Essays on Human Capital, Fertility, and Child Development

by

Seyed Mohammad Sadegh Eshaghnia

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2019 by the
Graduate Supervisory Committee:

Basit Zafar, Co-Chair
Esteban Aucejo, Co-Chair
Matthew Wiswall
Gustavo Ventura

ARIZONA STATE UNIVERSITY

May 2019

ABSTRACT

This dissertation consists of two chapters. The first chapter studies children's skill formation technology while endogenizing the maternal age and child investments. I estimate the effect of a mother's age at childbirth on her child's health, skill level, educational attainment, and adulthood earnings. There is a tradeoff between delaying childbirth to provide a more secure economic environment for mother and child versus the potential negative biological consequences for a child of having an older parent. I quantify this tradeoff. The results indicate that a five-year decrease in the maternal age of educated women, ceteris paribus, results in over 0.50 std increase in the childs skill level due to an increase in the childs ability to acquire skills. However, if one adjusts child investment according to individuals wage profile conditional on reduced maternal age, the average childs skill level decreases by 0.07 std. This reduction in childrens skill highlights the impact of lower inputs that children of younger mothers receive. The negative effect of foregone wages may be reduced through policy approaches. My policy analysis indicates implementing a two-year maternity leave policy that freezes mothers wages at the level before childbirth would reduce average maternal age at the first birth by about two years, while also increasing the average childs skill level by about 0.22 std and future earnings by over 6%.

In chapter two, I study the impact of females' perceptions regarding their future fertility behavior on their human capital investments and labor market outcomes. I exploit a natural experiment to study the causal effect of fertility anticipation on individual's investments in human capital. I use the arguably exogenous variation in gender mix of children as an exogenous shock to the probability of further fertility. I document that having two children of the same gender is associated with about 5% lower wages for the mother compared to having two children of the opposite sexes. Mothers with same-sex children perceive themselves as more likely to bear one more

child, and so less attached to the labor market, so invest less in human capital, and this is reflected in wages today.

To my family.

ACKNOWLEDGEMENTS

I am deeply indebted to MattewWiswall for his endless support and guidance. I am grateful to Basit Zafar, Esteban Aucejo, Gustavo Ventura, and MattewWiswall for their great advice and insightful comments. I benefited from many discussions with Gregory Veramendi and Daniel Silverman. I also thank Jerome Adda, Kvein Reffet, Nick Kuminoff, Alvin Murphy, Galina Vereshchagina, Stephie Fried, Seung Ahn, Francesco Agostinelli, Roozbeh Hosseini, Alexander Bick, Alejandro Manelli, Domenico Ferraro, Kelly Bishop, Berthold Herrendorf and all the seminar and workshop participants at the ASU General Economics Workshop and ASU PhD Seminar for their comments and feedback.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

Table Page

Table	Page

LIST OF FIGURES

Figure | Page

Chapter 1

CHILDHOOD INVESTMENTS, ABILITY, AND ENDOGENOUS MATERNAL AGE

1.1 Introduction

A remarkable change in the age of motherhood has occurred gradually in American society, as well as in other developed countries, over the last four decades: the trend is to postpone childbirth until a later age. In the US, the mean age of mothers at first birth has steadily increased from 21.4 years in 1970 to 26.6 years in 2016 (OECD (2018)).

One reason for this trend is the increase in female labor supply.[1] Postponing childbirth allows females to build human capital, which has returns in the labor market and presumably in investing in children. There is an extensive literature in economics emphasizing the role of parental investment in child outcomes (Del Boca *et al.* (2013); Cunha *et al.* (2010)).[2] Since parental financial resources increase with parental age (Powell *et al.* (2006)), mothers who delay parenthood will have more financial ability to invest in the childhood of their children.[3] Thus, children might benefit from

[1]Womens returns to experience have increased significantly since the 1970s (Blau and Kahn (1997) and Olivetti (2006)), and the labor force participation rate for US women has increased from 33.8% in 1950 to 56.7% in 2015 (OECD (2018)). There has been also a dramatic increase in married womens hours of labor market work during that period (Olivetti (2006)). Caucutt *et al.* (2002) argue that increasing returns to labor market experience have led highly-educated women to delay their fertility and take advantage of stronger prospects in both labor and marriage markets.

[2]Also see Dahl and Lochner (2012) for the role of family income; Blau (1999) for a more conservative estimate of the effect of family income on children's outcomes; and Cooper *et al.* (2013) for a literature review on the relationship between household financial resources and childrens outcomes.

[3]It is a well-known fact that the real wage of individuals increases over the life-

being born at a later maternal age. However, there are potential negative consequences of postponement on child outcomes. Namely, the risks of Down syndrome, childhood cancer, and autism increase with the maternal age (Durkin *et al.* (2008); Johnson *et al.* (2009), Yip *et al.* (2006)), as does the risk of a low birthweight infant (Restrepo-Méndez *et al.* (2015); Goisis and Sigle-Rushton (2014)). Delayed motherhood is associated with higher risk of diabetes in the child (Byrnes (2001); Cardwell *et al.* (2009)). Older maternal age is also associated with mental retardation that occurs in the absence of Down syndrome but accompanied by other neurological conditions that affect the central nervous system, such as epilepsy, cerebral palsy, or birth defects (Tearne (2015); Yeargin-Allsopp *et al.* (1995); Drews *et al.* (1995)). Postponing childbirth is also associated with reduced intelligence in the child (Bacharach and Baumeister (1998)).

Thus, women face a crucial trade-off when deciding at which age to bear a child: on the one hand, postponement of childbirth is beneficial to their children because individuals receive higher wages as they accumulate human capital, which means women will have more resources to invest in their children at later ages. On the other hand, postponing childbearing might have a negative impact on the child, as described above. Women have to balance these positive and negative effects when deciding on the timing of childbirth and their subsequent investments. In this paper, I develop a life-cycle model and estimate it using data from the Panel Study of Income Dynamics (PSID), which provides information on the parents decisions on the timing of fertility, labor supply, and child investments as well as data on childs outcomes. I

cycle, at least for the first half of the work-life-cycle. The life-cycle wage profile is usually hump-shaped with a maximum at around age forty (Chéron *et al.* (2013)). Theoretically, individuals accumulate human capital through work experience (learning by doing), and this leads to an increasing wage profile (see Polachek *et al.* (2008) for a theoretical framework, and Lagakos *et al.* (2018) for an empirical cross-country study).

estimate my mixed-integer stochastic dynamic programming model using the method of simulated moments. First, the model is solved by backward induction given a set of parameters. Then, it is simulated for individuals over their life cycle. I define a set of moments that capture the relationships between different variables in the sample. These moments can be computed for both simulated and observed data. Model parameters are chosen such that simulated moments replicate data moments. Using the estimated model, I analyze both negative and positive potential impacts of postponed childbirth on a childs skill level and investigate the extent to which negative effects can be compensated for by additional parental investment during childhood. I further conduct policy counterfactual analyses to evaluate the impact on child development of implementing a maternity leave policy in the US.

Investment in the child has at least two components: child goods (money investment) and the time that the mother spends with the child (time investment). There is a trade-off between the two; the more hours the mother spends with her child, the fewer hours she can work, and the less financial resources she has for child goods. However, the trade-off between money investment and time investment has a dynamic nature due to the mothers wage life-cycle; as her wage increases, the set of feasible pairs of time and money investments expands. Thus, a mother can make more investments of both types by delaying her motherhood. Moreover, the trade-off between the two types of investments is endogenous; the individuals wage profile depends on her human capital stock, and the human capital accumulation process is determined by intertemporal labor supply decisions. Females fertility decisions might affect their participation in the labor market, as they need to make time investments in their children after childbirth, and it might negatively impact their human capital accumulation (future wages). Lower future wages lead to lower capacity for both their own consumption and child goods expenditures. Hence, the timing of fertility and decisions concerning child investments and labor supply are endogenously and intertemporally determined.

In order to study the dynamics of child development and the role of maternal age in this process, it is necessary to endogenize all of the individuals decisions on the timing of fertility, labor supply, leisure, consumption, and child investments. My model builds on a rich history of dynamic life-cycle models.[4] In this model, an individuals human capital stock determines the wage evolution over her life-cycle.[5] In each period, individuals make labor supply decisions and might increase their stock of human capital by working in that period. They also decide on the timing of childbirth and receive utility from the skill levels of their children. If a woman decides to conceive a child in a particular period, she has to drop out of the labor force for that period. Thus, the opportunity cost of having a child is higher in early periods because of the forgone human capital that could have been accumulated through working in that period, loss of which will affect her wages in all subsequent periods (Imai and Keane (2004)). Moreover, the opportunity cost of having a child is higher for highly-educated women because of higher forgone wages. When a child is born, the woman chooses to optimally allocate her time among leisure, work, and childcare, and to allocate her money between consumption and expenditure on the child. If she has a

[4]See Heckman and MaCurdy (1980) for a dynamic life-cycle model of female labor supply; Eckstein and Wolpin (1989) for a discrete choice model of fertility, labor supply, and wages; and Keane and Wolpin (1997) and Lee (2005) for a dynamic life-cycle model of occupational choice, education, and labor supply for men. My work is also related to that of Imai and Keane (2004), Heckman *et al.* (1999), and Shaw (1989) who developed a life-cycle model of labor supply, consumption, and human capital accumulation. Also see Attanasio *et al.* (2008) for a life-cycle model of female labor force participation and savings. Finally, see Hotz and Miller (1988), Gayle *et al.* (2006), and Sheran (2007) for a life-cycle model of female labor supply and fertility. None of the studies mentioned here, however, considered the dynamics of child development.

[5]See Wmann (2003) and Folloni and Vittadini (2010) for a summary of extant studies on the effect of human capital on wage.

child earlier in her life-cycle, the child is more likely to have higher productivity in acquiring skills, but the net present value of forgone wages is higher, which indicates that there would be fewer resources available to invest in the child to increase the childs skill level. This trade-off leads more-educated women to have children later in their life-cycles.[6]

My estimation results show a negative effect of advanced maternal age at childbirth on the child's productivity. I find that, everything else constant, lowering maternal age of educated women by five years increases the average child's skill level by about 11% (0.50 std), and 15% (0.24 std) increase in the child's future earnings due to a higher child's ability to acquire skills. However, if I adjust child investment according to individuals' wage profile conditional on reduced maternal age, the average child's skill level decreases by 1.6%. This reduction in children's skill highlights the impact of lower inputs that children of younger mothers receive from their parents due to the wage life-cycle implied by a lower maternal age.

The negative effect of foregone wages may be reduced through policy approaches such as maternity leave. Implementing a nationwide maternity leave policy has been the subject of heated debate among both policymakers and researchers in the US, in part because the United States is an outlier in maternity leave provision (Rossin-Slater (2017a)). The Family and Medical Leave Act (FMLA) entitles eligible workers to take only 12 weeks of unpaid parental leave, a short period of time compared to most other developed countries; for instance, Germany and France have three-year maternity leave periods.[7] Previous studies have investigated the impacts of maternity leave policies on fertility rate, the mothers labor market outcomes, and the

[6] They face other trade-offs as well: conceiving a child earlier means lower opportunity cost for every hour spent with the child instead of working due to lower wages at early stages of the work-life-cycle.

[7] Klerman *et al.* (2012) reports that only 60% of workers in the US were eligible for the FMLA in 2012.

childs health (Lalive and Zweimüller (2009); Schönberg and Ludsteck (2014); Rossin (2011)). However, the potential impacts of maternity leave policies on the timing of birth and child development, have not been investigated. My policy analysis indicates that implementing a two-year maternity leave policy under which the mothers wage is frozen at its pre-pregnancy level lowers the mean maternal age by about two years and increases the mean childs test score by about 5% (0.22 std). This increase in mean child skills is due to both higher productivity (better health status) as a result of the reduction in maternal age and an increase in child investments as a result of the lower wage penalty following childbirth. Back-of-the-envelope calculations suggest that the change in the mean test score of children translates to an increase of about a 6.6% (0.10 std) in their future earnings. With regard to the impacts of maternity leave policies on mothers' labor market outcomes, the predictions of my simulation analysis are in line with the findings of a handful of papers that have investigated the impacts of six weeks of paid leave in California on labor market outcomes (Rossin-Slater (2017b); Baum (2003)). My policy analysis shows that implementing a two-year paid maternity leave policy that freezes wages to the level before the leave would increase the labor force participation of educated and non-educated mothers of six-year-old children by about 20% and 34%, respectively.

I also study the impacts of implementing a childcare subsidy program on mothers and children. My policy analysis indicates that although subsidizing childcare services does not significantly change the maternal age, it can significantly increase the human capital of both mothers and children. My analysis suggests that under a free childcare program, the labor force participation of educated and non-educated women would increase by about 45% and 40%, respectively. This finding is remarkably close to the finding of Barros *et al.* (2011), who evaluate the causal impact of a childcare program on the maternal labor market outcomes taking advantage of a lottery carried by the municipal government in 2007 in Brazil.

This is the first study to investigate the effects of timing of birth on childrens cognitive development while taking into account the impacts of endogenous parental investments in children. Researchers have generally focused on the effects of different types of investments on the childs skill level. In their models, the life-cycle generally starts when a child is born. They take the timing of childbirth as a given and study how the childs cognitive ability react to monetary investments, the time that parents spend with their children, childcare usage, and other inputs received from parents. Cunha and Heckman (2008) emphasize the importance of investments during childhood and show that investments made in the early years are more important than those made in later years. Del Boca *et al.* (2013) argue that parents time inputs are more productive for their childrens cognitive development than money expenditures. However, if one does not take into account the impact of maternal age, her estimate of child's skill formation technology might be misleading. To estimate the impact of child investments on child's skill level, researchers use the variations in children's test scores and child investments. Older parents tend to to have more resources to invest in their children, however, part of these greater investments only compensate for the negative effect of a later childbirth. Hence, by overlooking the impact of maternal age, one might underestimate the impact of investments on the child's skill level. I "correct" previous estimates of investment's impact by endogenizing the timing of birth and taking maternal-age effect into account.

Current child development models cannot evaluate the role of mother's age at childbirth on their child's skill level, as current literature neglects the maternal age effect. Despite this, some studies have assessed the effect of parents' age at childbirth on the child's skill level through a reduced-form approach (Goisis *et al.* (2017); Leigh and Gong (2010); Barclay and Myrskylä (2016)). The problem is that, though these researchers try to control confounding factors by adding explanatory variables such as parental characteristics such as education or income, they overlook the time and monetary investments that parents make in their children (and the fact that these investments are endogenous with respect to parental age); thus, the estimates from their models might be impacted by the positive selection to delayed parenthood. Overlooking child investments might lead a researcher to underestimate the magnitude of the negative maternal age effect. Older parents tend to invest more in their children due to greater parental resources, and there is a dynamic complementarity between child investments and child's skill in producing the next period's skills. My simulation experiment shows that neglecting such child investments leads one to significantly underestimate the magnitude of the negative health impacts of greater maternal age on child outcomes. In this paper, I address this problem by taking into account both mother's age at childbirth and investments made by parents in the child. Moreover, I improve upon the reduced-form approach used in previous papers by decomposing the negative and positive effects of delaying childbearing on the child's outcomes.

The paper is organized as follows. Section 2 describes the data for the empirical work, provides some descriptive statistics, and presents suggestive evidence concerning the impact of postponement of childbirth on child outcomes. Section 3 describes the model. Section 4 discusses some estimation issues. Section 5 presents the estimation results. Section 6 analyzes the effects of some counterfactual policies, and section 7 provides the conclusions. Details on the model and supplementary results are presented in the Supplementary Appendix.

1.2 Maternal Age and Child Development: Reduced-form Evidence

In this section, after describing the datasets that I use, I show some evidence obtained using reduced- form regressions for the role of the mothers age at childbirth in determining the childs skill level and also the physical characteristics of the child.

1.2.1 Data

Throughout this paper, I utilize the Panel Study of Income Dynamics (PSID) combined with its four Child Development Supplements (CDS), and five Transition into Adulthood Supplements (TAS). PSID is a longitudinal household survey. The study began in 1968 with a nationally representative sample of over 18,000 individuals in about 5,000 families in the US. PSID provides researchers with panel data on education, marriage, wealth, employment, income, health, expenditures, childbearing, and child development of the initial sample of individuals and their descendents. The CDS includes data on children and their extended families, which can be used to study the dynamics of human capital formation in children. Specifically, CDS collects extensive child-specific developmental data. The first wave of CDS in 1997 included up to two children per household who were between zero to twelve years old. These children were followed over three waves, ending in 2013-14. The CDS sample size in 1997 was approximately 3,500 children in 2,400 households. The first follow-up study, i.e. CDS-II, was conducted in 2002-03, when the children were between the ages of eight and 18. The next follow-up survey with these children, i.e. CDS-III, was conducted in 2007-08. No new children were added to the study until 2013-2014 (Hofferth *et al.* (1997)). In 2014, CDS was conducted with a new sample including all children in PSID households aged 0-17 years, with a sample size of 4333. The other supplement to the PSID, i.e. the Transition into Adulthood Supplements (TAS), is a follow-up survey of the original CDS children that was conducted when those children reached adulthood, i.e. age 18. TAS was initiated in 2005. Hence, in the first wave of TAS, the oldest CDS respondents were 18 to 20 years of age. TAS continued to follow up the CDS sample in their adulthood in 2007, 2009, 2011, 2013, and 2015.

I use PSID and CDS, which provide me with data on children and their family characteristics, and include information on parents' investments in their children. I use the childrens test scores on Letter-Word (LW) Identification test as a measure of child skill levels. The LW test is a subset of the Woodcock-Johnson Revised (WJ-R) test of achievement. LW test aims to measure reading identification abilities of children (Woodcock and Johnson (1989)). This measure of skill level has been used by previous papers in the literature (see Del Boca *et al.* (2016)). With the TAS dataset, I link each childs LW test score to their future educational attainment and wages, and show that the LW test score has a predictive power for the future educational attainment and wages of the child and therefore is a relevant measure of the child's skill, at least with respect to future educational attainment and wages.

1.2.2 Mother's Age at First Birth (AFB)

Figure (1.1) shows the distribution of mothers age at first childbirth (AFB) for non-college graduates and college graduates. A great deal of variation exists with respect to mothers age at first birth, both among educated and less-educated females. It is also clear that college graduates, on average, tend to have their first child at a later age; both mean age and median age at first birth increase with the mothers education level. While the median age for mothers with at most a high school education is 21, it increases to 28 for those with a graduate degree. Furthermore, while the fraction of mothers who conceived their first child after age 27 is 0.13 and 0.21 for mothers with fewer than 12 and 12 years of schooling, respectively, it increases to 0.47 and 0.59 for college graduates and women with a graduate degree, respectively.[8] For the purpose

[8]Table (A.2) in the Supplementary Appendix reports the mother's mean age and median age at first childbirth for different education levels.

of this model, I focus on women who did not conceive a child before age 18.[9,10]

Table (1.1) reports summary statistics for the subsample selected for the empirical work.[11] On average, college graduates have their first child later compared to non-college graduates. College graduates also have higher mean hourly wage, mean husbands hourly wage, and mean hours of market work. They also spend more time with their children, and their children outperform the children of non-college graduates in terms of LW test scores.

1.2.3 Mother's Age at First Birth and Child Outcomes

Using the PSID dataset, I provide some evidence on the impact of mother's age at birth on the following child's outcomes: (1) the probability of a low birthweight infant, (2) the child's skill level, and (3) child investments.[12]

Mother's Age at Birth and the Risk of Low Birthweight (LBW)

Note that birthweight is a child outcome that occurs immediately after birth and is not impacted by later parental investments. Therefore, I use birthweight data to identify an immediate effect of maternal age on child outcomes. Using the PSID dataset, I examine the role of mothers age at childbirth on the probability of low birthweight (LBW) of the child. LBW infants face many complications in their lives, some of

[9]In my dataset, about 90 percent of women did not conceive a child before age 18.

[10]Figure (A.1) in the Supplementary Appendix shows the distribution of the mother's age at first childbirth for the subsample selected for the estimation (women who have not conceived their first child before age 18). Figure (A.2) in the Supplementary Appendix shows the distribution of completed years of schooling for the same subsample against the alternative group (women who had their first child before the age of 18). It shows that women who conceived a child before age 18 are, on average, less-educated than those who did not.

[11]Table (A.1) in the Supplementary Appendix reports summary statistics for the full sample, which are similar to those of the selected subsample.

[12]According to the World Health Organization (WHO), an infant is low birthweight if it weighs less than 2,500 grams (5 pounds, 8 ounces) at birth (WHO (2014)).

which persist into adult life. LBW infants are more likely to suffer from weaknesses in attention and hyperactivity, anxiety and depression, and poor social skills (Hack *et al.* (2009)). These behavioral complications affect the cognitive outcomes of LBW infants. Ribeiro *et al.* (2011) reports attention problems and language development in LBW children, while Conley and Bennett (2000) reports that a LBW child is substantially less likely to graduate from high school by age 19; specifically, the probability of graduation is reduced by 74 percent, as compared with his or her siblings.

Table (1.2) reports the result of a linear probability regression of LBW on mothers age while controlling for child gender, race, year fixed-effects, mothers education level, paternal age at delivery, marital status, and family income. In order to capture a potential non-linear relationship between maternal age and the probability of LBW, I consider two specifications: first, a linear specification, and second, a quadratic form in which I add the square of maternal age as an additional regressor. Columns 1 and 2 of Table (1.2) present the linear and quadratic regression results, respectively, before controlling for family income. Columns 3 and 4 report the results when also controlling for family income at delivery. Both the first and the second power of the mothers age at delivery are significant at the five percent level. These results suggest that the risk of LBW in infants with respect to maternal age at delivery can be represented by both a U-shaped and a linear-shaped relationship.[13] There is a higher probability of LBW associated with adolescent pregnancy. Meanwhile, there is also a higher risk associated with delaying motherhood for most maternal

[13]The results still might be driven by selection based on unobservables, so I will be cautious in their interpretation. Figure (A.3) in the Supplementary Appendix depicts a binned scatterplot showing the U-shaped association between LBW and maternal age. The left part of the curve might be due to selection, as adolescent childbearing is concentrated among teenagers of low-income families, and low-income youth are at higher risk for nutrition problems that increase the risk of LBW.

ages, and this risk increases as the mothers age increases. Delaying pregnancy by one additional year is associated with about 0.47 to 0.60 percentage points higher risk of LBW (5 to 7 percent, compared to mean).[14] I also could use the variation in maternal ages between siblings to investigate the impacts of maternal age on the risk of LBW. Table (A.3) in the Supplementary Appendix shows the results of a fixed effect regression, in which the variation in maternal ages between sibling is used to estimate the impact of the mother's age at birth on the risk of LBW. The results are similar to the linear regression model of Table (A.3); delaying pregnancy by one additional year is associated with about 0.66 percentage points higher risk of LBW.

One might be concerned that this result is driven by LBW deliveries occurring above the age of 35, in which case the result would be misleading because there are no differences between maternal ages 20-25 and 25-30. To address this concern, I repeated the probability regression of LBW on mothers age with dummy variables representing five-year bins for the mothers age at first birth as follows: ages 15-20, ages 25-30, ages 30-35, ages 35-40, ages 40-45, and ages 45-50. Maternal ages between 20 and 25 were omitted, as this group provides the baseline for comparisons. Table (1.3) reports the results of this regression. The hypothesis that the previous regression results were driven by LBW at extreme maternal ages is rejected. Instead, the results suggest that an increase in maternal age from 20-25 to 25-30 is associated with a 4.3 percentage point (58%) increase in the risk of LBW.

One might also be concerned that the above-mentioned regression results are driven by selection based on the socioeconomic status. Table (1.4) reports the same regression analysis as before, but this time utilizing within-group variation in mothers age at delivery and the LBW status of her child. Panel A of Table (1.4) shows

[14] Table (A.4) reports the results when I do not deal with the issue of missing data on incomes.

that when considering the relationship between maternal age at delivery and the risk of LBW within educational strata, there remains a significant positive relationship between maternal age and LBW. The best fit for non-college graduates is a U-shaped specification, while the best fit for college graduates is a linear one.[15] Figure (1.2) separately shows the predicted probability of LBW at each maternal age for non-college graduates and college graduates, based on coefficients from the regression results of Panel A in Table (1.4). The predicted probability of LBW increases with maternal age for both groups, but is lower for college graduates, which might be due to better average health status among that group. Panel B of Table (1.4) shows that a similar pattern occurs when investigating the relationship within family income levels.[16]

In addition, I observe a negative association between infant birthweight and the mothers age at birth. These results are reported in the Supplementary Appendix.[17]

I furthermore provide in the Supplementary Appendix, a brief literature review on the effect of mothers age at delivery on the childs intelligence, risk of Down syndrome, risk of any chromosomal abnormality, and risk of autism.

[15]Figures (A.4) and (A.5) in the Supplementary Appendix depict binned scatterplots assuming quadratic and linear associations between LBW and maternal age within education levels, respectively. Tables (A.5) and (A.6) provide the regression results from quadratic and linear forms, respectively.

[16]Tables (A.5) and (A.6) provide the regression results for quadratic and linear forms, respectively. Figures (A.6) and (A.7) in the Supplementary Appendix depict the binned scatterplots assuming quadratic and linear associations between LBW and maternal age within income levels, respectively. For the above-median income families, the relationship tends to be an upward linear trend instead of U-shaped. This might suggest that the left part of the U-shaped curve in previous figures is due to selection into adolescent pregnancy based on income, as most teenage pregnancies occur among low-income families.

[17]Table (A.7) and Figure (A.8) in the Supplementary Appendix show the relationship between maternal age and the natural logarithm of the child's birthweight while controlling for the same set of variables used in previous regressions for the risk of LBW. The results suggest a negative association between birthweight and maternal age. See Black *et al.* (2007) for the impacts of birthweight on adulthood outcomes.

Mother's Age at Birth and Child's Skill Level

To investigate the impacts of AFB on child skill level, I use a linear specification as follows:

$$Y_{i,t^k} = \beta_1 age_i^p + \gamma Z_{i,t^k} + \alpha W_i + u_{i,t^k} \equiv BX_{i,t^k} + u_{i,t^k}, \tag{1.1}$$

where Y_{i,t^k} is the child outcome of interest for family i at child's age t^k, age_i^p is the maternal age at first birth for family i, β_1 is the coefficient of interest, W_i is a vector of child fixed-effect controls such as gender and race; Z_{i,t^k} is a vector of family controls such as the mother's education, father's age at childbirth, mother's marital status, and family income; and $E(u|X) = 0$.

The first column of Table (1.5) shows the result of OLS regressions of the log test score of the first child on the mother's age at childbirth using only data from one-child families.[18] The first column reports the regression result when I control only for the child's age; i.e. age-specific dummy variables are used to control for the effect of the childs age on their test score. The coefficient of the mothers age is positive, but statistically insignificant. The second and third columns present the results when controlling for other confounding factors: the mothers and fathers education, and the mothers and fathers wages, respectively. After controlling for the mothers and fathers education, the coefficient of interest becomes negative (changing from 0.001 to -0.005). When controlling for additional confounding factors, the coefficient of the mother's age at childbirth becomes greater in absolute value; the coefficient of

[18] Regression results including all families are reported in the Supplementary Appendix. Here, I restrict the sample to one-child families to avoid any sibling effects and/or parental skill effects. For a given mother, parental skills might be different when rearing the first child vs the second one. Mothers can accumulate parental skills by "practicing" parenthood when rearing the first child. Duncan *et al.* (2018), using a reduced-form approach, find that each additional year of delaying a second childbirth after the first one is associated with a 0.02 to 0.04 standard deviation increase in school achievement of the second child, which might be due to accumulating parenting skills over the first child. However, it also might be due to increasing parental resources over the mother's age.

interest changes to -0.014, and it is statistically significant at 5%.[19] This suggests that postponing motherhood for one additional year is on average associated with a 1.4 percent decrease in the childs test score, ceteris paribus.[20]

Note, however, that the estimate presented in Tables (1.5) are biased. If we have omitted variables that contribute to the effect of the mothers age at childbirth on the childs skill levelsuch as parental investments in the childand if these variables are correlated with the mothers age at childbirth, then estimates of the reduced-form OLS regressions will be biased. For example, if older parents spend more hours with their children, then under the maintained assumption that these time investments positively impact the childs skill level, the omitted variable bias causes the coefficient of interest in Table (1.5) to be underestimated (in terms of magnitude). In what follows, I show some evidence that the omitted variables are in fact correlated with the mothers age at childbirth.

Mother's Age at Birth and Child Investments

One approach to demonstrating the relationship between mothers age at childbirth and the amount of time she invests in her child is to regress the time spent by the mother with her first child on age at first birth and determine whether the coefficient

[19]In Table (1.5), I keep the number of observations fixed across regressions in different columns. Alternatively, I could utilize all observations in all regressions in different columns. In that case, going from the first column to the last, the number of observations declines as I add more confounding factors. This is because of missing observations on wages when a mother does not work in a given period. Table (A.8) in the Supplementary Appendix shows the mentioned alternative regression results. There might be a concern regarding selection due to the missing data on mothers' wages, so I prefer to keep the number of observations fixed across different regressions. Tables (A.9) and (A.10) report the regression results if I also include households with more than one child (so sibling effects might exist), when the number of observations are fixed and when the number of observations varies, respectively.

[20]In other words, delaying childbearing for one more year is on average associated with $\frac{1}{70}$ standard deviation decrease in the child's test score.

is positive. I use the weekly number of hours that the mother spends with her child as the dependent variable. Table (1.6) reports the result of this regression. As the first column shows, the coefficient of the mothers age at childbirth is 0.22; it is both positive and statistically significant at 1% after I control for race, the childs age, and family size. As I control for the mothers education in the second column, and for the mothers wage and hours of market work in the last column, the coefficient of interest does not change significantly; it slightly decreases to 0.19 and remains significant at the 1% level.[21,22]

These results show that there is a systematic relationship between time investment in children and the mother's age at childbirth. This might suggest that it is important to simultaneously model the timing of fertility and the investments made by parents during the childhood, as these two factors are both related to each other and important with respect to the childs skill formation.

In the next section, I develop a structural model that enables me to investigate the role of childbearing age in the child's skill formation.

1.3 Model

In this section, I develop a life-cycle model that will be used in my empirical work. Note that I model the problem from the woman's perspective. Each individual lives until age 60 (so she lives for 42 years within the model). The model assumes that

[21]Given the five-year gap between the childbearing ages of educated and non-educated women, the results of the above regression can be translated as follows: a simple calculation shows that, everything else constant, the time that an educated woman actively spends with her first child during their childhood, is on average 750 hours more than for a non-educated woman. (For this calculation, we assume that the development phase of children takes 15 years.) Note that time investments are very important for developing the child's cognitive ability, as argued by Del Boca *et al.* (2013), so I expect that this substantial difference in time investment between older parents (educated women) and younger parents (non-educated women) make a remarkable difference in terms of child's skill level.

[22]I keep the number of observations fixed through all regressions in Table (1.6).

a women can conceive a child until age 40, and she becomes infertile thereafter.[23] Each individual decides whether or not to have a child and the timing of childbirth, and she can have only one child. When a child is born, parents make human capital investments in the child over each period during the childs developmental phase. The developmental phase in this model begins when a child is born and lasts for $M (= 16)$ years.[24] Therefore, parents invest in their child for the first "M" years of the childs life. At age 17, the child leaves her parents and becomes independent. I model this problem as a finite horizon problem with discrete time, and each period in the model equals a year.

It is worth mentioning that I do not model the educational decisions of (potential) mothers here. However, I consider their educational decisions as exogenous in my analysis, and assume that if she chooses to continue her education in college, then she will remain in college for four years and will not conceive a child until age 22.

1.3.1 Timing of Events

The initial condition is as follows: the model considers women with no child at age 18 (I do not model adolescent pregnancy) who have already decided whether or not to continue their education in college.[25] A fraction of these individuals are married, the rest are single, and the probability of being married depends on the educational decision, i.e. whether to go to college or not.[26] Furthermore, they have no work

[23]According to the data, less than half a percent of mothers conceived their first child after age 40.

[24]I use the Letter-Word (LW) identification test score as a measure of the child's skill level. I have the children's LW test scores up to age 18. However, it seems that LW test scores in the dataset do not significantly change after age 16. Thus, I assume the development stage lasts for 16 years. Similarly, Del Boca *et al.* (2013) assume development stage of 15 years.

[25]According to the data, less than 10% of mothers conceived their first child before age 18.

[26]This probability is set according to the data.

experience.

At the beginning of each period, first, exogenous shocks to marital status are realized. With some probability, single women may get married and married women may get divorced; these probabilities depend on current marital status, the woman's age and education level, and the presence of children.[27] Next, nonlabor income shock and wage shock are realized. Both of these shocks are exogenous and i.i.d. across individuals and over time.

After observing these two shocks, if a woman is fertile and not in college, she decides whether or not to try to conceive a child. If she tries, then with probability $\pi_{i,t}$, she becomes pregnant and will be out of the labor force for at least one period (current period). The probability of getting pregnant, $\pi_{i,t}$, depends on the woman's age, t. I also allow for conception error: she faces a probability $\lambda(s)$ of getting pregnant whether she intends to or not, where s denotes the woman's education level. Then, if she is not pregnant, she chooses her labor supply. After employment status is established, she decides about consumption, leisure, and if there is a child whose age is under M years, also about child expenditures and the time to spend with the child. Figure (1.3) shows the timing of events in period t.

1.3.2 Utilities

I assume that utility is intertemporally separable, and the instantaneous utility is separable in consumption and leisure (see Del Boca *et al.* (2013)).[28] The utility of a woman in period "t," $u_{i,t}$, is derived from household consumption, $c_{i,t}^{HH}$, and the amount of leisure time, $l_{i,t}$. If a child exists at the beginning of period "t," then the

[27]See Klaauw (1996), Keane and Wolpin (2010), Francesconi (2002), and Adda *et al.* (2013) for studies that endogenize marriage decisions in a life-cycle model of female labor supply. They do not model child development, however.

[28]See Blundell *et al.* (2016a) for an intertemporally separable but instantaneously not separable utility function in consumption and leisure.

woman also has utility from her childs skill level, $k_{i,t}$. The women's age is indexed by "t." I assume a log linear utility function; hence, when no child exists, utility is defined over consumption and leisure as follows:

$$u^0(c_{i,t}^{HH}, l_{i,t}) = \alpha_1 log(c_{i,t}^{HH}) + \alpha_2 log(l_{i,t}),$$

As stated before, when a child exists, the woman also has a utility derived from her child's skill level. I assume that when the developmental phase is done (i.e. after the first "$M+1$" periods of the child's life), the child leaves her home and begins the next stage of her development. Thus, the flow utility derived from the child's skill level after the child leaves home is discounted by a multiplier ϕ, which is a free parameter to be estimated. Therefore, I have the following utility when a child exists:

$$\begin{aligned} u^1(c_{i,t}^{HH}, l_{i,t}, k_{i,t}, t^k) &= \alpha_1 log(c_{i,t}^{HH}) + \alpha_2 log(l_{i,t}) \\ &\quad + \alpha_3[I_{i,t^k \leq M} log(k_{i,t}) + \phi(1 - I_{i,t^k \leq M}) log(k_{i,M+1})], \end{aligned}$$

where $I_{i,t^k \leq M}$ is an indicator variable equal to one if the child's age is less than M, and zero otherwise. I normalize the impact factors, i.e, α's, such that $\sum_{j=1}^{3} \alpha_j = 1$.

Therefore, a woman receives the following (flow) utility at period t:

$$u(c_{i,t}^{HH}, l_{i,t}, k_{i,t}, bb_{i,t}, t^k) = bb_{i,t} u^1(c_{i,t}^{HH}, l_{i,t}, k_t, i, t^k) + (1 - bb_{i,t}) u^0(c_{i,t}^{HH}, l_{i,t}), \qquad (1.2)$$

where $bb_{i,t}$ is is an indicator variable equal to one if there exists a child at the beginning of the period t and zero otherwise, and $u^j(.)$ is defined in the previous equations $\forall j = 0, 1$.

1.3.3 Child's Skill Formation

Let t^k denote the child's age. I assume that the child's skill level at $t^k + 1$, k_{i,t^k+1} is produced by the current level of the child's skill, k_{i,t^k}; the time that the mother actively spends with the child, τ_{i,t^k}; and the money investment in the child, e_{i,t^k}. I use a Cobb-Douglas functional form to specify the child's skill formation:

$$k_{i,t^k+1} = f(R_{i,t^k}, \tau_{i,t^k}, e_{i,t^k}, k_{i,t^k}) = R_{i,t^k} k_{i,t^k}^{\gamma_{1,t^k}} \tau_{i,t^k}^{\gamma_{2,t^k}} e_{i,t^k}^{\gamma_{3,t^k}}, \tag{1.3}$$

where R_{i,t^k} is a function of the mother's age at birth and innate ability, and the child's age; i.e. $R_{i,t^k} = g(age_i^p, IQ_i^p, t^k)$. I use the following functional form to specify the relationship between the child's productivity and her mother's characteristics:

$$log(R_{i,t^k}) = log(A_{t^k}) + \epsilon_1(age_i^p) + \epsilon_2(age_i^p)^2 + \epsilon_3(IQ_i^p) + \upsilon_{i,t^k}, \tag{1.4}$$

In this specification, age_i^p is the mother's age at childbirth; ϵ_1 and ϵ_2 together define the effect of maternal age at delivery on the productivity factor; IQ_i is the mothers innate ability; and ϵ_3 captures the effect of the mothers innate ability on the productivity factor. Finally, υ_{i,t^k} is the productivity shock, which I assume is i.i.d. $N(0, \sigma_R^2)$. I use a Cobb-Douglas functional form to specify the skill formation technology. Other empirical studies have used linear, constant elasticity of substitution (CES) or Cobb-Douglass functional forms to specify the child's skill production function (see for example Cunha and Heckman (2008), Bernal (2008), Del Boca *et al.* (2013), and Lken *et al.* (2012)). I do not use a linear functional form for two reasons. First, it imposes the restriction that early and late investments are perfect substitutes; and second, I want to allow the inputs to interact in producing the output.

Assuming a log-linear production function would help make a complicated problem more tractable. Nonetheless, I also use the Cobb-Douglas functional form to specify the production function; this enables me to get closed-form solutions for expenditures and time investments in the child. However, note that choosing a Cobb-Douglas functional form to specify the childs skill production function, in combination with the log linear utility function, implies that time and monetary investments are independent of the childs skill level. This might not be true in reality; parents may consider the childs skill level when making decisions about human capital investment in their children.

I do not use a linear function to specify the productivity term because, based on the medical evidence, the marginal effect of the mothers age at childbirth on the childs innate ability might not be constant; instead, the marginal detrimental effect of maternal age may be higher as maternal age increases.

I model the effect of the parents age at childbirth on the childs innate ability by including the parents age in the productivity term. Indeed, I assume that human capital corresponds to any stock of knowledge, experience, and skills but has nothing directly to do with innate ability.

One might expect that including other regressors such as the mothers health status might attenuate my estimate of the maternal age coefficient. First of all, the reduced form regressions presented in the previous section suggest that maternal age plays a role in the childs skill formation, and here, I aim to investigate that role. Hence, the production function includes maternal age. Moreover, the results of Table 2 suggested that adding more and more variables makes the coefficient of the maternal age greater (in absolute value) and more precise. Note that I do not intend to (and do not need to) exclude the effects of other factors that affect the childs productivity, as long as these factors are driven by the mothers age or happen to occur as mothers get older.[29]

[29] As an example, consider the effect of the mothers health status on the childs productivity, which is omitted in my specification. Suppose that the mothers health status is negatively correlated with her age and might affect the childs productivity in some way. Now, if health status were exacerbated as the mother grows old (for example, some mutation might happen in the reproductive cells that negatively affects the childs productivity, and of course, that mutation is positively correlated with the mothers age), then I do want the coefficient of age in the productivity function to pick up such effects of health as well. Note that the individual faces a trade-off between childbearing versus continuing to participate in the labor market. As long as they do not have the power to avoid such negative effects of delayed childbearing on child outcomes, it is not important what factor actually makes older parents more likely to have a child who might have lower productivity. Interpreting the age coefficient this

1.3.4 Skills and Wages

Wages depend on education level and actual experience, which may depreciate when out of work; the depreciation rate potentially depends on education level, as in Blundell *et al.* (2016a).

I assume that individuals accumulate skills by working. For any given period, if they work, then their skills will be increased by one unit. Therefore, I have the following equation:

$$x_{i,t} = x_{i,t-1} + I_{h_{i,t} \geq \underline{h}} - (1 - I_{h_{i,t} \geq \underline{h}})\delta_{s,t} x_{i,t-1}, \tag{1.5}$$

where $x_{i,t}$ denotes human capital, which, in this model, I assume to be the number of (effective) years of work experience accumulated up to period t; $I_{h_{i,t} \geq \underline{h}}$ is an indicator variable equal to one if the individual works at least $\underline{h}$ hours; and $\delta_{s,t}$ is the depreciation rate of skills due to unemployment or career interruptions associated with the years that the individual does not work, which I allow to depend on education level, s (time invariant), and the woman's age, t, at which she drops out of the labor market. I allow the depreciation rate to depend on education level. This allows the model to capture the fact that while out of the labor force, less-educated individuals with routine jobs may loose less than educated individuals in more skilled (abstract) jobs.[30] I allow the depreciation rate to depend on the woman's age because in the first few years of the work-life-cycle, individuals might build their career, develop a network of professional contacts, show their commitment to work, and improve their reputation;

way, I might leave the age effect as a black box; I do not know if it is because health status becomes worse as people get older, because some mutations in genes are more likely to happen as more time passes, et cetera. I am interested in investigating the implications of this fact on both individuals decisions and child outcomes.

[30]I assume that the depreciation rate is constant across individuals. I know that it might be different across occupations; some women might select into child-friendly occupations, and the depreciation rate for skills might be lower in those occupations (see for example Adda *et al.* (2017)). However, I do not focus on this issue here.

thus work interruptions during this phase might have a detrimental, irreversible impact on the individual's human capital in its broad sense. In contrast, dropping out of the labor market for one/a few years later in the work-life-cycle may have only a mild negative impact on the individual's human capital. The above-mentioned specification of human capital allows the impacts of work interruption to decrease during the work-life-cycle, i.e. $\frac{\partial \delta_{s,t}}{\partial t} < 0$. Finally, I assume that a lower bound is needed for working hours in order to obtain an increase in the work experience of individuals. I set $\underline{h}$ to correspond to working four hours per workday.[31]

Following the Mincerian approach, I assume that the wage offer received depends on education level and skills. However, I specify the wage equations separately for educated and non-educated females:

$$log(w_{i,t}) = \eta_{0,s} + \eta_{1,s} x_{i,t} + \eta_{2,s} x_{i,t}^2 + \omega_{i,t}^s, \;\; s = H, C, \tag{1.6}$$

where subscript s denotes the individual's education level indicating non-college graduates if it takes H, and college graduates if C; $\eta_{0,s}$ is a constant; $\omega_{i,t}^s$ is the log wage shock, which is i.i.d. $N(0, \sigma_{w,s}^2)$; and $\eta_{1,s}$ and $\eta_{2,s}$ define the return to work experience for group s.

With the above-mentioned specification, I allow the return to experience to increase with education and decrease with the level of work experience. As I assume that if a woman decides to conceive a child, she has to drop out of the labor force for one period, this specification implies that the opportunity cost of childbearing is higher for more-educated women than for less-educated women. Moreover, the

[31]Indeed, $\underline{h}$ is the cut-off level for the lower bound of hours needed in every given period in order to change the years of work experience effectively. Note that without this threshold, the solution to labor supply is sometimes not well-defined because one might work for an infinitesimal amount (that converges to zero) just to take advantage of greater future wages through higher experience, so the solution does not exist. I set $\underline{h}$ to 20 hours per week. I assume that to get one unit of experience, the individual needs to work at least half as much as an average full-time worker.

opportunity cost of conceiving a child might be higher at the beginning of the work-life-cycle. The reason for the higher cost is that if a woman drops out of the labor force for one period early in her working life, even though the forgone wage is relatively low, her wages in all subsequent periods will be lower as a consequence of being out of the labor force and not accumulating human capital.[32]

1.3.5 Marriage and Divorce

I do not model marriage and divorce in this paper as choices of individuals. Instead, I incorporate them in the model as exogenous events.[33] I assume that the probability of getting married, $\Pi^m_{i,t}(age_{i,t}, bb_{i,t}, s_i)$, depends on woman's age, the presence of children, and her education level. If an individual is married in a given period, the probability of getting divorced, $\Pi^d_{i,t}(age_{i,t}, bb_{i,t}, s_i)$, again depends on the woman's age, the presence of children, and the education level. I use data to estimate these probabilities non-parametrically– i.e. I use the crude marriage (divorce) rates while conditioning on marital status, the woman's age and education level, and the presence of children.[34]

1.3.6 Husband's Earnings

I assume that the husband always works, so instead of his skill level, I use his age as a determinant of the earning equation. His earnings also depend on the wife's education level. That is, I allow the model to have assortative mating based on education level as observed in data (Chiappori *et al.* (2009)); hence, the spousal

[32] See Huggett *et al.* (2011) for heterogeneity in wage profiles, and Adda *et al.* (2013) for specifying human capital accumulation based on hours of market work.

[33] Chiappori *et al.* (2002) analyze the impacts of the marriage market and divorce legislation, which vary by state and over time, on household labor supply. Here, I abstract this issue.

[34] I assume that marriage probability does not depend on the individual's choices such as labor supply decisions, nor on their years of work experience.

earnings are as follows:

$$log(w_{i,t}^{sp}) = \eta_{0,s}^{sp} + \eta_{1,s}^{sp} age_{i,t}^{sp} + \eta_{2,s}^{sp} age_{i,t}^{sp2} + \omega_{i,t}^{s,sp}, \tag{1.7}$$

where subscript s denotes the wife's education level indicating non-college graduates if it takes H and college graduates if C; and $\omega_{i,t}^{s,sp}$ is the wage shock, which is assumed to be i.i.d. $N(0, \sigma_w^{s,sp2})$.

1.3.7 Conception

Based on medical evidence, the probability of conception decreases with the woman's age. Rosenthal and Khatamee (2002) estimate that at age 20, a woman has 90% chance of conceiving a child, but the probability declines to 70% at age 30, 6% at age 45, and almost zero after age 50. Therefore, I assume that in any given period, if the woman decides to conceive a child, she does so with probability π_t, where t denotes the woman's age. I assume that all women are capable of conceiving, but allow the probability to decline with age (see Rosenthal and Khatamee (2002)).[35] I use the probability function estimated by Rosenthal and Khatamee (2002) to specify π_t; this function is depicted in Figure (A.9) in the Supplementary Appendix. Moreover, I allow for conception error because in reality, some pregnancies are unwanted. This means that if a woman decides not to have a child in the next period, she still faces a probability of getting pregnant. This probability depends on her education level; it is λ_c if she is a college graduate and λ_h if she is a non-college graduate, where λ_c and λ_h are free parameters to be estimated. For this estimation, I use data on the fraction of unwanted live births in the US, conditional on the woman's education level, to recover λ_c and λ_h. Table (A.11) in the Supplementary Appendix provides

[35] I am aware that in the US, about 8% of women 15-29 years old have impaired fecundity. Some of them, however, can conceive a child after treatment (Chandra *et al.* (2005)).

some information about pregnancy intention in the US between 2006 and 2010.

1.3.8 Nonlabor Income

For the nonlabor income process, because I see a large number of observations with no nonlabor income in a given period, following Del Boca *et al.* (2013) I consider nonlabor income a truncated version of a latent variable process in levels. Specifically, I assume the following process to model the latent non-labor income in period t:

$$I^*_{i,t} = \mu^{nonlb} + \omega^{nonlb}_{i,t}, \tag{1.8}$$

where μ^{nonlb} is the mean and $\omega^{nonlb}_{i,t}$ is the disturbance term that is i.i.d. $N(0, \sigma^2_{nonlb})$.[36] The observed (actual) nonlabor income is as follows:[37]

$$I_{i,t} = \max(0, I^*_{i,t}). \tag{1.9}$$

1.3.9 Childcare Costs

I assume that if a mother wants to participate in the labor market in a given period, then the family should pay a fixed childcare cost equal to cc in that period unless the child is school-age. I also assume that cost does not depend on hours of

[36] The estimates of the mean and standard deviation of this process for one-child households are as follows: $\mu^{nonlb} = -14.12$ and $\sigma_{nonlb} = 376.16$.

[37] I assume that if a woman gets pregnant unintentionally (note that in this model women can conceive a child out of wedlock), and she is single (so has no husband's income), then she will receive a one-time transfer equal to "TR" in the period of pregnancy. In this model, if a woman is single at period "t," and the nonlabor income in that period is zero, then she never chooses to intentionally become pregnant in that period. This is because I assume logarithmic preference, and in the optimal solution, consumption cannot be zero. If she wants to get pregnant, she cannot work for that period; then, having no husbands income, no labor income, and no nonlabor income, there is nothing to consume. Naturally, there is a probability of unintentionally getting pregnant even if she is single and has no nonlabor income. For the problem to be well-defined in that situation, I assume a one-time transfer equal to "TR." However, it turns out that changing the value of TR has no impact on the individuals decisions, it only helps the process have well-defined solutions. Hence, I can fix TR at an arbitrary level

market work. Moreover, I assume that hours spent in day care (or family care) centers has no impact on the child's human capital accumulation process.

1.3.10 Budget Constraint

Each individual faces the following budget constraint in each period:

$$c_{i,t}^{HH} + e_{i,t} + ccI_{h_{i,t}>0}I_{i,t^k<8} = h_{i,t}w_{i,t} + \bar{h}w_{i,t}^{sp}m_{i,t} + I_{i,t}, \tag{1.10}$$

where $c_{i,t}^{HH}$ is household consumption. If the woman is single, I simply assume that it is equal to the individual's own consumption, $c_{i,t}$; if she is married, it is double, $2c_{i,t}$. Therefore, by deciding her own consumption, she decides about household consumption as well. In the above equation, $e_{i,t}$ denotes expenditures on a child's skill; cc is the fixed childcare cost if the mother works in the labor market; $I_{h_{i,t}>0}$ is an indicator equal to one if the mother works, and zero otherwise; $I_{i,t^k<8}$ is an indicator equal to zero if the mother has no children or her child is school-age, and 1 otherwise; $h_{i,t}w_{i,t}$ is the woman's labor income, and $\bar{h}w_{i,t}^{sp}$ is the spousal total earnings; $I_{i,t}$ is the nonlabor income; and $m_{i,t}$ is an indicator variable for marital status, equal to one if married and zero otherwise.[38]

Each individual faces a time constraint, as well. This constraint can be written as follows:

$$TT = l_{i,t} + h_{i,t} + \tau_{i,t}, \tag{1.11}$$

where TT is the total available time in each period; $l_{i,t}$ is the leisure time; $h_{i,t}$ is the hours of market work; and $\tau_{i,t}$ is the time that the mother actively spends with her child.

There is one more nontrivial constraint for this problem:

$$h_{i,t}b_{i,t} = 0,$$

[38] I assume that the husband works for $\bar{h}$ hours in each period no matter what.

where $b_{i,t}$ denotes pregnancy decisions. This constraint ensures that the woman cannot work during the period in which she gets pregnant.[39]

Finally, the model has nonnegativity constraints for choice variables $h_{i,t}, c_{i,t}, l_{i,t}, \tau_{i,t}$, and $e_{i,t}$.

1.3.11 *Recursive Formulation*

If a child exists, then a woman optimally chooses her labor supply, leisure, consumption, and child inputs to maximize her lifetime utility depending on marital status, nonlabor income, and the current level of the childs skill. If no child exists, she makes a pregnancy decision (if still fertile), and labor supply decisions as well. This problem can be formulated recursively. For simplicity of representation, suppose that there is no uncertainty about fertility decisions– i.e., $\pi(age^w) = 1$– and there is no conception error– i.e., $\lambda = 0$. With these assumptions, the value function for the individual's problem is as follows:

Let $V^0(t, S_t)$ denote the value of having no child at period t while the vector of state variables is S_t, and let $V^1(t, S_t)$ denote the value of having a child at period t while the vector of state variables is S_t. I measure both $V^0(t, S_t)$ and $V^1(t, S_t)$ at the beginning of period t, after all shocks are realized.[40] Then, I have the following Bellman equations:

$$V^1(t, S_t) = \max_{c_t^{HH}, e_t, \tau_t, l_t, h_t} \{u(c_t^{HH}, l_t, k_t, bb_t, age^c) + \beta E_t[V^1(t+1, S_{t+1})]\} \tag{1.12}$$

$$V^0(t, S_t) = \max_{c_t^{HH}, l_t, b_t, h_t} \{u(c_t^{HH}, l_t, k_t, bb_t, age^c) + \beta[b_t E_t(V^1(t+1, S_{t+1})) + (1-b_t) E_t(V^0(t+1, S_{t+1}))]\}, \tag{1.13}$$

[39]Regression results in Table (A.12) in the Supplementary Appendix suggest that this is not a very bad assumption.

[40]Note that in this model, they can have only one child.

where $\beta \in (0,1)$ is the discount factor, and E_t is the conditional expectation operation, which takes expectation with respect to the period t information set. This expectation is taken with respect to future wage shocks, marital status shocks, and nonlabor income shocks; b_t is a binary variable equal to 1 if the individual decides to conceive a child at period t and zero otherwise. Finally, $S_t = (m_t, s, x_t, age_t^p, k_t, \Omega_t)$ is the vector including state variables: m_t is the marital status; s is the level of education; x_t is the work experience at the beginning of period t; age_t^p is the age of the mother at childbirth[41]; k_t is the skill level of the child[42]; and Ω_t is a vector including all shocks (the wage shock for the wife and the husband, the non-labor income shock, and the pregnancy shock).

Note that the last Bellman equation is valid when the individual is fertile (i.e. $t \leq 40$), and thus able to decide whether or not to conceive a child. If $t > 40$ and no child exists, then since she can no longer get pregnant, the Bellman equation is as follows:

$$V^0(t, S_{i,t}) = \max_{c_{i,t}, l_{i,t}, h_{i,t}} \{u(c_{i,t}^{HH}, l_{i,t}, k_{i,t}, bb_{i,t}, t^k) + \beta E_t[V^0(t+1, S_{i,t+1})]\}. \tag{1.14}$$

[41] $age_t^p \in \{\varnothing, 18, 19, ..., 40\}$, where $\varnothing$ indicates that there is no child, so the age of mother when a child is born, is not defined.

[42] It equals nothing if no child exists.

Having the recursive formulation of the problem, since I have a finite horizon problem, I can solve the problem by starting with the final period and moving backwards. I begin by solving the last-period problem of the individual numerically and then going on recursively to find the value function for the individual for any possible values of the state vector. In what follows next, I briefly discuss the solution to the individual's problem modeled in the previous subsection.

Child Investments

According to equation (1.3), by assuming a Cobb-Douglass form for the skill formation technology, it is obvious that whenever a child exists, I can never have any corner solutions to the household input choice problem during the investment period (i.e. the development stage).[43] However, the model allows for corner solutions for labor supply– i.e. the hours worked may be zero in any given period. I can write the conditional factor demands for child inputs (expenditures on the child and the time spent with the child), where I condition on the labor supply choices, nonlabor income, and the husband's income (if she is married). These conditional factor demands also depend on the child's existence and the child's age when a child exists. The conditional factor demand for child inputs at child's age t^k, is as follows:

$$\tau^*_{i,t^k}(h_{i,t^k}) = (TT - h_{i,t^k})\frac{\varphi_{2,t^k}}{\alpha_2 + \varphi_{2,t^k}}, \tag{1.15}$$

$$e^*_{i,t^k}(h_{i,t^k}) = (h_{i,t^k}w_{i,t^k} + \bar{h}w^{sp}_{i,t^k}m_{i,t^k} + I_{i,t^k} - ccI_{h_{i,t^k}>0}I_{i,t^k<8})\frac{\varphi_{3,t^k}}{\alpha_1 + \varphi_{3,t^k}}, \tag{1.16}$$

[43]If any factor takes zero in any given period, then because of the Cobb-Douglass technology function, the child's skill level will be zero at all the subsequent periods, and by log linear utility function $u_{i,t} \to -\infty$ whenever $k_{i,t} \to 0$.

where

$$\varphi_{j,t^k} = \beta\gamma_{j,t^k}\xi_{i,t^k+1}, j = 2, 3 \tag{1.17}$$

The sequence $\{\xi_{i,t^k}\}_{t^k=1}^{M+1}$ is defined (backwards-) recursively as

$$\xi_{i,M+1} = \psi\alpha_3$$

$$\xi_{i,M} = \alpha_3 + \beta\gamma_{1,M}\xi_{i,M+1}$$

$$\vdots$$

$$\xi_{i,m} = \alpha_3 + \beta\gamma_{1,m}\xi_{i,m+1}$$

$$\vdots$$

$$\xi_{i,1} = \alpha_3 + \beta\gamma_{1,1}\xi_{i,2}$$

where $\psi_i = \phi(\beta + \beta^2 + \cdots + \beta^{60-age_i^p-M})$.

It is worth mentioning that $\xi_{i,m}$ is the marginal utility of (log) child quality to the household in period t, i.e., $\xi_{i,m} = \partial V_m(S_{i,t^k})/\partial ln(k_{i,t^k})$.

Labor Supply

I derived the solutions to child inputs conditional on hours of market work. Now, the optimal solution to the labor supply is conditional on the existence of a child and on the pregnancy decision as well.

Working more than the cut-off level, $\underline{h}$, increases the individual's work experience by one unit, but any changes to hours of market work, as long as the individual works more than the cut-off level has no additional impacts on her work experience and so future wages. Also, any changes to hours of market work, as long as the individual works less than the cut-off level does not change her future work experience and so future wages. However, changing hours of market work from below the cut-off level to above it effectively increases the individual's work experience and so changes future wage possibilities. In other words, the decisions on labor supply at the extensive margin, i.e. whether to work or not, affect future wage possibilities through the return to work experience, but at the intensive margin, the future return to hours of market work is constant no matter how many hours she work (as long as it is at least equal to the threshold). Hence, the individual makes decision on discrete choices here. In other words, there are two corner solutions here: 1) $h = 0$ so she does not work, and 2) $h = \underline{h}$ so she works the least hours needed to get one more unit of experience in period t. There are four cases in regard with the child existence and the fertility status:

First Case: a child exists whose age is less than M: The individual maximizes her life-time utility by choosing hours of market work, $h_{i,t}$:

$$\max_{h_{i,t}} \alpha_1 log(h_{i,t}w_{i,t} + \bar{h}w_t^{sp}m_{i,t} + I_{i,t} - ccI_{h_{i,t}>0}I_{i,t^k<8}) + \alpha_2 log(TT - h_{i,t} - \tau(h_{i,t}))$$
$$+ \alpha_3 \phi log(k_{i,t}) + \beta E_t[V^1(t+1, S_{i,t+1})], \tag{1.18}$$

where V^1 is the value function described in the previous section. Remember that I already derived the optimal child inputs conditional on hours of market work and I compute the value function recursively by starting from the last period and going backwards in time. So for every single value of $h_{i,t}$, I can compute the individual's life-time utility.

Now, to derive the first order condition (FOC) for this maximization problem, one should note that the continuation value, $V^1(t+1, S_{i,t+1})$ is not continuous at the cut-off level, $\underline{h}$, because the work experience (which is a state variable of function V^1) is a step function with respect to $h_{i,t}$. If $h_{i,t} \in [0, \underline{h})$, then $V^1(t+1, S_{i,t+1}) = V^1(t+1, ., ., x_{i,t} - \delta_{s,t}x_{i,t}, ., .)$, which means that the state variable "work experience" will be depreciated by the trophy rate δ; and if $h_{i,t} \in [\underline{h}, TT)$, then $V^1(t+1, S_{i,t+1}) = V^1(t+1, ., ., x_{i,t}+1, ., .)$, which means that the state variable "work experience" will be one unit more tomorrow.

Hence, the continuation value, $V^1(t+1, S_{i,t+1})$ is a step function in $h_{i,t}$ and there is a kink at $\underline{h}$. This means that $V^1(t+1, S_{i,t+1})$ is not differentiable with respect to $h_{i,t}$ at the kink point, which is $\underline{h}$, but it is differentiable with respect to $h_{i,t}$ over both intervals $(0, \underline{h})$ and $(\underline{h}, TT)$, and the derivative is zero over both intervals. Hence, in order to find the optimal solution to $h_{i,t}$, I consider three cases:

1) $h_{i,t} \neq \underline{h}$: in this case, since the $V^1(t+1, S_{i,t+1})$ is differentiable with respect to $h_{i,t}$, I can derive the FOC by taking derivative of the life-time utility, i.e. expression (17), with respect to h, which yields the following (after I solve it for $h_{i,t}$):[44]

$$\hat{h}^1_{i,t} = \frac{w_{i,t}TT(\alpha_1 + \varphi_{3,t}) - (\alpha_2 + \varphi_{2,t})(h_{i,t}w_{i,t} + \bar{h}w^{sp}_t m_t + I_{i,t})}{w_{i,t}(\alpha_1 + \alpha_2 + \varphi_{3,t} + \varphi_{2,t})}, \tag{1.19}$$

where $h^1_{i,t}$ is the interior solution to the maximization problem. So the life-time utility of the individual is equal to:

$$\alpha_1 log(\hat{h}^1_{i,t}w_{i,t} + \bar{h}w^{sp}_t m_{i,t} + I_{i,t} - ccI_{h_{i,t}>0}I_{i,t^k<8}) + \alpha_2 log(TT - \hat{h}^1_{i,t} - \tau(h_{i,t}))$$
$$+ \alpha_3 \phi log(k_{i,t}) + \beta E_t[V^1(t+1, S_{i,t+1})],$$

Where the state variable of work experience in the state vector, $x_{i,t}$, should be updated correspondingly with respect to the value of $\hat{h}^1_{i,t}$.

2) $h_{i,t} = \underline{h}$: in this case, the life-time utility of individual is simply equal to the following:

$$\alpha_1 log(\underline{h}w_{i,t} + \bar{h}w^{sp}_t m_{i,t} + I_{i,t} - ccI_{h_{i,t}>0}I_{i,t^k<8}) + \alpha_2 log(TT - \underline{h} - \tau(h_{i,t})) + \alpha_3 \phi log(k_{i,t})$$
$$+ \beta E_t[V^1(t+1, S_{i,t+1})],$$

[44]The derivation is shown in the Supplementary Appendix.

Where the state variable of work experience in the state vector, $x_{i,t}$, should be updated correspondingly with respect to the value of $h_{i,t} = \underline{h}$, so $x_{t+1} = x_{i,t} + 1$.

3) $h_{i,t} = 0$: in this case, the life-time utility of individual is simply equal to the following:

$$\alpha_1 log(\bar{h} w_t^{sp} m_{i,t} + I_{i,t}) + \alpha_2 log(TT - \tau(h_{i,t})) + \alpha_3 \phi log(k_{i,t}) + \beta E_t[V^1(t+1, S_{i,t+1})],$$

Where the state variable of work experience in the state vector, $x_{i,t}$, should be updated correspondingly with respect to the value of $h_{i,t} = 0$ and the depreciation rate. The optimal solution to the labor supply can be found easily by comparing the life-time utility to the individual under three cases described above.

To put it in a nutshell, the optimal choice for labor supply is as follows:

$$h^*_{i,t} = \arg \max_{h_{i,t} \in \{0, \max\{0, \hat{h}^1_{i,t}\}, \underline{h}\}} \alpha_1 log(h_{i,t} w_{i,t} + \bar{h} w_t^{sp} m_{i,t} + I_{i,t}) + \alpha_2 log(TT - h_{i,t} - \tau(h_{i,t}))$$
$$+ \alpha_3 \phi log(k_{i,t}) + \beta E_t[V^1(t+1, S_{i,t+1})],$$

where $S_{i,t} = (m_{i,t}, s_i, x_{i,t}, age^p_{i,t}, k_{i,t}, \Omega_{i,t})$ is the vector including state variables, and $V^1(t+1, S_{i,t+1})$, as defined earlier, indicates the value function when a child exists at the beginning of period $t+1$ and the state vector is $S_{i,t+1}$. When solving the problem backwards in time, in a given period t, I know the value of $V^1(t+1, S_{i,t+1})$ for all possible values that the state vector might have. Note that when $h_{i,t}$ is zero or less than the cut-off level, the next period's experience will be discounted by $\delta_{s,t-1}$, and when $h_{i,t}$ is greater than the threshold the next period's experience will be one more unit higher than today's, no matter what the value to $h_{i,t}$ is.

Second Case: a child exists, but older than M: In this case, since the development stage is already done, then, hours of market work does not play any role in determining current or future child's skill levels. Hence, hours of market work is determined by comparing the values to corner solutions and the interior solution while the interior solution is determined only by the trade-off between consumption and leisure (no dynamics exist in this case). The procedure is similar to the previous case, but in this case the interior solution, which is derived using the FOC, is as follows:[45]

$$\hat{h}^0_{i,t} = \frac{\alpha_1 w_{i,t} TT - \alpha_2(\bar{h} w^{sp}_{i,t} m_{i,t} + I_{i,t})}{w_{i,t}(\alpha_1 + \alpha_2)}, \tag{1.20}$$

Hence, the optimal hours of market work is as follows:

$$h^*_{i,t} = \arg \max_{h_{i,t} \in \{0, \max\{0, \hat{h}^0_{i,t}\}, \underline{h}\}} \alpha_1 log(h_{i,t} w_{i,t} + \bar{h} w^{sp}_t m_{i,t} + I_{i,t}) + \alpha_2 log(TT - h_{i,t})$$
$$+ \varphi \alpha_3 \phi log(k_{i,M+1}) + \beta E_t[V^1(t+1, S_{i,t+1})],$$

Third Case: no child exists and the fertility period is over: In this case, there is no child, and since the individual is older than 40 years of old, she cannot decide to get pregnant, so no pregnancy decision is made. Similar to the previous case, hours of market work is determined by comparing the values to corner solutions and the interior solution while the interior solution is determined only by the trade-off between consumption and leisure (no dynamics exist in this case). The solution in this case is similar to the previous case:

$$h^*_{i,t} = \arg \max_{h_{i,t} \in \{0, \max\{0, \hat{h}^0_{i,t}\}, \underline{h}\}} \alpha_1 log(h_{i,t} w_{i,t} + \bar{h} w^{sp}_t m_{i,t} + I_{i,t}) + \alpha_2 log(TT - h_{i,t}))$$
$$+ \beta E_t[V^0(t+1, S_{i,t+1})],$$

Fourth Case: no child exists and the individual is still in the fertility window: In this case, the optimal choice for labor supply can be derived conditional on the pregnancy choice. First, suppose that she decides to get pregnant. Then, she cannot work in this period if she actually gets pregnant.[46] Hence there is no choice

[45]The derivation is shown in the Supplementary Appendix.

[46]For the simplicity of representation, I ignore the uncertainty about the pregnancy outcome here.

to make on hours of market work, and it is equal to zero: $h^* = 0$. Now, if she decided not to conceive a child in the current period hours of market work is determined by comparing the values to corner solutions and the interior solution while, similar to the previous case, the interior solution is determined only by the trade-off between consumption and leisure (no dynamics exist in this case). Hence, the optimal choice for labor supply can be derived as follows:

$$h^*_{i,t} = \arg \max_{h_{i,t} \in \{0, \max\{0, \hat{h}^0_{i,t}\}, \underline{h}\}} \alpha_1 log(h_{i,t} w_{i,t} + \bar{h} w_t^{sp} m_{i,t} + I_{i,t}) + \alpha_2 log(TT - h_{i,t}))$$
$$+ \beta E_t[V^0(t+1, S_{i,t+1})],$$

Pregnancy Choices

The solutions to child inputs described above are derived conditional on existence of a child; upon arrival of a child, mothers make optimal decisions regarding the labor supply and child inputs. In a given period, if a fertile individual has no child, then she has to decide whether to conceive a child, or to delay childbearing for one additional period, and also make decision on how much to work if not pregnant. To make optimal childbearing decisions, she compares the lifetime utility of conceiving a child in the current period (conditional on making optimal decisions regarding the child inputs and labor supply thereafter) to the lifetime utility of not conceiving a child (conditional on making optimal decisions in the following periods). Hence, the individual makes the decision weather or not to conceive a child in the current period, t, according to the following maximization:

$$V^0(t, S_{i,t}) = \max\{V^0_{nobb}(t, S_{i,t}), V^0_{bb}(t, S_{i,t})\},$$

where V_{bb}^0, i.e. the value to decide to get pregnant, is as follows:

$$V_{bb}^0(t, S_{i,t}) = \alpha_1 log(\bar{h} w_t^{sp} m_{i,t} + I_{i,t}) + \alpha_2 log(TT)$$
$$+ \beta E_t[V^1(t+1, S_{i,t+1})],$$

where the value of state variables in $S_{i,t+1}$ is updated; specifically, $bb_{i,t+1} = 1$ (i.e. it changes from zero to one), and $age^p_{i,t+1} = t$ (i.e. it changes from nothing to t).[47] Note that in this case, she cannot work so $h_{i,t} = 0$, and the income comes from the husband's work and non-labor income. Moreover, the value to the work experience decreases due to depreciation, $x_{i,t+1} = (1 - \delta_{s,t})x_{i,t}$.

The value to decide not to get pregnant, i.e. $V_{nobb}^0(t, S_{i,t})$ in the above-mentioned maximization is as follows:

$$V_{nobb}^0(t, S_{i,t}) = \max_{h_{i,t} \in \{0, \hat{h}^0_{i,t}, \underline{h}\}} \alpha_1 log(h_t w_t + \bar{h} w_t^{sp} m_{i,t} + I_{i,t}) + \alpha_2 log(TT - h_{i,t})$$
$$+ \beta E_t[V^1(t+1, S_{i,t+1})],$$

Note that the value of all state variables in $S_{i,t+1}$ is the same as $S_{i,t}$ except (probably) for the work experience.

1.4 Estimation

In this section, I first explain the model specification. I discuss the functional forms that I have chosen and the implications of these modeling choices for the final results. Then I review the sources of heterogeneity and uncertainty in the model that help me recover the model parameters. After that, I discuss the parameter identification: I assume that the specification represents the data generating process, and the issue is how to recover the model parameters using observed data. Finally, I describe the estimation method.

[47] It has been equal to nothing while entering this period.

1.4.1 The Sources of Heterogeneity and Uncertainty

This model has two sources of heterogeneity. Individuals differ in their education levels, and they face different probabilities of unwanted pregnancies according to those levels. They also differ in their marital status. I use completed years of schooling to denote education level in the estimation. To be more precise, individuals differ in their educational decisions by age 18.

Except for these differences, all individuals face the same problem. Individuals also face uncertainty regarding wage shocks, productivity factor shocks, and nonlabor income shocks. All of these shocks are independently distributed across individuals and over time. Individuals also face shocks to their marital status in each period; these shocks depend on the woman's age, the presence of children, and her previous marital status. Moreover, individuals face uncertainty about the outcomes of their pregnancy decisions. If a woman decides to get pregnant at age y, she will get pregnant with probability $\pi(y)$, and if she decides not to get pregnant, then she may still get pregnant with probability $\lambda(s)$. Exogenous variations in nonlabor income, wage, marital status, pregnancy, and education level in the data help to recover the model parameters.

1.4.2 Model Specification

In this section, I discuss my modeling choices.

I allow the elasticity parameters in the child's skill production function, i.e. γ_{1,t^k}, γ_{2,t^k}, γ_{3,t^k}, to change with age of the child. However, to keep the number of parameters that I need to estimate as small as possible, I assume the following functional form for modeling how the elasticity parameters in Equation (1.3) change with the child's age, t^k:

$$\gamma_{j,t^k} = \exp(\gamma_{j,0} + \gamma_{j,1} t^k), j = 1, 2, 3 \tag{1.21}$$

I use a similar approach for modeling the depreciation rate in Equation (1.5), $\delta(s,t)$ as follows:

$$\delta_{s,t} = \exp(\delta_{s,0} + \delta_{s,1} t), s = H, C \tag{1.22}$$

where subscript H denotes non-college graduates, and subscript C indicates that the individual is a college graduate.

Furthermore, I allow the intercept of the productivity factor to depend on the age of the child as follows:

$$A_{t^k} = \exp(A_{0,0} + A_{0,1} t^k), j = 1, 2, 3 \tag{1.23}$$

In this paper, I model the effect of the mother's age at childbirth on the child's innate ability by including the parent's age in the productivity factor term. I assume that initial skill is homogeneous because every child is born with almost the same endowment in terms of knowledge, skills, and experience, but with different innate abilities. Lower innate ability driven by the age effect might affect the childs ability to absorb investments (both time and monetary) and to translate them into human capital. For example, suppose that two children have the same stock of human capital at period t. If child 1 has a greater innate ability (say a higher productivity factor) than child 2, and they receive the same amount of inputs, child 1 will have a greater stock of human capital at period t+1 compared to child 2 because the first child can better absorb the investment and transform it into human capital.

To specify the productivity factor term, I use a log-quadratic functional form over childbearing age. Medical evidence shows that the marginal effect of the mother's age at childbirth on the child's innate ability is not constant. The marginal detrimental effect of maternal age is greater as maternal age increases. Moreover, some evidence shows that it has an inverse-U shape. There are also some risks associated with pregnancy before age 20. The specification I chose is able to capture such a pattern.

1.4.3 Identification

In this subsection, I discuss the econometric identification issues. First, I assume that the functional-form assumptions represent the true data generation process of the population exactly. Then, I show how the model parameters could be recovered using the observed data in a relatively simple manner. The estimator I actually use for recovering the primitive model parameters has practical and theoretical advantages over the proposed identification method below. Nonetheless, the following discussion helps one develop intuition to better understand the key sources of identifying information given the model I specified earlier.

I begin with the determination of preference parameters, and the parameters in the skill formation technology. Here, I face some challenges: first, household consumption and child expenditures are not observed in my dataset. The second issue concerns missing data in the dataset. For each child in each survey year, I observe the time spent with the child and the child's test scores, but I do not observe the child's test scores in successive years– i.e. I do not observe "k_{i,t^k}" and "k_{i,t^k+1}" at the same time. I can see the child's skill level only in the next survey, which is five years later. Therefore, I have observations on "k_{i,t^k}" and "k_{i,t^k+5}." This makes it difficult to estimate parameters for the skill formation technology.

Let's first restrict our attention to the skill formation technology parameters in the child's skill formation function specified in Equations (2) and (3). In the skill formation technology specified earlier, the observation period for each household is defined by the age of their child, i.e. period t begins when the child turns age t^k.

Under the maintained assumptions, if the child of household i turns age t^k with a skill level of k_{i,t^k}, then, her skill level at the beginning of the next period, k_{i,t^k+1}, is determined by the following equation:

$$\begin{aligned} lnk_{i,t^k+1} &= lnA_{t^k} + \epsilon_1 age_i^p + \epsilon_2 (age_i^p)^2 + \epsilon_3 IQ_i^p + exp(\gamma_{1,0} + \gamma_{1,1} t^k) lnk_{i,t^k} \\ &+ exp(\gamma_{2,0} + \gamma_{2,1} t^k) ln\tau_{i,t^k} + exp(\gamma_{3,0} + \gamma_{3,1} t^k) lne_{i,t^k} + \upsilon_{i,t^k}, \\ &\equiv X(age_i^p, IQ_i^p, t^k, \tau_{i,t^k}, e_{i,t^k}, k_{i,t^k}; \gamma) + \upsilon_{i,t^k}; i = 1, 2, ..., N, \end{aligned}$$

where N is the number of observations. The disturbances, υ_{i,t^k}'s, are independently distributed across households and over time. The IQ of mothers is proxied by their level of education, and for each child, I can observe the education level of the mother and the mother's age at birth, age_i^p. Given the assumptions I made, the skill formation technology parameters, γ's, could be estimated using a non-linear least squares (NLS) estimator as follows:

$$\hat{\gamma}_{NLS} = \underset{\gamma}{\text{argmin}} \sum_{i}^{N} (lnk_{i,t^k+1} - X(age_i^p, IQ_i^p, t^k, \tau_{i,t^k}, e_{i,t^k}, k_{i,t^k}; \gamma))^2$$

In order to recover the technology parameters in this manner, I need the standard full rank condition on the matrix X. The full rank condition requires that not all mothers have the same education level, not all mothers have the same age at first birth, and not all mothers choose the same level of investments in their child. These conditions are easily satisfied in the actual dataset. Moreover, since I specified the parameters as a monotone function of the child's age, I need at least two children of different ages in my dataset in order for the full rank condition to be satisfied.

A challenge that one faces when estimating skill formation technology in the manner just discussed, is that the child goods expenditures, e_{i,t^k}'s, are not directly observable in the dataset. However, under the maintained assumptions, child goods expenditures is the only decision variable that depends on the family's income. As previously shown in Equation (1.15), the child goods expenditures is a fraction of the family income, and that fraction is a function of preference parameters, skill formation technology parameters, and the mother's age at childbirth. As family incomes are observable for each observation period, Equation (1.15) can be used to back out child goods expenditures given preference parameters. Hence, the skill formation technology parameters could be estimated using a NLS estimator given the preference parameters.

Another challenge faced in estimating the technology parameters using a NLS estimator is missing data in the children's test scores. The concern here is that child investments and skill levels are observable, but child supplement interviews have only been conducted every five years, so for a given child, skill levels cannot be observed for two successive years. Hence, for family i, both k_{i,t^k} and k_{i,t^k+1} cannot be observed at the same time. The estimation method I actually use is explained in the next section in detail. Using the derived solutions to child investments under the specified model, the child skill levels in each period can be simulated for every realization of exogenous shocks to marital status, wages, and non-labor incomes for a given set of parameters. Hence, for an arbitrary set of model parameters, it is possible to solve each individual's problem, and simulate the path of their child's skill levels. With the simulated data, I can compute sample characteristics for the simulated dataset and compare them to the actual sample characteristics.[48] The next step is to change the set of assumed model parameters and repeat the process, i.e. simulate a dataset based on the assumed parameters, compute the simulated sample characteristics, and compare it to the actual sample characteristics. This process is continued until the

[48]Also, before computing the simulated sample characteristic, I generate the simulated dataset S times, while each time drawing a random number for the random shock variables according to their data generating process. In this way, when computing the sample characteristics, I can effectively take the effect of random shocks out by integrating over the different random shock realizations.

model parameters are tuned such that the simulated sample characteristics match the actual sample characteristics.

In this section, I discussed how I could use a NLS approach to estimate the technology parameters when given the preference parameters while ignoring missing data issue. Now, I discuss how to estimate the preference parameters using the dataset when given the technology parameters. There are five preference parameters: the impact factors for consumption, leisure, and child's skill level, i.e. α's; the discount factor for when the child reaches age 18 and leaves home, i.e. ϕ; and the time discount factor, β. After normalizing the impact factors by assuming that $\sum_{1=1}^{3} \alpha_i = 1$, only four parameters effectively remain to be estimated. I already derived optimal solutions for the labor supply and child inputs conditional on labor supply choices. The labor supply choices and time investments in children are observable in my dataset, and I can use this information to back out the preference parameters. I need two individuals whose ages at childbirth are different and who both chose non-corner solutions for labor supply in at least one period. This condition is satisfied in my dataset. Hence, using Equations (14) and (16), which specify the optimal solutions for child time investment and labor choices, and the actual choices observed in the dataset for these two hypothetical individuals, I can recover the four preference parameters given the skill formation technology parameters.

As shown here, even though the consumption and child goods expenditures are not observed in my dataset, the dataset is rich enough to still enable me to recover the primitive parameters of the specified model.

The process for nonlabor income is assumed to be exogenous in this model. Separately, I can estimate this process, described in equations (7) and (8), using cross-section data on the households' nonlabor income. Hence, parameters in the nonlabor income process can be identified outside of the model.

What remains is the determination of parameters in the wage equations. Although the wage offer process can be estimated outside of the model, the problem with directly estimating the wage process is that there are nonrandomly missing observations due to corner solutions for labor supply. Indeed, when the individual is not working, I cannot observe the wage offer. These missing wage offers are not randomly distributed because the higher the wage offer, the higher is the probability that the woman works in that period. In other words, the larger the error term of the wage process, the higher the probability that the wage offer is observable. Wage offers are observed only for women who work for wages, so a naive estimator would be biased. Hence, I cannot estimate the wage process using OLS.[49,50] Therefore, I estimate the parameters of the wage processes simultaneously with other parameters of the model using the estimation method that is described in the next section. Using the parametric distribution assumptions about the wage offers, explained in the model section, I can identify the parameters describing the wage offer distribution alongside other parameters of the model.

In the next section, I explain the estimation strategy. But first, I will explain intuitively why I am able to recover the preference parameters and the technology parameters even without observing expenditures and consumption. According to the model, expenditures are always a fraction of total income (Equation (1.15)), which is observable, and this fraction is a function of the model parameters. Hence, because I observe total income and the child's skill level, exogenous variations in the wage shock and nonlabor income help me recover the impact of expenditures on the child's skill level– i.e. γ_3, provided that the model is correctly specified. Consumption, again, is

[49]I also used a Monte Carlo simulation to see if I could obtain consistent coefficients for the wage process using OLS regression when ignoring those nonrandomly missing data. My simulation results suggest that I cannot.

[50]This problem is known as incidental truncation in the econometrics literature.

a fraction of total income, and this fraction is a function of the model parameters. Therefore, the model reveals the relationship between consumption and income, and this helps me recover the impact of consumption on the individual's utility. It is worth thinking about the two extreme cases. If an individual cares only about her consumption– i.e. puts zero weight on leisure and the child's skill level– then she never decides to conceive a child. On the other hand, if she cares only about the child's skill level, then she spends all of her resources on her child, consumes nothing, and never uses her free time for enjoyment (leisure). Observations on the time that individuals spend with their children and the evolution of the child's skill level help to recover the consumption impact parameter, α_1. The next section provides a more detailed explanation illustrating how the model parameters could be recovered using the observed data, even though I do not observe consumption and child expenditures, and even without having observations on childs skill levels for successive years.

1.4.4 Estimator

I consider two education levels, four-year college graduate and non-college graduate. A fraction of individuals decide to continue their education at college. I set the fraction of non-college graduates and college graduates according to data from Panel Study of Income Dynamics (PSID). The PSID data also show that, for each group of individuals, some percentage are single at age 18, and the others are married. These fractions define the initial distribution of heterogeneous individuals in the model.[51]

In total, the model has 40 parameters to be estimated simultaneously: eleven parameters of the technology function–i.e. ϵ_1, ϵ_2, ϵ_3, two for A, two for γ_1, two for γ_2,

[51]According to the data, 46% of women continue their education in college. Provided that they have decided not to continue their education in college, the probability of her being single is 0.85, and of being married is 0.15. In contrast, for a woman that has decided to go to college, the probability of being single is 0.95, and of being married is 0.05.

and two for γ_3; four preferences parameters– i.e. α_1, α_2, ϕ, and β; sixteen coefficients of the wage functions– i.e. four for δ, two for η_0, two for η_1, two for η_2, two for η_0^{sp}, two for η_1^{sp}, and two for η_2^{sp}; two probability parameters for conception errors, λ_c and λ_h; and five distributional parameters for wage shocks and productivity shocks– i.e. two for σ_w, two for σ_w^{sp}, and σ_R^2; and two parameters for the distribution of initial child skill. I use the Method of Simulated Moments (MSM) estimator to recover the primitive parameters of the model simultaneously, which is described below.[52]

The data I use provide detailed information about the households in the survey. They include some characteristics of parents, such as the ages at childbirth, the mother's education, hours worked, their accepted wages, household income, and hours spent with their child. They also include some measures of the child's skill level at different ages during the development period.

In order to implement the method of simulated moments, I first define a set of moments (sample characteristics) that capture the relationships between different variables in the sample for each year in which the survey was conducted and also between surveys in different years. Let this set be denoted by M. To find the estimator in this method, I utilize the simulation method. This means that, using my model, conditional on the parameter vector Ω, I can solve the individuals' problem throughout their life-cycle and compute the moments I already have defined for this population. I denote this moment vector by $\tilde{M}(\Omega)$. The MSM estimator of the parameter vector Ω can be defined as follows:

$$\hat{\Omega} = \arg\min_{\Omega} (M - \tilde{M}(\Omega))' W (M - \tilde{M}(\Omega)), \tag{1.24}$$

where W is a symmetric positive-definite weighting matrix. I use the bootstrapping method to define this weighting matrix. Indeed, by resampling the data, I simply

[52]See Lerman and Manski (1981), McFadden (1989), Pakes and Pollard (1989), Gourieroux *et al.* (1993), and Gallant and E. Tauchen (1996) for some references.

set W equal to the inverse of the covariance matrix of M.[53] It turns out that this estimator is consistent.[54]

The moment vector that I use includes the average and standard deviation of mothers' ages at the birth of the first child (if they have one), for both college and non-college graduates; the average and standard deviation of test scores at each child's age; the average and standard deviation of hours of work for mothers; the average and standard deviation of child investment hours at each child's age; the average and standard deviation of accepted wages and the correlation in wages across parents; and contemporaneous and lagged correlations among the observed labor supply, time with children, child's skill level, wages, and income. In order to compute the standard errors of the parameter estimates, I use the bootstrapping method. I generate new datasets using sampling data over individuals with replacement. Then, I implement the method of simulated moments to reestimate the model parameters using the new data sets.[55] Finally, I calculate the standard errors of the parameter estimates.

1.5 Model Estimates

1.5.1 Parameter Estimates

In this section, I present the results of the estimation of the theoretical model– namely, the estimated parameters and sample fit of the model. Then, I report the results of the variance decomposition of the child's skill level at the end of the devel-

[53] I did not use the asymptotically optimal weighting matrix because of computational cost. Also, Altonji and Segal (1994) examined the finite-sample performance of the asymptotically optimal weighting matrix and concluded that the asymptotically optimal weighting estimator is seriously biased in small samples.

[54] See Cameron and Trivedi (2005) for a discussion of the properties of this estimator.

[55] While reestimating parameters using the new bootstrapped datasets, I keep fixed the random draws for wage shocks, marital status shocks, nonlabor income shocks, and productivity shocks. I also hold the starting values constant when reestimating model parameters for the bootstrapped datasets.

opment phase. Finally, I discuss some counterfactual policies regarding the timing of birth.

Preference Parameters

Preference parameters and the corresponding standard errors are presented in Table (1.7). The preference parameters α_1, α_2, and α_3 correspond to consumption, leisure, and the child's skill level, respectively. These parameters determine the relative importance of those factors in the utility function. A 1% increase in consumption has the highest impact on the utility function ($log(0.42)$), compared to a 1% increase in household leisure or in the child's skill level. The leisure impact is equal to 0.27, and the child's skill impact is equal to 0.31. To get a better understanding of what these numbers mean, it might be helpful to think about the following two extreme cases: first suppose that $\alpha_3 = 0$. In this case, the woman does not care about her child's skill level at all. In the context of this model, this means that she would get no enjoyment from having a child. Therefore, she never decides to conceive a child; no child will be conceived intentionally. On the other hand, suppose that $\alpha_3 = 1$, which implies that $\alpha_1 = \alpha_2 = 0$. In this case, the woman cares only about the child's skill level and the utility is derived only through the child's skill level. In this case all resources are devoted to the child and nothing would be consumed by the household, i.e, no leisure and no consumption occurs at all.

Production Technology Parameters

Production technology parameters are presented in Table (1.8). It seems that of all inputs, time investment has the strongest effect on skill formation, especially during the early years of the child's life. The elasticity of the next period's skill level with respect to time investment is much greater than that of money investment and the current period's level of skill. Figure (1.4) depicts the technology parameters against child age. The results suggest that while the impact of time investments decreases with increasing child age, the impacts of child goods expenditures and the current level of human capital become larger.

Table (1.8) also shows the estimate of the effect of the mother's age at childbirth on the child's productivity. Both $\hat{\epsilon_1}$ and $\hat{\epsilon_2}$ are negative. However, $\hat{\epsilon_2}$ is close to zero, meaning that the second-order effect is relatively small. The estimate of the first-order effect, ϵ_1, is equal to -0.015. If I assume that ϵ_2 is zero, then, according to equations (2) and (3), I can interpret the results as follows: all else held constant, one year increase in maternal age at childbirth decreases the next period's skill level by about 1.5% with respect to the child-age-specific average of test scores. Given that the development phase lasts for 15 years, the detrimental effect of postponement of childbirth on child's skill level seems to be considerable. Figure (A.10) shows how the productivity of the child changes with the mother's age at childbirth.

Table (1.8) also reports the standard error of these estimated parameters.

Depreciation Rate Parameters

Table (1.9) provides the estimated parameters regarding depreciation rates. Figure (1.5) depicts how the depreciation rates of one period of absence from the labor market change over the woman's life-cycle by education level. The results suggest that the deprecation rates decrease with age, and that the rate for educated women is higher than that of less-educated women. Figure (1.5) shows that for college graduates, the depreciation rate ranges from over 12% to about 3%, while it ranges from 6% to about 4% for non-college graduates. Previous papers provide a wide range of estimates for depreciation rates. Blundell *et al.* (2016b) document that the human capital for UK women depreciates between 5.7% and 11.0% a year when being out of work. Olivetti

(2006) finds that a woman with zero labor supply between ages 20 and 29 loses about 50% of her human capital by age 30. Guner *et al.* (2018) find yearly depreciation rates of 2.5% for unskilled women and 5.6% for skilled women. Using data from the German chemical sector, Gerst and Grund (2017) estimate the effect of career interruption on wages as about 11% per year of interruption. Moreover, Jacobsen and Levin (1995) find that there is a 14% reduction in women's wages in the case of intermittent labor force attachment. Caucutt *et al.* (2002) finds that if a woman does not supply any labor when she is young, she experiences about a 10% decline in her wages ten years later. Kleven *et al.* (2018) show from the Danish administrative data that the hourly wages of women drop permanently by about 10% right after the birth of their first child. Finally, to estimate the human capital depreciation rate, Ejrnæs and Kunze (2013) exploit arguably exogenous variation in the time out of the labor force induced by the expansion in parental leave in West Germany; they find that mothers wages after returning to work decrease by 3.4% and 5.8% per year of leave for low-skilled and medium-skilled mothers, respectively. None of these studies, however, examined how the wage penalty associated with career interruption due to a childbirth changes over the female's life-cycle. That said, Miller (2011) uses biological fertility shocks to instrument for age at first birth, and finds that motherhood delay leads to a substantial increase in wages of 3% per year of delay; this advantage is largest for college-educated women and those in professional and managerial occupations.

Finally, Table (1.10) provides estimated parameters regarding the earnings functions and unintended pregnancy.

Within-sample Fit

Table (1.11) reports the values of different moments calculated using the data and compares them to corresponding values from the simulated data using estimated parameters. The first two columns correspond to highly-educated women and the last two correspond to less-educated women.

The first moment in Table (1.11) is the mean age at childbirth. It can be seen that the model is a good fit for mean age at first childbirth in both highly-educated and less-educated women. The second row reports the standard deviation of the mother's age at first childbirth. The standard deviation of the simulated data is lower than that observed in the real data. The model is also a good fit on the test scores at the last period of development. Hours worked per week match well between the data and the model.

Figures (1.6) and (1.7) show the sample fit of the average child's test score over the whole child development phase for children of less-educated and educated women, respectively.

Figures (1.8) and (1.9) show the sample fit of the distribution of age at first childbirth for less-educated and educated women, respectively. The simulated distribution of maternal age at birth matches the data distribution well for both less-educated women and educated women.

Discussion

In this section, I relate my results to previous papers in the literature. Cunha and Heckman (2008) and Cunha *et al.* (2010) provide some empirical evidence that tends to support the idea that child goods expenditures during early childhood are very effective, and to a large extent determine future child quality.

Del Boca *et al.* (2013) examines the impacts of child goods expenditures and time investments in children by estimating the child's skill formation technology within a standard household model. They model the household decisions on child investments and labor supplies, leaving the age of parents exogenous. Their findings suggest that even though early childhood expenditures might be more important than those made later in the child's life; however, the effects of child goods expenditures are very limited at any stage of the child's life; time investments made by parents have more effective impact on developing skills in the child. The findings of this paper tend to support the findings of Del Boca *et al.* (2013), in which, time investments play a bigger role compared to the child goods expenditures. However, my findings suggest that the effect of child goods expenditures could be underestimated in models that do not account for decision of household with respect to the timing of childbearing. Households decide on both the timing of birth and child investments, and these decisions are related to each other. As explained earlier in this paper, educated women tend to have their first child later in life, and holding education constant, older parents, on average, make more investments in their children. Moreover, children born to older individuals may be more likely to be less productive. Neglecting these age effects may lead one to underestimate the effects of time investments and money investments. In fact, the returns from investments made by educated parents, who are, on average wealthier and older, might not be as large as for those made a few years earlier, when they were younger. Hence, if we ignore the role of parental age in the childs skill formation technology, we might underestimate the effects of child investments.

In this paper, as in many previous papers (including those mentioned in this section), I did not allow for saving/opportunity in the model developed for my empirical work. One reason is the computational difficulty related to the huge space of state variables in this model. I expect, however, that extending the model to incorporate saving opportunity would not significantly affect childbearing decisions. In the framework presented in this paper, saving opportunity does not relieve the main trade-off that women face between the childs productivity factor and the resources available to invest in the child. They cannot avoid the postponement of childbearing that accompanies accumulating financial resources during early work-life-cycles. Moreover, their wages are relatively low in early periods, which makes it hard for them to save in order to make future investments in their child. It is more likely that individuals would rather have an opportunity to borrow in the early periods to smooth their consumption. If they could transfer their future resources to early periods, then they would have more resources to invest in the child during early periods to increase her skill level; hence, they would not be required to delay their childbearing to later ages at which they would have higher wages.

Nonetheless, it is useful to think about how not including saving could make my estimates biased. First of all, if saving was allowed, people could invest part of their savings in their children, and people with higher incomes could save more. When estimating the impact of child goods expenditures on the child's skill level, I use the variation of incomes between families. Since I did not allow saving to play a role in my model, the impact of child goods expenditures might be overestimated; the variation in child goods expenditures may have actually been higher than what is implied by the differences in incomes between families (because saving is used to increase child goods expenditures). Also, it is possible that neglecting saving opportunities makes me underestimate the effect of maternal age at delivery on the child's skill level. I use variation in the mother's age at birth to assess the effect of the mother's age on the child's skill level while taking into account the effects of income differential due to the wage life-cycle. However, if I overlook the effect of saving on child goods expenditures, I might have not completely accounted for how higher incomes (and savings) made by older parents can compensate for negative effects. This may introduce a downward bias to my estimate of the magnitude of the negative impacts of mother's age on child skill level. I could enrich the model to include borrowing opportunities, as well.

However, I know that at early periods, individuals usually do not have high credit scores, and their access to the credit market is limited. The model presented here can be considered an extreme case in which the borrowing upper bound is zero. Therefore, the results are unlikely to change significantly if I extend the model to include saving and borrowing opportunities, provided there are tight credit constraints in the early stages of the work-life-cycle.

As discussed earlier, previous papers usually have used a reduced-from regression to estimate the impacts of maternal age on child cognitive ability or child's test scores by controlling for socioeconomic characteristics of parents such as parental education, single parenthood, and a measure of family income (see for example Goisis *et al.* (2017), Barclay and Myrskylä (2016), and Leigh and Gong (2010)). The results of my simulation exercise reported in Table (A.13) in the Supplementary Appendix suggest that neglecting child investments significantly biases one's estimate of maternal age on child outcomes even after controlling for those parental socioeconomic characteristics, and the bias leads one to underestimate the magnitude of the negative health impacts of greater maternal age on child outcomes, ceteris paribus. My results suggest that neglecting child investments biases my estimate of the maternal age effect on child's skill level because older parents tend to invest more in their children due to greater parental resources, and there is a dynamic complementarity between child investments and child's skill in producing the next period's skills. Hence, the estimate of the maternal age effect in a log-linear specification is biased if one neglects the child investments, and this statement is true even after controlling for the mother's and the father's wages, hours of market work, and their marital status.[56]

[56]In section 2, I provided a reduced-form estimate of the maternal age effect on child's test scores using observational data (see Table (1.5)). As explained before, those estimates might be biased due to omitted variables, i.e. child investments, which are positively correlated with maternal ages. To illustrate the omitted variable bias problem, I use the simulated data in order to estimate the impact of maternal

1.6 Simulation Experiments

1.6.1 Decomposition Analysis

As explained earlier, previous studies emphasized the impacts of both child goods expenditures and time investments on the child's skill level. Other studies highlighted the effects of maternal age on child outcomes. Hence, females face a trade-off once deciding when to bear a child: the younger the parents, the greater the expected productivity of the child, but the lower their resources to invest in the child. One contribution of this paper is to separate out the two effects by modeling all endogenous choices on the timing of birth, child goods expenditures, and time investments in the child. In order to decompose the effects of maternal age at childbirth on the child's skill from the impacts of child investments, I design some simulation experiments. I use the estimated parameters presented in the previous section throughout all experiments. The decomposition exercise is as follows:

First: I hold investments in the child fixed at their level in the baseline model, and change the maternal age of educated mothers to 22.[57] Then, assuming that the maternal age is 22, I simulate the model and evaluate child skill levels at different ages while using the investments from the baseline model to update the child's skill

age on child's test scores through a reduced-form approach similar to what I used in Table (1.5). Table (A.13) reports regression results of such an exercise. Column 1 of Table (A.13) shows the results when I control for the mother's education and marital status, and also the child's age. The coefficient of the maternal age is equal to -0.014. Column 2 reports the results when I also control for the family income when the child was 2 years of old; the coefficient of interest changes to -0.015. Column 3 shows the results when I also control for the mother's hourly wages, mother's hours of market work, and father's hourly wages. This specification is similar to the regression results reported earlier in Table (1.5). The coefficient of interest in this case is equal to -0.017. Finally, Column 4 of of Table (A.13) shows the results when child investments in the past periods are also controlled for; the coefficient of interest changes to -0.019.

[57]This is the median age of first childbirth for less-educated women.

level in each period. In other words, I assume that investments in the child are not affected by the change that I made with respect to maternal ages, i.e. the reduction of maternal age from the baseline to 22.

Second: In the second experiment, I require the maternal age of educated females to be 22 while allowing child goods expenditures to decrease according to that hypothetically enforced parental age. Then, I evaluate the child's skill level at different ages of the child. In other words, I assume that everyone conceives a child at age 22, and child investments are reduced accordingly with respect to the mother's wages. Assuming that the maternal age is 22, the implied mother's wages during the early stages of childhood development are lower due to both lower work experience and higher depreciation rate at age 22.

Third: In the last experiment, in addition to the decrease in money investment, I also reduce time investments in children according to the results of Table (1.6); with a five-year decrease in the maternal age, time investments are reduced by 3%.

The difference in child skill between exercise 1 and the benchmark is due to the five-year difference in maternal age; this demonstrates the intrinsic effect of maternal age on child skill. The difference in child skill between exercise 1 and step 3 illustrates the impact of differential investments originating from different birth timing.

The results of the decomposition exercise are shown in Table (1.12). They suggest that both choices of the parent, i.e. childbearing age and child inputs, are extremely important for developing the child's human capital. Comparing the benchmark to exercise 1 suggests that reducing the maternal age of educated women without changing actual child investments would increase the final stage human capital of the child by over 11 percent, which is about 0.5 standard deviations. Also, comparing exercise 3 to the benchmark reveals that reducing child investments to the hypothetical level that educated women would have made if they conceived a child five years earlier would decrease the final test score of the children by about 1.6% on average, which is about 0.12 standard deviations. From exercises 2 and 3, it can be seen that even though the five-year decrease in maternal age gives children the advantage of a better health status compared to the benchmark, their final skill level is still lower. This is because lower child investments are made by the age-shifted mothers. Hence, it seems that the investment effects are dominant over health effects for a marginal mother.

In order to get a better sense for the magnitude of the mothers age effect on the childs skill level, it might be useful to compare the result to the value added by school teachers. Published estimates of the average standard deviation for gains in student achievement within a single grade attributable to higher value-added teachers at a given school range from 0.13 to 0.17 (Hanushek and Rivkin (2010)). Hence, the effect of a five-year decrease in maternal age on the childs skill level is comparable to the impact of having a higher value-added teacher for between three and four years.

1.6.2 Policy Analysis

In this section, I analyze the impacts of a wide variety of maternity leave polices, child care subsidy programs, and transfer policies on human capital of mothers and their children.

Maternity Leave Policies

Implementing a nationwide maternity leave policy has been the subject of heated debate among both policymakers and researchers in the US, in part because the United States is an outlier in maternity leave provision (Rossin-Slater (2017a)). The Family and Medical Leave Act (FMLA) entitles eligible workers to take only 12 weeks of unpaid parental leave, a short period of time compared to most other developed countries; for instance, Germany and France have three-year maternity leave periods. It is worth mentioning that less than 60% of private sector workers in the US were eligible for the FMLA in 2012 (Klerman *et al.* (2012)). Previous studies have investigated the impacts of maternity leave policies on fertility rate, the mothers labor market outcomes, and the childs health (see Lalive and Zweimüller (2009); Schönberg and Ludsteck (2014); Rossin (2011)). However, the potential impacts of maternity leave policies on the timing of birth and child development, have not been investigated.[58]

I implement an unpaid maternity leave policy under which employers must guarantee a woman on maternity leave a return to her old job after two years. It also freezes her wages at the level before childbirth. In the context of the model, this means that the depreciation rate is zero. In this case, the only impact of childbearing on wages would be the forgone wages associated with the year of lost experience right after childbirth. I analyze the impacts of such a policy on the timing of births, child investments, and child skill.

I also consider some counterfactual policies in regard to paid maternity leave. I define two types of paid maternity leave policies: 1) mothers are paid for two years after birth while not working, but depreciation to a decrease in their future wages when going back to work, and 2) mothers are paid for two years after birth, and the wage for each individual is the same as its value before childbirth, which means that the depreciation rate for being out of the labor market during pregnancy is zero.

Table (1.13) shows the results of the above-mentioned maternity leave policies. Columns (1) and (4) of Table (1.13) shows the result of an unpaid maternity leave policy under which the wage is frozen at its level before childbirth. Results suggest that implementing such a policy could decrease the childbearing age of educated

[58]Lalive and Zweimüller (2009) is an exception. They evaluate the effects of expansion of paid leave in Austria in 1990 on the likelihood of a closely spaced second child. They show that increasing paternal leave from one to two years is associated with a 15% increase in the likelihood of having another child within three years.

and less-educated females, on average, by over 6% (less than 2 years) and about 2% (less than half a year), respectively, or equivalently, by about 0.3 and 0.07 standard deviations of maternal age at first birth in the benchmark. It also reduces the variation of childbearing ages among both groups. As a result, the final test scores of children increase by about 5% (0.25 standard deviations) and 3% (0.15 standard deviations) for college graduates and non-college, respectively. Hence, such a maternity leave policy, which avoids the wage penalty by freezing wages, can increase the children's human capital by a sizable amount. This increase stems from both having more financial resources available to invest in the child and reducing the negative productivity effects of delayed childbearing.

Columns (2) and (5) show the results for the first paid maternity leave policy, under which individuals are paid while on leave but the wage is allowed to decrease according to the depreciation rate. The results suggest the effect of this policy on the average maternal age is similar to the result of the unpaid maternity leave policy. The reduction in maternal ages improves the test scores of children. However, the improvement is not as big as the previous policy. The reason is that under this paid maternity leave policy, there is a depreciation of human capital associated with being out of the labor market during pregnancy, and it negatively affects future wages of mothers and so child goods expenditures. Under the second paid leave policy, individuals are incentivized to bear in early periods due to elimination of depreciation associated with being out of the labor market for pregnancy. Columns (3) and (6) of Table (1.13) show that implementing such a policy, on average, would lower the average maternal age by about 8% (over two years) and 3% (over half a year) for educated and less-educated females, respectively, or equivalently, by about 0.4 and 0.12 standard deviations of maternal age at first birth in the benchmark. Consequently, it would increase the final test scores of children by about 6% (0.3

standard deviations) and 4% (0.2 standard deviations) for college graduates and non-college, respectively. This is attributable to both the lower maternal age at birth and the higher future wages conditional on birth compared to the pre-policy era.

It is worth noting that while impacts of the discussed paid and unpaid maternity leave polices are qualitatively similar in terms of the maternal age at first birth and the average children's test scores, they have completely different implications for the human capital path of mothers. I focus on the first two previously explained maternity leave policies to elaborate the issue. When the depreciation rate is zero under the unpaid maternity leave policy, i.e. the first policy, the mean labor force participation, work experience, hours of market work, and hourly wages all increase. It is interesting that the mean time investments goes down in this case, which is compensated with greater child goods expenditures, however. Of course, both mothers and children would be better off under this policy. In contrast, under the second policy, the mean labor force participation, work experience, hours of market work, and hourly wages all decrease. The reason is that they decide to have a child earlier to take advantage of a child with higher ability and the transfer at childbirth. But, they face a depreciation rate associated with the years out of the market when pregnant, which negatively affects their wages. The optimal decision following the birth under this policy would be to increase the leisure and time investments in the child, which compensates the lower child goods expenditures. Again, both mothers and children would be better off.

With regard to the impacts of maternity leave policies on mothers' labor market outcomes, the predictions of my simulation analysis are in line with the findings of a handful of papers that have investigated the impacts of six weeks of paid leave in California on labor market outcomes. Using a difference-in-differences design, Baum (2003) show the FMLA has increased the probability that eligible mothers return to work at their pre-childbirth jobs by about %30.[59] Their finding is in line with my finding of the impacts of unpaid maternity leave policy on mothers' labor force participation six years after childbirth; columns 1 and 4 of Table (1.13) show that implementing two-year unpaid maternity leave policy would increase the labor force participation of educated and non-educated mothers of six-year-old children by about 20% and 33%, respectively.

A few studies has investigated the impacts of California's paid leave policy, which provides six weeks of paid leave. Rossin-Slater *et al.* (2013) provide evidence that California's paid family leave increased the work hours of employed mothers of one to three-year-old children by 10% to 17% and that their labor incomes have risen by a similar amount. Baum and Ruhm (2016) provides evidence that California's paid leave policy increased the likelihood that mothers return to work by a year after birth and raised maternal hours and weeks of work by 11% to 19% during the second year of the childs life. These findings are in line with my finding of the impacts of unpaid maternity leave policy on mothers' labor market outcomes; columns 3 and 6 of Table (1.13) show that implementing a two-year paid maternity leave policy that freezes wages to the level before the leave would increase the labor force participation of educated and non-educated mothers of six-year-old children by about 20% and 34%, respectively.

Schönberg and Ludsteck (2014) study the impact of five expansions in leave coverage in Germany on German mothers' labor supply. They find that with regard to the long-run effects of the expansions in leave coverage, four out of the five expansions in leave coverage had almost no impact on mothers employment rates and labor income three to six years after arrival of children. In all of these four reforms the job protec-

[59] See Rossin-Slater (2017b) for a review on studies on maternity and family leave policies.

tion period is as long as or exceeds the maternity benefit period. By contrast, for the other reform, which extended the maternity benefit period beyond the job protection period the result is different; they find this reform discouraged up to 4% of mothers from returning to work when their child was six years old, and lowered their labor income by roughly 8%. This finding is consistent with my finding of the impact of paid maternity leave on mothers' labor market outcomes (column 2 of Table (1.13)).

Childcare Subsidy Program

I also implement a counterfactual policy under which families receive subsidies for their childcare costs. Previous studies have investigated the impacts of childcare policies on fertility rate, mother's labor market outcomes, and child health (see Bauernschuster *et al.* (2016); Baker *et al.* (2008); Havnes and Mogstad (2011)). This is the first study, however, to investigate the potential impacts of childcare policies on the timing of birth, and through it on child skill level.

Table (1.14) shows the results of the above-mentioned childcare subsidy policy. The first to third columns provide the results of implementing the childcare policy when 50%, 75%, and 100% of childcare costs are paid by the government. It turns out that subsidizing childcare costs does not significantly change the maternal age. Even though the effect of the mentioned childcare subsidy policy on maternal age at first birth is not substantial, the policy would increase the final test scores of children by around one-twentieth and one-tenth of a standard deviation for college graduates and non-college graduates, respectively; this is because of higher financial resources for child investment stemming from lower childcare costs and higher participation rates. As Table (1.14) shows, providing a subsidy on child care costs would decrease the reservation wage of mothers, so the labor force participation goes up; the observed accepted wages go down. Consequently, women's work experiences increase (under

75% and 100% subsidies), which has positive impacts on the child goods expenditures.[60] My analysis suggests that under a free childcare program, the labor force participation of educated and non-educated women would increase by about 45% and 40%, respectively. This finding is remarkably close to the finding of Barros *et al.* (2011), who evaluate the causal impact of a childcare program on the maternal labor market outcomes taking advantage of a lottery carried by the municipal government in 2007 in Rio de Janeiro, Brazil. Barros *et al.* (2011) find that access to free publicly provided child care services led to a very large increase in the use of care (from 51% to 94%), a considerable increase in mothers' employment (from 36 to 46 percent), and an almost doubling in the employment of mothers who were not working before the lottery took place (from 9% to 17%).

Transfer Policy to Households

Finally, I investigate the impacts of two types of monetary transfer to the household: (1) a $250 transfer per week in the form of nonlabor income to each family with a child in the development stage, and (2) a transfer targeted to children by providing 250 dollars' worth of child goods to the household each week during the development stage. Table (1.15) reports the results. Mean maternal age at first birth does not change significantly under any policy. However, due to the increase in financial resources provided by the transfer to families, the average test score increases by about 2% and 4% for children of college graduates and non-college graduates, respectively. The impact is greater for non-college graduates because they have fewer financial resources. This transfer policy reduces the inequality between high income families (college graduates) and low-income families (non-college graduates). The impact of the second

[60]It is worth mentioning that I set the childcare cost to its lower bond, and that might explain the mild effect obtained from the childcare subsidy policy.

policy is larger because the entire transfer is invested in the children; In contrast to the untargeted policy, families cannot use the targeted transfer to smooth their consumption or leisure. However, the feasibility of implementing such a policy would be under question. Under the second policy, the mean test score at the end of the development stage increases by 11% and 15% for children of college graduates and non-college graduates, respectively. These results suggest that the impact of a transfer policy to households is substantially greater for less-educated mothers, which might be due to them having fewer financial resources to invest in their children if they work as much as educated women do, and also having fewer hours they can spend with their children if they increase their hours of market work in order to match the earnings of educated women.

1.6.3 Child's Skills, Educational Attainment, and Future Earnings

In this section, I link the Child Development Supplements (CDS) to the Adulthood Supplements (TAS) to identify the relationship between the specific measure of children's skill that I used in this paper, i.e. Letter-Word identification test score, and the future educational attainment and earnings of those children.

Table (1.16) reports the regression results that link LW test scores to the future educational attainment of the CDS sample of children. Figure (1.10) shows the binned scatterplot that depicts a strong relationship between the LW test scores and the probability of having a college degree at age 24. It suggests that after controlling for the mother's education level and race, and for the gender of the child, a ten percent increase in LW test score at ages 16-18 corresponds to an 0.63 percentage point (2.0%) increase in the probability of having a bachelor's degree seven to twelve years after taking the LW test. Thus, the LW test score has a predictive power for the future educational attainment of the child and therefore is a relevant measure of the child's skill, at least with respect to her future educational attainment. Moreover, using a back-of-the-envelope calculation, I can estimate the monetary benefits of the previously-mentioned counterfactual policies.

Table (1.17) reports the regression results that link LW test scores to the future annual earnings of the CDS sample of children. Figure (1.11) shows the binned scatterplot that depicts a strong relationship between LW test score and the future annual earnings of the child. The results suggest that after controlling for the mother's education level and race, and for the birth order, age, education level, and gender of the child, a ten percent increase in LW test score at ages 16-18 corresponds to increases of 6.6% and 9.8% in annual earnings based on OLS and fixed-effect regressions, respectively, seven to twelve years after taking the LW test.[61] Note that this impact of the LW test score is in addition to its impact on educational attainment. The results suggest that the LW test score has a predictive power for the future income of the child, and thus is a relevant measure of child skill, at least with respect to her skills later in the labor market.

Now, using some simple back-of-the-envelope calculations, I evaluate the impacts of the maternity leave policies discussed earlier in this section in terms of a money-metric measure. I focus only on educated individuals here. First, it is worth noting that the fixed-effect regression results in column 3 of Table (1.17) suggest that a 10% increase in LW test score increases annual earnings 7-14 years later by about 9.8%, and this is in addition to the positive impact of increased LW test score on educational attainment. Second, from Table (1.16), a 10% increase in LW test score increase the probability of getting a four-year college degree by about 4.8 percentage point (16%).

Now, I use Zimmerman (2014) to find the causal impact of getting a four-year

[61]I observe a similar relationship between the LW test score and the future *hourly wages*. Figure (A.11) in the Supplementary Appendix depicts such a relationship when I use hourly wages instead of annual earnings.

college degree on thd future earnings. Zimmerman (2014) studies the impact of admission to a low-ranked public university on the future earnings of academically marginal students. For a marginal admission, he estimated earnings gains of 22% between eight and 14 years after high school completion. Here, I use Zimmerman's finding to set a lower bound for the average return to earnings of obtaining a four-year college degree. In Zimmerman's paper, the average return includes students who were admitted to college but later dropped out. Moreover, his result is based on data from a non-competitive public university. Hence, it seems that I can set 22% as conservative estimate, i.e. a lower bound, for the return to completing four-year college program. Hence, a 10% increase in mean LW test score increases the future earnings of children, on average, by about $22 * 0.16 \approx 3.5\%$, and does so only through increasing the probability of earning a four-year college degree that positively impacts their earnings. Thus, the 4.96% increase in the LW test scores of children under the unpaid maternity leave policy can be translated to a $3.5*0.5 \approx 1.8\%$ increase in future wages through the channel of college attainment. Moreover, according to Table (1.16), that 4.96% increase can be translated to a $0.5 * 9.8 \approx 4.9\%$ increase in future earnings on top of the impact of higher educational attainment. In total, an 4.96% increase in the mean LW test scores of children resulting from the unpaid maternity leave policy can be considered as an increase of about $4.9 + 1.8 = 6.7\%$ in the mean future earnings of the children.[62] Regarding the paid maternity policy, after a similar calculation, the

[62] I can also apply some back-of-the-envelope calculations for the monetary cost of implementing such a maternity leave policy. While implementing this policy is costless for the government, it may affect the profit of firms because they have to pay the same wage to mothers who take the maternity leave and come back to work even though their human capital is not the same as before (due to deprecation rate associated with being out of work). Thus, one might consider this policy as a redistribution from firms to working mothers, which benefits children of these mothers (6.7% increase in future earnings). One way to look at the cost to firms is to see how the wage bills change after implementing the policy. Implementing an unpaid maternity leave policy would increase the wage bills (defined as hours of market work times the hourly wage) by

change in the mean children's test score can be translated to a 2.9% increase in annual earnings.[63]

1.7 Conclusion

In this paper, using reduced-form regressions, I presented evidence of the negative relationship between mother's age at childbirth and child's skill level. However, I also showed that mothers' ages are correlated with the amount of time that mothers spend with their children; thus, the reduced-form estimate of the pure effect of mother's age at childbirth on child's skill level is biased. Then, I developed a structural model to investigate the effect of delayed childbearing on the child's skill level. I studied the problem women face in choosing the timing of childbirth while taking into account possible negative and positive effects of delayed childbearing on the child's skill level. I developed a life-cycle model of child development while endogenizing the timing of childbearing to assess the effect of maternal age at childbirth on the child's skill level. The model also helped me determine the effect of the mother's age versus those of all the inputs that a child receives from their parents during the child development process. Knowing that advanced maternal age at childbirth may be associated with negative effects on the child's skill level, individuals choose the timing of childbearing,

about 10%. It is worth noting that mothers also benefit from the policy through both higher wages (so more consumption), and higher child skill levels. These benefits, however, have not been captured by my cost-benefit analysis here.

[63] I can also apply some back-of-the-envelope calculations for the monetary cost of implementing such a paid maternity leave policy. In this case, there is a cost only during the leave because the firm does not compensate for depreciation rates in the periods following childbirth, when the mother goes back to work. The post-policy maternal age is about 25. On average, the government has to pay only once for each agent, amounting to twice her annual earnings at age 25. If I suppose that each individual works on average for 30 years, then the cost burden of the policy is equal to $2 * \frac{1}{30} \approx 6.7\%$ of their annual earnings. Again, we should note that mothers also benefit from the policy through both higher wages (so more consumption), and higher child skill levels.

make labor supply decisions, and provide time and money inputs into the child's skill formation process during the development period.

In the model presented in this paper, each individual can increase her work experience by one unit in each period through working in that period. She also decides the timing of childbirth. Individuals receive utility from the skill level of their children. If a woman decides to conceive a child in a given period, she has to drop out of the labor force for that period. Therefore, the opportunity cost of having a child at early periods is higher because the forgone human capital will affect her wages in all subsequent periods. Moreover, the opportunity cost of having a child is higher for highly-educated women because their forgone wages are greater. When a child is born, the woman chooses to allot her time optimally among leisure, work, and childcare and to allocate her money between consumption and child goods expenditures. Therefore, when it comes to the timing of childbirth, women face a trade-off: if they have a child earlier in their life-cycle, the child is more likely to have higher productivity in acquiring skills, but the net present value of forgone wages is higher, which indicates there would be fewer resources available to invest in the child during early childhood. This trade-off leads more educated women to have children later in their life-cycles, which is consistent with the studied data.

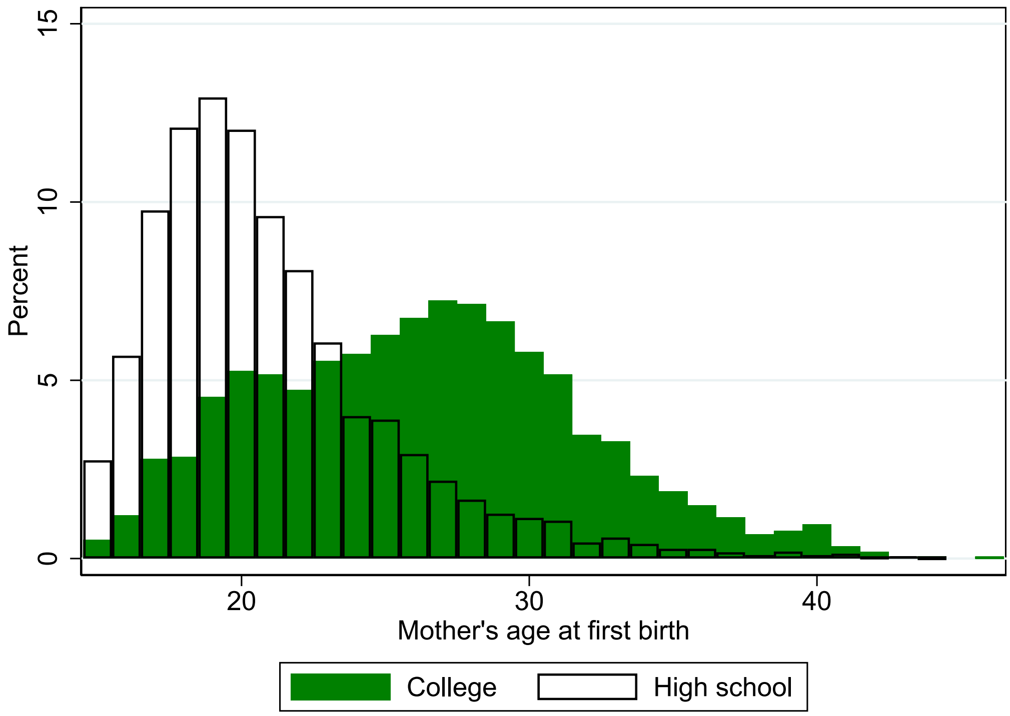
I estimated the model and found a negative effect of advanced maternal age at childbirth on the child's productivity. With everything else constant, delaying maternity age by five years decreases the child's skill level by about 11%. However, consistent with the data, the model predicts that even though highly-educated women delay childbearing, the skill levels of their children at the end of the development process are, on average, actually higher than those of the children of less-educated women. This finding highlights the effects of inputs that a child receives from her parents during childhood.

In this paper, I showed that implementing a maternity leave policy that freezes the wages of mothers decreases maternal age at first birth by about two years and increases the human capital of children by about 5%, which leads them to earn 6.6% more in adulthood. I also investigated the effects of a childcare subsidy program, which pays a portion of childcare costs to families. The effect of the policy depends on childcare costs and the amount of subsidy. Overall, the policy might not be as effective as the described maternity leave policy in terms of lowering maternal age at first birth and boosting the human capital of children.

Future work should address the role of credit constraint in the timing of fertility and child investments. First, there is no opportunity for borrowing and saving in this model. If I allowed for saving and borrowing, I could investigate the effects of credit constraints on the timing of childbirth and child skill levels. It would also be interesting to investigate the effect of uncertainty concerning labor and nonlabor incomes on the timing of childbearing and child skill levels within the framework presented in this paper.

In this paper, I analyzed the impacts of implementing a maternity leave policy on the timing of first birth and child skill levels. From a policy perspective, it is also important to study the impacts of such a policy on the demand side of the labor market. The model developed in this paper can be extended to analyze the overall impacts of a maternity leave policy in a general equilibrium context. Moreover, the impacts of a maternity policy are heterogeneous across women depending on their education levels. Hence, it is also interesting to investigate how implementing such a policy might change women's educational decisions.

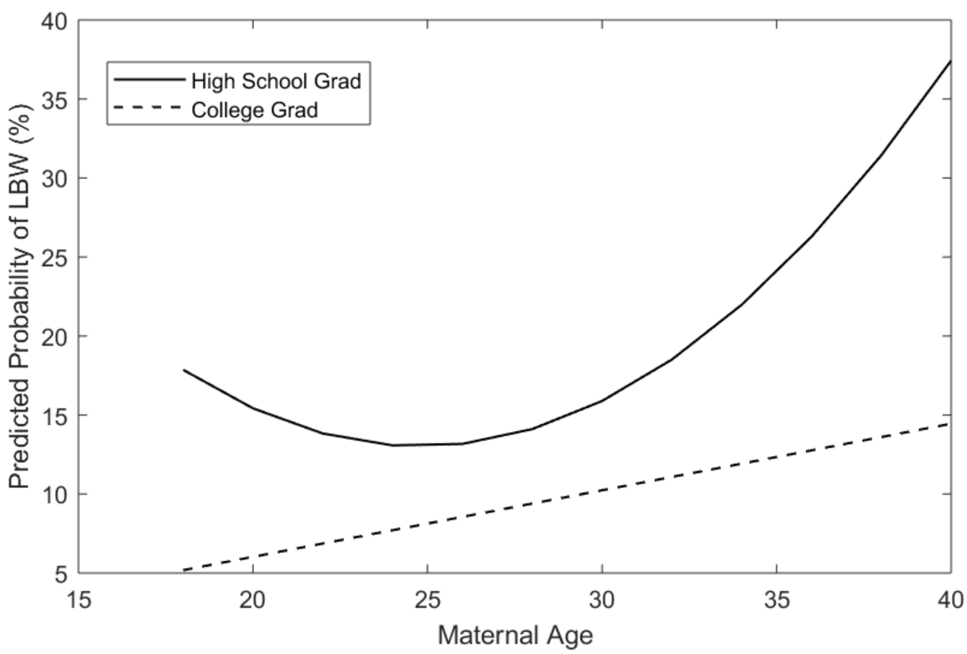
Figure 1.1: Distribution of Age at First Childbirth by Education Level



Notes: This graph shows the distribution of age at first birth for women aged 40 or above in 2015. High school level is defined as 12 completed years of schooling or less. College graduates are defined as 16 completed years of schooling or more.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Figure 1.2: The Risk of Low Birthweight and the Maternal Age



Notes: Low birthweight indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
High school level is defined as 12 completed years of schooling. College graduates are defined as 16 completed years of schooling.
The probability of low birthweight at each maternal age is calculated base on the coefficient estimates reported in Panel A of Table (1.4).

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Figure 1.3: Timing of events in period t

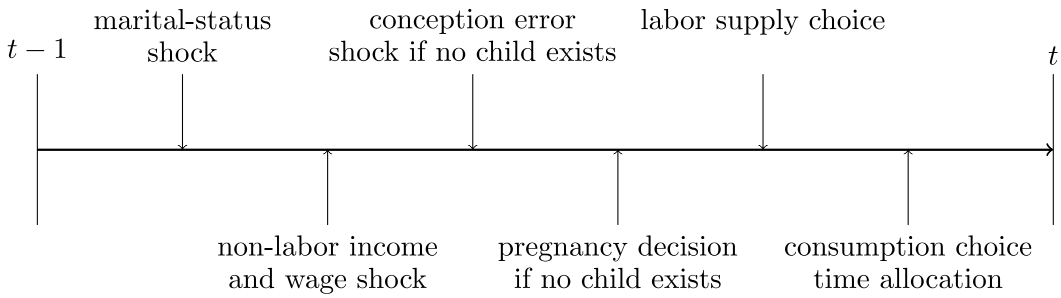
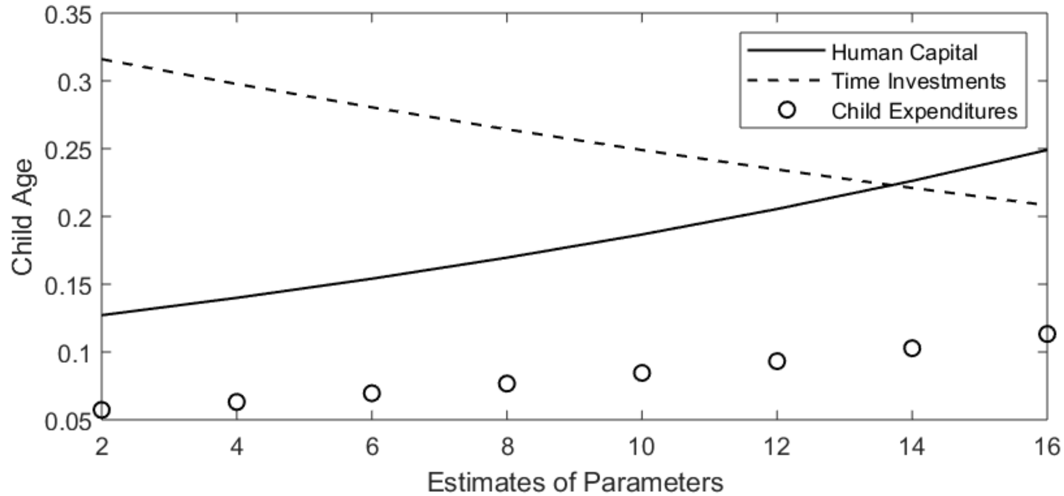
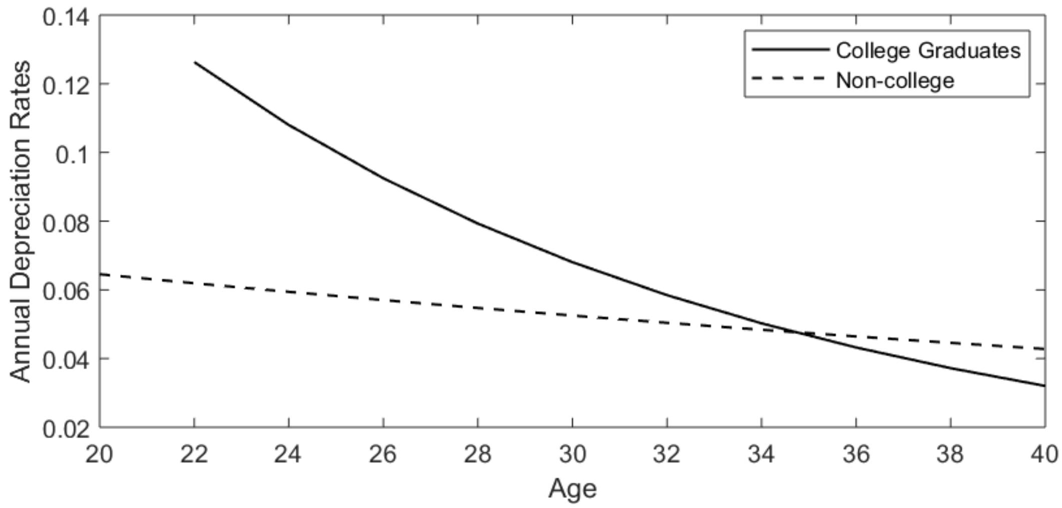


Figure 1.4: Estimated Technology Parameters by Child Age



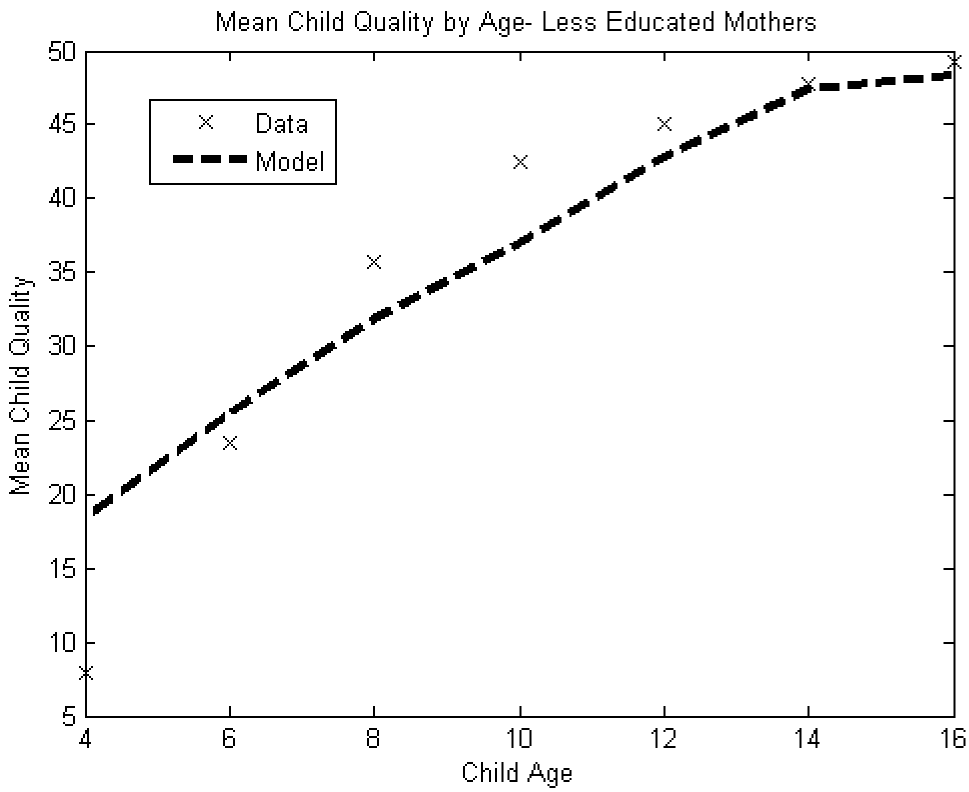
Notes: This graphs estimated parameters by child age (from Table (1.8)).

Figure 1.5: Estimated Depreciation Parameters by Mother's Age



Notes: This graphs estimated effect of career interruption on human capital by women's age (from Table (1.9)).
It shows the estimated depreciation rate of human capital associated with one year being out of the labor market.

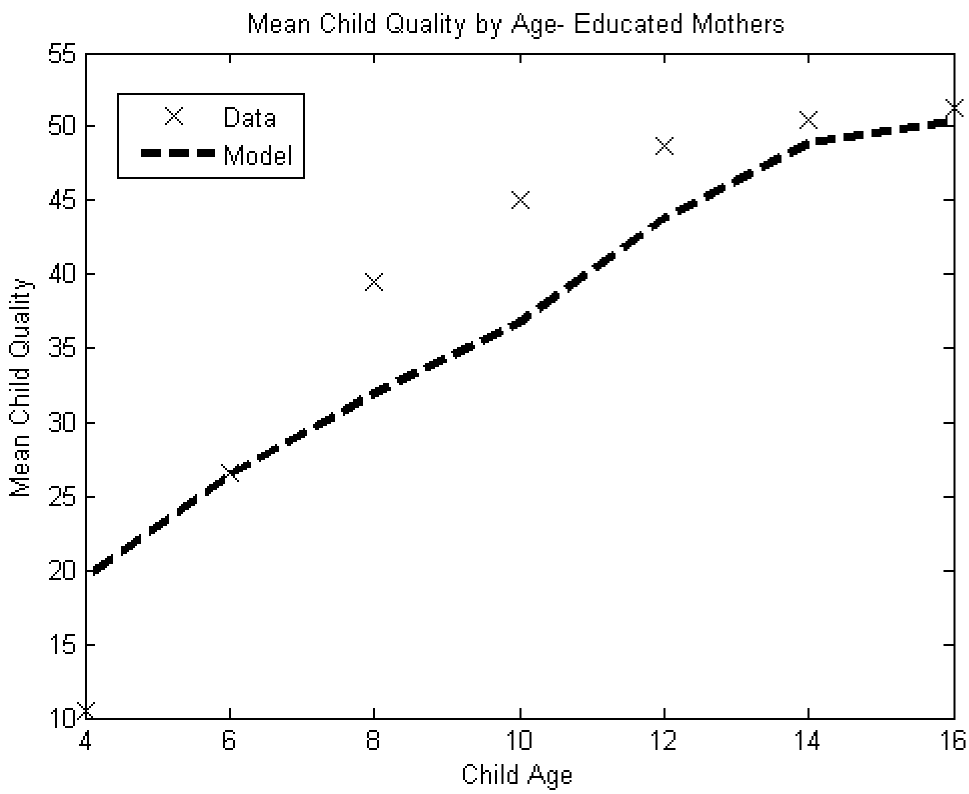
Figure 1.6: Sample Fit of Average Childs Letter Word Score (Non-college Graduates)



Notes: Data is actual data from sample of non-college mothers with one child. Simulated is the model prediction at estimated parameters given above.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1967-2015 PSID core data.

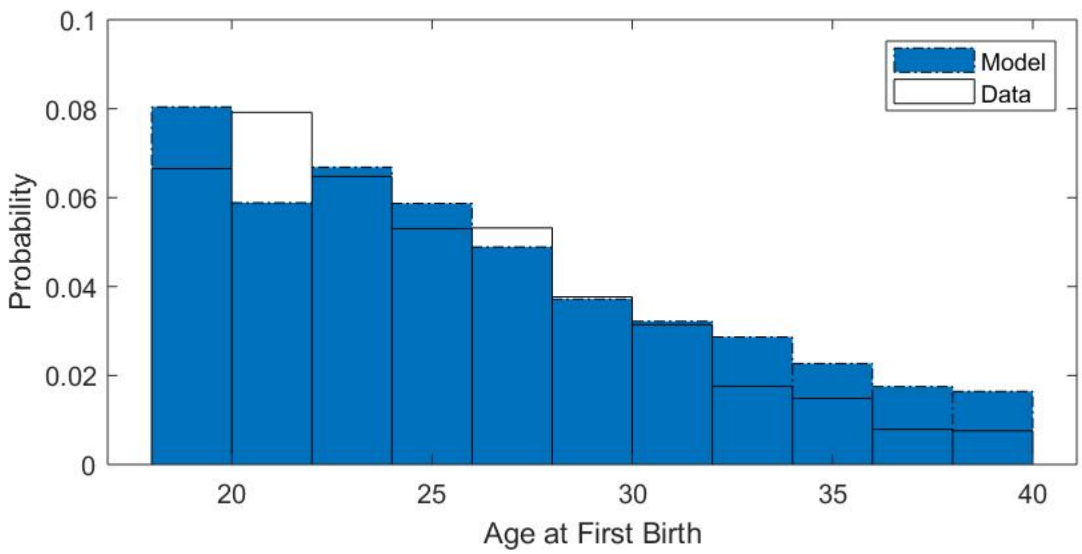
Figure 1.7: Sample Fit of Average Childs Letter Word Score (College Graduates)



Notes: Data is actual data from sample of college mothers with one child. Simulated is the model prediction at estimated parameters given above.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1967-2015 PSID core data.

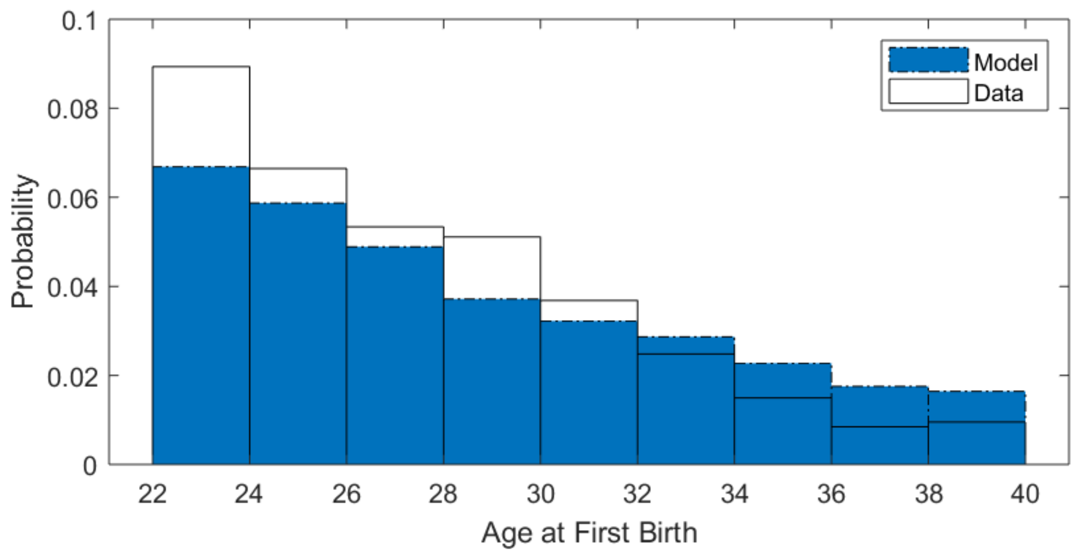
Figure 1.8: Sample Fit of Distribution of Age at First Birth (Non-college Graduates)



Notes: Data is actual data from sample of non-college mothers with at least one child in 2015. Simulated is the model prediction at estimated parameters given above.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1967-2015 PSID core data.

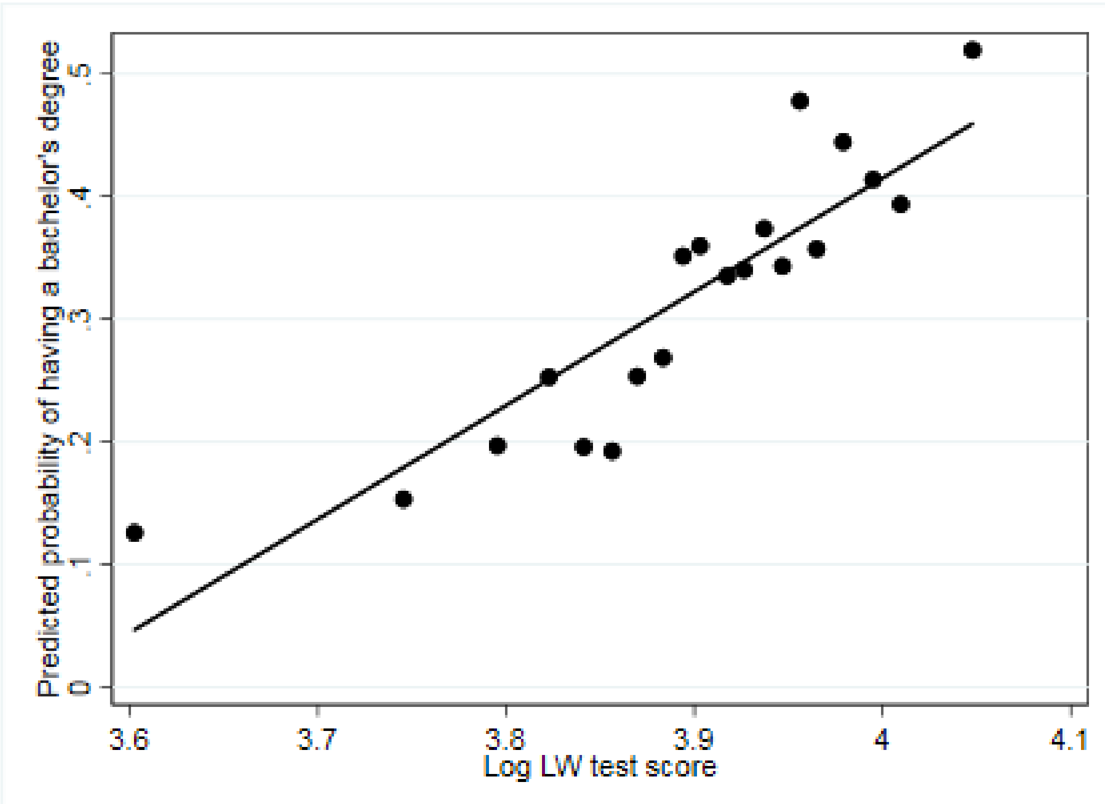
Figure 1.9: Sample Fit of Distribution of Age at First Birth (College Graduates)



Notes: Data is actual data from sample of college mothers with at least one child in 2015. Simulated is the model prediction at estimated parameters given above.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1967-2015 PSID core data.

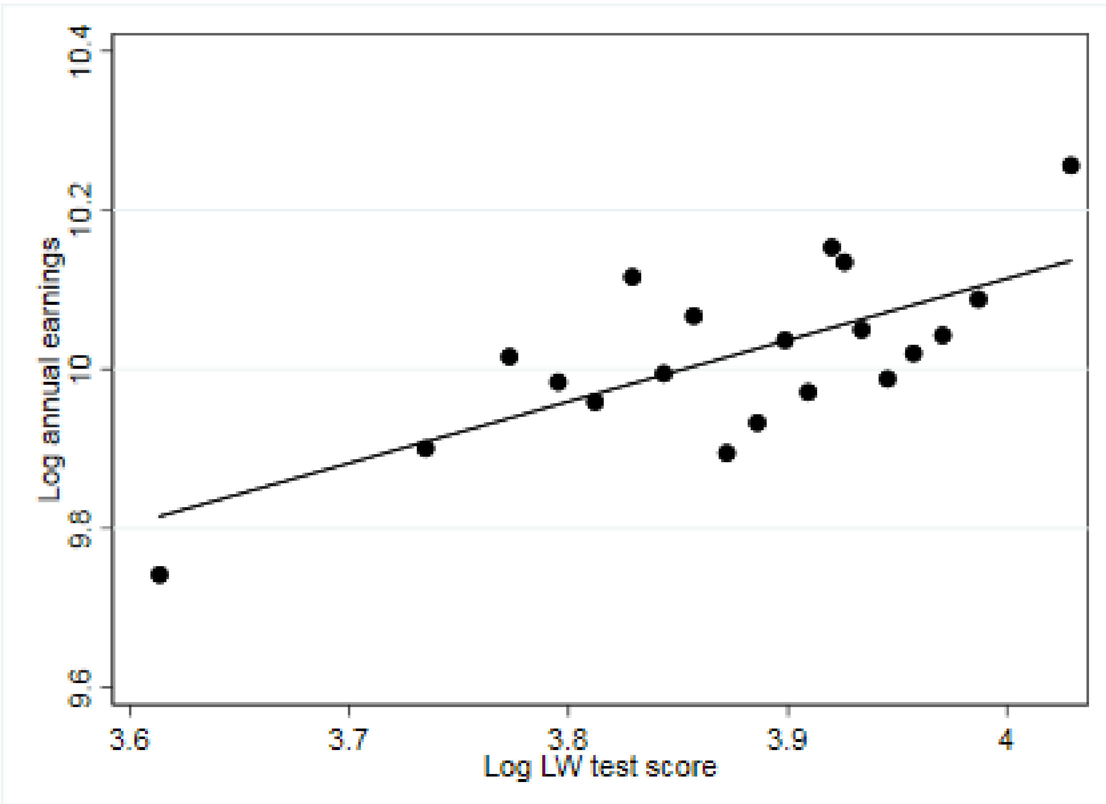
Figure 1.10: The Letter-Word Test Score and Future College Attainment



Notes:This Graph shows the relationship between children's Letter-Word (LW) test scores at age 16-18 and their future college attainment, 7 to 12 years later, i.e. when they were between 23 to 30 years of old.
The x-axis variable is LW test score of children in CDS dataset. Children were between ages 16-18 when they took the LW test.
The y-axis variable show the predicted probability of having a college degree in 2015, when the children were between 23 to 30 years of old.
The binned scatterplot is shown here. The plot shows the regression in column (2) of Table (1.16), binned into 20 bins.

Source: CDS dataset (2002-2007) combined with TAS dataset (2005-2007-2009-2011-2013-2015).

Figure 1.11: The Letter-Word Test Score and Future Earnings



Notes:This Graph shows the relationship between children's Letter-Word (LW) test scores at age 16-18 and their future earnings, 7 to 12 years later, i.e. when they were between 23 to 30 years of old.
The x-axis variable is LW test score of children in CDS dataset. Children were between ages 16-18 when they took the LW test.
The y-axis variable is the predicted log annual earnings in 2014, when the children were between 23 to 30 years of old.
The binned scatterplot is shown here. The plot shows the regression in column (1) of Table (1.17), binned into 20 bins.

Source: CDS dataset (2002-2007) combined with TAS dataset (2015).

Table 1.1: Summary Statistics

	Non-college	**College Graduates**
Variable	**Mean**	**Mean**
First child LW test score	34.9	37.9
	(15.4)	(15.5)
First child's age	9.9	9.5
	(4.1)	(4.0)
Mother's age at first birth	22.4	27.2
	(5.5)	(5.2)
Mother's years of schooling	11.7	15.5
	(1.43)	(1.2)
Father's hourly wage ($)	17.6	24.6
	(10.7)	(17.3)
Mother's hourly wage ($)	12.7	16.7
	(8.6)	(11.8)
Mother's hours worked (per week)	16.9	22.8
	(17.9)	(22.8)
Father's hours worked (per week)	38.0	39.6
	(17.4)	(15.1)
Moms' time with child (hrs/week)	33.6	35.3
	(19.0)	(18.9)

Note: This Table shows summary statistics of the subsample of observations for which age at first birth was after 18.
Standard deviations are reported in parenthesis.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1968-2015 PSID core data.

Table 1.2: Linear Regression of Low Birthweight (LBW) on Maternal Age

	Probability of LBW (%)			
	Quadratic	**Linear**	**Quadratic**	**Linear**
	(1)	(2)	(3)	(4)
Mother's age at 1st birth (age_i^p)	-2.16**	0.47***	-1.71*	0.60***
	(1.00)	(0.16)	(1.01)	(0.16)
Age square (age_i^{p2})	0.05**	–	0.04**	–
	(0.02)	–	(0.02)	–
Mean dependent variable	8.24	8.24	8.24	8.24
Child's gender dummy	Yes	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Education level dummies	Yes	Yes	Yes	Yes
Paternal age dummies	Yes	Yes	Yes	Yes
Marital status dummies	Yes	Yes	Yes	Yes
Family income	No	No	Yes	Yes
Observations	3169	3169	3169	3169
Adjusted R^2	0.02	0.02	0.03	0.04

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = \beta_1 age_i^p + \beta_2 age_i^{p2} + \alpha W_i + u_i$, for the quadratic specification in columns (1) and (3). $100 * LBW_i = \beta_1 age_i^p + \alpha W_i + u_i$, for linear specification in columns (2) and (4).
Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
Income levels are measured at the year in which the first child is born.
In columns 3 and 4, the missing data on income is treated as follows: I plugged in an arbitrary value (-9) for all missing data cases on incomes, and I included in the regression a dummy variable coded 1 if data in the original variable was missing (i.e. a value has been plugged in for missing data), 0 otherwise. Table A.4 in the Supplementary Appendix shows the case when I do not adjust for the missing data on incomes, and the number of observations decreases.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015)

Table 1.3: Low Birthweight Risk Regression Using Age Intervals

	Probability of LBW (%)		
	(1)	(2)	(3)
15 <Maternal age at first birth ≤ 20	2.54	1.78	1.78
	(2.46)	(2.50)	(2.50)
25 < Maternal age at first birth ≤ 30	2.86	3.71*	4.27**
	(2.10)	(2.15)	(2.17)
30 < Maternal age at first birth ≤ 35	10.53***	11.65***	12.86***
	(2.82)	(2.88)	(2.95)
35 < Maternal age at first birth ≤ 40	13.51***	14.84***	16.68***
	(4.34)	(4.40)	(4.51)
40 < Maternal age at first birth ≤ 45	25.45**	26.52**	27.59**
	(11.13)	(11.23)	(11.23)
45 < Maternal age at first birth ≤ 50	90.94***	92.89***	97.84***
	(27.91)	(27.83)	(27.94)
Race dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Paternal age dummies	Yes	Yes	Yes
Marital status dummies	Yes	Yes	Yes
Education level dummies	No	Yes	Yes
Family income	No	No	Yes
Observations	1580	1580	1580
Adjusted R^2	0.02	0.03	0.04

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = I_{15<AFB\leq 20} + I_{25<AFB\leq 30} + I_{30<AFB\leq 35} + I_{35<AFB\leq 40} + I_{40<AFB\leq 45} + I_{45<AFB\leq 50} + \alpha W_i + u_i$, where $I_{25<AFB\leq 30}$ is an indicator variable equal to one if mother's age at first birth is between 20 and 25 and zero otherwise, and W is the vector of control variables.
Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (2500 grams).
The baseline group for comparison is mothers whose age at first birth are between 20 and 25, which is the omitted dummy in the above regression. The mean dependent variable for the baseline group is 7.3.
Income levels are measured at the year in which the first child is born.
Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Table 1.4: Linear Regression of LBW on Maternal Age

Panel A: By Education Level		
	Non-college	**College Graduates**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at 1st birth (age_i^p)	-4.10*	0.42**
	(2.17)	(0.20)
Age square (age_i^{p2})	0.09**	–
	(0.04)	–
Mean dependent variable	11.30	7.03
Observations	777	657
Adjusted R^2	0.07	0.03

Panel B: By Income Level		
	Below-Median	**Above-Median**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at 1st birth (age_i^p)	-5.26**	1.06**
	(2.45)	(0.31)
Age square (age_i^{p2})	0.12***	–
	(0.05)	–
Mean dependent variable	11.23	8.27
Child's gender dummy	Yes	Yes
Race dummies	Yes	Yes
Year dummies	Yes	Yes
Education dummies	Yes	Yes
Observations	800	800
Adjusted R^2	0.04	0.05

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = \beta_1 age_i^p + \beta_2 age_i^{p2} + \alpha W_i + u_i$, for the quadratic specification in column (1).
$100 * LBW_i = \beta_1 age_i^p + \alpha W_i + u_i$, for linear specification in column (2).
Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
Estimates in Panel A are controlled for family income, and estimates in Panel B are controlled for education level.
Estimates in Panel A are represented for two levels of education, separately. The first column reports the result for non-college graduates (i.e. 12 years of schooling), and the second column shows the results for college graduates (i.e. 16 years of schooling).
Income levels are measured at the year in which the first child is born.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 1.5: Regression Results Including Only One-child Families

	(1) **ln(Score)**	(2) **ln(Score)**	(3) **ln(Score)**
Mother's age at 1st birth (age_i^p)	**0.001**	**-0.005**	**-0.014****
	(0.003)	(0.004)	(0.006)
Mother's years of schooling		0.040***	0.030**
		(0.015)	(0.015)
Father's years of schooling		0.010	-0.001
		(0.011)	(0.014)
Father's hourly wage			0.007**
			(0.003)
Father's hours worked (per week)			0.005**
			(0.002)
Mother's hourly wage			0.006*
			(0.003)
Mother's hours worked (per week)			-0.003
			(0.002)
Child's age dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Marital status dummies	No	No	Yes
Paternal age dummies	No	No	Yes
Child's gender dummy	No	No	Yes
Observations	206	206	206
Adjusted R^2	0.87	0.88	0.89

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $Ln(score)_{i,t} = \beta_1 age_i^p + \gamma Z_{i,t} + \alpha W_i + u_{i,t}$,
The dependent variable is the natural logarithm of Letter-Word Identification Test score. I only use the data on the first child of mothers.
Child's age is controlled in all above equations using age-specific dummy variables.
The coefficient of interest can be translated to $\frac{1}{70}$ standard deviation of LW test score, i.e, Z-score is equal to -0.014.
The coefficient of interest can be translated to $\frac{1}{20}$ of average increase in LW test score when a child becomes a year older.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews.

Table 1.6: Regression Results: Time Investment Per Week

	(1) Hours/week (τ)	(2) Hours/week (τ)	(3) Hours/week (τ)
Age at 1st birth (age_i^p)	**0.22*** **	**0.20*** **	**0.19*** **
	(0.07)	(0.07)	(0.08)
Years of schooling		0.20	0.30
		(0.21)	(0.22)
Mother's hourly wage			-0.01
			(0.05)
Mother's work hours per week			-0.15***
			(0.03)
Child's age dummies	Yes	Yes	Yes
Year fixed effect dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Number of children dummies	Yes	Yes	Yes
Marital status dummies	No	No	Yes
Child's gender dummy	No	No	Yes
Observations	1726	1726	1726
Adjusted R^2	0.16	0.16	0.17

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $\tau_{i,t} = \beta_1 age_i^p + \gamma Z_{i,t} + \alpha W_i + u_{i,t}$,
The dependent variable is the weekly hours that the mother spends with her child.
I only use the data on the first child of mothers.
Mean dependent variable is 34.
Child's age is controlled in all above equations using age-specific dummy variables.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews.

Table 1.7: Preference Parameter Estimates

Parameter	Description	Value	SE
α_1	Consumption impact	0.420	0.0064
α_2	Parents leisure impact	0.275	0.0035
α_3	Child's skill level impact	0.305	0.0228
φ	Child's skill multiplier after final period	0.026	0.0005

Table 1.8: Technology Parameter Estimates

Parameter	Description	Value	SE
$\gamma_{1,0}$	Last periods child quality intercept	-2.0625	0.0050
$\gamma_{2,0}$	Mother's time intercept	-1.1520	0.0083
$\gamma_{3,0}$	Child expenditures intercept	-2.8612	0.0111
$\gamma_{1,1}$	Last periods child quality slope	0.0960	0.0010
$\gamma_{2,1}$	Mother's time slope	-0.0596	0.0016
$\gamma_{3,1}$	Child expenditures slope	0.0974	0.0010
ϵ_1	Parameter in productivity factor equation	-0.0150	0.0012
ϵ_2	Parameter in productivity factor equation	-0.0004	0.0001
ϵ_3	Parameter of prod. wrt mother's innate ability	0.0710	0.0047
$A_{0,0}$	Constant term in skill formation technology	1.3813	0.0053
$A_{0,1}$	Constant term in skill formation technology	0.0803	0.0003
σ_R	Standard deviation of shocks to A	0.0086	0.0002

Table 1.9: Depreciation Parameter Estimates

Parameter	Description	Value	SE
$\delta_{H,0}$	Depreciation rate intercept for non-college grads	-2.29	0.1188
$\delta_{C,0}$	Depreciation rate intercept for college grads	-1.01	0.0459
$\delta_{H,1}$	Depreciation rate slope for non-college grads	-0.04	0.0017
$\delta_{C,1}$	Depreciation rate slope for college grads	-0.06	0.0032

Table 1.10: Wage Parameter Estimates and Pregnancy Parameter Estimates

Parameter	Description	Value	SE
$\eta_{0,C}$	Intercept of the log hourly wages (college grads)	2.816	0.0355
$\eta_{1,C}$	Return to education (college grads)	0.110	0.0067
$\eta_{2,C}$	Coefficient of sq. experience (college grads)	-0.006	0.0002
$\eta_{0,H}$	Intercept of log hourly wages (non-college grads)	2.431	0.1035
$\eta_{1,H}$	Return to education for college graduates	0.050	0.0013
$\eta_{2,H}$	Coefficient of sq. experience (non-college grads)	-0.006	0.0002
σ_{ω_C}	Std. dev. of log wage shocks (college grads)	0.549	0.0251
σ_{ω_H}	Std. dev. of log wage shocks (non-college grads)	0.853	0.0645
λ_c	Prob. of unintended births (college grads)	0.079	0.0022
λ_h	Prob. of unintended births (non-college grads)	0.193	0.0374

Table 1.11: Sample Fit

Moment	**College Graduates**		**Non-college**	
	Data	**Model**	**Data**	**Model**
Mean age at the first birth	26.2	26.5	22.4	21.6
Mean test score at the last period	50.3	50.6	48.5	48.0
Mean wage at 2 years of experience	13.3	14.4	11.3	12.2
Mean wage at 8 years of experience	18.4	19.7	14.1	14.9
Mean hours worked per week	26.7	23.1	24.2	21.8
Mean time with the 1st child	21.9	30.5	21.7	30.9
Fraction of the 1st births after age 27	0.47	0.47	0.15	0.09
Fraction of unintended births	0.17	0.15	0.40	0.34
$Corr(w^{sp}, k)$	0.19	0.24	0.19	0.23
$Corr(age^{ch}, k)$	0.85	0.71	0.85	0.56
$Corr(age^{child}, \tau)$	-0.44	-0.69	-0.47	-0.71
$Corr(w, k)$	0.14	0.27	0.14	0.14
$Corr(x, k)$	0.38	0.50	0.37	0.28
$Corr(age^{p}, k_{i,t} - k_{t-5})$	-0.08	-0.06	0.02	-0.01

Notes: Data is actual data from sample of non-college mothers with at least one child in 2015.
Simulated is the model prediction at estimated parameters given above.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1967-2015 core data.

Table 1.12: Decomposition Exercise

	Percent Change from Baseline		
	Exercise (1)	**Exercise (2)**	**Exercise (3)**
Mean test score at age 16	11.4	-0.8	-1.6
Child goods	fixed	↓	↓
Time investments	fixed	fixed	↓
Age at first birth (AFB)	22	22	22

Notes: This Table decomposes the maternal age effect on the child's skill level into the negative health effect of the maternal aging and the positive impact of higher child investments. College graduates are considered for this exercise.
Exercise 1: First, I lower the maternal age of each educated mother from the benchmark (the fitted simulated dataset in which the mean AFB is 27) to age 22. Hence, Exercise 1 assumes that everyone conceives a child at age 22. But the investments are the same as the benchmark (no reduction in investments as a result of a reduction in maternal age). Thus, investments are those investments that older parents would make, but the maternal age is 22.
Exercise 2 assumes that everyone conceive a child at age 22, and child investments are reduced accordingly with respect to the mother's wages. Child good expenditures are reduced in proportion to the reduce in the mother's wage if the mother would conceive a child at age 22, which means that the wage is lower due to both lower work experience and the high depreciation rate at age 22 .
Exercise 3 is the same as Exercise 2, but this time the time investments are also reduced according to the results of Table (1.6). Time investments are reduced by 3%.
Numbers in this Table show the average of percentage changes from the baseline over the children in the simulated data (first changes are calculated for each individual, and then, the average change is calculated).

Table 1.13: Counterfactual Maternity Leave Policies

Moment	**College Graduates**			**Non-college**		
	(1) Unpaid	**(2) Paid**	**(3) Paid (no dep)**	**(4) Unpaid**	**(5) Paid**	**(6) Paid (no dep)**
	Percent Change from Baseline					
Age at first birth	-6.40	-7.89	-7.91	-1.72	-3.04	-3.04
Fraction of births after 35	-51.92	-61.54	-67.31	-9.09	-36.36	-36.36
Mean test score at age 16	4.96	2.16	5.87	3.04	1.03	3.71
Mean time w/ child at age 6 of children	-6.58	13.35	-6.59	-9.59	0.76	-9.75
Mean child expenditures at age 6 of children	13.95	-32.42	10.94	29.44	-5.44	27.41
Fraction who works at age 6 of children	20.53	-45.19	19.75	33.23	-4.27	33.84
Mean hours worked at age 6 of children	22.06	-44.76	22.09	32.04	-2.54	32.58
Mean leisure at age 6 of children	-6.58	13.35	-6.59	-9.59	0.76	-9.75
Mean consumption at age 6 of children	13.94	-32.47	10.93	29.44	-5.44	27.41
Mean wage at age 6 of children	13.17	-49.25	10.01	31.54	-3.59	29.83
Mean experience at age 6 of children	-30.08	-58.56	-34.00	6.33	-10.35	2.89

Notes: All values are the percentage change from the baseline values given in prior tables.
Unpaid maternity leave policy: Under the unpaid maternity leave policy, mothers can leave their job for 2 years. They can return to their work while receiving the same wage as before the childbirth.
Paid maternity leave policy: Under this policy, mothers can leave their job for 2 years and receive their previous wage. They can return to their work after 2 years. Their wages, however, might be lower when back due to depreciation.
Paid (no depreciation) maternity leave policy: Mothers leave their job for 2 years and receive their previous wage. They can return to their work while receiving the same wage as before the childbirth (there is no depreciation).
Fraction of births after age 35 in the benchmark is equal to 0.13 and 0.04 for college grads and non-college grads, respectively.
For each policy, first changes are calculated for each individual, and then, the average change is calculated.

Table 1.14: Counterfactual Childcare Subsidy Policy

Moment	**College Graduates**			**Non-college Graduates**		
Subsidy rate	**50%**	**75%**	**100%**	**50%**	**75%**	**100%**
	Percent Change from Baseline					
Age at first birth	-0.10	0.08	0.10	0.00	0.00	-0.10
Fraction of births after 35	-5.77	-1.92	-5.77	0.00	0.00	-9.09
Mean test score at age 16	0.04	0.74	1.15	0.96	1.52	2.19
Mean time w/ child at age 6 of children	-1.66	-6.49	-9.06	-3.39	-5.14	-7.18
Mean child expenditures at age 6 of children	-1.00	10.53	15.43	6.50	10.72	16.22
Fraction who works at age 6 of children	9.22	31.99	45.73	14.94	26.52	40.15
Mean hours worked at age 6 of children	5.58	21.77	30.38	11.31	17.18	23.99
Mean leisure at age 6 of children	-1.66	-6.49	-9.06	-3.39	-5.14	-7.18
Mean consumption at age 6 of children	-0.99	10.54	15.44	6.50	10.72	16.22
Mean wage at age 6 of children	-11.19	-5.82	-7.72	-1.94	-3.29	-4.06
Mean experience at age 6 of children	-1.96	10.37	14.45	5.70	9.39	13.84
Mean utility	2.27	3.30	3.91	3.30	4.36	5.45

Notes: All values are the percentage change from the baseline values given in prior tables.
50% subsidy policy: Under this policy, the government provides 50% subsidy on the child care cost.
75% subsidy policy: Under this policy, the government provides 70% subsidy on the child care cost.
100% subsidy policy: Under this policy, the government provides 100% subsidy on the child care cost (free public child care system).
Fraction of births after age 35 in the benchmark is equal to 0.13 and 0.04 for college graduates and non-college graduates, respectively.
For each policy, first changes are calculated for each individual, and then, the average change is calculated.

Table 1.15: Counterfactual Transfer Policy

Moment	College Graduates		Non-college Graduates	
	(1) Untargeted	(2) Targeted	(3) Untargeted	(4) Targeted
	Percent Change from Baseline			
Age at first birth	0.68	-0.66	0.50	-0.37
Fraction of births after 35	-4.08	-24.49	0.00	0.00
Mean test score at age 16	1.97	11.97	4.40	15.75
Mean time w/ child at age 6 of children	6.31	0.50	10.83	0.80
Mean child expenditures at age 6 of children	6.77	198.78	8.91	269.76
Fraction who works at age 6 of children	-5.45	-0.89	-14.38	-1.89
Mean hours worked at age 6 of children	-14.74	-1.17	-27.08	-2.00
Mean leisure at age 6 of children	6.31	0.50	10.83	0.80
Mean consumption at age 6 of children	6.77	-2.41	8.91	-1.83
Mean wage at age 6 of children	1.80	-2.57	5.53	-1.20
Mean experience at age 6 of children	-1.32	-2.47	-5.57	-2.47

Notes: All values are the percentage change from the baseline values given in prior tables.
Untargeted policy: a $250 transfer per week in form of nonlabor income to each family with a child during the development stage.
Targeted policy: a transfer targeted to children by providing 250 dollars worth of child goods to the household each week.
Fraction of births after age 35 in the benchmark is 0.13 and 0.04 for college graduates and non-college graduates, respectively.
For each policy, first changes are calculated for each individual, and then, the average change is calculated.

Table 1.16: Regression Results- LW Test Score and Educational Attainment

	Pooled OLS		Fixed-effect using siblings	
	(1) **High school**	**(2)** **Bachelor**	**(3)** **High school**	**(4)** **Bachelor**
Log LW test score	**0.52*****	**0.77*****	**0.63*****	**0.43*****
	(0.06)	(0.10)	(0.13)	(0.13)
Mean dependent variable	0.94	0.30	0.94	0.30
Gender	Yes	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes	Yes
Mother's education dummies	Yes	Yes	Yes	Yes
Maternal age dummies	Yes	Yes	Yes	Yes
Birth order dummies	Yes	Yes	Yes	Yes
Observations	1354	1354	1354	1354
Adjusted R^2	0.14	0.22	0.26	0.19

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $hs_{i,2014} = \beta_1 ln(LW_{i,16yrs}) + \alpha W_i + u_i$,
$college_{i,2014} = \beta_1 ln(LW_{i,16yrs}) + \alpha W_i + u_i$,
This Table shows the relationship between children's Letter-Word (LW) test scores at age 16-18 and their future educational attainment, 7 to 12 years later, i.e. when they were between 23 to 30 years of old.
The regressor in this Table is LW test score of children in CDS dataset. Children were between ages 16-18 when they took the LW test.
In the first column, the dependent variable, i.e. $hs_{i,2014}$, is a binary variable equal to 1 if the individual has a high school degree, and 0 otherwise. In the second column, the dependent variable, i.e. $college_{i,2014}$, is a binary variable equal to 1 if the individual has a bachelor's degree, and 0 otherwise. The dependent variable is related to the individual's educational status in year 2014, when the children were between 23 to 30 years of old.
The last two columns report the family fixed-effect regression results. I use the within-family variations in educational attainments and LW test scores, i.e. the variations in LW test scores and educational attainments between the siblings in a family, to examine the relationship between the LW-test score and the educational attainment.
A one-percent increase in LW test score is associated with 0.52 or 0.63 percentage point increase in the probability of getting a high school degree based on pooled OLS or fixed-effect, respectively.
A one-percent increase in LW test score is associated with 0.77 or 0.43 percentage point increase in the probability of getting a bachelore's degree based on pooled OLS or fixed-effect, respectively.
Source: CDS dataset (2002-2007) combined with TAS (2005-2007-2009-2011-2013-2015).

Table 1.17: Regression Results- LW Test Score and Future Earnings

	Pooled OLS		Fixed-effect
	(1)	**(2)**	**(3)**
	Ln(earnings)	**Ln(earnings)**	**Ln(earnings)**
Log LW test score	**0.75****	**0.66****	**0.98*****
	(0.31)	(0.32)	(0.50)
Gender	Yes	Yes	Yes
Age dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Mother's education dummies	Yes	Yes	Yes
Maternal age dummies	Yes	Yes	Yes
Birth order dummies	Yes	Yes	Yes
Education level	No	Yes	Yes
Observations	408	408	408
Adjusted R^2	0.08	0.12	0.77

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $\ln(earnings_{i,2014}) = \beta_1 ln(LW_{i,16yrs}) + \alpha W_i + u_i$,
This Table shows the relationship between children's Letter-Word (LW) test scores at age 16-18 and their future earnings, 7 to 12 years later, i.e. when they were between 23 to 30 years of old.
The regressor in this Table is LW test score of children in CDS dataset. Children were between ages 16-18 when they took the LW test.
The dependent variable is the log annual earnings in all regressions. All earnings are related to year 2014, when the children were between 23 to 30 years of old.
The last two columns report the family fixed-effect regression results. I use the within-family variations in educational attainments and LW test scores, i.e. the variations in LW test scores and annual earnings between the siblings in a family, to examine the relationship between the LW-test score and the annual earnings.
A one-percent increase in LW test score is associated with 0.66 or 0.98 % increase in the annual earnings based on pooled OLS or fixed-effect regressions, respectively, and this is on top of the impacts of LW test score on the annual earnings.
Source: CDS dataset (2002-2007) combined with TAS dataset (2015).

Chapter 2

GENDER MIX OF CHILDREN, MOTHERS EARNINGS, AND THE GENDER WAGE GAP

2.1 Introduction

The gender wage gap (GWG) has been extensively studied over the past few decades. In accounting for the observed differences in earnings between males and females, early approaches to explaining the GWG focused on the role of human capital (schooling and work experience), the family division of labor, compensating wage differentials, gender differences in occupations/industries, and discrimination (Blau and Kahn (2017)). These studies emphasize the role of differences in preferences, comparative advantage, and job characteristics. Some recent papers highlight the role of motherhood in explaining the gender gap in pay (Kleven *et al.* (2018); Cukrowska-Torzewska and Lovasz (2016); Adda *et al.* (2017)). Among these studies, Kleven *et al.* (2018) show that the arrival of children creates a gap in earnings. They find that 80% of the GWG in Denmark for the year 2013 can be attributed to the arrival of children. Gallen (2018) studies the relationship between the gender productivity gap, the pay gap, and motherhood in Denmark. The study finds that 75% of the GWG can be accounted for by productivity differences between men and women. However, while the earnings gap coincides with the productivity gap for mothers, the productivity gap is so small that it cannot explain the GWG for young women without children, suggesting there is a discrimination against young, non-mother female workers.[1]

[1]Researchers tend to use regression analysis to decompose the gap in pay into that which can be attributed to observable differences between males and females workers and a residual is usually perceived as discrimination; see Blau and Kahn (2017) for a

If a sizable portion of the gap in pay between males and females is due to the arrival of children as Kleven *et al.* (2018) find, or if it can be attributed to the decline in future productivity of women expected for childbirth (and/or interruption of women's work as a consequence of childbearing), then one might ask the following question: Assuming perfect foresight for individuals, how might this anticipated *child penalty*" affect females' decisions to invest in human capital, and thereby their earnings? Most of the previous studies have focused on the ex-post effects of children given choices made before arrival of children, and less is known about the ex-ante effects of fertility plans on women's earnings.[2] In this paper, I aim to shed light on the influence of female workers' expectations regarding their future fertility behavior on their current investment in human capital through on-the-job training. This study is designed to investigate the extent to which there is a wage differential between female workers, if any, based on their expectations about future fertility decisions. The question I ask is For women at child-bearing ages, what is the effect of the anticipation of motherhood in terms of earnings and other labor market outcomes?

In this paper, I utilize variation in the gender mix of children to investigate the role of the female's expectation regarding future fertility in guiding her human capital investments and earnings. The genders of children are generally exogenously determined, i.e. their assignment is a quasi-random event. As soon as a second child is born, the gender mix of children comes into play, and it exogenously changes the likelihood of future fertility. Parental preferences for variety in their children's genders make mothers of children of the same gender more likely to conceive another child. This exogenous heterogeneity in the likelihood of further fertility helps me identify

survey of the literature.

[2]Adda *et al.* (2017) is an exception. They assess the contribution of fertility to the gender wage gap within a dynamic model of career choice, human capital accumulation, and labor supply decisions. However, they mainly focus on the effects of fertility on occupational choices.

the impacts of fertility expectation on a woman's earnings.

I use the Panel Study of Income Dynamics (PSID), National Longitudinal Survey of Youth 1997 (NLSY97), and National Longitudinal Survey of Youth 1979 (NLSY79) datasets to investigate the role of anticipation of motherhood on females' earnings. I track females' labor market decisions (i.e. labor force participation, hours of market work, and occupations/industries) and their labor market outcomes (i.e, labor income and hourly wages) over their life-cycle. The datasets also provide rich information on each individual's fertility-related behavior (such as the number of children, timing of births, and the marital status), as well as each individual's age, race, and educational attainment. Thus, I observe how each individual's labor market outcome changes over time, both before and after each childbirth. I consider the gender mix of children as a treatment, mothers with children of different genders as the control group, and mothers of children of the same sex as the treatment group. Then, in an event study framework, I investigate the effects of the quasi-random assignment of gender mix on the treatment group.

I find that having two children of the same gender (two boys or two girls) is associated with between 4% and 6% lower hourly wages for the mother after the second birth compared to having two children of opposite sexes, regardless of the number of children the mother will ultimately have. I find that parental preferences for variety make having a third child about 12% more likely when the first two are of the same gender. I argue that mothers with two boys or two girls perceive themselves more likely to bear one more child, so are less attached to the labor market compared to mothers with one boy and one girl. The wage differential between mothers of same-gender children and opposite-gender children exists even for women who never conceived a third child. Having a third child is probabilistic; women who will have a third child do not know this fact beforehand with certainty. However, mothers who have two children of the same gender face a higher probability of conceiving a third child. In other words, when it comes the the impact of sibling genders on their mother's earnings, the mother's beliefs about the likelihood of future fertility matters, not the actual future fertility. I also provide support for this hypothesis using data on individuals' beliefs concerning their future fertility plans.

In order to model the impact of a woman's expectation regarding her future fertility on human capital investments, I follow Polachek *et al.* (2008); I assume that in each period, each individual maximizes her expected present value of life-time earnings by allocating optimally her resources to human capital investments. In a given period, actual earnings are potential earnings minus investment costs. Human capital investments increase the individual's human capital stock and future earnings power. The expected lifetime labor force participation influences an individual's human capital investments. The higher the expected labor force participation, the more gains from investments in human capital, the higher the amount of investments, and higher wages. Thus, the higher probability of future fertility, the lower expected life-time labor force participation, the lower gains from investing in human capital, the lower wages. This model explains how an exogenous shock to the probability of future fertility, driven by the gender mix of the first two children, affects one's human capital investment decisions, so wages.

One might argue that the consequences of having a third child might not be as significant of those of a first childbirth and therefore individual's beliefs about the likelihood of their future fertility are not relevant for their human capital investments. To address this concern, using a difference-in-differences event study approach, I find that labor force participation declines by about 30% right after a third childbirth compared to the year before the third pregnancy. Also, for those who continue to participate in the labor market after a their third childbirth, hours of market work decrease by about 30%. Hence, conceiving a third child is associated with sizable reductions in labor force participation at both extensive and intensive margins, and the higher the probability of a third childbirth, the higher is the probability of a reduction in labor force participation in future. Anticipation of this discontinuous labor force participation affects individuals' decisions on human capital investment and on-the-job training. Hence, on average, mothers of two children of the same gender have less incentive to invest in their own human capital and face lower wages compared to mothers of children with opposite genders.

Kuziemko *et al.* (2018) hypothesize that when women are making key human capital decisions, they underestimate the impact of motherhood on their future labor supply; they do not anticipate the substantial and persistent "motherhood penalty". Results of this paper might suggest that women *do* anticipate the consequences of the arrival of children. However, the results also might be reconciled with the findings of Kuziemko *et al.* (2018) in the following sense: in this paper I use data on the gender mix of children and therefore all women in my sample already have two children, which means they have faced with the consequences of childbirth on their labor market outcomes in the past and learned from it. Hence, they can anticipate the consequences of an additional childbirth on their labor market outcomes.

Also, it is worth mentioning that the gender mix of children has been widely used as an Instrumental Variable (IV) to study the effect of family size on labor market outcomes. The results of this paper cast doubt on the validity of using this IV in many applications within the literature.

2.2 Data

I utilize the Panel Study of Income Dynamics (PSID) and four waves of its Child Development Supplement (CDS). The PSID is a longitudinal survey of a nationally-representative sample of U.S. families. It was initiated in 1968 including about 4,800 families. Data was collected for those families and their descendants in different waves over 49 years, through 2017. Currently, PSID collects information on almost 25000 individuals in more than 10000 families. The dataset includes characteristics of parents such as their age at childbirth, education level, hours worked, accepted wages, household income, and hours spent with children (see Dascola *et al.* (2015) for the PSID Main Interview User Manual, and Hofferth *et al.* (1997) for the user's guide to the CDS.). I also use the National Longitudinal Survey of Youth 1997 (NLSY97) and the National Longitudinal Survey of Youth 1979 (NLSY79). The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. These individuals were interviewed annually through 1994, and biennially thereafter. The NLSY97 consists of a nationally representative sample of about 9,000 youths who were 12 to 16 years old in 1996. The first wave was conducted in 1997. For this paper, I mainly use data from PSID in the empirical analysis. I use the NLSY79 and NLSY97 datasets for robustness checks with respect to the choice of dataset.

Table (2.1) provides summary statistics for the PSID data used throughout the paper. Notably, the median number of children at age 40 for mothers whose first two children have different genders is two, while the median is three for mothers whose first two children have the same gender. More precisely, 53% of mothers with two children of the same gender will rear another child, compared to 47% of mothers with two children of different genders.

2.3 Evidence from the PSID

The question posed in the first section has implications for the gender mix of children, which can be tested by observational data. Under three assumptions that:

(1) parental preferences for variety make having a third child more likely when the first two have the same gender, i.e. there is a desire to have at least one of each gender;

(2) the gender of children is assigned exogenously; and

(3) females, while investing in their own human capital, might take their future fertility decisions (and their consequences) into account,

I hypothesize that there might be a human capital differential between females with two children of the same gender and females with two children of different genders, and this differential could be reflected in their hourly wages.

I test the first assumption using the observatiobal data. Specifically, I test whether or not the gender mix of children has predictive power for the total number of children that a woman will have.

In Tables (2.2), I use different models to estimate the effect of having two children of the same sex on the probability of having further children. The dependent variable is an indicator variable, which is equal to one if the individual has her third child by age 40, and zero otherwise. The first column of Table (2.2) shows the results from a linear probability model. The coefficient of interest is positive and statistically significant at the one percent confidence level. Mothers who have two children of the same gender are about six percentage points (i.e. about 12%) more likely to have further children by the end of their childbearing years, i.e. age 40. The second and third columns show the results from using probit and logit models, respectively. The coefficient of interest is again positive and statistically significant at the one percent level. All told, these results show that the gender mix of children is informational concerning the future fertility of women. Tables (B.1) in the Supplementary Appendix show the results of using indicators for the gender of children when a mother has two children with the same gender, i.e. having two boys might impact the probability of having further children differently than having two girls. The results suggest that the effect of having two boys is similar to the effect of having two girls. Again, mothers who have then children of the same gender are about 12% more likely to have further children by the end of their childbearing age.[3]

Moreover, in the dataset that I use, the individuals are explicitly asked whether they want to have more children in future. Hence, there is data on individuals' beliefs concerning their future fertility plans. Restricting the sample to mothers with exactly two children, it can be seen that those with two children of the same sex are more likely to show interest in having more children, and this is significantly different from mothers of children with different genders. These results are reported in Tables (2.3). Similar to the ex-post analysis in Table (2.2), I use linear probability, probit, and logit models to investigate the effects of child gender on ex-ante fertility decisions.[4] The results from these models are similar to the ex-post analysis. However, the coefficients of interest are bigger in ex-ante case.

The results in Tables (2.2) and (2.3) show that I can confirm the first assumption based on both individuals' actual fertility behavior and their beliefs regarding their future fertility plans.

As the fertility rate has declined over the past four decades, one might expect that nowadays, not many people decide to have more than two children. To address this concern, I look at the distribution of family sizes in the US over the past four decades. Figure (B.2) in the Supplementary Appendix shows that in 1976, 65% of mothers age

[3]Figure (B.1) in the Supplementary Appendix shows how the marginal effect of the gender mix of children on future fertility decisions has changed over the past few decades for different cohorts. All point estimates are significantly different from zero. Overall, mothers who have two children of the same gender are 8 to 13% more likely to have further children by the end of their childbearing age, compared to mothers with children of different genders.

[4]Table (B.2) in the Supplementary Appendix shows the results from using different indicators for mothers with two boys and mothers with two girls. The results are similar to Table (2.3).

40 to 44 had at least three children, while in 1994 and 2014 that became 36% and 38%, respectively. The highest frequency of children in 1976 was four, followed by three, two, and one, and the median number of children was three. However, in the mid-nineties the median number of children was two and the highest frequency is two children, followed by three, one and four children. This pattern has remained stable thereafter. Hence, it seems that four decades ago, the marginal mother was deciding between having three or more children, while nowadays, the marginal mother is deciding between having two or three children. Therefore, the gender mix of children might play a larger role in determining the likelihood of rearing a third child.

Assuming that parents have no control over the gender of their child, I exploit exogenous variation in the gender mix of children to examine the effect of the likelihood of future fertility on labor market outcomes.[5]

2.3.1 Regression Analysis

I use a Mincerian wage equation in order to investigate the impact of children gender mix on mothers' hourly wages:

$$log(w_{it}) = \beta_0 + \beta_1 pot_{it} + \beta_2 pot_{it}^2 + \beta_3 d_{it} + \beta_4 same_i + \epsilon_{it}, \tag{2.1}$$

where w_{it} denotes the hourly wage of individual i at time t; pot_{it} is their potential experience; d_{it} is the education level measured as completed years of schooling; $same_i$ is an indicator variable that is one if the first two children are of the same gender, and zero otherwise; and ϵ_{it} is a zero conditional mean error term, which is assumed to be i.i.d., normally distributed over individuals, and time-invariant.[6] I control for year fixed-effects, race, etc. by adding dummy variables to the regression specified

[5]It is unlikely that they deliberately terminate pregnancies using abortion when the gender of their child is realized. This is because abortion might have negative impacts on both their health and their future fertility outcomes.

[6]The potential experience is defined as age minus years of schooling minus six.

above. The coefficient of interest is β_4, which captures the effect of children with the same sex on their mother's wages.

I also use a similar setting to examine the impact of the gender mix of children on other aspects of mothers' labor market outcomes, including labor force participation, labor income, and hours of market work. The results are reported in the next subsection.

2.3.2 Results

Table (2.4) reports the regression results for Equation (2.1). The criteria used for sample selection are as follows: I only use observations on women. The sample is restricted to women in PSID who are of childbearing age (between ages 21 and 40). I also restrict the analysis to the subsample of women who already have two children. Then, the sample is selected based on the following time restrictions: at least two years after the second childbirth and before a third childbirth. Hence, all observations on wages are from women with two children where the age of the youngest child is greater than two. These individuals might end up having two or three children, but I only use observations corresponding to periods in which only two children exist.[7] I also focus on women with greater than high school education as these women are working in positions that need some skills and doing tasks that require some knowledge, training, and human capital. These individuals are more likely to improve their skills through on-the-job training and investing in their human capital, the better to get promoted. The dependent variable is the natural logarithm of hourly wage. I control for age effects using age-specific dummy variables. I also control for year fixed-effects in all regressions. Moreover, I control for the number of children that they will have by age

[7]This is because I do not want the coefficient of the gender mix of children to pick up the effects of a third child on female wages.

40 as this number might reveal some information about their types. The coefficient of interest, i.e. the coefficient on $same_i$, is negative, statistically significant, and sizable, -0.04. This coefficient does not change significantly as I control for additional confounding factors. In column 2, the state fixed effects are controlled for. Column 3 reports the results when I control for occupation and industry fixed effects as well. Hours of market work and ages at births are also controlled as confounding factors in columns 4 and 5, respectively. As column 5 shows, after controlling for all confounding factors, the coefficient of interest is -0.05, statistically significant at the five percent level. This result is economically significant as the order of magnitude of the effect is comparable to the return to one more additional year of schooling at college, i.e, 0.07 according to the results in column 5.

Table (B.3) in the Supplementary Appendix shows the results when I do not control for the number of children that each individual ultimately will have. These results are similar in magnitude to the results in Table (2.4), which might suggest that women who will have a third child do not know this fact beforehand with certainty. In other words, having a third child is probabilistic. However, mothers who have two children of the same gender face a higher probability of conceiving a third child. I also provide support for the hypothesis that a third child is probabilistic by restricting the sample to individuals who end up having only two children as of age 40. Table (2.5) reports the results, which still shows the gender mix effect.

Table (B.4) in the Supplementary Appendix reports the results when using different indicators for having two boys versus having two girls. It seems that the effects of these gender combinations are not significantly different.

Table (B.5) in the Supplementary Appendix reports the results when I select the sample based on a narrower range of childbearing ages, i.e. individuals between age 21 and 35. In this case, the coefficient of interest increases (in absolute value) to

-0.06, after controlling for all confounding factors in the last column. It is worth mentioning that the coefficient is more precisely estimated in this case, and it is statistically significant at the 1% level in all columns.

As explained earlier, in the dataset that I use, the individuals are explicitly asked whether they want to have more children in future. I use the data on individuals' beliefs concerning their future fertility plans along with their hourly wages to find the impact of uncertainties about future fertility on hourly wages of these women. In the survey, women are asked about their future fertility as follows: "How sure are you that you will not have any (more) children? (1) sure or very sure, (2) fairly sure or hope not to, (3) will have a child/not sure/do not know". I define dummy variables corresponding to answers to this question regarding future fertility. Then, I use a regression similar to equation (2.1), in which, instead of using gender mix of children, I use dummy variables that indicate individual's uncertainties about her future fertility behavior. Table (2.6) shows the results. The benchmark in the regression, i.e. omitted dummy variable, is when the individual has mentioned that she will have a child in future, or she is not sure, or she does not know about her future fertility plans (answer (3) to the question above). It turns out that after controlling for the education level, race, year fixed-effect, state fixed-effect, and hours of market work, individuals who are certain that they will not have any (more) children, on average, have hourly wages 7% higher than those who are uncertain about having a child in future. It is also interesting that individuals who are fairly sure that they will not have a child in future, on average, have hourly wages 5% higher than those who are uncertain about having a child in future. The coefficient for the latter group is not significant however, which is probably due to small sample size. This finding lends support to my previously-stated hypothesis that females, while investing in their own human capital, might take their uncertainty about future fertility into account. Human capital differential across females with different degrees of uncertainty is reflected in their hourly wages.

2.3.3 Robustness Checks

Table (B.6) in the Supplementary Appendix shows the regression results of Equation (2.1) with the same sample selection criteria used for regression in Table (2.4), except for one difference: I also include observations corresponding to periods with three children while controlling for the number of children. The results are similar to the previous results.

Table (B.7) in the Supplementary Appendix shows the regression results of Equation (2.1) when wage level, instead of log wage, is used as the dependent variable. The coefficient of interest is still significant at the five percent level and is similar in magnitude to those suggested by the regressions in Table (2.4).

Table (B.8) in the Supplementary Appendix shows the regression results of Equation (2.1) with the same sample selection criteria used for regression in Table (2.4), except for one difference: I restrict the sample to males instead of females. In this case, the coefficient of interest is almost zero in all regressions no matter which confounding factors are controlled for. This is consistent with my hypothesis regarding the effect of fertility plans and the ex-ante effects of children on labor market outcomes, as we know that females and not males are those usually affected in terms of their labor market outcomes when rearing children.

I also have done a robustness check with respect to the choice of dataset: I use the NLSY79 and NLSY97 datasets as alternatives and repeat the analysis. The results of this set of robustness checks are presented in the Supplementary Appendix.

As another exercise, Table (B.9) in the Supplementary Appendix shows the regression results of Equation (2.1) with the same sample selection criteria used for regression in Table (2.4), except for one difference: I restrict the sample to women who completed their fertility period with two children. The results suggest that the gender mix of children plays no role in terms of wages when uncertainty about future fertility is resolved.

Then, I look for some potential explanations for the negative association between the hourly wages of mothers and having children of the same gender. I look at other aspects of the mothers' labor market decisions to see if there is a significant difference between mothers with children of the same gender and those of different genders. Table (B.10) in the Supplementary Appendix, calculated using a probit model, suggests that there is no difference between the two groups in labor force participation at the extensive margin, i.e. whether to work or not. Moreover, Table (B.11) in the Supplementary Appendix shows that, conditional on being working, there is no significant difference between the two groups at the intensive margin, i.e. hours of market work. It seems that when it comes to decide on whether to work and how much to work, the gender mix of children plays no role. Also, Table (B.12) in the Supplementary Appendix provides evidence that there is no significant difference between the two groups in terms of the hours they spend with their children. Hence, the wage differential across the two groups cannot be attributed to different decisions on labor force participation at the extensive or intensive margins, and it also cannot be attributed to how they spend their time while at home.

2.4 Model

Polachek *et al.* (2008) provides a theoretical framework in which anticipated discontinuous labor force participation affects individuals' decisions on human capital investment (see also Weiss and Gronau (1981)). In the case of an anticipated discontinuous labor force participation, the present value of the marginal gain from a unit of investment is initially lower. Moreover, investments do not need to decline monotonically for the discontinuous worker. This implies that discontinuous workers not only invest less over their lifetime, but also their investment need not monotonically decline.

The objective function of each individual is as follows:

$$\max J = \int_0^T [N_t - s_t] w(K_t) K_t e^{-rt} dt$$

subject to the following constraint on the rate of change of capital stock:

$$\dot{K} = Q_t - \delta(t, K) K_t = f[s_t, k_t, x_t] - \delta(t, K_t) K_t = b_0 s_t^{b_1} K_t^{b_2} X_t^{b_3} - \delta K_t$$

where K_t denotes the amount of human capital at time t, $w(K_t)$ is the wage rate per unit of human capital, s_t denotes the percent of total available time spent investing in human capital in period t, N_t is the percent of total available time spent in labor force participation including the investment time, r_t is the sum of the rate of discount and rate of depreciation of human capital stock, x_t denotes the goods used in the production of human capital, and $\delta(t, K_t)$ is the annual rate of depreciation of capital stock.

The Hamiltonian equation for this problem are as follows:

$$H = w_0 (N_t - s_t) e^{-rt} K_t + \lambda b_0 s_t^{b_1} K_t^{b_1}$$

The first order conditions for this problem is as follows:

$$\frac{\partial H}{\partial s_t} = -w_0 K_t e^{-rt} + \lambda b_0 b_1 s_t^{b_1 - 1} k_t^{b_1} = 0,$$

$$\dot{\lambda} = -w_0 (N_t - S_t) e^{-rt} - \lambda b_0 b_1 s_t^{b_1} k_t^{b_1 - 1},$$

$$\dot{K} = Q_t = b_0 s_t^{b_1} K_t^{b_1}.$$

It is worth mentioning that λ_t represents the marginal returns on human capital investment in period t, and it can be shown that:

$$\lambda = [\frac{w_0}{b_0 b_1}] K_t^{1-b_1} s_t^{1-b_1} e^{-rt},$$

$$\dot{\lambda} = -w_0 N_t e^{-rt} \leq 0,$$

$$\lambda(t_0) = \int_0^T w_0 N_\tau e^{-r\tau} d\tau,$$

$$\lambda(t) = \int_t^T w_0 N_\tau e^{-rt} d\tau.$$

In the continuous participation case, I would have: $N(\tau) = N(t)$ for all $t \leq \tau \leq T$. However, when there is an anticipated discontinuous participation, $N(t)$ is changing over time, and the marginal benefit of investment in human capital is not monotonically decreasing. Hence, if females with two children of the same gender expect a higher likelihood of a further pregnancy, and so higher unemployment, then their investment pattern could be different over their entire remaining life cycle from mothers with two children of opposite genders, even before any actual childbearing.

In the next section, I use a Difference-In-Differences (DID) approach to show (1) how the labor market outcomes of the mother will be affected by a third childbirth in terms of labor force participation (at both intensive and extensive margins) and hourly wages, and (2) how these anticipated changes in labor market outcomes following a future third childbirth affect a woman's human capital investments and wages today, i.e. *before* a third childbirth is actually realized.

2.5 Empirical Approach

Following Kleven *et al.* (2018), I use an event study approach to identify post-child effects after nonparametrically controlling for education level, time trends, and age effects. I denote by t = 0 the year in which the individual has her third child and index all years relative to that year.

I follow individuals from four years before childbirth through ten years after. I study labor market outcomes as a function of event time.

$$Y_{ist} = \sum_{j \neq -2} \alpha_j.\boldsymbol{I}[j = t] + \sum_{k} \beta_k.\boldsymbol{I}[k = age_{is}] + \sum_{y} \gamma_y.I[y = s] + \sum_{e} \theta_e.I[e = d] + v_{ist}, \tag{2.2}$$

where Y_{ist} is the labor market outcome of individual i at year s and at event time t. Parameter α_j, i.e. the event time coefficient, captures the effect of the third childbirth on labor market outcomes in the *j-1* year after the third pregnancy, β_k captures the age fixed effect when age is equal to k, β_k represents the fixed effect of year k, and θ_e's parameters capture the effects of education level. I control for education level, d, using dummy variables. Finally, v_{ist} is the zero conditional mean error term. Note that the event time coefficients measure the impact of a third childbirth relative to the year just before pregnancy.

2.6 Results

I use Equation (2.2) to find the impact of the third childbirth on a variety of labor market outcomes. The results are reported in the following subsection.

2.6.1 Impacts of Third Childbirth on Labor Market Outcomes

Figure (2.1) depicts the impact on labor income of conceiving a third child. It is clear that labor income tends to decrease during pregnancy and right after childbirth, and it takes about five years for it to converge to the level before the pregnancy. The reduction in labor income is huge; about 40% decrease in labor income relative to the year before the pregnancy occurs right after childbirth.[8] To identify the driving forces behind this reduction in labor income, I look at the changes in labor force participation, hours worked, and hourly wages. Figure (2.2) shows the impact of the third childbirth on labor market participation at the extensive margin. The results suggest that the third childbirth is associated with a decrease of over 15 percentage points in the probability of working right after childbirth, relative to the year just before the pregnancy, which is equivalent to about 25% reduction in labor force participation.[9] Moreover, if I restrict my attention to the intensive margin by looking at women who work after childbirth or during pregnancy, it is clear that these individuals work fewer relative to the years before pregnancy. As Figure (2.3) shows, the third childbirth is associated with a reduction of about seven hours in market work per week (about 20% decrease relative to the baseline) for individuals who participated in the labor force, and it takes about seven years for them to converge to their previous levels.[10] Finally, Figure (2.4) shows that the third childbirth is associated with a decrease in hourly wage rate that tends to be both sizable (between 10 to 20%) and persistent.[11].[12]

[8]I excluded the observations corresponding to periods after a fourth childbirth. So my estimates are not contaminated by introducing the effects of further childbearing decisions.

[9]I excluded the observations corresponding to periods after a fourth childbirth so that my estimates are not contaminated by introducing the effects of further childbearing decisions.

[10]Again, I excluded the observations corresponding to periods after a fourth childbirth. So my estimates are not contaminated by introducing the effects of further childbearing decisions.

[11]Again, I excluded the observations corresponding to periods after a fourth childbirth so that my estimates are not contaminated by introducing the effects of further childbearing decisions.

[12]These reductions in labor force participation at the intensive and extensive margins, and in hourly wages following childbirth might be due to discrimination against mothers with infants, or it might be the case that mothers with a new-born child optimally decide to drop out of the labor market or reduce the hours of market work to take care of their child. I do not take a stand on any side here. I only assume that they have perfect foresight over these changes in labor market outcomes following the

2.6.2 DID approach

In order to get more precise estimates of the impact of third childbirth on the labor market outcomes of mothers, I exploit a difference-in-differences (DID) event study design that uses women who never conceived a third child. Following Kleven *et al.* (2018), I assign a placebo third childbirth to women who did not actually rear a third child, drawing from the distribution of age at third childbirth among individuals who did have a third one.[13] I follow a DID approach in which I treat those individuals with an assigned placebo third childbirth as the control group, and the women with a third child as the treatment group. Then, using the event study approach, I investigate the effects of the third childbirth on mothers' labor market outcomes by comparing the changes in the variable of interest before and after the childbirth across the treatment and control group.[14] Figure (2.5) depicts labor income trends for the control group and the treatment group before and after the third childbirth.[15] The pre-treatment trend is not common between the two groups, so the parallel trend assumption is violated. Figure (2.6) depicts labor force participation rates for the control group and the treatment group before and after the third childbirth. In this, the control group and the treatment group share a common pre-treatment group.[16] The results suggest that bearing a third child is associated with a decrease of over 20 percentage points in the labor force participation rate, which is a 30% decrease from the baseline.

---

third childbirth.

[13] I repeat this process 50 times.

[14] It is similar to what was described in the previous section; however, I do not control for age effects, year fixed effects, or education level. This is because I take advantage of using the DID approach; by differencing between the control group and the treatment group, those fixed effects will be taken out (as long as parallel pre-treatment trends exist).

[15] Figure (B.3) in the Supplementary Appendix shows the difference between the control and the treatment group.

[16] Figure (B.4) in the Supplementary Appendix shows the difference between control and treatment groups.

Figure (2.7) shows that the parallel pre-treatment trend between the treatment group and the control group is maintained when the outcome variable is hours of market work conditional on having worked.[17] This suggests that bearing a third child is associated with a decrease of over ten hours' reduction in hours worked, which is about a 30% decrease. Finally, Figure (2.8) shows the results when the outcome variable is log hourly wage. Obviously, the pre-treatment trends are different across groups, and it is not clear from these figures how wage rate changes with the third childbirth[18].[19]

2.6.3 Gender Mix of Children and Mothers' Earnings

In this section, I follow the event study approach described using Equation (2.1) to investigate the effects of the gender mix of children on the mothers' labor market outcomes, with modification. Here, I define $t = 0$ as the year in which the individual has her second child and index all years relative to that year. Hence, I study the outcome variables before and after the second childbirth, i.e. when the gender mix of children is determined by the second childbirth. In fact, I assume that the gender of the second child is exogenously determined, i.e. it is a quasi-random event. As soon as the second child is born, the gender mix of children is realized (0 if the gender of the second child is different from that of the first one, or 1 if the genders of the two children match). I consider the gender mix of children as the treatment, mothers with children of different genders as the control group, and mothers of children of the same sex as the treatment group. Then, I investigate the effects of the quasi-random

[17] Figure (B.5) in the Supplementary Appendix shows the difference between control and the treatment group.

[18] Figure (B.6) in the Supplementary Appendix shows the difference between control and the treatment group.

[19] Based on what is said, one can conclude that the parallel trend assumption in labor income analysis is violated due to there being different trends in wage rates between the control group and the treatment group.

treatment assignment on the treatment group.

Figure (2.9) depicts the results when the outcome variable is annual labor income, which suggest that the treatment group tends to have lower labor income in the post-treatment period. In order to identify the driving forces behind this pattern, I look at the effect of treatment assignment on labor force participation, hours of market work, and the wage rate.

With respect to attachment to the labor market, Figure (2.10) shows the results when I focus on the effect of treatment on labor force participation. These results suggest that the treatment has no significant effect on labor force participation. Also, Figure (2.11) depicts the effect of treatment on hours worked; it suggests that there is no significant effect. Overall, Figures (2.10) and (2.11) indicate that my treatment does not affect the extent to which the treated individuals are attached to the labor market.

Finally, Figure (2.12) shows the results when the outcome variable is log hourly wages. While the wage rates are not significantly different between the control group and the treatment group in the pre-treatment period, they are significantly different in the post-treatment period. Mothers of children with the same sex tend to have lower hourly wages. Note that in all analyses in this subsection, I exclude observations after a third child is born. Hence, the wage differential between treatment and control groups cannot be attributed to a difference in the number of children across the two groups.

This result confirms the regression results presented earlier in Tables (2.4) and (2.5). The wage differential is much more evident if I restrict the sample to individuals with at least a high school degree. Figure (2.13) depicts this case. Mothers of children with the same gender earn significantly less per hour of market work than those with two children of different genders. This result lends support to my hypothesis regarding the effect of future fertility decisions on investment in human capital; as explained before, I expect that the link between human capital and hourly wages is stronger for those tasks that need knowledge, skill, and experience; and highly educated women are more likely to work at positions that necessitate such tasks.

2.7 Implications for Gender Wage Gap

The above analysis suggests that the ex-ante effects of children on mothers' wages might be an important factor in the wage gap between men and women. Previous literature has mostly have focused on the ex-post effects of children on mothers' labor market outcomes. However, the evidence provided in this paper lends support to the hypothesis that women's perception of their likelihood of future fertility might significantly contribute to the gender wage gap through decreasing incentives to invest in their human capital.

Given the hypothesis, I expect the gender wage gap, driven by uncertainty about future fertility, to be higher for individuals at prime childbearing ages. At the end of the fertility period, the likelihood of future fertility goes down to zero,, the uncertainty about future futility is resolved, and the degree of attachment to the labor market increases, leading these individuals to invest more in their human capital; this can reduce the gender wage gap. I test this hypothesis using separate linear regressions for young workers (21 to 35 years old) and old workers (36 years or older). Tables (2.7) and (2.8) show the results. While the gender wage gap in the first group is between 22% and 24%, it decreases to between 18% and 19% for those who have completed their fertility phase and resolved the uncertainty about future fertility.

2.8 Conclusion

In this paper, I documented the new finding of a statistically significant relationship between the gender mix of children and their mother's wage. Specifically, having two children of the same gender (two boys or two girls) is associated with between 4% and 6% lower hourly wages for the mother compared to having two children of the opposite sex. My results passed various alternative specifications. I discussed several potential explanations that could explain the relationship I observed in the data.

Moreover, I used exogenous variation in the gender mix of children to study the effect of future fertility decisions on current investment in human capital, and thereby wage rates. I showed that if the two first children are of the same sex, then the likelihood of conceiving a third child is significantly higher, which is consistent with the findings of previous papers. This increase might be due to parental preference for variety in child genders. I found that right after a third childbirth labor force participation rate declines by about 30% compared to the year before the third pregnancy. Also, for those who continue to participate in the labor market after their third childbirth, hours of market work decrease by about 30%. Hence, conceiving a third child is associated with sizable reductions in labor force participation at both extensive and intensive margins, and the higher the probability of a third childbirth, the higher probability of reduced labor force participation in future. I used a theoretical framework to illustrate that this anticipated reduced attachment to the labor market, perceived by mothers with children of the same sex, leads them to invest less in their human capital using on-the-job training, compared to mothers of children of different genders; this lower investments in human capital will be reflected in their wage rates. My Difference-In-Differences analysis showed that the exogenous gender mix of children has a sizable impact on female wages per se, i.e. not only through the impacts on family size (the number of children) but also probably through different investments in human capital due to differential likelihood of future fertility.

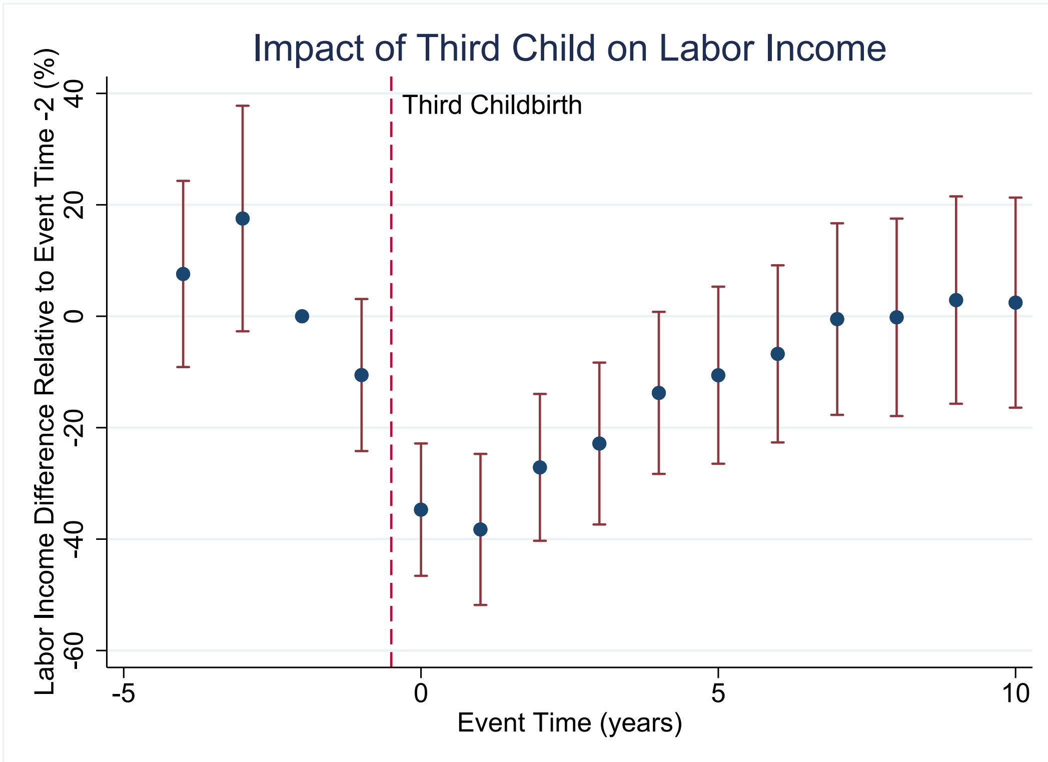
One important implication of the findings of this paper is that the ex-ante effects of fertility decisions on the gender wage gap are greater than previously thought. Previous papers have focused on the impact of motherhood on women's labor market outcomes while neglecting the impact of anticipated fertility decisions on human capital investments. The results of this paper suggest that women's expectation regarding their future fertility might play an important role in their human capital accumulation.

It is worth mentioning that since Angrist and Evans (1996), the gender mix of children has been widely used as an IV to study the effect of family size on labor market outcomes. However, the results of this paper cast doubt on the validity of using this IV in that context and in many other applications within the labor market literature.

Future work should address the welfare and policy implications of the findings of this paper. From a policy perspective, it is interesting to study how individuals' human capital investments might be affected by provision of a high quality publicly subsidized day care. Implementing a childcare subsidy program might help females to remain in the labor force after childbirth. Knowing that they can take advantage of childcare services for their child, women perceive themselves as being more attached to the labor market in the future, so they have incentive to invest more in their human capital long before any actual childbirth is realized. Moreover, the return to human capital investments potentially depends on education level. The impact of a subsidized childcare policy might be heterogeneous across women depending on their education levels. Hence, it would also be interesting to investigate how implementing a childcare policy might change women's educational decisions.

2.9 Figures

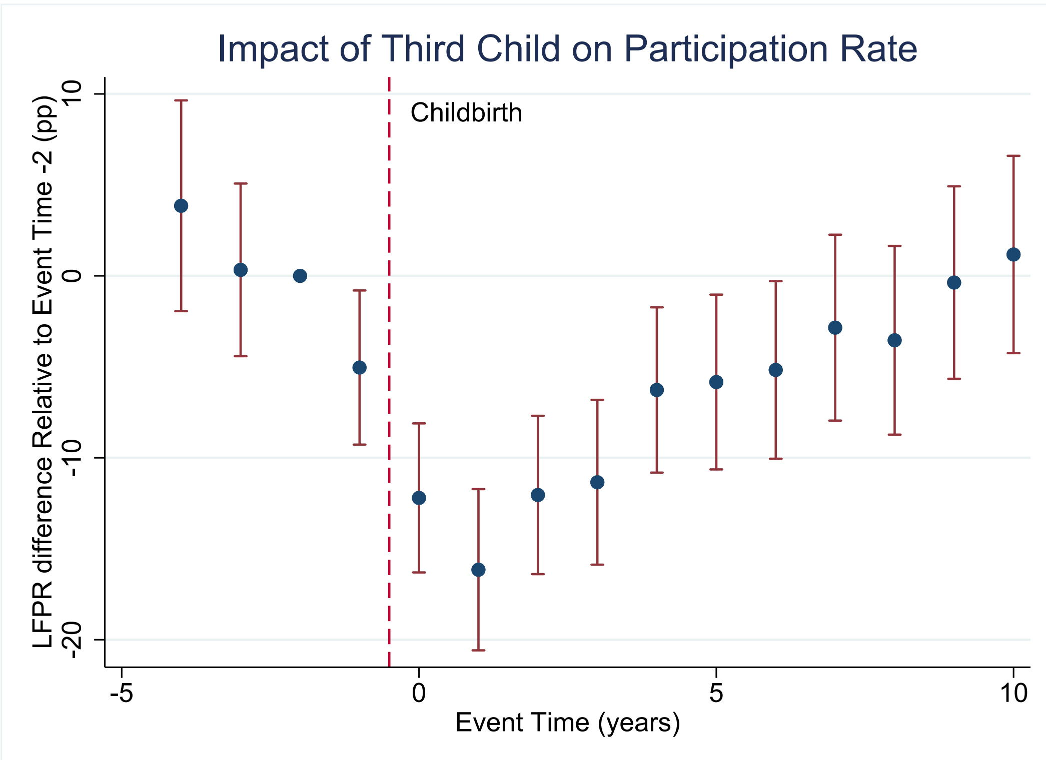
Figure 2.1: Arrival of a Third Child and Mother's Labor Income



Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the annual income. Cohort fixed-effect, age effect, race effect, and education effects are controlled using dummy variables.
Event time coefficients measure the impact of the third childbirth relative to the year just before pregnancy in terms of percentage changes.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

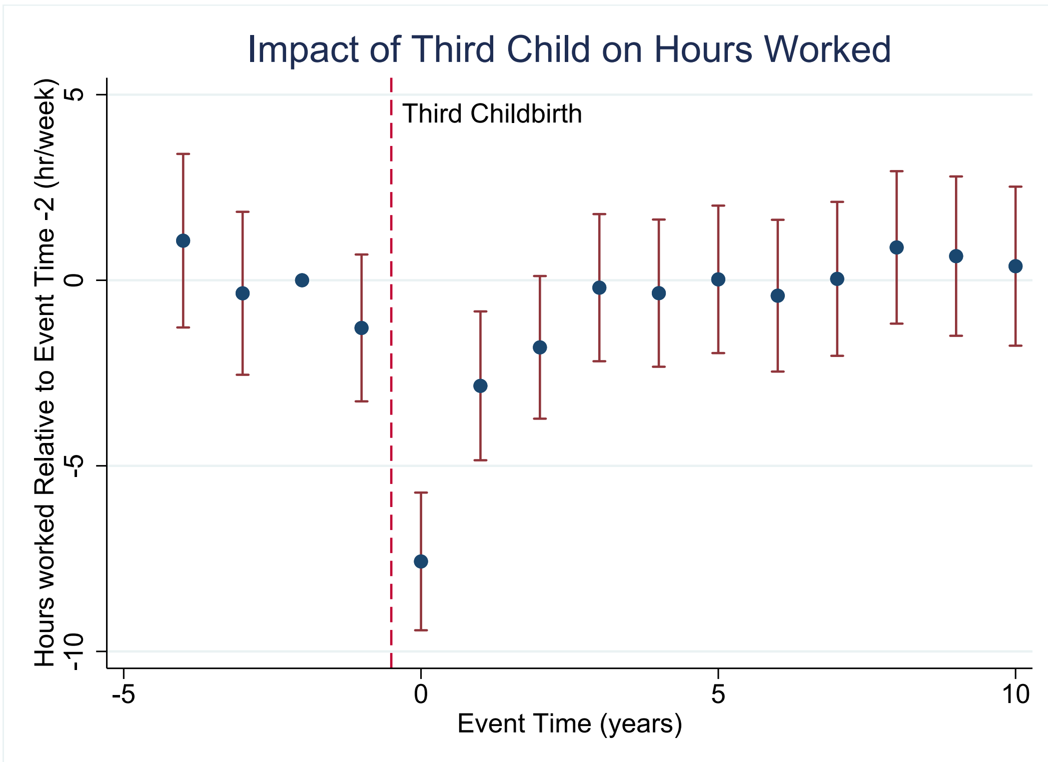
Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Figure 2.2: Arrival of a Third Child and Mother's Participation



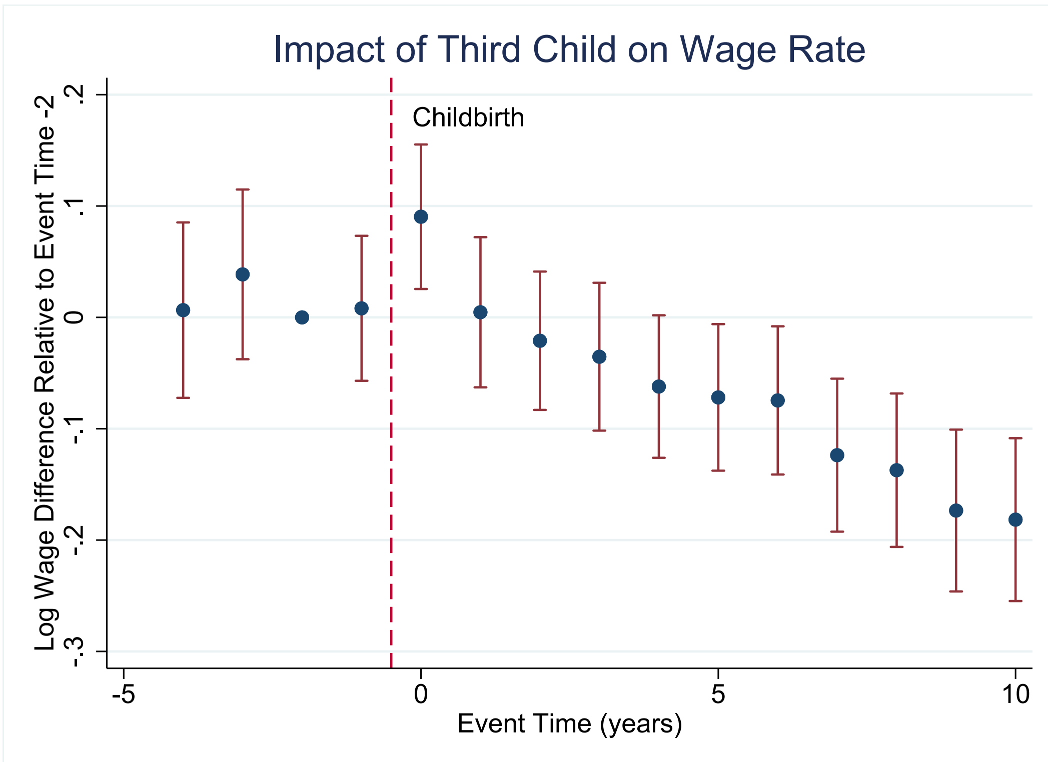
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the labor force participation status. Cohort fixed-effect, age effect, race effect, and education effects are controlled using dummy variables.
Event time coefficients measure the impact of the third childbirth relative to the year just before pregnancy in terms of percentage point changes.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.3: Arrival of a Third Child and Mother's Work Hours



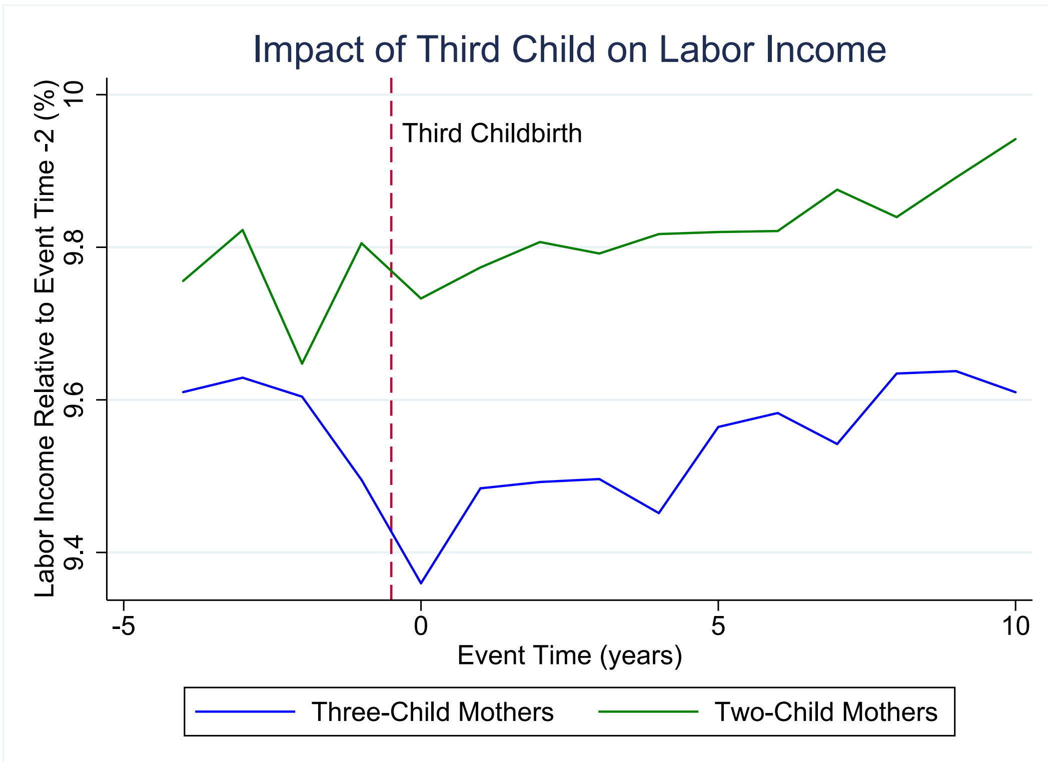
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the hours of market work conditional on participating in the labor market. Cohort fixed-effect, age effect, race effect, and education effects are controlled using dummy variables.
Event time coefficients measure the impact of the third childbirth relative to the year just before pregnancy in terms of hours of work per week.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.4: Arrival of a Third Child and Mother's Wage Rate



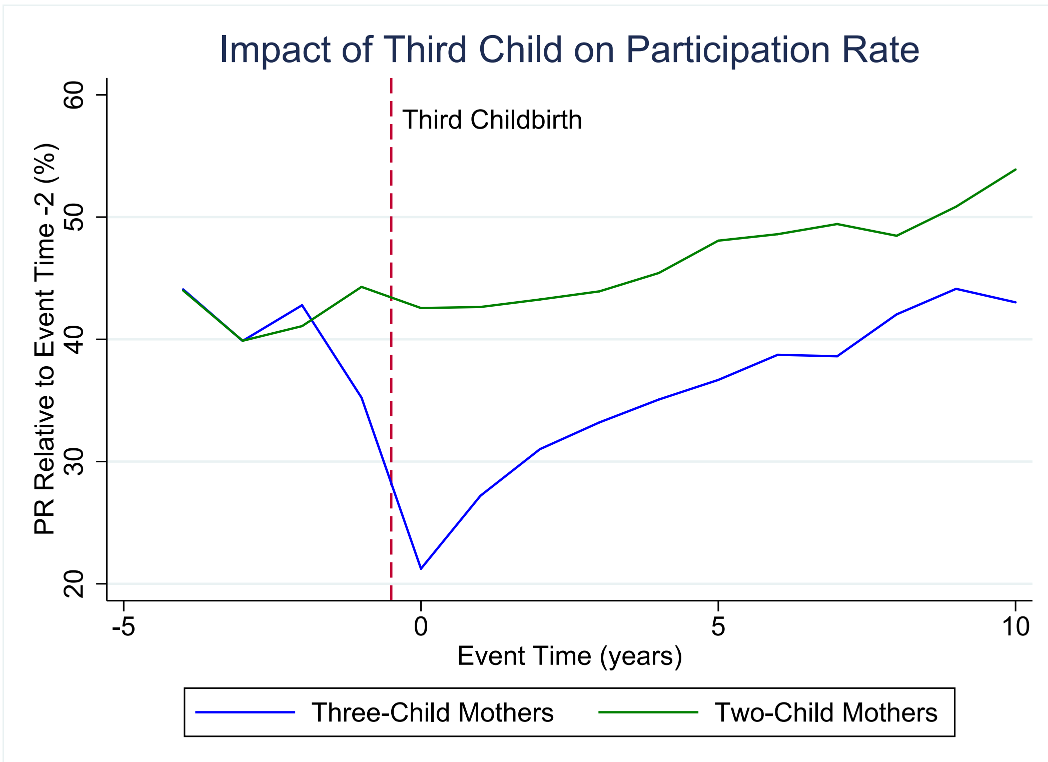
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the hourly wage conditional on participating in the labor market. Cohort fixed-effect, age effect, race effect, and education effects are controlled using dummy variables.
Event time coefficients measure the impact of the third childbirth relative to the year just before pregnancy in terms of fraction of changes.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.5: Arrival of a Third Child and Mother's Labor Income (Placebo Control)



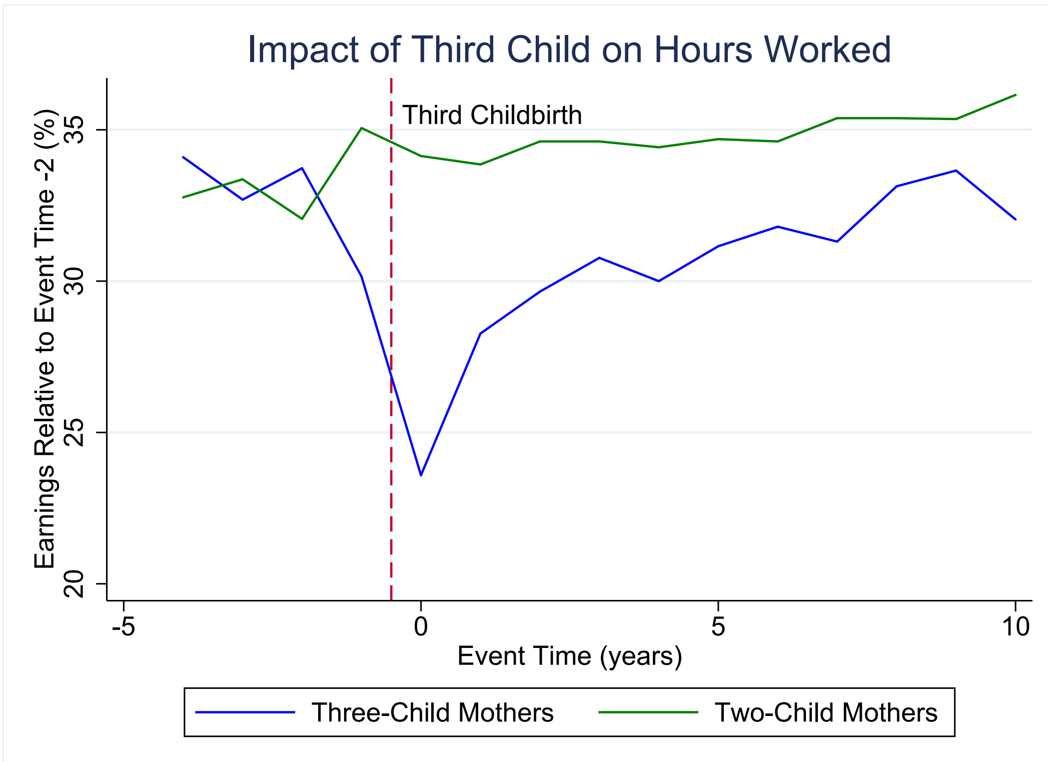
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the log annual income.
Here, I run the specification (2.1) separately for mothers with two children and mothers with three, while assigning placebo birth date for the third childbirth to the first group. The placebo birth time is randomly assigned based on the distribution of observed timing of third birth among women with three children.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.6: Arrival of a Third Child and Mother's Participation (Placebo Control)



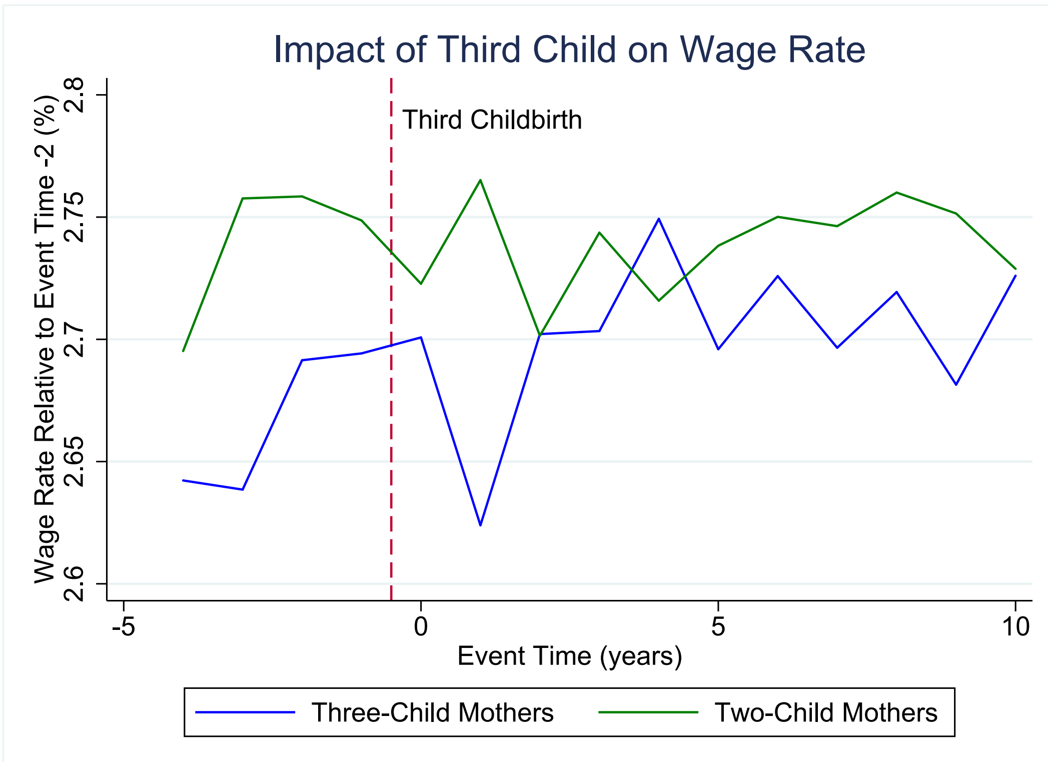
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the percentage of participants in the labor market.
Here, I run the specification (2.1) separately for mothers with two children and mothers with three, while assigning placebo birth date for the third childbirth to the first group. The placebo birth time is randomly assigned based on the distribution of observed timing of third birth among women with three children.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.7: Arrival of a Third Child and Mother's Work Hours (Placebo Control)



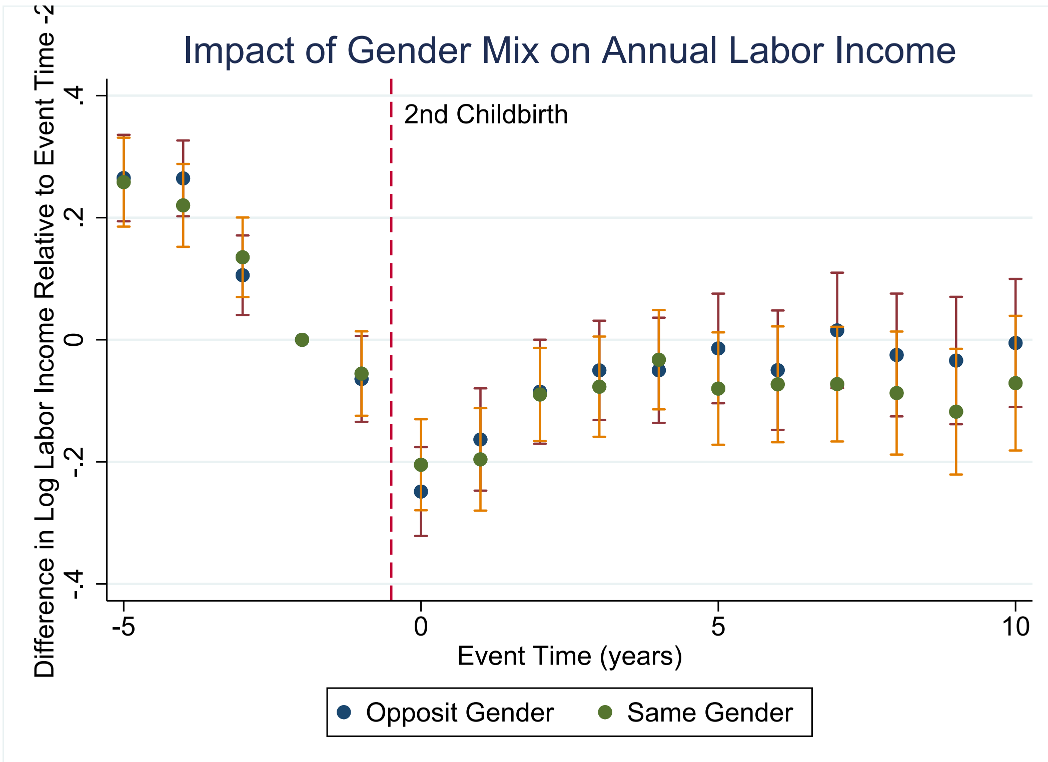
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the mean hours of market work per week conditional on participating in the labor market.
Here, I run the specification (2.1) separately for mothers with two children and mothers with three, while assigning placebo birth date for the third childbirth to the first group. The placebo birth time is randomly assigned based on the distribution of observed timing of third birth among women with three children.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.8: Arrival of a Third Child and Mother's Wage Rate (Placebo Control)



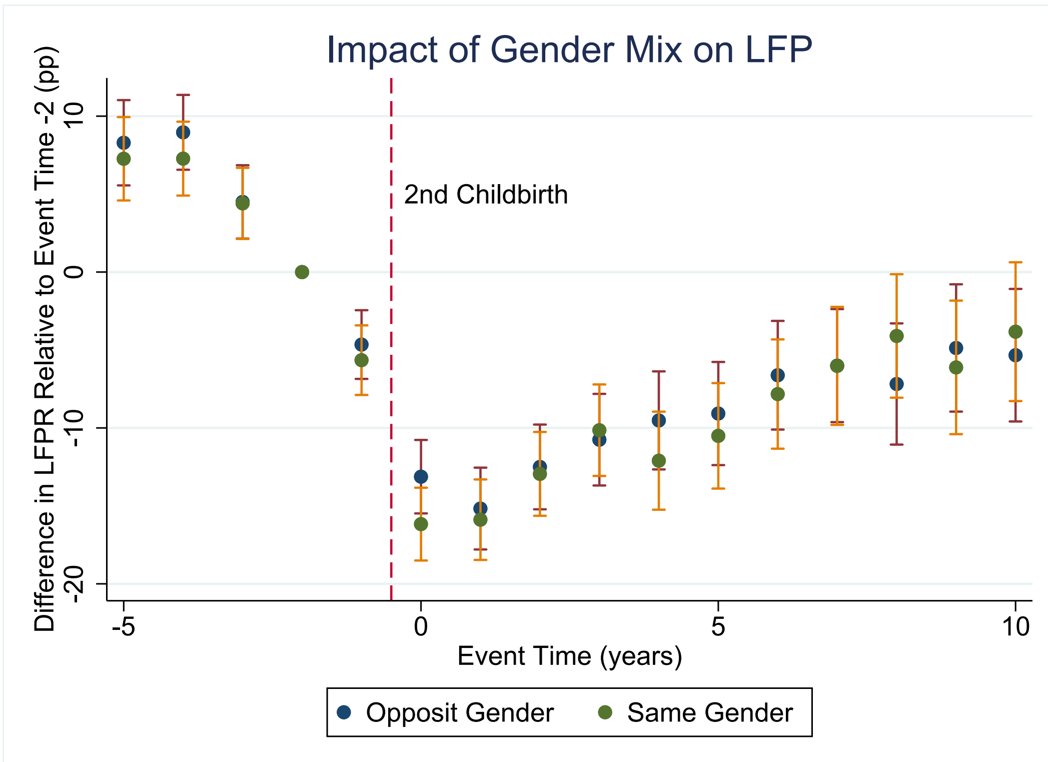
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the third childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the log hourly wage conditional on participating in the labor market.
Here, I run the specification (2.1) separately for mothers with two children and mothers with three, while assigning placebo birth date for the third childbirth to the first group. The placebo birth time is randomly assigned based on the distribution of observed timing of third birth among women with three children.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.9: Gender Mix of Children and Mother's Labor Income



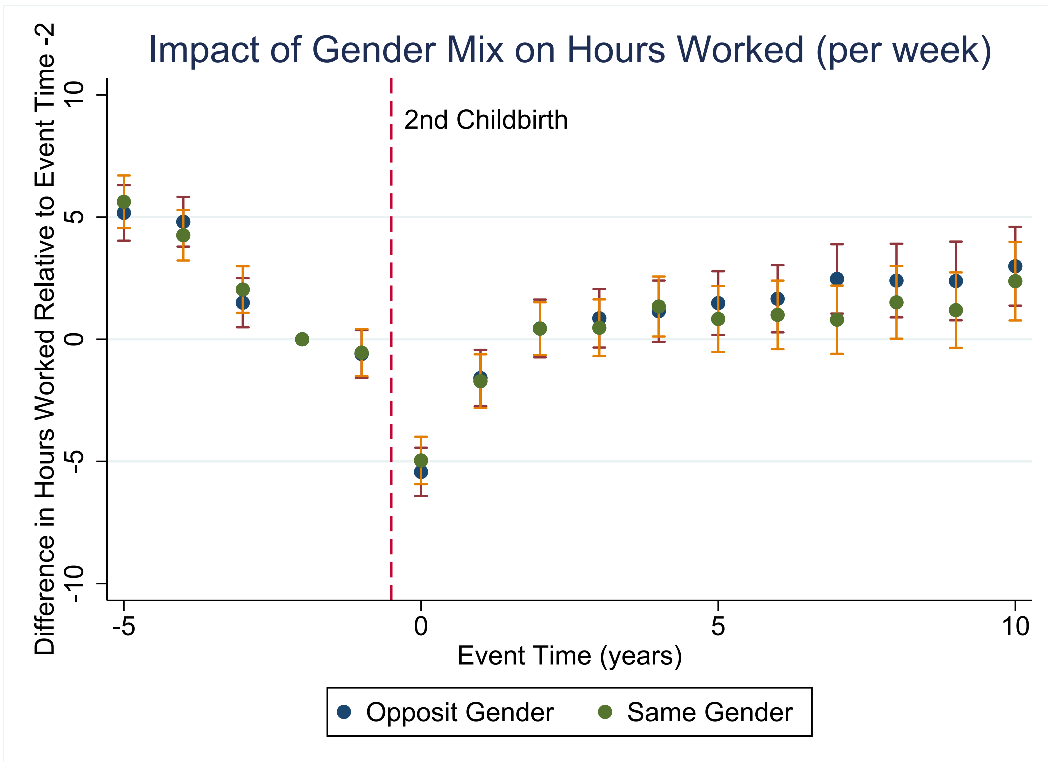
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the second childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the difference in log annual income relative to the year before the second pregnancy, conditional on participating in the labor market.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.10: Gender Mix of Children and Mother's Participation



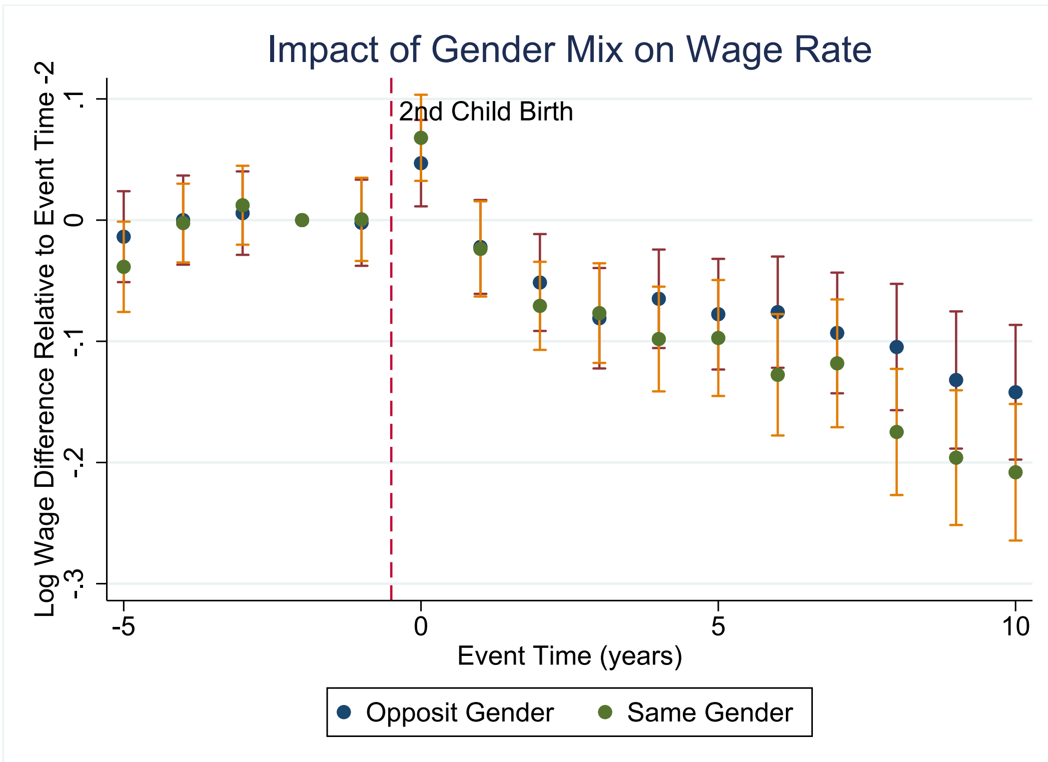
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the second childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the difference in the labor force participation rate relative to the year before the second pregnancy, conditional on participating in the labor market.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.11: Gender Mix of Children and Mother's Work Hours



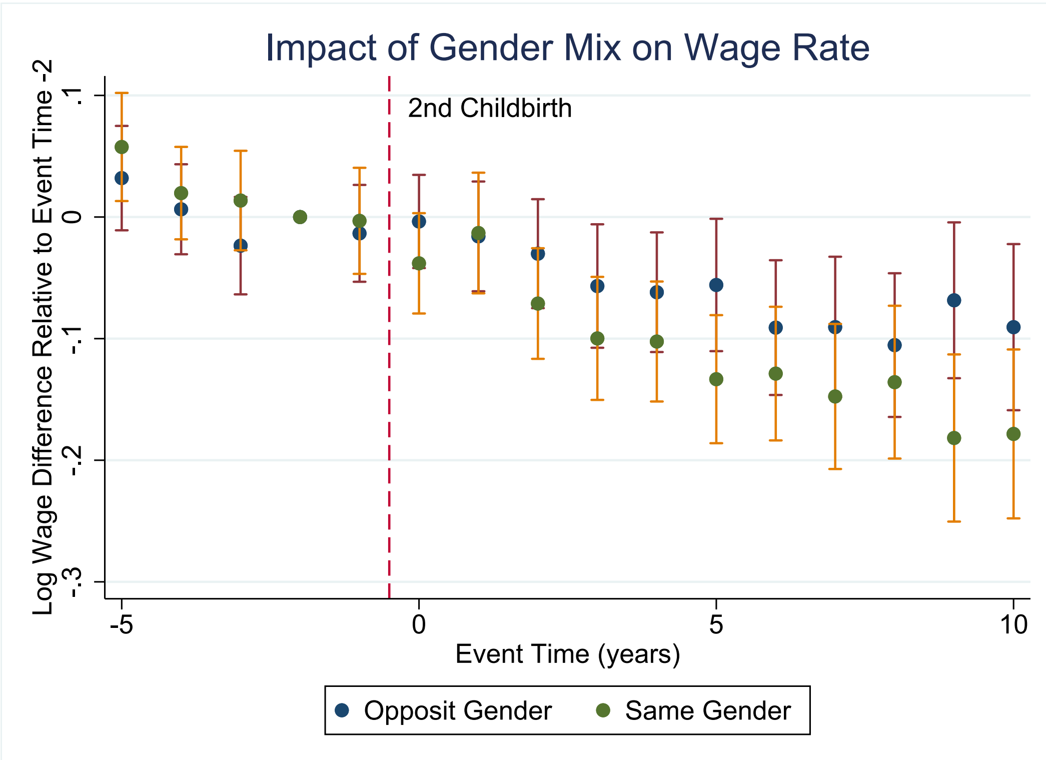
Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the second childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the difference in the hours of work (per week) relative to the year before the second pregnancy, conditional on participating in the labor market.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.12: Gender Mix of Children and Mother's Wage Rate



Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the second childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the difference in log hourly wage relative to the year before the second pregnancy, conditional on participating in the labor market.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Figure 2.13: Gender Mix of Children and Mother's Wage (High School Graduates)



Notes: This figure shows event time coefficients estimated from equation (2.1), 5 years before the second childbirth, and 10 years after.
The outcome of interest, i.e. the dependent variable, is the difference in log hourly wage relative to the year before the second pregnancy, conditional on participating in the labor market.
Sample is restricted to women who completed high school education.
The 95% confidence intervals, shown by vertical line around each point, are based on robust standard errors.

Table 2.1: Summary Statistics- Year 2015

Variable	Statistic	All	Same sexes	Opposite	p-value
Years of schooling	Mean:	13.13	13.14	13.12	0.66
	Median:	13.00	13.00	13.00	
	Std Dev:	(2.56)	(2.54)	(2.59)	
Labor force participation	Mean:	0.67	0.67	0.68	0.31
	Median:	1.00	1.00	1.00	
	Std Dev:	(0.47)	(0.47)	(0.47)	
Hours worked (per week)	Mean:	21.16	21.07	21.28	0.71
	Median:	22.88	22.62	23.00	
	Std Dev:	(18.38)	(18.38)	(18.39)	
Time with child (hr/week)	Mean:	28.83	27.70	29.83	0.25
	Median:	27.04	24.92	28.42	
	Std Dev:	(20.29)	(20.15)	(20.39)	
Hourly wage ($)	Mean:	16.85	16.98	16.71	0.56
	Median:	13.47	13.54	13.22	
	Std Dev:	(11.20)	(11.66)	(10.72)	
Annual Labor income ($)	Mean:	15.17	15.18	15.16	0.97
	Median:	9.85	10.06	9.15	
	Std Dev:	(17.78)	(17.89)	(17.66)	
Mother's age	Mean:	49.88	50.12	49.63	0.02
	Median:	51.00	51.00	51.00	
	Std Dev:	(13.25)	(13.16)	(13.33)	
Mother's age at 1st birth	Mean:	22.11	22.12	22.09	0.65
	Median:	21.00	21.00	21.00	
	Std Dev:	(4.63)	(4.63)	(4.62)	
Mother's age at 2nd birth	Mean:	25.62	25.64	25.61	0.74
	Median:	25.00	25.00	25.00	
	Std Dev:	(5.15)	(5.16)	(5.14)	
Mother's age at 3rd birth	Mean:	27.40	27.47	27.30	0.16
	Median:	27.00	27.00	27.00	
	Std Dev:	(5.19)	(5.22)	(5.15)	
Marital status	Mean:	0.64	0.63	0.65	0.09
	Median:	1.00	1.00	1.00	
	Std Dev:	(0.48)	(0.48)	(0.48)	
Number of children	Mean:	2.87	2.88	2.85	0.11
	Median:	3.00	3.00	2.00	
	Std Dev:	(1.15)	(1.13)	(1.18)	
Same genders	Mean:	0.51			
	Median:	1.00			
	Std Dev:	(0.50)			
Fraction with 3 children	Mean:	0.51	53	47	
Number of observations		14486	7365	7095	

Notes: Income and wages are normalized to 2000 dollar.
Sample restricted to females with at least two children at 2015.
Income and wages are normalized to 2000 dollar.
Incomes above $100000 are dropped. Hourly wages less than $7 are dropped.

Table 2.2: The Gender Mix of Children and (Ex Post) Fertility Behavior

	Third child (OLS) (1)	Third child (Probit) (2)	Third child (Logit) (3)
Same gender	0.06***	0.14***	0.23***
	(0.01)	(0.03)	(0.05)
Constant	0.81***	1.09**	1.89*
	(0.09)	(0.53)	(1.04)
Observations	7601	7601	7601
Adjusted R^2	0.04	0.03	0.03
Race	Y	Y	Y
Education Level	Y	Y	Y

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample have been restricted to individuals older than 40 years of old at 2014.
I also restrict the analysis to the subsample of women who have at least two children.

Table 2.3: The Gender Mix of Children and (Ex Ante) Fertility Decisions

	Third child (OLS) (1)	Third child (Probit) (2)	Third child (Logit) (3)
Same Gender	0.08***	0.35***	0.61***
	(0.03)	(0.12)	(0.22)
Constant	0.03	-2.62***	-4.88***
	(0.09)	(0.61)	(1.32)
Observations	1249	1249	1249
Adjusted R^2	0.11	0.15	0.15
Age	Y	Y	Y
Race	Y	Y	Y
Education	Y	Y	Y
Year	Y	Y	Y

The question is asked from females in PSID between 1969 to 1972 & 1976.

Dependent variables take 1 if the individual expects more children, 0 otherwise.

The Sample is restricted to women in PSID between age 21 and 40 with 2 children.

Robust standard errors are reported in parentheses in all specifications.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.4: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.08***	0.08***	0.07***	0.07***	0.07***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Potential experience square	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.09***	0.09***	0.06***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same sex	**-0.04****	**-0.04****	**-0.04****	**-0.04****	**-0.05****
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	1.10***	0.90***	1.55***	1.62***	1.41***
	(0.25)	(0.26)	(0.28)	(0.29)	(0.32)
Observations	4871	4753	4561	4561	4540
Adjusted R^2	0.18	0.23	0.30	0.31	0.33
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample is restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
I control for the number of children that each individual will have in all regressions.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.5: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.07***	0.06***	0.07***	0.06***	0.06***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Potential experience square	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.08***	0.08***	0.06***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same sex	**-0.04***	**-0.04***	**-0.04****	**-0.04***	**-0.05****
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	1.26***	1.19***	1.67***	1.70***	1.49***
	(0.19)	(0.21)	(0.28)	(0.29)	(0.31)
Observations	4501	4390	4218	4218	4198
Adjusted R^2	0.19	0.23	0.30	0.30	0.34
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample is restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to individuals who will end up having two children.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.6: Regression Results Using Beliefs about Future Fertility

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.060***	0.060***	0.060***	0.060***	0.057***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Potential experience square	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
1-No more child (sure/very sure)	**0.098*** **	**0.100*** **	**0.069***	**0.066***	**0.074** **
	(0.032)	**(0.031)**	**(0.037)**	**(0.037)**	**(0.035)**
2-No more child (fairly sure/hope not to)	**0.082**	**0.081**	**0.041**	**0.029**	**0.055**
	(0.061)	**(0.061)**	**(0.065)**	**(0.067)**	**(0.063)**
Constant	1.886***	1.808***	1.790***	1.686***	1.916***
	(0.050)	(0.105)	(0.106)	(0.131)	(0.246)
Observations	611	611	611	611	611
Adjusted R^2	0.31	0.31	0.31	0.32	0.38
Education	Y	Y	Y	Y	Y
Race	N	Y	Y	Y	Y
Year	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
State	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
This table shows the impact of uncertainty about future fertility on hourly wages of women.
Women are asked about their future fertility: "How sure are you that you will not have any (more) children?"
(1) sure or very sure, (2) fairly sure or hope not to, (3) will have a child/not sure/do not know.
The benchmark in the above regression, i.e. omitted dummy, is the last choice, i.e. (3).
The Sample is restricted to women in PSID younger than 40.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.7: Mincerian Wage Equation- OLS Results for Older Individuals

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)
Female	**-0.18***** **(0.03)**	**-0.18***** **(0.02)**	**-0.18***** **(0.03)**	**-0.19***** **(0.03)**
Potential experience	-0.08 (0.07)	-0.08 (0.06)	-0.04 (0.06)	-0.04 (0.06)
Potential experience square	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Years of schooling	0.03 (0.03)	0.03 (0.03)	0.02 (0.03)	0.02 (0.03)
Constant	3.20*** (1.24)	3.10*** (1.17)	2.49** (1.18)	2.54** (1.19)
Observations	4279	4174	3809	3809
Adjusted R^2	0.11	0.15	0.23	0.23
Age	Y	Y	Y	Y
Race	Y	Y	Y	Y
Year	Y	Y	Y	Y
Region	N	Y	Y	Y
Occupation	N	N	Y	Y
Hours worked	N	N	N	Y

Standard errors in parentheses

The potential experience is defined as age minus years of schooling minus six.

The Sample restricted to people older than 35 years of old in PSID.

Sample restricted to individuals whose education levels are higher than high school.

Sample restricted to individuals who work at least 20 hours per week.

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table 2.8: Mincerian Wage Equation- OLS Results for Younger Individuals

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)
Female	**-0.24***** **(0.01)**	**-0.23***** **(0.01)**	**-0.22***** **(0.01)**	**-0.24***** **(0.01)**
Potential experience	-0.00 (0.00)	-0.02*** (0.00)	-0.01*** (0.00)	-0.02*** (0.00)
Potential experience square	0.00 (0.00)	0.00*** (0.00)	0.00*** (0.00)	0.00*** (0.00)
Years of schooling	0.09*** (0.00)	0.09*** (0.01)	0.06*** (0.01)	0.06*** (0.01)
Constant	1.52*** (0.09)	1.53*** (0.10)	1.86*** (0.12)	2.06*** (0.12)
Observations	19772	19189	18140	18140
Adjusted R^2	0.18	0.21	0.27	0.28
Age	Y	Y	Y	Y
Race	Y	Y	Y	Y
Year	Y	Y	Y	Y
Region	N	Y	Y	Y
Occupation	N	N	Y	Y
Hours worked	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample restricted to individuals in PSID between age 21 and 35.
Sample restricted to individuals whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Adda, J., C. Dustmann, C. Meghir and J.-M. Robin, "Career progression, economic downturns, and skills", Working Paper 18832, National Bureau of Economic Research, URL `http://www.nber.org/papers/w18832` (2013).

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

Altonji, J. G. and L. M. Segal, "Small sample bias in gmm estimation of covariance structures", Working Paper 156, National Bureau of Economic Research, URL `http://www.nber.org/papers/t0156` (1994).

Angrist, J. D. and W. N. Evans, "Children and their parents' labor supply: Evidence from exogenous variation in family size", Tech. rep., National bureau of economic research (1996).

Attanasio, O., H. Low and V. Snchez-Marcos, "Explaining changes in female labor supply in a life-cycle model", American Economic Review **98**, 4, 1517–52, URL `http://www.aeaweb.org/articles?id=10.1257/aer.98.4.1517` (2008).

Bacharach, V. R. and A. A. Baumeister, "Effects of maternal intelligence, marital status, income, and home environment on cognitive development of low birthweight infants", Journal of Pediatric Psychology **23**, 3, 197–205 (1998).

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

Barclay, K. and M. Myrskylä, "Advanced maternal age and offspring outcomes: Reproductive aging and counterbalancing period trends", Population and development review **42**, 1, 69–94 (2016).

Barros, R., P. Olinto, T. Lunde and M. Carvalho, "The impact of access to free childcare on womens labor market outcomes: evidence from a randomized trial in low-income neighborhoods of rio de janeiro", in "World Bank Economists Forum", (2011).

Bauernschuster, S., T. Hener and H. Rainer, "Children of a (policy) revolution: The introduction of universal child care and its effect on fertility", Journal of the European Economic Association **14**, 4, 975–1005 (2016).

Baum, C. L., "The effects of maternity leave legislation on mothers' labor supply after childbirth", Southern Economic Journal pp. 772–799 (2003).

Baum, C. L. and C. J. Ruhm, "The effects of paid family leave in california on labor market outcomes", Journal of Policy Analysis and Management **35**, 2, 333–356 (2016).

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

Black, S. E., P. J. Devereux and K. G. Salvanes, “From the cradle to the labor market? the effect of birth weight on adult outcomes”, The Quarterly Journal of Economics **122**, 1, 409–439 (2007).

Blau, D. M., “The effect of income on child development”, The Review of Economics and Statistics **81**, 2, 261–276, URL `https://doi.org/10.1162/003465399558067` (1999).

Blau, F. D. and L. M. Kahn, “Swimming upstream: Trends in the gender wage differential in the 1980s”, Journal of labor Economics **15**, 1, Part 1, 1–42 (1997).

Blau, F. D. and L. M. Kahn, “The gender wage gap: Extent, trends, and explanations”, Journal of Economic Literature **55**, 3, 789–865, URL `http://www.aeaweb.org/articles?id=10.1257/jel.20160995` (2017).

Blundell, R., M. Costa Dias, C. Meghir and J. Shaw, “Female labor supply, human capital, and welfare reform”, Econometrica **84**, 5, 1705–1753, URL `https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11576` (2016a).

Blundell, R., M. Costa Dias, C. Meghir and J. Shaw, “Female labor supply, human capital, and welfare reform”, Econometrica **84**, 5, 1705–1753 (2016b).

Byrnes, G., “Maternal age and risk of type 1 diabetes in children: Flawed analysis invalidates conclusions”, BMJ: British Medical Journal **322**, 7300, 1489 (2001).

Cameron, A. C. and P. K. Trivedi, *Microeconometrics: methods and applications* (Cambridge university press, 2005).

Cardwell, C. R., L. C. Stene, G. Joner, M. K. Bulsara, O. Cinek, J. Rosenbauer, J. Ludvigsson, M. Jané, J. Svensson, M. J. Goldacre *et al.*, “Maternal age at birth and childhood type 1 diabetes: a pooled analysis of 30 observational studies”, Diabetes (2009).

Caucutt, E. M., N. Guner and J. Knowles, “Why do women wait? matching, wage inequality, and the incentives for fertility delay”, Review of Economic Dynamics **5**, 4, 815–855 (2002).

Chandra, A., G. M. Martinez, W. D. Mosher, J. C. Abma and J. Jones, “Fertility, family planning, and reproductive health of us women; data from the 2002 national survey of family growth”, (2005).

Chéron, A., J.-O. Hairault and F. Langot, “Life-cycle equilibrium unemployment”, Journal of Labor Economics **31**, 4, 843–882 (2013).

Chiappori, P.-A., B. Fortin and G. Lacroix, “Marriage market, divorce legislation, and household labor supply”, Journal of political Economy **110**, 1, 37–72 (2002).

Chiappori, P.-A., M. Iyigun and Y. Weiss, “Investment in schooling and the marriage market”, American Economic Review **99**, 5, 1689–1713, URL `http://www.aeaweb.org/articles?id=10.1257/aer.99.5.1689` (2009).

Conley, D. and N. G. Bennett, "Is biology destiny? birth weight and life chances", American Sociological Review pp. 458–467 (2000).

Cooper, K., K. Stewart *et al.*, "Does money affect childrens outcomes?", Tech. rep., Centre for Analysis of Social Exclusion, LSE (2013).

Cukrowska-Torzewska, E. and A. Lovasz, "Are children driving the gender wage gap? comparative evidence from poland and hungary", Economics of Transition **24**, 2, 259–297 (2016).

Cunha, F. and J. J. Heckman, "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation", The Journal of Human Resources **43**, 4, 738–782, URL `http://www.jstor.org/stable/40057370` (2008).

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

Dahl, G. B. and L. Lochner, "The impact of family income on child achievement: Evidence from the earned income tax credit", American Economic Review **102**, 5, 1927–56, URL `http://www.aeaweb.org/articles?id=10.1257/aer.102.5.1927` (2012).

Dascola, M., V. Freedman, N. Insolera, F. Pfeffer, K. McGonagle and N. Sastry, "Psid main interview user manual: Release 2015", (2015).

Del Boca, D., C. Flinn and M. Wiswall, "Household choices and child development", Review of Economic Studies **81**, 1, 137–185 (2013).

Del Boca, D., C. Flinn and M. Wiswall, "Transfers to households with children and child development", The Economic Journal **126**, 596, F136–F183 (2016).

Drews, C. D., M. Yeargin-Allsopp, P. Decoufl and C. C. Murphy, "Variation in the influence of selected sociodemographic risk factors for mental retardation.", American Journal of Public Health **85**, 3, 329–334, URL `https://doi.org/10.2105/AJPH.85.3.329`, pMID: 7892914 (1995).

Duncan, G. J., K. T. Lee, M. Rosales-Rueda and A. Kalil, "Maternal age and child development", Demography pp. 1–27 (2018).

Durkin, M. S., M. J. Maenner, C. J. Newschaffer, L.-C. Lee, C. M. Cunniff, J. L. Daniels, R. S. Kirby, L. Leavitt, L. Miller, W. Zahorodny *et al.*, "Advanced parental age and the risk of autism spectrum disorder", American Journal of Epidemiology **168**, 11, 1268–1276 (2008).

Eckstein, Z. and K. I. Wolpin, "Dynamic labour force participation of married women and endogenous work experience", The Review of Economic Studies **56**, 3, 375–390, URL `http://www.jstor.org/stable/2297553` (1989).

Ejrnæs, M. and A. Kunze, "Work and wage dynamics around childbirth", The Scandinavian Journal of Economics **115**, 3, 856–877 (2013).

Folloni, G. and G. Vittadini, “Human capital measurement: a survey”, Journal of economic surveys **24**, 2, 248–279 (2010).

Francesconi, M., “A joint dynamic model of fertility and work of married women”, Journal of Labor Economics **20**, 2, 336–380, URL `http://www.jstor.org/stable/10.1086/338220` (2002).

Gallant, A. and G. E. Tauchen, “Which moments to match?”, **12**, 657–681 (1996).

Gallen, Y., “Motherhood and the gender productivity gap”, (2018).

Gayle, G.-L., R. A. Miller *et al.*, “Life-cycle fertility and human capital accumulation”, Tepper School of Business. Paper **128** (2006).

Gerst, B. and C. Grund, “Career interruptions and current earnings: The role of interruption type, compensation component, and gender”, (2017).

Goisis, A., D. C. Schneider and M. Myrskylä, “The reversing association between advanced maternal age and child cognitive ability: evidence from three uk birth cohorts”, International journal of epidemiology **46**, 3, 850–859 (2017).

Goisis, A. and W. Sigle-Rushton, “Childbearing postponement and child well-being: a complex and varied relationship?”, Demography **51**, 5, 1821–1841 (2014).

Gourieroux, C., A. Monfort and E. Renault, “Indirect inference”, Journal of Applied Econometrics **8**, S85–S118, URL `http://www.jstor.org/stable/2285076` (1993).

Guner, N., R. Kaygusuz and G. Ventura, “Child-related transfers, household labor supply and welfare”, (2018).

Hack, M., H. G. Taylor, M. Schluchter, L. Andreias, D. Drotar and N. Klein, “Behavioral outcomes of extremely low birth weight children at age 8 years”, Journal of developmental and behavioral pediatrics: JDBP **30**, 2, 122 (2009).

Hanushek, E. A. and S. G. Rivkin, “Generalizations about using value-added measures of teacher quality”, American Economic Review **100**, 2, 267–71 (2010).

Havnes, T. and M. Mogstad, “Money for nothing? universal child care and maternal employment”, Journal of Public Economics **95**, 11-12, 1455–1465 (2011).

Heckman, J. J., L. Lochner and C. Taber, “Human capital formation and general equilibrium treatment effects: a study of tax and tuition policy”, Fiscal Studies **20**, 1, 25–40 (1999).

Heckman, J. J. and T. E. MaCurdy, “A life cycle model of female labour supply”, The Review of Economic Studies **47**, 1, 47–74 (1980).

Heffner, L. J., “Advanced maternal agehow old is too old?”, New England journal of medicine **351**, 19, 1927–1929 (2004).

Hofferth, S., P. E. Davis-Kean, J. Davis and J. Finkelstein, “The child development supplement to the panel study of income dynamics: 1997 user guide”, Ann Arbor: Survey Research Center, Institute for Social Research, University of Michigan (1997).

Hotz, V. J. and R. A. Miller, “An empirical analysis of life cycle fertility and female labor supply”, Econometrica **56**, 1, 91–118, URL http://www.jstor.org/stable/1911843 (1988).

Huggett, M., G. Ventura and A. Yaron, “Sources of lifetime inequality”, American Economic Review **101**, 7, 2923–54, URL http://www.aeaweb.org/articles?id=10.1257/aer.101.7.2923 (2011).

Imai, S. and M. P. Keane, “Intertemporal labor supply and human capital accumulation”, (2004).

Jacobsen, J. P. and L. M. Levin, “Effects of intermittent labor force attachment on women’s earnings”, Monthly Lab. Rev. **118**, 14 (1995).

Johnson, K. J., S. E. Carozza, E. J. Chow, E. E. Fox, S. A. Horel, C. C. McLaughlin, B. A. Mueller, S. E. Puumala, P. Reynolds, J. V. Behren and L. G. Spector, “Parental age and risk of childhood cancer: a pooled analysis.”, Epidemiology **20 4**, 475–83 (2009).

Keane, M. P. and K. I. Wolpin, “The career decisions of young men”, Journal of Political Economy **105**, 3, 473–522, URL http://www.jstor.org/stable/10.1086/262080 (1997).

Keane, M. P. and K. I. Wolpin, “The role of labor and marriage markets, preference heterogeneity, and thewelfare system in the life cycle decisions of black, hispanic, and white women”, International Economic Review **51**, 3, 851–892, URL http://www.jstor.org/stable/40784808 (2010).

Klaauw, W. V. D., “Female labour supply and marital status decisions: A life-cycle model”, The Review of Economic Studies **63**, 2, 199–235, URL http://www.jstor.org/stable/2297850 (1996).

Klerman, J. A., K. Daley and A. Pozniak, “Family and medical leave in 2012: Technical report”, Cambridge, MA: Abt Associates Inc (2012).

Kleven, H., C. Landais and J. E. Søgaard, “Children and gender inequality: Evidence from denmark”, Tech. rep., National Bureau of Economic Research (2018).

Kuziemko, I., J. Pan, J. Shen and E. Washington, “The mommy effect: Do women anticipate the employment effects of motherhood?”, Tech. rep., National Bureau of Economic Research (2018).

Lagakos, D., B. Moll, T. Porzio, N. Qian and T. Schoellman, “Life cycle wage growth across countries”, Journal of Political Economy **126**, 2, 797–849, URL https://doi.org/10.1086/696225 (2018).

Lalive, R. and J. Zweimüller, "How does parental leave affect fertility and return to work? evidence from two natural experiments", The Quarterly Journal of Economics **124**, 3, 1363–1402 (2009).

Lee, D., "An estimable dynamic general equilibrium model of work, schooling, and occupational choice", International Economic Review **46**, 1, 1–34 (2005).

Leigh, A. and X. Gong, "Does maternal age affect children's test scores?", Australian Economic Review **43**, 1, 12–27 (2010).

Lerman, S. and C. Manski, *On the Use of Simulated Frequencies to Approximate Choice Probabilities* (The MIT Press, 1981).

Livingston, G., "Childlessness falls, family size grows among highly educated women", Pew Research Center **7** (2015).

Lken, K. V., M. Mogstad and M. Wiswall, "What linear estimators miss: The effects of family income on child outcomes", American Economic Journal: Applied Economics **4**, 2, 1–35, URL `http://www.aeaweb.org/articles?id=10.1257/app.4.2.1` (2012).

McFadden, D., "A method of simulated moments for estimation of discrete response models without numerical integration", Econometrica **57**, 5, 995–1026, URL `http://www.jstor.org/stable/1913621` (1989).

Miller, A. R., "The effects of motherhood timing on career path", Journal of population economics **24**, 3, 1071–1100 (2011).

OECD, "SF2.3 Age of mothers at childbirth and age-specific fertility", Tech. rep., Organization for Economic Co-operation and Development (OECD), URL `http://www.oecd.org/els/soc/SF`$_{23A}ge_{m}others_{c}hildbirth.pdf$(2018).

Olivetti, C., "Changes in women's hours of market work: The role of returns to experience", Review of Economic Dynamics **9**, 4, 557–587 (2006).

Pakes, A. and D. Pollard, "Simulation and the asymptotics of optimization estimators", Econometrica **57**, 5, 1027–1057, URL `http://www.jstor.org/stable/1913622` (1989).

Polachek, S. W. *et al.*, "Earnings over the life cycle: The mincer earnings function and its applications", Foundations and Trends® in Microeconomics **4**, 3, 165–272 (2008).

Powell, B., L. C. Steelman and R. M. Carini, "Advancing age, advantaged youth: Parental age and the transmission of resources to children", Social Forces **84**, 3, 1359–1390, URL `http://dx.doi.org/10.1353/sof.2006.0064` (2006).

Restrepo-Méndez, M. C., D. A. Lawlor, B. L. Horta, A. M. Matijasevich, I. da Silva dos Santos, A. B. Menezes, F. C. Barros and C. G. Victora, "The association of maternal age with birthweight and gestational age: A cross-cohort comparison", in "Paediatric and perinatal epidemiology", (2015).

Ribeiro, L. A., H. D. Zachrisson, S. Schjolberg, H. Aase, N. Rohrer-Baumgartner and P. M. Magnus, "Attention problems and language development in preterm low-birth-weight children: Cross-lagged relations from 18 to 36 months", in "BMC pediatrics", (2011).

Rosenthal, M. S. and M. A. Khatamee, *The fertility sourcebook: everything you need to know* (McGraw-Hill Contemporary, 2002), third edn.

Rossin, M., "The effects of maternity leave on children's birth and infant health outcomes in the united states", Journal of health Economics **30**, 2, 221–239 (2011).

Rossin-Slater, M., "Maternity and family leave policy", Tech. rep., National Bureau of Economic Research (2017a).

Rossin-Slater, M., "Maternity and family leave policy", Tech. rep., National Bureau of Economic Research (2017b).

Rossin-Slater, M., C. J. Ruhm and J. Waldfogel, "The effects of california's paid family leave program on mothers leave-taking and subsequent labor market outcomes", Journal of Policy Analysis and Management **32**, 2, 224–245 (2013).

Schönberg, U. and J. Ludsteck, "Expansions in maternity leave coverage and mothers labor market outcomes after childbirth", Journal of Labor Economics **32**, 3, 469–505 (2014).

Shaw, K. L., "Life-cycle labor supply with human capital accumulation", International Economic Review **30**, 2, 431–456, URL http://www.jstor.org/stable/2526656 (1989).

Sheran, M., "The career and family choices of women: A dynamic analysis of labor force participation, schooling, marriage, and fertility decisions", Review of Economic Dynamics **10**, 3, 367–399 (2007).

Tearne, J., "Older maternal age and child behavioral and cognitive outcomes: A review of the literature", **103** (2015).

Weiss, Y. and R. Gronau, "Expected interruptions in labour force participation and sex-related differences in earnings growth", The Review of Economic Studies **48**, 4, 607–619, URL http://www.jstor.org/stable/2297200 (1981).

WHO, "Global nutrition targets 2025: Low birth weight policy brief", Tech. rep., World Health Organization (WHO) (2014).

Woodcock, R. W. and M. B. Johnson, "Tests of achievement, standard battery (form b)", Chicago: Riverside (1989).

Wmann, L., "Specifying human capital", Journal of Economic Surveys **17**, 3, 239–270, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6419.00195 (2003).

Yeargin-Allsopp, M., C. D Drews, P. Decoufl and C. C Murphy, "Mild mental retardation in black and white children in metropolitan atlanta: A case-control study", **85**, 324–8 (1995).

Yip, B. H., Y. Pawitan and K. Czene, "Parental age and risk of childhood cancers: a population-based cohort study from sweden", International Journal of Epidemiology **35**, 6, 1495–1503, URL http://dx.doi.org/10.1093/ije/dyl177 (2006).

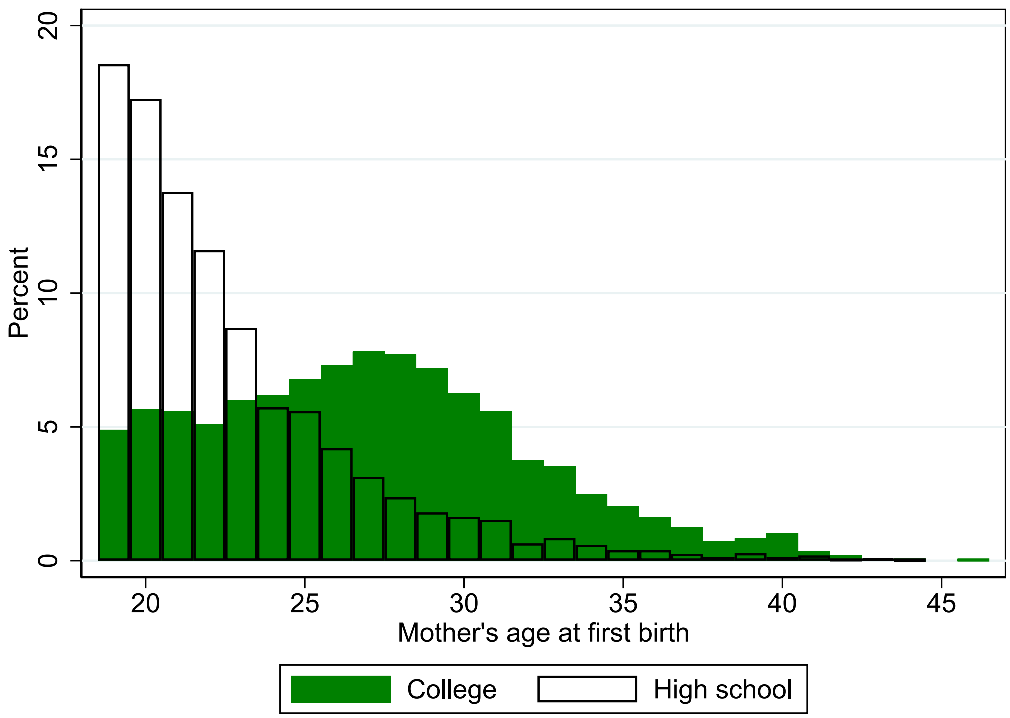
Zimmerman, S. D., "The returns to college admission for academically marginal students", Journal of Labor Economics **32**, 4, 711–754, URL https://doi.org/10.1086/676661 (2014).

APPENDIX A

SUPPLEMENTARY APPENDIX I

A.1 Extra Figures

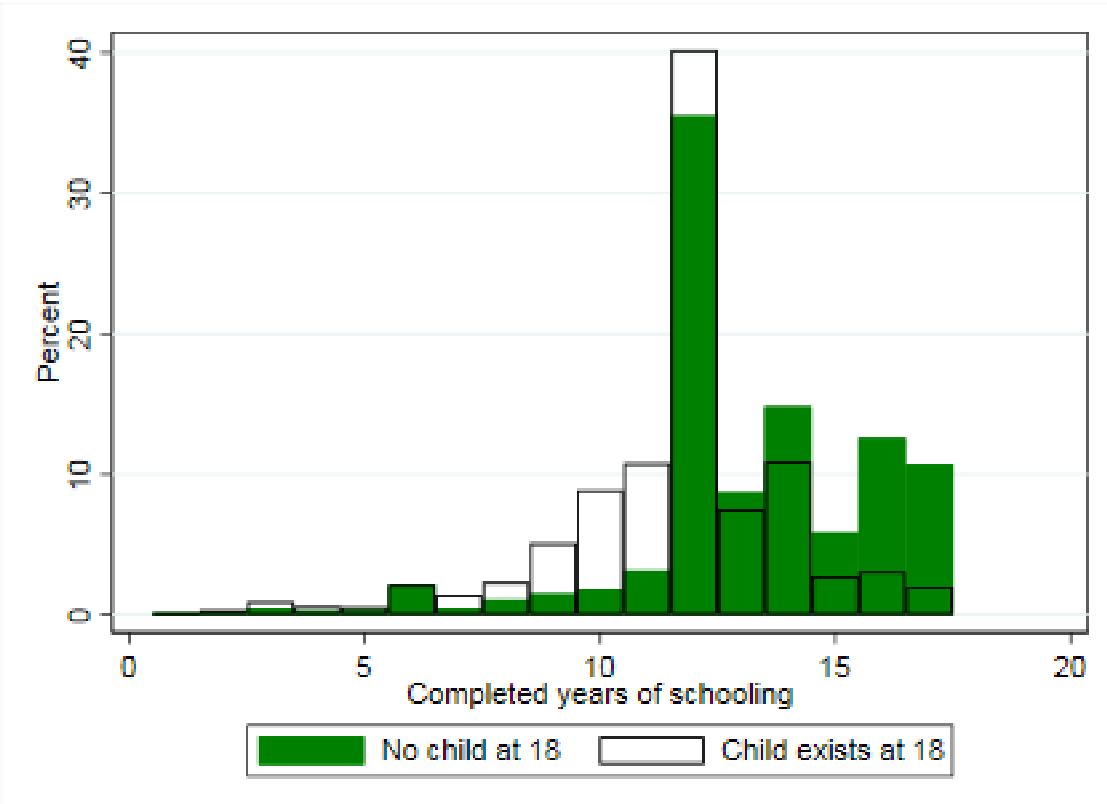
Figure A.1: Distribution of First-birth Age for Women with No Child at Age 18



Notes: This graph shows the distribution of age at first birth for women aged 40 or above in 2015 whose first birth was after age 18.
High school level is defined as 12 completed years of schooling or less. College graduates are defined as 16 completed years of schooling or more.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Figure A.2: Distribution of Completed Years of Schooling

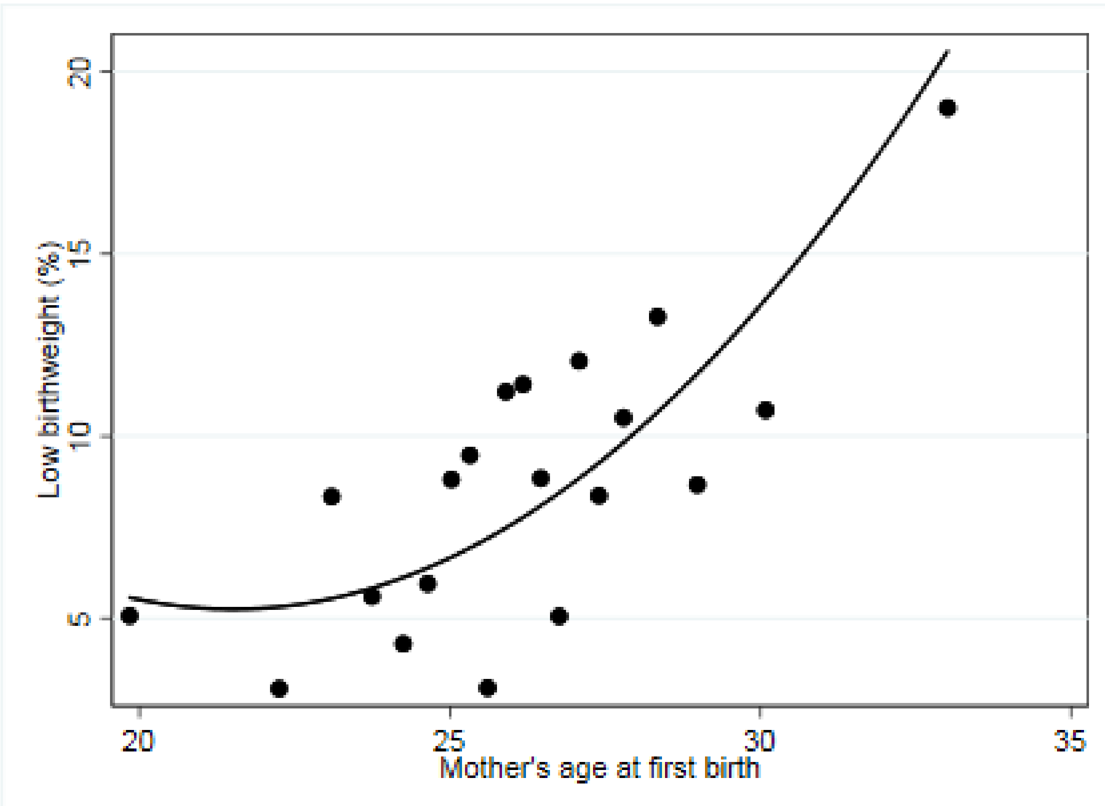


Source: PSID-CDS.

Notes: This graph shows the distribution of years of schooling for women aged 40 or above in 2015 whose first birth was after age 18.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

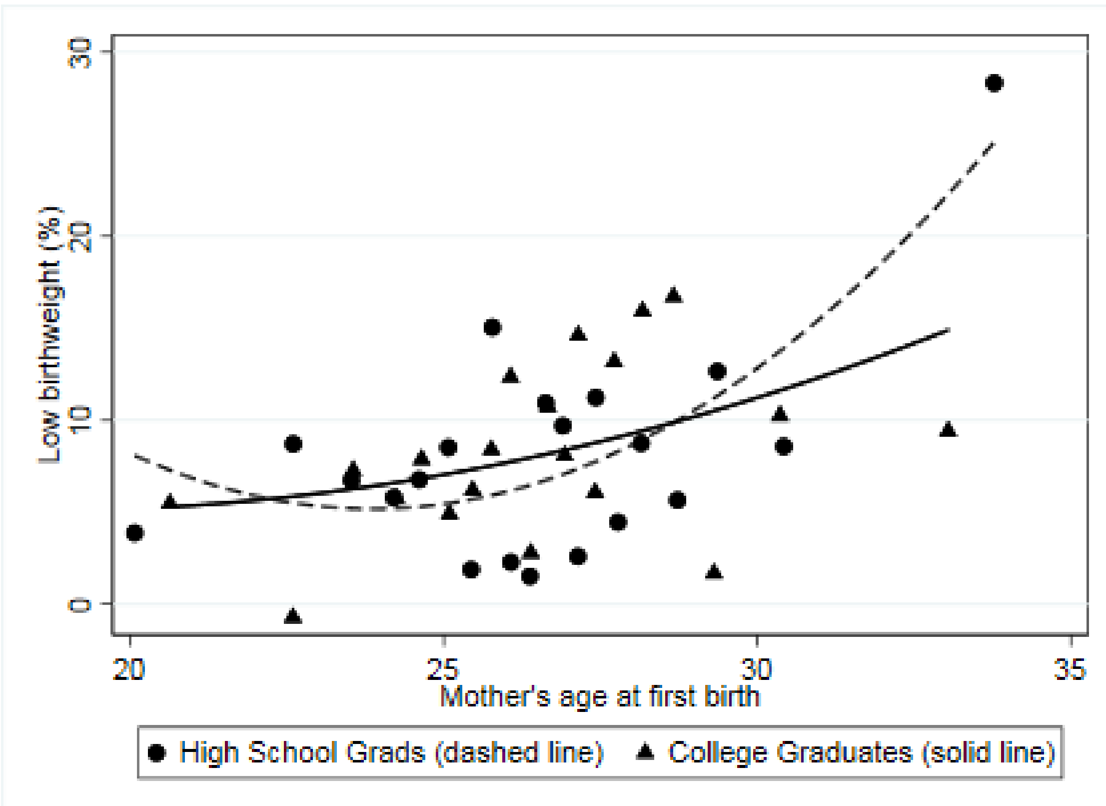
Figure A.3: Maternal Age at Delivery and Prob. of Low Birthweight (LBW)



Notes: Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
The binned scatterplot is shown here. The plot shows the regression in column (1) of Table (1.2), binned into 20 bins.
The probability of low birthweight is controlled for paternal age, infant's gender, race, mother's education and marital status at birth, and year fixed effects.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

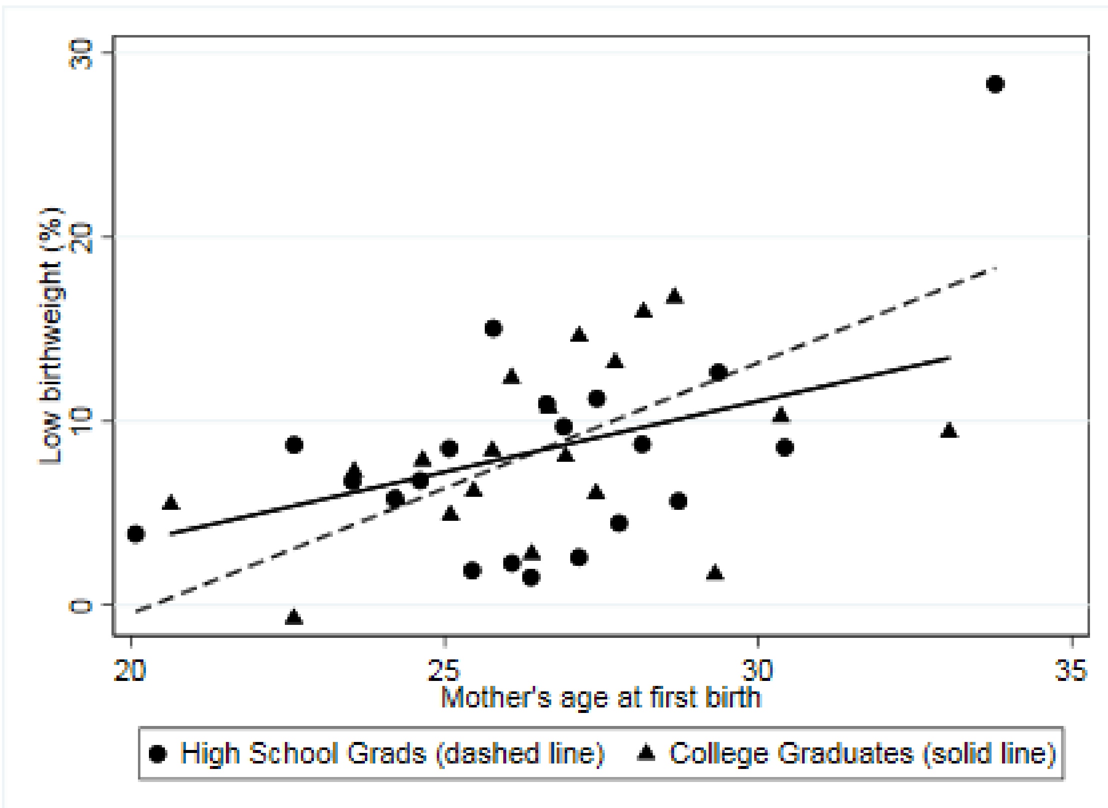
Figure A.4: LBW and Maternal Age, by Education Level (Quadratic From)



Notes: Low birthweight indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
The binned scatterplot is shown here. The plot shows the regression in columns (1) and (2) of of Panel A of Table (A.5), binned into 20 bins.
The probability of low birthweight is controlled for the infant's gender, race, and year fixed effects.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

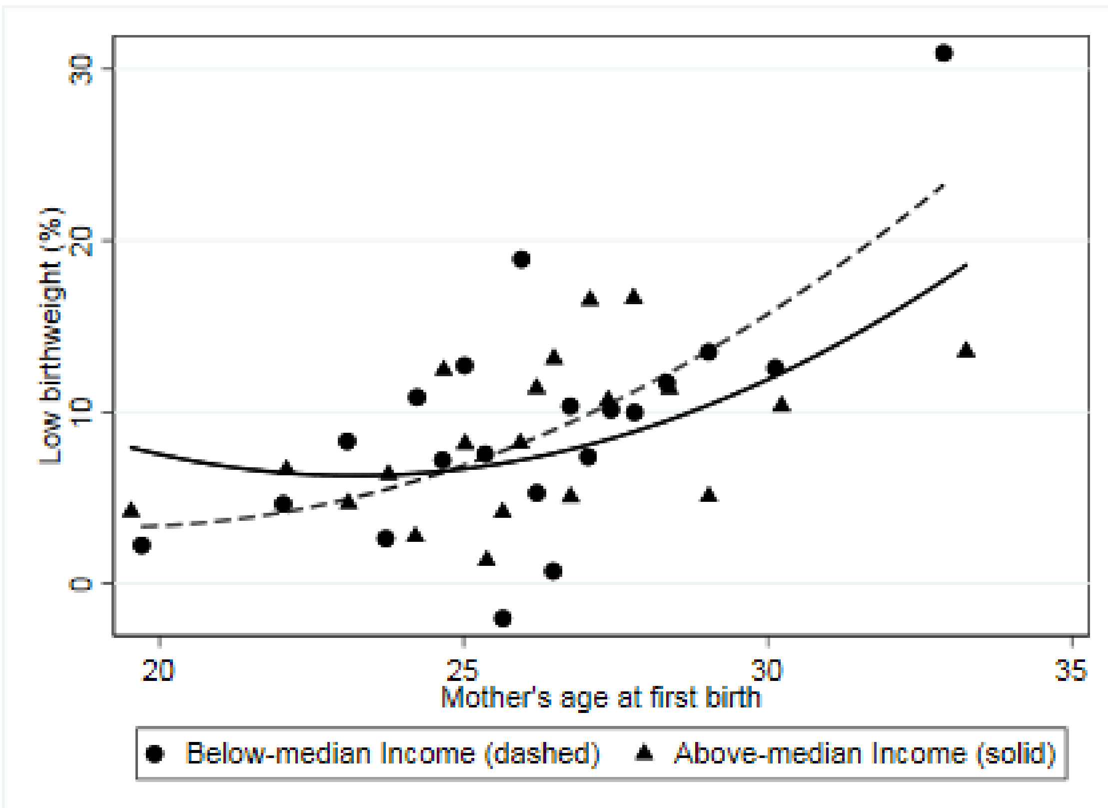
Figure A.5: LBW and Maternal Age, by Education Level (Linear Form)



Notes: Low birthweight indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
The binned scatterplot is shown here. The plot shows the regression in columns (1) and (2) of of Panel A of Table (A.6), binned into 20 bins.
The probability of low birthweight is controlled for the infant's gender, race, and year fixed effects.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

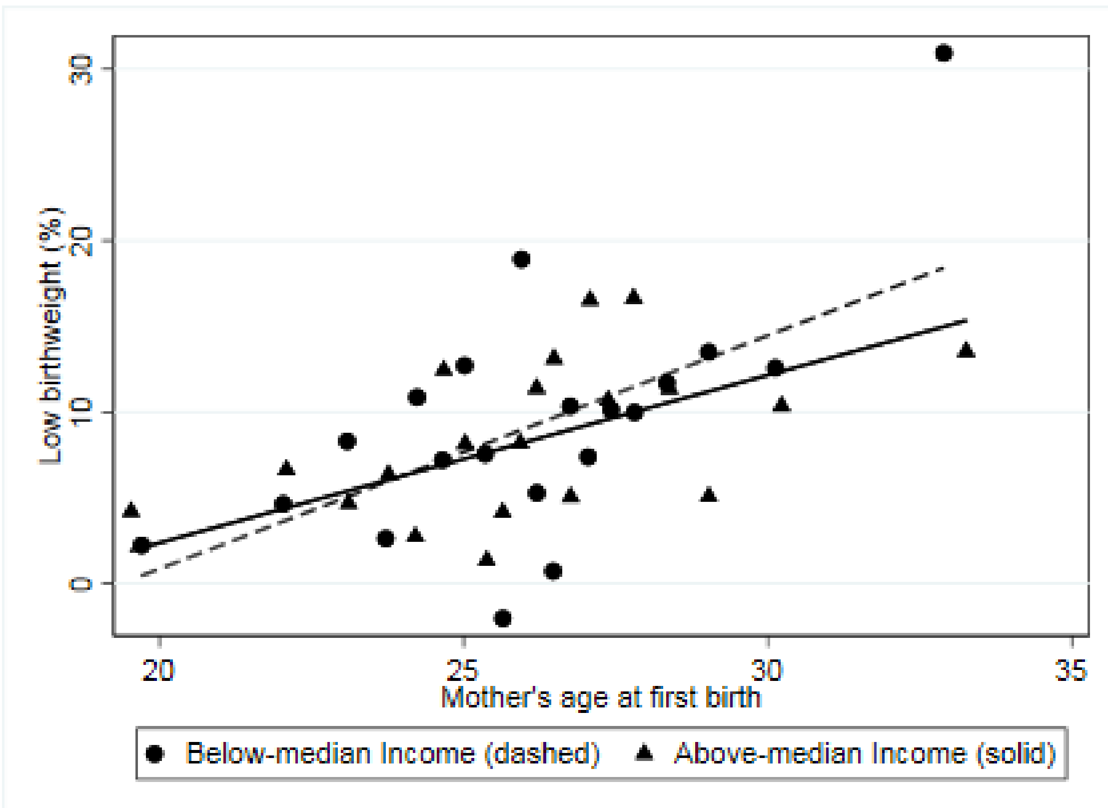
Figure A.6: LBW and Maternal Age, by Income Level (Quadratic From)



Notes: Low birthweight indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
The binned scatterplot is shown here. The plot shows the regression in columns (1) and (2) of of Panel B of Table (A.5), binned into 20 bins.
The probability of low birthweight is controlled for the infant's gender, race, and year fixed effects.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

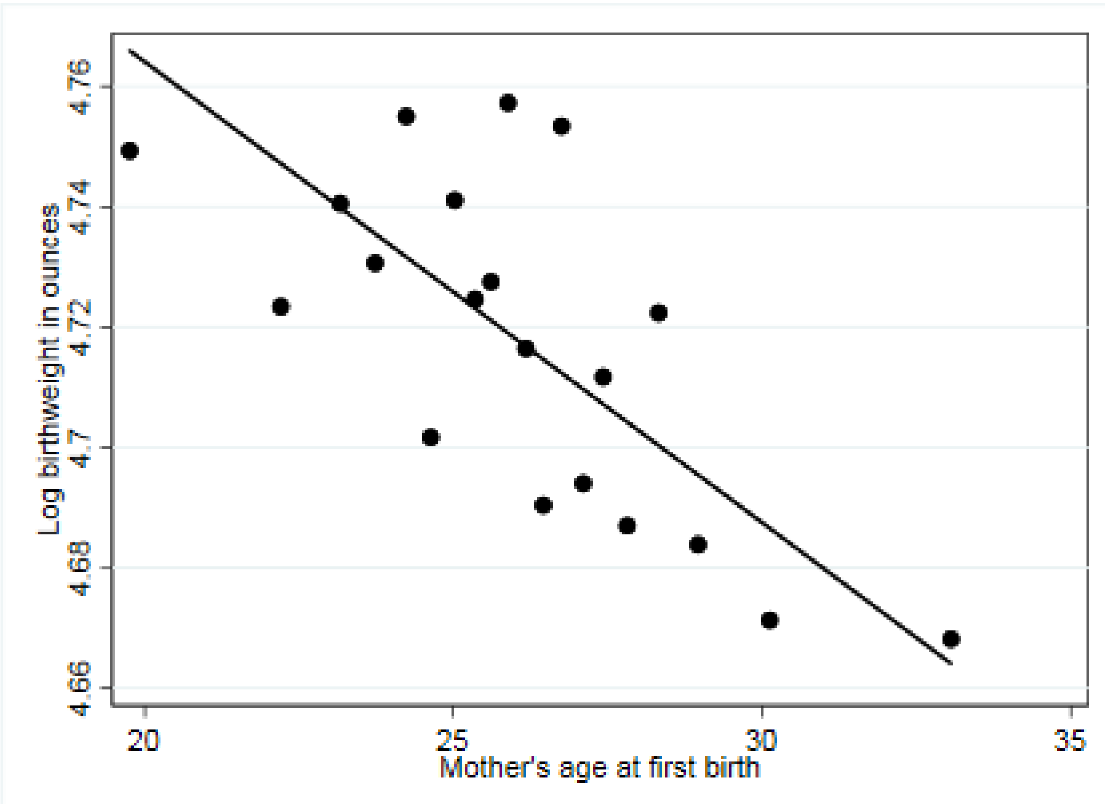
Figure A.7: LBW and Maternal Age, by Income Level (Linear From)



Notes: Low birthweight indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).
The binned scatterplot is shown here. The plot shows the regression in columns (1) and (2) of of Panel B of Table (A.5), binned into 20 bins.
The probability of low birthweight is controlled for the infant's gender, race, and year fixed effects.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

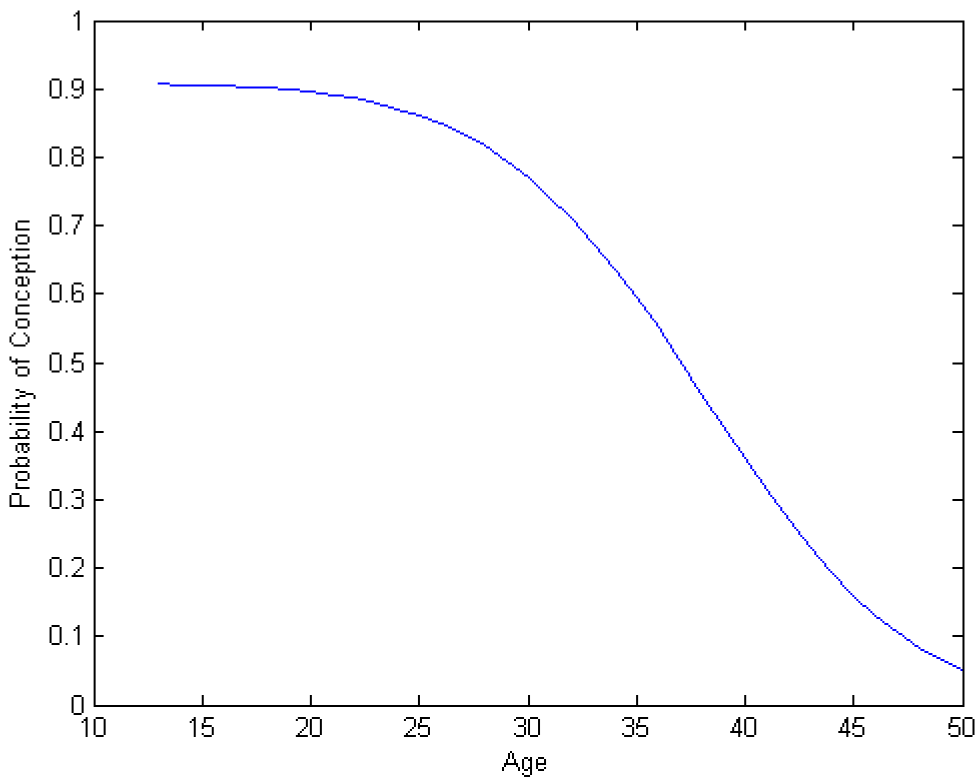
Figure A.8: Mother's Age at First Birth and the Log Birthweight of the Infant



Notes: The binned scatterplot is shown here. The plot shows the regression in column (1) of Table (A.7), binned into 20 bins.
The dependent variable, i.e, log birthweight, is controlled for paternal age, infant's gender, race, mother's education and marital status at birth, and year fixed effects.

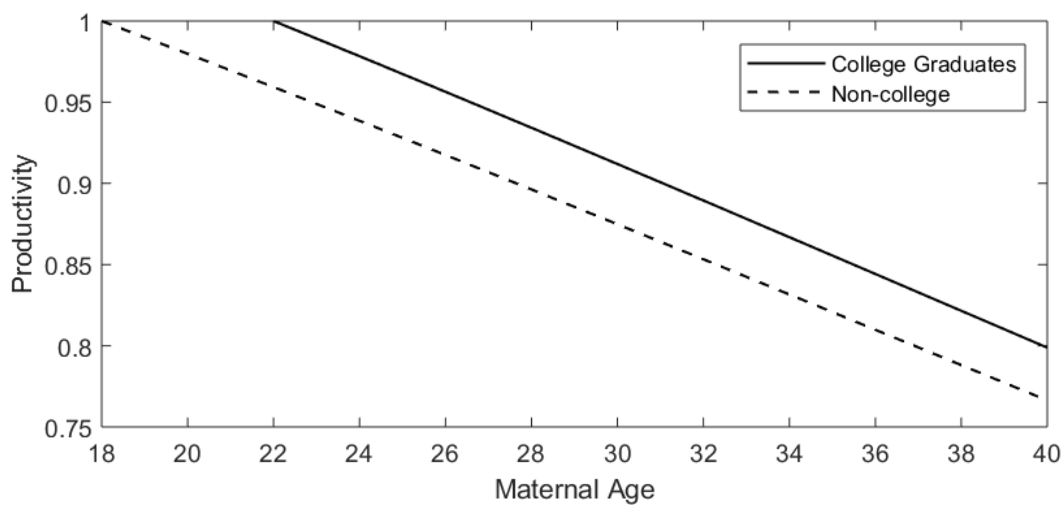
Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Figure A.9: Conception Probability over the Life-cycle



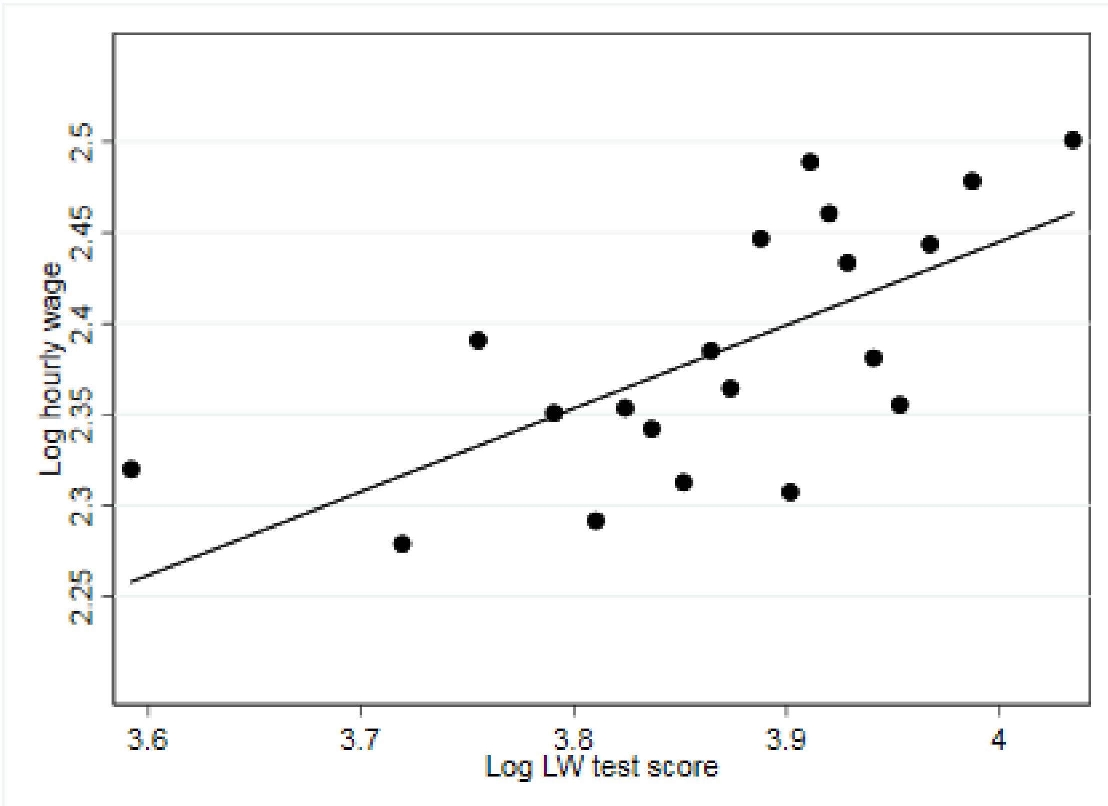
Notes: This graph shows how the conception probability declines over a female's age.
Source: Rosenthal and Khatamee (2002)

Figure A.10: Estimated Productivity Parameters by Child Age



Notes: This graphs estimated productivity of the child by mother's age at childbirth (from Table (1.8)). All values for college graduates are normalized to the child's productivity of the maternal age 22, and all values for non-college graduates are normalized to the child's productivity of the maternal age 18.

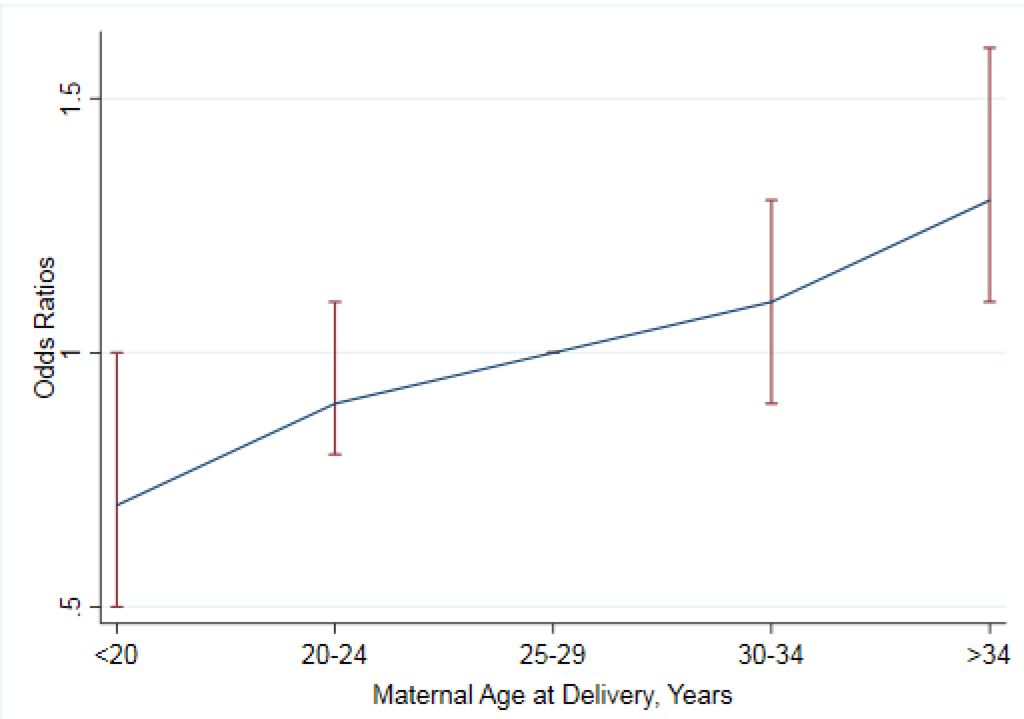
Figure A.11: The Letter-Word Test Score and Future Hourly Wages



Notes:This Graph shows the relationship between children's Letter-Word (LW) test scores at age 16-18 and their future wages, 7 to 12 years later, i.e. when they were between 23 to 30 years of old. The x-axis variable is LW test score of children in CDS dataset. Children were between ages 16-18 when they took the LW test.
The y-axis variable is the predicted hourly wage in 2014, when the children were between 23 to 30 years of old.
The binned scatterplot is shown here. Observations are binned into 20 bins.

Source: CDS dataset (2002-2007) combined with TAS dataset (2015).

Figure A.12: Maternal Age and the Risk of Autism



Notes: The odds ratios are controlled for paternal age, birth order, gender, race, education, gestational age, and birthweight.
The odds ratios are normalized with respect to the benchmark group, which is maternal age between 25-29.

Source: Durkin *et al.* (2008)

A.2 Extra Tables

Table A.1: Summary Statistics

	Non-college Grads	**College Grads**
Variable	**Mean**	**Mean**
First child LW test score	34.7	37.8
	(15.5)	(15.5)
First child's age	10.3	10.3
	(4.1)	(4.1)
Mother's age at first birth	22.2	25.8
	(4.8)	(5.5)
Mother's Completed years of schooling	11.6	15.4
	(1.49)	(1.2)
Father's hourly wage ($)	17.9	23
	(10.7)	(15.4)
Mother's Hourly wage ($)	13.5	17.9
	(8.7)	(11.6)
Mother's Hours worked (per week)	24.8	27.8
	(29.9)	(22.8)
Father's hours worked (per week)	40.7	40.5
	(28.4)	(20.8)
Moms' total time with child (hrs/week)	35.1	36.4
	(20.0)	(18.9)

Note: This Table shows summary statistics of the data before dropping women whose age at first birth were before 18.
Standard deviations are shown in parenthesis.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1968-2015 PSID core data.

Table A.2: Mean and Median Age at First Birth over Education Level

	Mean	**Median**	**Std**	**Fraction after 27**
0-12 years of schooling	22.4	21	4.2	0.13
14 years of schooling	23.8	23	4.6	0.21
16 years of schooling	26.7	27	4.9	0.47
Graduate degree	27.9	28	5.0	0.59

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews and 1968-2015 PSID core data.

Table A.3: Regression Results for Low Birthweight (LBW), Fixed-Effect

	(1)	(2)
Mother's age at birth	**0.66***	**0.66***
	(0.38)	(0.37)
Child's gender dummy	Yes	Yes
Year dummies	Yes	Yes
Marital status dummies	Yes	Yes
Birth order	Yes	Yes
Family income	No	Yes
Observations	6153	6153
Adjusted R^2	0.02	0.02

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: These results are obtained using a fixed effect regression, in which the variation in maternal ages between sibling is used to estimate the impact of the mother's age at birth on the risk of LBW.

$100*LBW_{i,j} = \beta_1 age^p_{i,j} + \beta_2 \alpha W_j + u_{i,j}$, where i denotes mothers and j denotes children of mother i.

Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).

Income levels are measured at the year in which the birth occurs.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015)

Table A.4: Linear Regression of Low Birthweight (LBW) on Maternal Age

	Prob. of LBW (%)			
	Quadratic	**Linear**	**Quadratic**	**Linear**
	(1)	(2)	(3)	(4)
Mother's age at first birth (age_i^p)	-2.16**	0.47***	-2.47*	1.15***
	(1.00)	(0.16)	(1.47)	(0.23)
Age square (age_i^{p2})	0.05**	–	0.06**	–
	(0.02)	–	(0.02)	–
Child's gender dummy	Yes	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Education level dummies	Yes	Yes	Yes	Yes
Paternal age dummies	Yes	Yes	Yes	Yes
Marital status dummies	Yes	Yes	Yes	Yes
Family income	No	No	Yes	Yes
Observations	3169	3169	1600	1600
Adjusted R^2	0.02	0.02	0.03	0.04

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = \beta_1 age_i^p + \beta_2 age_i^{p2} + \alpha W_i + u_i$, for the quadratic specification in columns (1) and (3)

$100 * LBW_i = \beta_1 age_i^p + \alpha W_i + u_i$, for linear specification in columns (2) and (4)

Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams).

Income levels are measured at the year in which the first child is born.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015)

Table A.5: Linear Regression of LBW on Maternal Age, Quadratic Specification

Panel A: By Education Level

	Non-college	**College Graduates**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at 1st birth (age_i^p)	-4.10*	-1.79
	(2.17)	(1.40)
Age square	0.09**	0.04
	(0.04)	(0.02)
Observations	777	1898
Adjusted R^2	0.07	0.03

Panel B: By Income Level

	Below-Median	**Above-Median**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at 1st birth (age_i^p)	-5.26**	-0.85
	(2.45)	(2.35)
Age square	0.12***	0.02
	(0.05)	(0.04)
Child's gender dummy	Yes	Yes
Race dummies	Yes	Yes
Year dummies	Yes	Yes
Education dummies	Yes	Yes
Observations	800	800
Adjusted R^2	0.04	0.05

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = \beta_1 age_i^p + \beta_2 age_i^{p\,2} + \alpha W_i + u_i$,
Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams). Estimates in Panel A are controlled for family income, and estimates in Panel B are controlled for education level.
Income levels are measured at the year in which the first child is born.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Table A.6: Linear Regression of LBW on Maternal Age

Panel A: By Education Level		
	Non-college	**College Graduates**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at 1st birth	0.45	0.42**
	(0.33)	(0.20)
Observations	777	1898
Adjusted R^2	0.05	0.02

Panel B: By Income Level		
	Below-Median	**Above-Median**
	Prob. of LBW (%)	**Prob. of LBW (%)**
Mother's age at first birth	1.12***	1.06***
	(0.39)	(0.31)
Child's gender dummy	Yes	Yes
Race dummies	Yes	Yes
Year dummies	Yes	Yes
Education dummies	Yes	Yes
Observations	800	800
Adjusted R^2	0.03	0.05

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: $100 * LBW_i = \beta_1 age_i^p + \alpha W_i + u_i$,
Low birthweight (LBW) indicates that the baby is born weighing less than 5 pounds, 8 ounces (i.e. 88 ounces, or 2500 grams). Estimates in Panel A are controlled for family income, and estimates in Panel B are controlled for education level.
Income levels are measured at the year in which the first child is born.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Table A.7: Linear Regression of Log Birthweight (BW) on Maternal Age

	Log BW	Log BW
Mother's age at first birth	**-0.004***	**-0.008****
	(0.001)	(0.002)
Child's gender dummy	Yes	Yes
Race dummies	Yes	Yes
Year dummies	Yes	Yes
Education Level dummies	Yes	Yes
Paternal Age dummies	Yes	Yes
Marital Status dummies	Yes	Yes
Income	No	Yes
Observations	2758	1609
Adjusted R^2	0.03	0.04

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Income levels are measured at the year in which the first child is born.
$ln(BW)_i = \beta_1 age_i^p + \alpha W_i + u_i$, for linear specification in column (2).
Birthweights are measured in ounces.

Source: PSID Family-level Data (1967-2015) combined with Childbirth and Adoption History (1985-2015).

Table A.8: Regression Results Including Only One-child Families

	ln(Score)	**ln(Score)**	**ln(Score)**
Mother's age at 1st birth (age^p_{it})	**0.009*****	**0.006****	**-0.014****
	(0.002)	(0.002)	(0.006)
Mother's years of schooling		0.023***	0.030**
		(0.008)	(0.015)
Father's years of schooling		0.019**	-0.001
		(0.008)	(0.014)
Father's hourly wage			0.007**
			(0.003)
Father's hours worked (per week)			0.005**
			(0.002)
Mother's hourly wage			0.006*
			(0.003)
Mother's hours worked (per week)			-0.003
			(0.002)
Child's age dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Marital status dummies	No	No	Yes
Paternal age dummies	No	No	Yes
Child's gender dummy	No	No	Yes
Observations	832	543	206
Adjusted R^2	0.86	0.87	0.89

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $ln(score_{i,t}) = \beta_1 age^p_i + \gamma Z_{i,t} + \alpha W_i + u_{i,t}$,
The dependent variable is the natural logarithm of Letter-Word Identification Test score. I only use the data on the first child of mothers.
Child's age is controlled in all above equations using age-specific dummy variables.
The number of observations is not constant across columns due to missing data on wages and education levels.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews.

Table A.9: Regression Results Including All Households (Constant # of Obs.)

	ln(Score)	ln(Score)	ln(Score)
Mother's age at 1st birth (age^p)	**0.007*** **	**0.004***	**-0.000**
	(0.002)	(0.002)	(0.003)
Mother's years of schooling		0.013**	0.014**
		(0.006)	(0.006)
Father's years of schooling		0.016***	0.014***
		(0.005)	(0.005)
Father's hourly wage			0.002
			(0.001)
Father's hours worked (per week)			0.000
			(0.001)
Mother's hourly wage			-0.000
			(0.002)
Hours worked (per week)			-0.001*
			(0.001)
Child's age dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Number of children dummies	Yes	Yes	Yes
Marital status dummies	No	No	Yes
Paternal age dummies	No	No	Yes
Child's gender dummy	No	No	Yes
Observations	762	762	762
Adjusted R^2	0.85	0.86	0.86

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $ln(score_{i,t}) = \beta_1 age_i^p + \gamma Z_{i,t} + \alpha W_i + u_{i,t}$,
The dependent variable is the natural logarithm of Letter-Word Identification Test score. I only use the data on the first child of mothers.
Child's age is controlled in all above equations using age-specific dummy variables.
The number of observations is held constant across columns.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews.

Table A.10: Regression Results Including All Households

	ln(Score)	ln(Score)	ln(Score)
Mother's age at 1st birth (age_i^p)	**0.011***	**0.006***	**-0.000**
	(0.001)	(0.001)	(0.003)
Mother's years of schooling		0.014***	0.014**
		(0.004)	(0.006)
Father's years of schooling		0.017***	0.014***
		(0.003)	(0.005)
Father's hourly wage			0.002
			(0.001)
Father's hours worked (per week)			0.000
			(0.001)
Mother's hourly wage			-0.000
			(0.002)
Hours worked (per week)			-0.001*
			(0.001)
Child's age dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Race dummies	Yes	Yes	Yes
Number of children dummies	Yes	Yes	Yes
Marital status dummies	No	No	Yes
Paternal age dummies	No	No	Yes
Child's gender dummy	No	No	Yes
Observations	3077	2117	762
Adjusted R^2	0.85	0.86	0.86

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: $ln(score_{i,t}) = \beta_1 age_i^p + \gamma Z_{i,t} + \alpha W_i + u_{i,t}$,
The dependent variable is the natural logarithm of Letter-Word Identification Test score. I only use the data on the first child of mothers.
Child's age is controlled in all above equations using age-specific dummy variables.
The number of observations is not constant across columns due to missing data on wages and education levels.

Source: PSID-CDS combined sample from 1997, 2002, 2007, and 2014 interviews.

Table A.11: Percentage of Births Intended at Conception, by Education of Mother

Statistics	Percentage
Less than high school diploma	59
High school diploma	60
College degree	83

Source: U.S. Department of Health and Human Services, 2012

Table A.12: Linear Regression of Work Experience

	Years of Work Experience
Potential Experience	0.601***
	(0.006)
Years of Schooling	0.651***
	(0.013)
1. Family size	-0.664***
	(0.076)
2. Family size	-1.790***
	(0.100)
3. Family size	-3.121***
	(0.145)
4. Family size	-4.772***
	(0.217)
5. Family size	-7.001***
	(0.284)
Race dummies	Yes
Birth year dummies	Yes
Marital status dummies	Yes
Demographics	Yes
Observations	166076
Adjusted R^2	0.64

Robust standard errors clustered at the individual level are reported in parentheses.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Income levels are measured at the year in which the first child is born.

Source: Source: NLSY79 (1979-2015) combined with NLSY97 (1997-2015).

Table A.13: Regression Results Using the Simulated Data

	(1) **ln(Score)**	(2) **ln(Score)**	(3) **ln(Score)**	(4) **ln(Score)**
Mother's age at 1st birth (age_i^p)	**-0.014***** (0.001)	**-0.015***** (0.001)	**-0.017***** (0.001)	**-0.019***** (0.001)
Mother's years of schooling	0.025*** (0.002)	0.025*** (0.002)	0.031*** (0.002)	0.029*** (0.002)
Mother's hourly wage			0.005*** (0.000)	0.004*** (0.000)
Mother's hours worked (per week)			-0.003*** (0.000)	-0.003*** (0.000)
Child's age dummies	Yes	Yes	Yes	Yes
Marital status dummies	Yes	Yes	Yes	Yes
Family income at age 2 of the child	No	Yes	Yes	Yes
Father's hourly wage	No	No	Yes	Yes
Child investments in past periods	No	No	No	Yes
Observations	3300	3300	3300	3300
Adjusted R^2	0.65	0.65	0.71	0.77

Notes: $Ln(score)_{i,t} = \beta_1 age_i^p + \gamma Z_{i,t} + \alpha W_i + \eta I_{i,t'} + u_{i,t}$,
where $I_{i,t'}$ denotes the vector of past investments.
The dependent variable is the natural logarithm of Letter-Word Identification Test score.
Child's age is controlled in all above equations using age-specific dummy variables.
Source: The simulated dataset is used for this exercise.

Table A.14: Regression of Child IQ at 36 Months on Independent Variables

	Child IQ (1)	Child IQ (2)	Child IQ (3)	Child IQ (4)
Maternal IQ	0.47***	0.47***	0.33***	0.26***
	(0.04)	(0.04)	(0.05)	(0.08)
Maternal Age		**-0.02**	**-0.24***	**-0.24***
		(0.19)	(0.14)	(0.14)
Family Income			0.40***	0.35***
			(0.05)	(0.05)
Quality of Home				0.60***
				(0.14)
Constant	47.70***	47.97***	56.60***	43.04***
	(3.19)	(4.05)	(4.08)	(4.18)
Observations	453	453	453	453
R^2	0.28	0.28	0.36	0.38

Standard errors are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The sample restricted to preterm low-birth-weight children.

Source: Bacharach and Baumeister (1998)

Table A.15: Risk of Down Syndrome and Chromosomal Abnormalities at Live Birth

Maternal age at delivery (yr)	**Risk of Down Syndrome**	**Risk of Any Chromosomal Abnormality**
20	$\frac{1}{1667}$	$\frac{1}{526}$
25	$\frac{1}{1200}$	$\frac{1}{476}$
30	$\frac{1}{952}$	$\frac{1}{385}$
35	$\frac{1}{378}$	$\frac{1}{192}$
40	$\frac{1}{106}$	$\frac{1}{66}$
45	$\frac{1}{30}$	$\frac{1}{21}$

Source: Heffner (2004)

A.3 Literature Review on Biological Impacts of Maternal Age

There are many potential negative consequences of postponement of childbirth on child outcomes, which are documented in previous papers. The risks of Down syndrome, childhood cancer, and autism increase with the maternal age (Durkin *et al.* (2008); Johnson *et al.* (2009); Yip *et al.* (2006)). A ten-year increase in maternal age is associated with a 20% to 30% increase in Autism Spectrum Disorder risk. Also, the risk of a low birthweight infant increases with maternal age (Restrepo-Méndez *et al.* (2015); Goisis and Sigle-Rushton (2014)). Delayed motherhood is associated with a higher risk of diabetes in the child (Byrnes (2001); Cardwell *et al.* (2009), after reviewing 30 observational studies, conclude that, on average, a five-year increase in maternal age is associated with a 5% to 10% increase in the risk of childhood type 1 diabetes. Older maternal age has also been linked with mental retardation in the absence of Down syndrome but accompanied by other neurologic conditions such as epilepsy, cerebral palsy, or birth defects affecting the central nervous system (Tearne (2015); Yeargin-Allsopp *et al.* (1995); Drews *et al.* (1995)).

Postponing childbirth is associated with reduced intelligence in the child (Bacharach and Baumeister (1998)). Table (A.14) shows the results of Bacharach and Baumeister (1998). After controlling for the maternal IQ, the family income, and the quality of home, there is a significant negative association between the maternal age at delivery and the child's IQ at 36 months, suggesting that postponing childbirth is associated with reduced intelligence in the child. Since they directly control for the mother's IQ in their regressions, one might be less worry about the role of selection in timing of births based on the innate ability of mothers.

With respect to the risk of Down syndrome, Table (A.15) in the Supplementary Appendix shows how the risk of Down syndrome and the risk of any chromosomal abnormality increase with maternal ages. The statistics are based on the crude frequency of the chromosomal abnormality in live births (see Heffner (2004)).

Figure (A.12) is depicted based on the regression results reported by Durkin et al. (2008). It shows the positive association between maternal age at delivery and the risk of autism after controlling for the paternal age, race, birth order, and gender of the child, mothers' education level, gestational age, and birthweight. Based on these results, the risk of autism increases by maternal age at delivery by a sizable amount.

APPENDIX B

SUPPLEMENTARY APPENDIX 2

B.1 Robustness Checks

In this appendix, I use other datasets to investigate whether my regression results in Table (2.4) are robust to my choice of dataset. I use NLSY79 and NLSY97 as alternative datasets and follow a similar approach presented in section 3.1. However, using these mentioned alternative datasets, I take advantage of having more precise information on the work experience of individuals and their cognitive ability as well. The latter is because I observe the AFQT score for each individual. Hence, the regression used for the analysis using these alternative datasets is as follows:

$$log(w_{it}) = \beta_0 + \beta_1 AFQT_i + \beta_2 exp_{it} + \beta_3 exp_{it}^2 + \beta_4 d_{it} + \beta_5 same_i + \epsilon_{it}, \qquad \text{(B.1)}$$

where w_{it} denotes hourly wage of individual i at at time t, $AFQT_i$ denotes the individual's AFQT percentile score, exp_{it} is the years of work experience[1], d_{it} is the education level measured as completed years of schooling, $same_i$ is an indicator variable which is one if the first two children are of the same gender, and zero otherwise, and ϵ_{it} is a zero conditional mean error term, which is assumed to be i.i.d and normally distributed over individuals and assumed to be time-invariant. I control for year fixed effects, race, etc by adding dummy variables to the regression specified above. The coefficient of interest is β_5, which captures the effect of children with the same sex on their mother's wages.

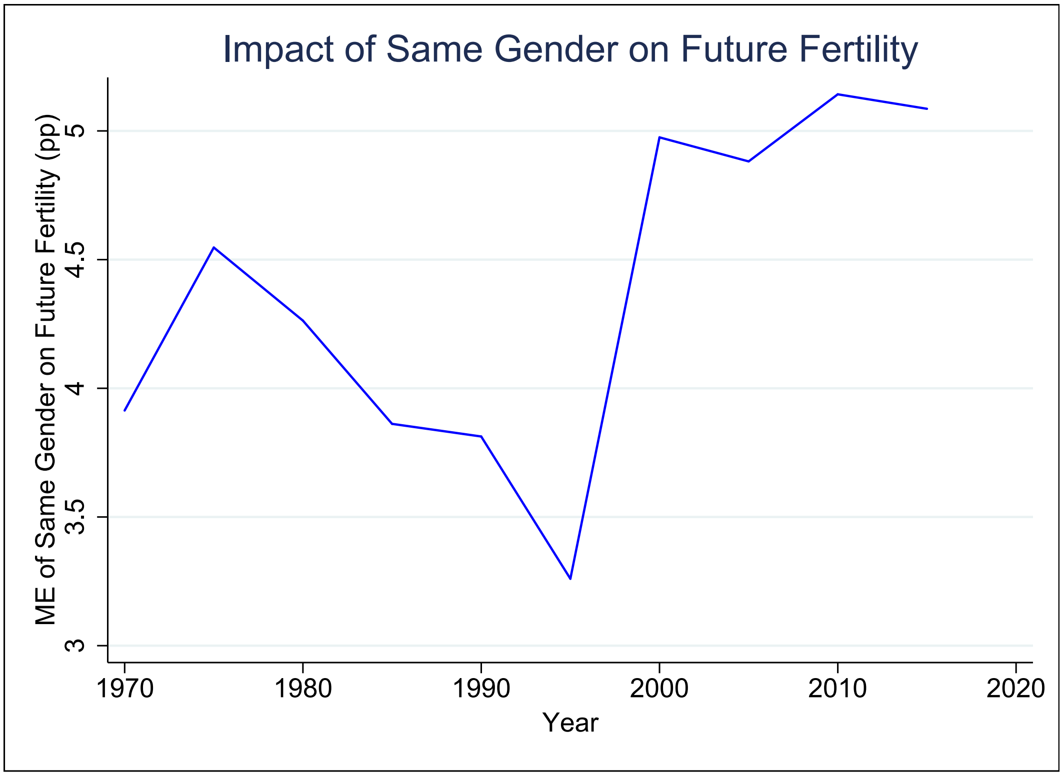
Results are presented in Tables (B.14) through (B.17) when I use different criteria for the sample selection procedure.

The results are less robust compared to my original analysis using PSID data. I conjecture that this is due to the sample size issue as the number of observations is considerably lower compared to the baseline dataset, and this makes my estimate less precise. However, in general, the results provide some evidence for the effects on the mother's hourly wage of the gender mix of her children.

[1] It is calculated as the number of years (since age 18) in which the individual has worked at least for 1500 hours.

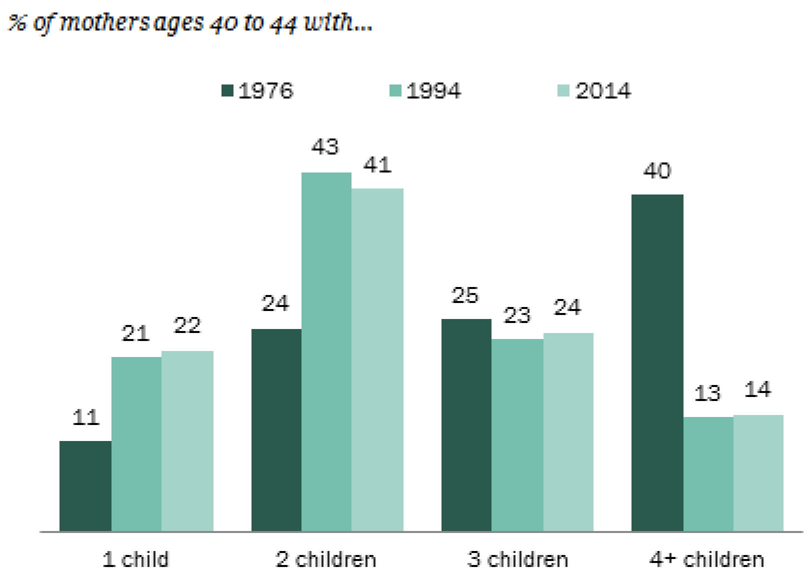
B.2 Extra Figures

Figure B.1: Gender Mix of Children and Observer Fertility Rate over Time



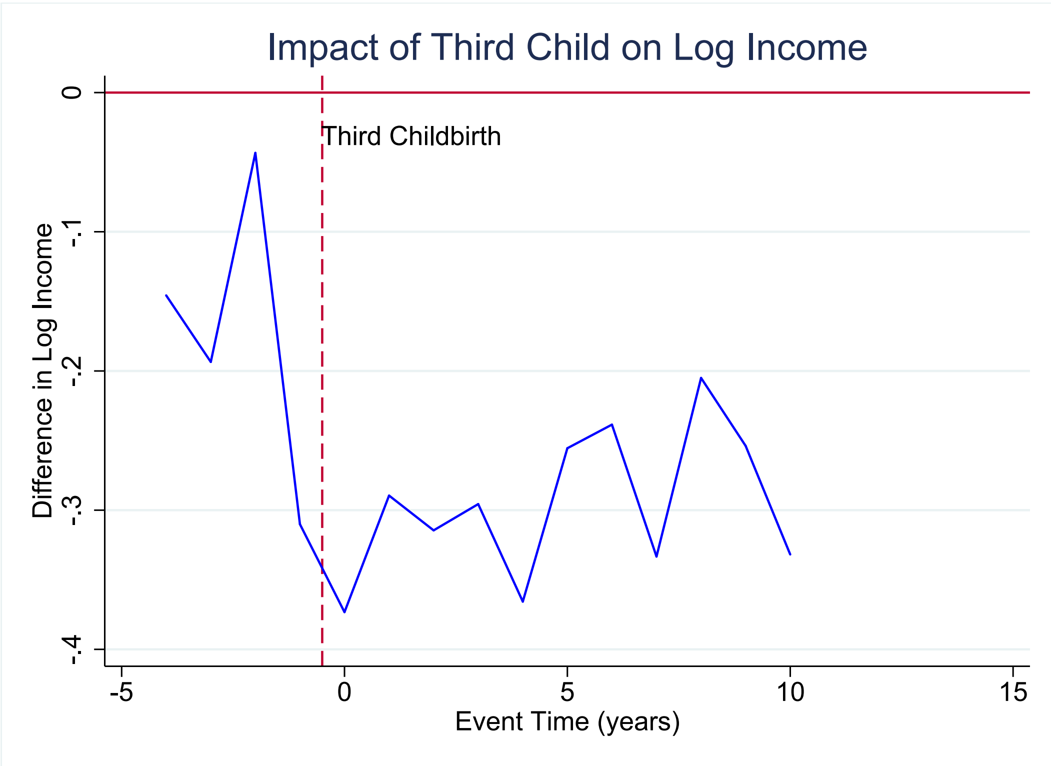
Notes: This figure reports the marginal effect of having two children of the same gender on future fertility based on ex post fertility outcomes.
Notes: A probit model, similar to column 2 of Table (2.2) has been used to estimate the marginal effect of having same-gender children on future fertility.
Notes: A probit model, similar to column 2 of Table (2.2) has been used to estimate the marginal effect of having same-gender children on future fertility.
Notes: In order to estimate the marginal effect, for each year, I use the fertility outcomes of all women above 40 years old in that year.
Notes: The abrupt increase in 1997 might be due to the refresher sample added to the PSID data set in 1997.

Figure B.2: Family Size Distribution in the US



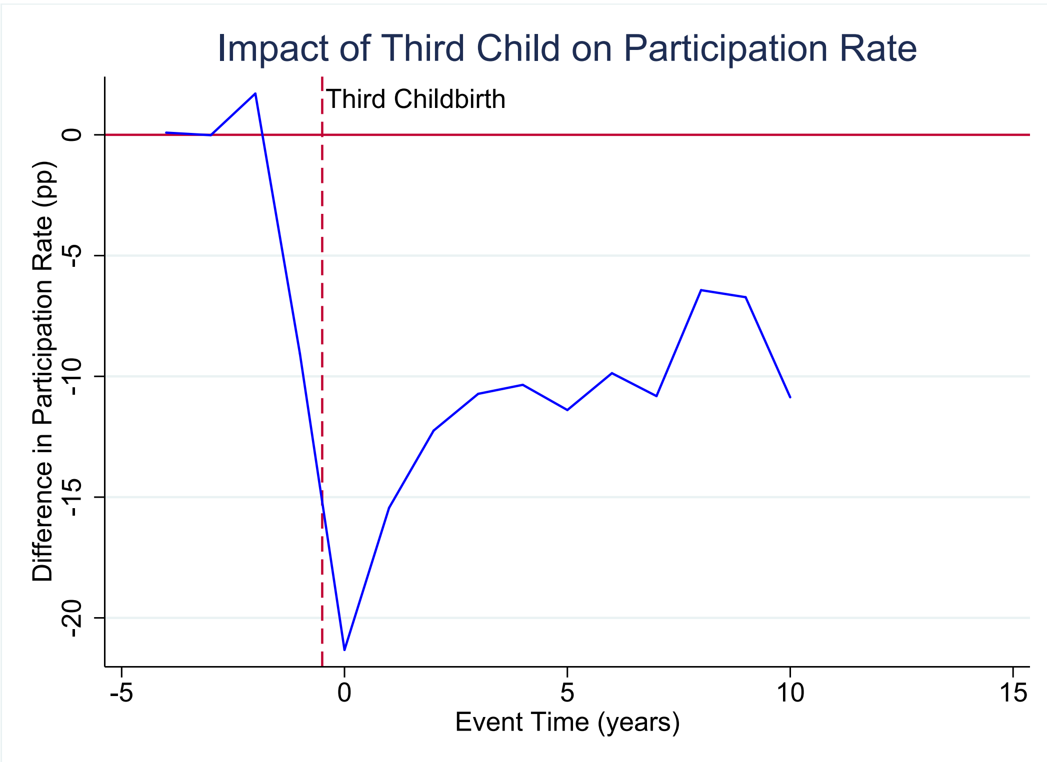
Source: Livingston (2015)

Figure B.3: Impact of Arrival of a Third Child on Mother's Labor Income



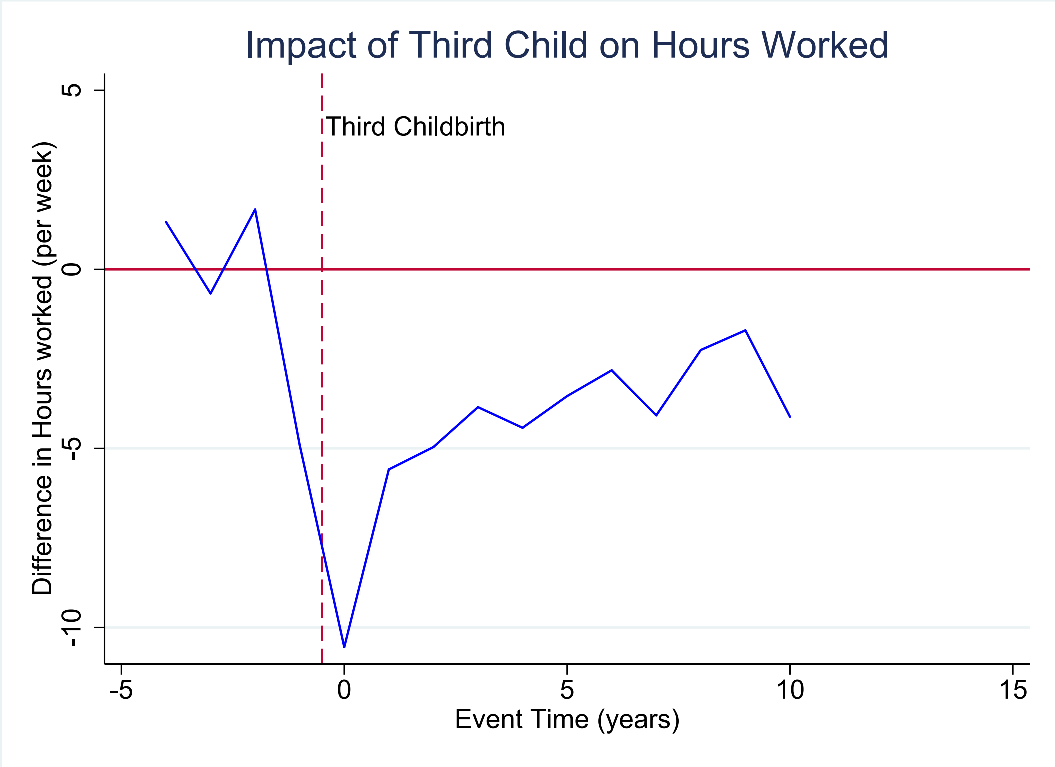
Notes: This figure shows the difference between the control group and the treatment group in Figure (2.5) 5 years before the third childbirth, and 10 years after.

Figure B.4: Impact of Arrival of a Third Child on Mother's Participation



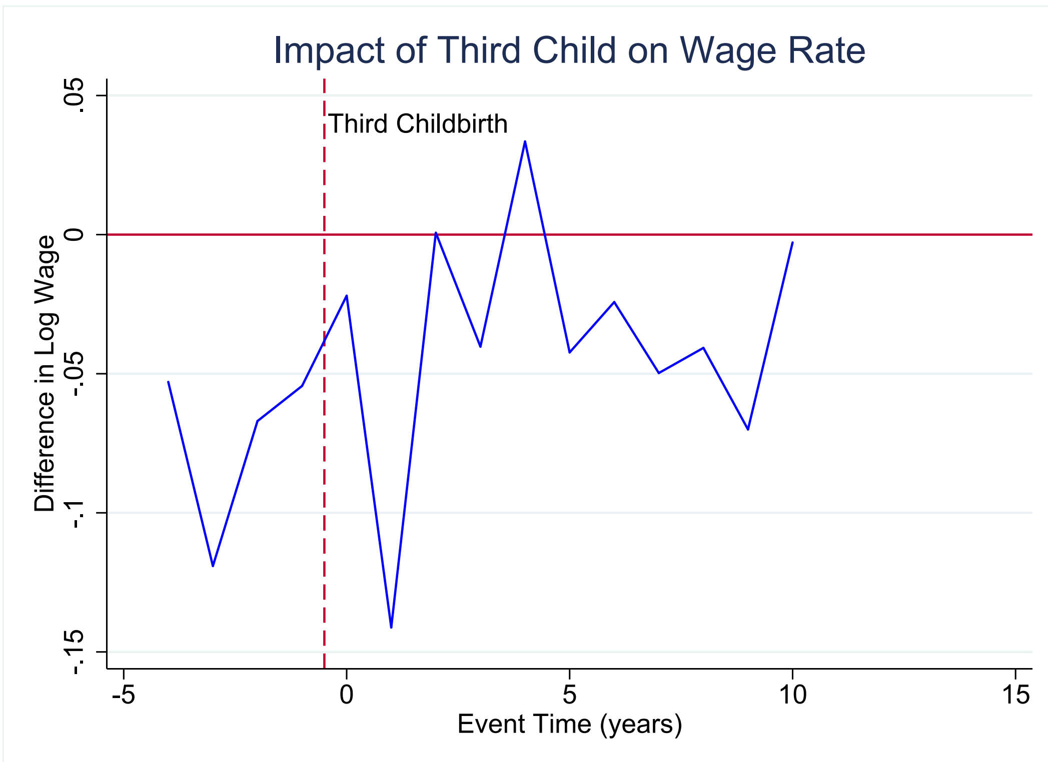
Notes: This figure shows the difference between the control group and the treatment group in Figure (2.6) 5 years before the third childbirth, and 10 years after.

Figure B.5: Impact of Arrival of a Third Child on Mother's Work Hours



Notes: This figure shows the difference between the control group and the treatment group in Figure (2.7) 5 years before the third childbirth, and 10 years after.

Figure B.6: Impact of Arrival of a Third Child on Mother's Wage Rate



Notes: This figure shows the difference between the control group and the treatment group in Figure (2.8) 5 years before the third childbirth, and 10 years after.

B.3 Extra Tables

Table B.1: The Gender Mix of Children and (Ex Post) Fertility Behavior

	Third child (OLS) (1)	Third child (Probit) (2)	Third child (Logit) (3)
Two boys	0.05***	0.14***	0.22***
	(0.01)	(0.04)	(0.06)
Two girls	0.06***	0.15***	0.25***
	(0.01)	(0.04)	(0.06)
Constant	0.81***	1.09**	1.89*
	(0.09)	(0.53)	(1.04)
Observations	7601	7601	7601
Adjusted R^2	0.04	0.03	0.03
Race	Y	Y	Y
Education Level	Y	Y	Y

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Sample have been restricted to individuals older than 40 years of old at 2014.
I also restrict the analysis to the subsample of women who have two children.

Table B.2: The Gender Mix of Children and (Ex Ante) Fertility Decisions

	Third child (OLS) (1)	Third child (Probit) (2)	Third child (Logit) (3)
Two boys	0.08**	0.35**	0.63**
	(0.03)	(0.14)	(0.26)
Two girls	0.07**	0.34**	0.59**
	(0.04)	(0.15)	(0.27)
Constant	0.03	-2.62***	-4.88***
	(0.09)	(0.61)	(1.32)
Observations	1249	1249	1249
Adjusted R^2	0.11	0.15	0.15
Age	Y	Y	Y
Race	Y	Y	Y
Education	Y	Y	Y
Year	Y	Y	Y

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The question is asked from females in PSID between 1969-1972 & 1976.
Dependent variables take 1 if the agent expects more children, 0 otherwise.
The Sample is restricted to women between age 21 and 40 with 2 children.

Table B.3: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.08***	0.08***	0.07***	0.07***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Potential experience square	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.09***	0.09***	0.06***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same sex	**-0.04****	**-0.04****	**-0.04****	**-0.04****	**-0.04****
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	1.12***	0.91***	1.56***	1.63***	1.43***
	(0.24)	(0.26)	(0.28)	(0.28)	(0.31)
Observations	4871	4753	4561	4561	4540
Adjusted R^2	0.18	0.23	0.30	0.31	0.33
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample is restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.4: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.08***	0.08***	0.07***	0.07***	0.07***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Potential experience square	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.09***	0.09***	0.06***	0.06***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Two boys	**-0.04***	**-0.04***	**-0.04***	**-0.04***	**-0.04****
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Two girls	**-0.04**	**-0.04***	**-0.05****	**-0.04****	**-0.05****
	(0.03)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	1.09***	0.91***	1.56***	1.62***	1.42***
	(0.25)	(0.27)	(0.29)	(0.29)	(0.32)
Observations	4871	4753	4561	4561	4540
Adjusted R^2	0.18	0.23	0.30	0.31	0.33
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample is restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.5: Minceria Wage Equation- OLS Results (A Narrower Range of Ages)

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.09***	0.08***	0.07***	0.07***	0.09***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Potential experience square	-0.00***	-0.00***	-0.00***	-0.00***	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.06***	0.07***	0.04**	0.04***	0.04***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Same sex	**-0.07*** **	**-0.08*** **	**-0.07*** **	**-0.07*** **	**-0.06*** **
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	1.56***	1.18***	1.77***	1.91***	1.85***
	(0.31)	(0.32)	(0.36)	(0.36)	(0.40)
Observations	2551	2489	2368	2368	2633
Adjusted R^2	0.15	0.21	0.29	0.29	0.30
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample restricted to women in PSID between age 21 and 35.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.6: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	0.06***	0.06***	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Potential experience square	-0.00***	-0.00***	-0.00**	-0.00**	-0.00*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.10***	0.10***	0.07***	0.07***	0.06***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same sex	**-0.04****	**-0.05****	**-0.04****	**-0.04****	**-0.05****
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	0.92***	0.50***	0.89***	1.03***	1.11***
	(0.18)	(0.19)	(0.26)	(0.27)	(0.28)
Observations	3732	3640	3444	3444	3422
Adjusted R^2	0.14	0.19	0.27	0.27	0.29
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample restricted to women in PSID between age 21 and 35.
I also restrict the analysis to the subsample of women who already have two children.
Sample includes observations after a third child while controlling for the number of children.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.7: Mincerian Wage Equation- A Robustness Check

	Wage (1)	Wage (2)	Wage (3)	Wage (4)	Wage (5)
Potential experience	1.63***	1.49***	1.41***	1.40***	1.24***
	(0.25)	(0.25)	(0.23)	(0.23)	(0.23)
Potential experience square	-0.07***	-0.06***	-0.07***	-0.06***	-0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Years of schooling	1.43***	1.45***	1.03***	1.07***	1.05***
	(0.16)	(0.17)	(0.18)	(0.18)	(0.21)
Same sex	**-0.71****	**-0.71****	**-0.72****	**-0.69****	**-0.80****
	(0.35)	**(0.34)**	**(0.32)**	**(0.32)**	**(0.32)**
Constant	-10.06***	-13.06***	-2.83	-1.42	-2.54
	(3.35)	(3.54)	(4.65)	(4.80)	(5.43)
Observations	4871	4753	4561	4561	4540
Adjusted R^2	0.18	0.22	0.28	0.29	0.33
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample restricted to women in PSID between age 21 and 35.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.8: Mincerian Wage Equation- OLS Results (Males)

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Age	0.10***	0.09***	0.09***	0.11***	0.07*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Age square	-0.00***	-0.00**	-0.00**	-0.00***	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.07***	0.07***	0.05***	0.05***	0.04***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same gender	**-0.00**	**-0.00**	**-0.00**	**-0.00**	**-0.00**
	(0.02)	**(0.02)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	0.02	0.22	0.76	0.73	1.40**
	(0.49)	(0.51)	(0.57)	(0.59)	(0.65)
Observations	4535	4391	4137	3829	3826
Adjusted R^2	0.13	0.16	0.19	0.21	0.24
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses

The Sample restricted to males in PSID between age 21 and 40.

I also restrict the analysis to the subsample of males who already have two children.

Sample is restricted to two years after the second childbirth and before a third childbirth.

Sample restricted to men whose education levels are higher than high school.

Sample restricted to individuals who work at least 20 hours per week.

Robust standard errors clustered at the individual level are reported in parentheses.

* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.9: OLS Mincerian Regressions - When Fertility Period Completed

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
Potential experience	-0.01	0.00	-0.01	0.01	0.01
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Potential experience square	0.00	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.14***	0.14***	0.11***	0.11***	0.10***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same sex	**0.04**	**0.03**	**0.03**	**0.03**	**0.03**
	(0.03)	**(0.03)**	**(0.02)**	**(0.02)**	**(0.02)**
Constant	0.55	0.28	0.67	0.51	1.31**
	(0.54)	(0.56)	(0.55)	(0.57)	(0.56)
Observations	6098	5934	5696	5696	5656
Adjusted R^2	0.14	0.18	0.26	0.28	0.30
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Family size	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The potential experience is defined as age minus years of schooling minus six.
The Sample restricted to women in PSID between older than 40 years of old.
I also restrict the analysis to the subsample of women who have two children.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.10: Probit Model of Labor Force Participation

	LFP (1)	LFP (2)	LFP (3)	LFP (4)
Age	-0.02	-0.04	-0.15	-0.17
	(0.07)	(0.07)	(0.12)	(0.14)
Age square	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.03	0.05**	0.01	0.03
	(0.02)	(0.02)	(0.04)	(0.04)
Same gender	**-0.08**	**-0.11***	**-0.04**	**-0.00**
	(0.06)	**(0.06)**	**(0.07)**	**(0.07)**
Constant	-0.13	-0.18	3.95*	4.10*
	(1.11)	(1.19)	(2.09)	(2.33)
Observations	8692	8420	6073	5969
Race	Y	Y	Y	Y
Year	Y	Y	Y	Y
Family size	Y	Y	Y	Y
State	N	Y	Y	Y
Occupation	N	N	Y	Y
Age at births	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a 3rd one.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.11: OLS Regressions of Hours Worked (per week)

	Hours (1)	Hours (2)	Hours (3)	Hours (4)
Age	-0.03	0.01	-0.06	0.19
	(0.41)	(0.43)	(0.42)	(0.46)
Age square	0.00	-0.00	0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Years of schooling	-0.23	-0.28**	-0.33**	-0.16
	(0.14)	(0.14)	(0.15)	(0.15)
Same gender	**0.45**	**0.39**	**0.23**	**0.30**
	(0.32)	**(0.32)**	**(0.31)**	**(0.31)**
Constant	36.20***	37.28***	40.58***	31.88***
	(7.67)	(7.52)	(7.65)	(7.93)
Observations	5557	5424	5153	5131
Adjusted R^2	0.03	0.05	0.07	0.08
Race	Y	Y	Y	Y
Year	Y	Y	Y	Y
Family size	Y	Y	Y	Y
State	N	Y	Y	Y
Occupation	N	N	Y	Y
Age at births	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.12: OLS Regression of Time Spent with Children (Hours per week)

	Hours (1)	Hours (2)	Hours (3)	Hours (4)	Hours (5)
Age	7.01	6.04	3.89	2.56	-4.38
	(4.26)	(4.57)	(8.53)	(9.35)	(35.04)
Age square	-0.09	-0.08	-0.04	-0.02	0.19
	(0.06)	(0.07)	(0.12)	(0.13)	(0.27)
Years of schooling	0.54	0.68	-0.26	0.27	1.21
	(0.85)	(0.93)	(2.01)	(2.26)	(2.62)
Same gender	**0.12**	**0.46**	**1.62**	**0.67**	**-0.98**
	(1.83)	**(1.98)**	**(4.86)**	**(5.45)**	**(9.31)**
Constant	-96.48	-83.49	-36.52	-20.77	7.26
	(72.81)	(79.39)	(147.43)	(163.75)	(713.93)
Observations	373	363	132	125	125
Adjusted R^2	0.13	0.11	-0.17	-0.30	-0.07
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Age of Child	Y	Y	Y	Y	Y
State	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in PSID between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than high school.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.13: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
AFQT	0.01***	0.01***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Experience	0.05**	0.04*	0.05**	0.06**	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Experience square	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.06***	0.06***	0.09***	0.09***	0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Same gender	**-0.04**	**-0.06**	**-0.04**	**-0.03**	**-0.04**
	(0.04)	**(0.05)**	**(0.04)**	**(0.04)**	**(0.04)**
Constant	2.05***	2.72***	2.05***	1.85***	1.58***
	(0.39)	(0.50)	(0.62)	(0.61)	(0.56)
Observations	1152	960	951	951	947
Adjusted R^2	0.19	0.21	0.32	0.34	0.39
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Region	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in NLSY79 and NLSY97 between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than 15 years of schooling.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.14: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
AFQT	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Experience	0.04***	0.03**	0.04**	0.05***	0.05***
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Experience square	-0.00	-0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.07***	0.07***	0.09***	0.09***	0.09***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Same gender	**-0.02**	**-0.03**	**-0.02**	**-0.02**	**-0.01**
	(0.03)	**(0.04)**	**(0.03)**	**(0.03)**	**(0.03)**
Constant	1.20***	1.25***	1.32***	1.33***	0.89**
	(0.16)	(0.21)	(0.29)	(0.30)	(0.36)
Observations	1845	1570	1549	1549	1544
Adjusted R^2	0.22	0.24	0.35	0.36	0.39
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Region	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in NLSY79 and NLSY97 between age 21 and 40.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than 14 years of schooling.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.15: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
AFQT	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Experience	0.02*	0.02**	0.03**	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Experience square	0.00	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.06***	0.06***	0.07***	0.07***	0.07***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Same gender	**-0.06**	**-0.08***	**-0.07***	**-0.07***	**-0.09****
	(0.04)	**(0.04)**	**(0.04)**	**(0.04)**	**(0.04)**
Constant	1.83***	2.43***	2.39***	2.36***	1.84***
	(0.36)	(0.38)	(0.50)	(0.52)	(0.51)
Observations	2331	2134	2104	2104	2100
Adjusted R^2	0.21	0.22	0.33	0.34	0.38
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Region	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in NLSY79 and NLSY97 between age 21 and 50.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than 15 years of schooling.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.16: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
AFQT	0.00***	0.00***	0.00**	0.00**	0.00**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Experience	0.02	0.02	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Experience square	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.07***	0.07***	0.07***	0.07***	0.07***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Same gender	**-0.08***	**-0.09***	**-0.11***	**-0.11***	**-0.13***
	(0.05)	**(0.05)**	**(0.04)**	**(0.04)**	**(0.04)**
Constant	1.22***	1.31***	2.35***	2.52***	1.80***
	(0.35)	(0.34)	(0.36)	(0.41)	(0.48)
Observations	1607	1600	1572	1572	1572
Adjusted R^2	0.16	0.19	0.34	0.35	0.40
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Region	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in NLSY79 and NLSY97 between age 40 and 60.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than 15 years of schooling.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$

Table B.17: Mincerian Wage Equation- OLS Results

	Log(w) (1)	Log(w) (2)	Log(w) (3)	Log(w) (4)	Log(w) (5)
AFQT	0.00***	0.00***	0.00***	0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Experience	0.02**	0.03**	0.03**	0.03***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Experience square	0.00	0.00	-0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Years of schooling	0.06***	0.06***	0.07***	0.07***	0.07***
	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Same gender	**-0.06**	**-0.07***	**-0.07***	**-0.07***	**-0.09****
	(0.04)	**(0.04)**	**(0.04)**	**(0.04)**	**(0.04)**
Constant	1.64***	2.21***	2.30***	2.31***	1.83***
	(0.39)	(0.40)	(0.51)	(0.55)	(0.52)
Observations	2660	2461	2425	2425	2421
Adjusted R^2	0.20	0.21	0.33	0.33	0.38
Age	Y	Y	Y	Y	Y
Race	Y	Y	Y	Y	Y
Year	Y	Y	Y	Y	Y
Region	N	Y	Y	Y	Y
Occupation	N	N	Y	Y	Y
Hours worked	N	N	N	Y	Y
Age at births	N	N	N	N	Y

Standard errors in parentheses
The Sample restricted to women in NLSY79 and NLSY97 between age 21 and 60.
I also restrict the analysis to the subsample of women who already have two children.
Sample is restricted to two years after the second childbirth and before a third childbirth.
Sample restricted to women whose education levels are higher than 15 years of schooling.
Sample restricted to individuals who work at least 20 hours per week.
Robust standard errors clustered at the individual level are reported in parentheses.
* $p < .1$, ** $p < .05$, *** $p < .01$