. Taking into account various mechanisms, the net effect of the cash-for-care reform among preexisting children of low-education mothers was negative and small: a reduction of reading score by 1.14 percent (0.03 s.d.).

Our findings have important policy implications, especially at a time when there is an increasing advocacy in many countries for more subsidized child care at early ages. Although our analysis was conducted in the Norwegian context where the quality of formal child care is relatively high, it allows us to broadly investigate the effective design of child care policy by conducting counterfactual policy evaluations. Our policy simulation indicates that child care policies are costly to the government through both increased expenditure on existing children and additional expenditure on new children due to fertility effects. Child care policies should also take into account the potential consequences for the well-being of children. For instance, our policy simulation suggests that if the full cash-for-care program is changed to a partial program in which workers are ineligible for benefits, reading scores among children of low-education mothers will improve (although mothers will face larger work disincentives). Effective design of child care policies should therefore strike a fine balance between impacts on life-cycle decisions of women and effects on children's development.

REFERENCES

- Aarbu, K. O. and T. O. Thoresen (2001), "Income responses to tax changes—evidence from the Norwegian tax reform." *National Tax Journal*, 54, 319–335. [676]
- Adda, J., C. Dustmann, and K. Stevens (2017), "The career costs of children." *Journal of Political Economy*, 125 (2), 293–337. [660]
- Attanasio, O., S. Cattan, E. Fitzsimons, C. Meghir, and M. Rubio-Codina (2015), "Estimating the production function for human capital: Results from a randomized control trial in Colombia." Discussion paper, National Bureau of Economic Research. [682]
- Attanasio, O., C. Meghir, and A. Santiago (2011), "Education choices in Mexico: Using a structural model and a randomized experiment to evaluate PROGRESA." *Review of Economic Studies*, 79, 37–66. [661]

- Averett, S. L., H. E. Peters, and D. M. Waldman (1997), “Tax credits, labor supply, and child care.” *Review of Economics and Statistics*, 79 (1), 125–135. [660]
- Baker, M., J. Gruber, and K. Milligan (2008), “Universal child care, maternal labor supply, and family well-being.” *Journal of Political Economy*, 116 (4), 709–745. [660]
- Baker, M. and K. Milligan (2010), “Evidence from maternity leave expansions of the impact of maternal care on early child development.” *Journal of Human Resources*, 45 (1), 1–32. [660]
- Bennett, J. and C. P. Talyer (2006), “Starting strong II: Early childhood education and care.” OECD. [660]
- Bernal, R. (2008), “The effect of maternal employment and child care on children’s cognitive development.” *International Economic Review*, 49, 1173–1209. [661, 662, 666, 668, 670, 671, 678, 682, 684, 693]
- Bernal, R. and M. P. Keane (2011), “Child care choices and children’s cognitive achievement: The case of single mothers.” *Journal of Labor Economics*, 29 (3), 459–512. [660]
- Bettinger, E., T. Hægeland, and M. Rege (2014), “Home with mom: The effects of stay-at-home parents on children’s long-run educational outcomes.” *Journal of Labor Economics*, 32 (3), 443–467. [669]
- Black, S. E., P. J. Devereux, K. V. Løken, and K. G. Salvanes (2014), “Care or cash? The effect of child care subsidies on student performance.” *Review of Economics and Statistics*, 96 (5), 824–837. [673, 676]
- Blau, D. and J. Currie (2006), “Pre-school, day care, and after-school care: Who’s minding the kids?” *Handbook of the Economics of Education*, 2, 1163–1278. [659, 660]
- Blau, D. M. (1999), “The effect of child care characteristics on child development.” *Journal of Human Resources*, 34, 786–822. [660]
- Blau, D. M. and P. K. Robins (1988), “Child-care costs and family labor supply.” *The Review of Economics and Statistics*, 374–381. [660]
- Blundell, R., M. C. Dias, C. Meghir, and J. Shaw (2016), “Female labour supply, human capital and welfare reform.” *Econometrica*, 84, 1705–1753. [661]
- Carneiro, P., K. V. Løken, and K. G. Salvanes (2015), “A flying start? Maternity leave benefits and long-run outcomes of children.” *Journal of Political Economy*, 123 (2), 365–412. [660]
- Chan, M. K. (2013), “A dynamic model of welfare reform.” *Econometrica*, 81, 941–1001. [661, 669]
- Connelly, R. (1992), “The effect of child care costs on married women’s labor force participation.” *The Review of Economics and Statistics*, 74, 83–90. [660]
- Cools, S., J. H. Fiva, and L. J. Kirkebøen (2015), “Causal effects of paternity leave on children and parents.” *The Scandinavian Journal of Economics*, 117 (3), 801–828. [663]

- Cunha, F., J. J. Heckman, and S. M. Schennach (2010), "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica*, 78 (3), 883–931. [661]
- Dahl, G. B. and L. Lochner (2012), "The impact of family income on child achievement: Evidence from the earned income tax credit." *American Economic Review*, 102 (5), 1927–1956. [660]
- Nusche, D., L. Earl, W. Maxwell, and C. Shewbridg (2011), *OECD Reviews of Evaluation and Assessment in Education OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, Vol. 2011. OECD Publishing. [675]
- Del Boca, D., C. Flinn, and M. Wiswall (2014), "Household choices and child development." *The Review of Economic Studies*, 81 (1), 137–185. [662, 668, 669, 670]
- Drange, N. (2015), "Crowding out dad? The effect of a cash-for-care subsidy on family time allocation." *Nordic Journal of Political Economy*, 40, 1–29. [669, 698]
- Drange, N., T. Havnes, and A. M. Sandsør (2016), "Kindergarten for all: Long run effects of a universal intervention." *Economics of Education Review*, 53, 164–181. [673]
- Drange, N. and M. Rege (2013), "Trapped at home: The effect of mothers' temporary labor market exits on their subsequent work career." *Labour Economics*, 24, 125–136. [663]
- Dustmann, C. and U. Schönberg (2012), "Expansions in maternity leave coverage and children's long-term outcomes." *American Economic Journal: Applied Economics*, 4 (3), 190–224. [660]
- Eckstein, Z. and K. I. Wolpin (1989), "Dynamic labour force participation of married women and endogenous work experience." *The Review of Economic Studies*, 56 (3), 375–390. [668, 670]
- Fang, H. and D. Silverman (2009), "Time-inconsistency and welfare program participation: Evidence from the NLSY." *International Economic Review*, 50, 1043–1077. [671]
- Fernald, A., V. A. Marchman, and A. Weisleder (2013), "SES differences in language processing skill and vocabulary are evident at 18 months." *Developmental science*, 16 (2), 234–248. [696]
- Ferrall, C. (2012), "Explaining and forecasting results of the self-sufficiency project." *Review of Economic Studies*, 79, 1495–1526. [661]
- Francesconi, M. (2002), "A joint dynamic model of fertility and work of married women." *Journal of Labor Economics*, 20 (2), 336–380. [660, 668, 670]
- Gayle, G.-L., L. Golan, and M. A. Soytas (2014), "What accounts for the racial gap in time allocation and intergenerational transmission of human capital." Working Paper, Department of Economics, Washington University. [662]
- Gayle, G.-L. and R. A. Miller (2006), "Life-cycle fertility and human capital accumulation." Working paper, Carnegie Mellon University. [660]

- Gregg, P., E. Washbrook, C. Propper, and S. Burgess (2005), "The effects of a mother's return to work decision on child development in the UK." *The Economic Journal*, 115 (501), F48–F80. [660]
- Griffen, A. S. (forthcoming), "Evaluating the effects of child care policies on children's cognitive development and maternal labor supply." *Journal of Human Resources*, In Press. <http://doi.org/10.3368/jhr.54.3.0315.6988R1>. [662, 679]
- Hægeland, T., O. Raaum, and K. G. Salvanes (2005), "Pupil achievement, school resources and family background." IZA Working Paper. [685]
- Hart, B. and T. R. Risley (1995), *Meaningful Differences in the Everyday Experience of Young American Children*. Paul H Brookes Publishing. [696]
- Havnes, T. and M. Mogstad (2011), "No child left behind: Subsidized child care and children's long-run outcomes." *American Economic Journal: Economic Policy*, 3, 97–129. [660, 663]
- Heckman, J. and B. Singer (1984), "A method for minimizing the impact of distributional assumptions in econometric models for duration data." *Econometrica*, 52, 271–320. [667]
- Heckman, J. J. and S. Mosso (2014), "The economics of human development and social mobility." *Annu. Rev. Econ.*, 6 (1), 689–733. [685]
- Hotz, V. J. and R. A. Miller (1988), "An empirical analysis of life cycle fertility and female labor supply." *Econometrica*, 56, 91–118. [660]
- Kaboski, J. P. and R. M. Townsend (2011), "A structural evaluation of a large-scale quasi-experimental microfinance initiative." *Econometrica*, 79 (5), 1357–1406. [661]
- Keane, M. P. and K. I. Wolpin (2001), "The effect of parental transfers and borrowing constraints on educational attainment." *International Economic Review*, 42 (4), 1051–1103. [671]
- Keane, M. P. and K. I. Wolpin (2010), "The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women." *International Economic Review*, 51 (3), 851–892. [677, 686]
- Liu, H., T. Mroz, and L. Adair (2009), "Parental compensatory behaviors and early child health outcomes in Cebu, Philippines." *Journal of Development Economics*, 90 (2), 209–230. [684]
- Ljones, O. (1979), "Kvinneres yrkesdeltaking i Norge: en kryssløpsstudie (Female Labor Activity in Norway). Statistics Norway." [660]
- Michalopoulos, C., P. K. Robins, and I. Garfinkel (1992), "A structural model of labor supply and child care demand." *Journal of Human Resources*, 27, 166–203. [660]
- Moffitt, R. (1984), "Profiles of fertility, labour supply and wages of married women: A complete life-cycle model." *The Review of Economic Studies*, 51 (2), 263–278. [660]

- Mogstad, M. and C. Pronzato (2012), "Are lone mothers responsive to policy changes? Evidence from a workfare reform in a generous welfare state." *The Scandinavian Journal of Economics*, 114 (4), 1129–1159. [674]
- Rauan, E. (2013), *Undersøking om foreldrebetaling i barnehager, januar 2013 (Household payments in kindergartens, January 2013)*. Official Statistics Norway. Report 2013/28. [676]
- Ribar, D. C. (1992), "Child care and the labor supply of married women: Reduced form evidence." *Journal of Human Resources*, 27, 134–165. [660]
- Ribar, D. C. (1995), "A structural model of child care and the labor supply of married women." *Journal of Labor Economics*, 13, 558–597. [660]
- Schøne, P. (2004), "Labour supply effects of a cash-for-care subsidy." *Journal of Population Economics*, 17 (4), 703–727. [698]
- Statistics Norway (2001), *Arbeidskraftundersøkelsen. Yrkesveiledning (Labour Force Survey 2001)*. Official Statistics Norway Report C748. [660]
- Todd, P. E. and K. I. Wolpin (2003), "On the specification and estimation of the production function for cognitive achievement." *The Economic Journal*, 113 (485), F3–F33. [661, 684]
- Todd, P. E. and K. I. Wolpin (2006), "Assessing the impact of a school subsidy program in Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility." *American Economic Review*, 96, 1384–1417. [661]
- Todd, P. E. and K. I. Wolpin (2007), "The production of cognitive achievement in children: Home, school, and racial test score gaps." *Journal of Human Capital*, 1 (1), 91–136. [661]
- Van der Klaauw, W. (1996), "Female labour supply and marital status decisions: A life-cycle model." *The Review of Economic Studies*, 63 (2), 199–235. [670]

Co-editor Petra E. Todd handled this manuscript.

Manuscript received 18 September, 2015; final version accepted 15 August, 2017; available online 22 September, 2017.