RIETI Discussion Paper Series 14-E-016

Fertility and Maternal Labor Supply in Japan: Conflicting policy goals?

Andrew S. GRIFFEN
University of Tokyo

NAKAMURO Makiko
Keio University

INUI Tomohiko
RIETI

Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
http://www.rieti.go.jp/en/

RIETI Discussion Paper Series 14-E-016

April 2014

Fertility and Maternal Labor Supply in Japan: Conflicting policy goals?*

Andrew S. GRIFFEN（University of Tokyo）

NAKAMURO Makiko （Keio University）

INUI Tomohiko （RIETI）

Abstract

Using panel data on Japanese mothers, this paper estimates the impact of fertility on maternal labor supply using twins as an instrument for the total number of children. We find that having twins actually has a longer term positive impact on maternal labor force participation. To understand this result, we present evidence that spacing effects and the cost of children are particularly salient in Japan and differ in important ways between twins and non-twin families of the same size. Implications for fertility and labor supply policy in Japan are discussed.

Keywords: Low fertility, Maternal labor supply, Multiple birth, Causal impact

JEL classification: J01, J08, J13, J22

1Andrew S. Griffen, University of Tokyo, Email: griffen@e.u-tokyo.ac.jp; Makiko Nakamuro,Keio University, Email: makikon@sfc.keio.ac.jp; Tomohiko Inui, Nihon University and RIETI, Email:inui.tomohiko@nihon-u.ac.jp. This study was conducted as part of a project entitled "Analysis on ServiceIndustries: Productivity, Economic Welfare, and Policy Evaluation" undertaken by the Research Institute of Economy, Trade and Industry (RIETI). We gratefully acknowledge the financial support from Grant-In-Aid for Scientific Research (A), titled "The Assessments of the Quality and the Productivity of Non-marketable Services" Research Representative: Takeshi Hiromatsu, No. 3243944. Masahisa Fujita,Masayuki Morikawa, Kyoji Fukao, Keiichiro Oda and Yoko Konishi provided extremely helpful comments and discussion that greatly improved the paper. The authors would also like to thank the Ministry of Health, Labour andWelfare for granting permission to use the Longitudinal Survey of Newborns in the 21st Century. Any remaining errors are our own.

I. Introduction

In 2012, newly elected Prime Minister Shinzo Abe of Japan began a series of audacious policy reforms in an attempt to reverse two decades of sluggish or non-existent economic growth in Japan. While most are familiar with the Keynesian macroeconomic policies of *Abenomics*, less well known are policies to tackle issues of low fertility and lack of support for working mothers that are also considered potentially harmful to Japan's long run economic future.[2] Even after a 16 year upward trend, Japan's current fertility rate of 1.39 births per 1000 women is among the lowest in the developed world and demographers are forecasting a decline of 27% of the Japanese population by 2050 (Takahashi et al., 2003). Such large changes in population put tremendous pressure on traditional pay-as-you-go social security systems, which has led to the development of policies to increase both fertility and maternal labor supply. However, a possible unintended consequence of higher fertility rates is lower maternal labor force participation, which is why many modern pro-natal policies, including those in Japan, often operate through maternity leave laws or child care supports for working mothers. Unfortunately, the design of these joint fertility-labor supply policies in Japan are made in an almost complete vacuum of information about the responsiveness of Japanese mothers' labor supply to fertility.

A major difficulty in understanding the impact of fertility on maternal labor supply is that both are endogenous decisions made by households so that identifying variables that exogenously affect fertility but not labor supply is challenging. Two kinds of clever natural experiments have been used to address this problem. The first, developed by Rosenzweig and Wolpin (1980a,b), is the occurrence of a twin birth, which under some assumptions corresponds to an unexpected exogenous increase in the number of children. The second is the gender composition of earlier born children, which uses the randomness of gen-

[2]A complete discussion of Prime Minster Abe's policy proposals is contained in a document entitled "Japan Revitalization Strategy - Japan is Back" disseminated through the Prime Minster and his Cabinet's Public Relations office. Specific policy proposals include expanding slots in child care centers and increasing maternity leave from 1.5 to 3 years. Recently these policies have received push-back from business groups so the current status of their implementation is uncertain.

der together with parental preferences for sex ratio balance (Angrist and Evans, 1998) or for males (Lee, 2008).[3] Most results from this literature find a sizable negative impact of number of children on maternal labor supply, which is consistent with and can potentially explain secular decreases in fertility and increases in female labor supply in the 20th century.[4]

One criticism of this literature is the lack of focus on formal models, which has obscured some of the strong and implicit assumptions that these IV methods require. Rosenzweig and Wolpin (2000) develop a model of maternal labor supply and fertility to show that the validity of twins and sex-preference instruments for fertility rests not only on the randomness of twinning or child gender but also on strong assumptions about the structure of mothers' utility functions, the impact of children's ages and the costs of children in directly affecting maternal labor supply decisions. For example, in the Rosenzweig-Wolpin model, the occurrence of twins can actually increase maternal labor supply even though previous estimates from the US typically have not found evidence of such effects. Such an effect can occurs if either the spacing of births directly affects labor supply through preferences or if child costs vary by the spacing of births and consumption and leisure are non-separable in the utility function. The results from the Rosenzweig-Wolpin model are quite intuitive. First, for some child expenditures, such as clothing, twins require two of

[3]Unlike the United States where preference for gender balance seems to exist among some subset of the population, sex-preference for sons is the so-called "Asian instrument" for fertility developed by Lee (2008) for Korea. For Japan, Kureishi and Wakabayashi (2011) present evidence of historical sex-preference for sons that faded out over time. We tried using the sex preference for males as an instrument but our first-stage results suggested that a first born daughter is not a strong predictor of additional fertility in Japan, which is consistent with the findings of Kureishi and Wakabayashi (2011) that sex-preference for sons is no longer a common feature of Japanese society.

[4]Because modern families are small by historical standards, the majority of twins occur at the first birth so that twins papers typical identify the effect on labor supply from increasing from one to two children. Angrist and Evans (1998) use the sex ratio of the first two children to identify the impact of going from two to three children and compare their estimates with the impact of a twins instrument. Lee (2008) also examines the margin between 1 and 2 children by using preferences for males to predict subsequent fertility. As compared to the twins estimators, the preference estimators require that preference for sex balance or for males be independent of preferences for work or leisure. In Japan, 95% of families with children have either 1 or 2 children so the margin between 1 and 2 children is important for policy. It is more challenging to identify effects on the margin between 0 and 1 children but identification strategies have looked at miscarriage (Hotz et al., 2005) and samples of women seeking fertility treatments (Cristia, 2008).

everything whereas for two siblings of different ages, especially those of the same gender, families can reuse items. Depending on preferences for consumption and leisure mothers may decide to work more if faced with higher child costs. Secondly, if there is a strong preference for mothers to stay home with young children, then an additional important difference between twin and non-twin families of the same size is that twins will enter school simultaneously. Whereas for two different aged siblings there will be always a trailing younger sibling that may give non-twins mothers an added incentive to stay home.

In this paper, we use the occurrence of a twins birth to investigate the impact of fertility on maternal labor supply using a long panel of Japanese families. There are several unique results in the Japanese case as compared to the previous literature. Our main finding is that while an exogenous increase in fertility as a result of twins initially lowers maternal labor supply, mothers with twins actually begin to supply *more* labor than non-twins mothers from around the time of school entry. This higher labor supply persists through the school age years so that on average twins mothers have significantly higher labor supply than non-twins mothers over the course of our panel. To understand this "S-shaped" impact of fertility on labor supply, we draw on the intuition of the Rosenzweig-Wolpin model and show two pieces of suggestive evidence to understand the result. First, even conditional on family size, the spacing of children in Japan, and specifically the presence of younger children in the household, strongly affects maternal labor. Holding family size and other observable variables constant, the presence of a child less than age six in the household increases the negative association between the number of children and maternal labor supply by approximately 50%. Furthermore, the timing of twins mothers re-entry into the labor market occurs around the school entry age, which is consistent with the back to school hypothesis. Second, we show that expenditures per child differ sharply by twins and non-twins families even conditional on the number of children and other observable characteristics. On average, twins families spend approximately ¥37,471 ($374) more per child per year than non-twins families with the same number of children. In addition, the pattern of child care expenditures also follows this S-shape through children's age with families spending more on older children than younger children, which again can potentially explain our labor supply findings.

Our paper makes two important contributions. First, to the best of our knowledge, this is the first paper to investigate the joint determination of fertility and labor supply in Japan. According to our findings, compared to other countries, the total number of children seems less important in affecting Japanese maternal labor supply decisions than child costs and spacing of children within households. Existing policies should specifically incorporate the importance of these costs when targeting maternal labor supply. Our paper also contributes to a broader literature on maternal labor supply in Japan. Given the evidence of a sharp decline in maternal labor supply after giving birth,[5] previous research using Japanese data has paid considerable attention to the determinants of maternal labor force around the timing of birth including investigating the impacts of maternity leave policies (Shintani, 1998; Maruyama, 2001; Senda, 2002; Shigeno and Matsuura, 2003; Imada and Ikeda, 2004), government sponsored job continuity subsidies (Asai, 2013), childcare costs (Oishi, 2003), co-residence with maternal parents or in-laws (Sasaki, 2002), and the formation of maternal work preferences based on own mother's employment status during childhood (Tanaka, 2008; Kawaguchi and Miyazaki, 2009).

Second, unlike most previous work on twins and labor supply that uses cross-sectional census data or repeated cross-sections at low frequency 10 year intervals, we use a yearly panel data set. This allows us to map out the effects over time for the same mothers and to avoid the potential for cohort bias. Interestingly the original Rosenzweig and Wolpin (1980a) paper found a pattern of impacts similar to our result with initial negative impacts on labor supply followed by positive and offsetting effects later in the lifecycle, which lent support to the Mincer (1962) hypothesis of females choosing lifetime labor supply.

[5]According to the Longitudinal Survey of Newborns in the 21st Century, which we use in our empirical analysis, only 25 percent of mothers remained in the labor force one year and half after childbirth while 55 percent of them were in employment one year before childbirth in Japan.

Subsequent research in the US has only found negative impacts, which fade over time as children age (Bronars and Grogger, 1994; Jacobsen et al., 1999; Cáceres-Delpiano, 2006; Vere, 2011). One possible explanation is that whereas Rosenzweig and Wolpin (1980a) use a retrospective panel, more recent papers use cross sectional census data. In fact, Bronars and Grogger (1994) highlight the possibility of misleading cohort effects by comparing cohorts across 10 year census waves to synthetic cohorts from a single cross section. However, it appears that the difference in these findings has not been investigated. A second possible explanation for the difference of our results compared to the US is the role of differences in maternal preferences and child cost differences in Japan. Future research on international comparisons on the impacts of spacing and child expenditures on maternal labor supply could potentially shed light on these issues.

The main finding of this research, the S-shaped causal impact of fertility on maternal labor supply, has profound implications for fertility and labor supply policy in Japan. In particular, it indicates that the widely accepted simple negative relationship between fertility and maternal labor supply does not characterize the trade-off faced in Japan. Our empirical findings imply that current policy agendas to both increase the fertility rate and maternal labor force participation in Japan are actually *not* in conflict.

The rest of the paper is organized as follows. In section II, we develop a model of fertility and labor supply à la Rosenzweig and Wolpin (2000). Section III describes the data, section IV presents the econometric model, and we discuss the empirical results in section V. Finally, in section VI, we conclude.

II. Model

In this section, we develop a version of the labor supply-fertility model from Rosenzweig and Wolpin (2000) (henceforth RW) that has features applicable to our data set and estimation. The motivation of their paper is the idea that many natural experiment based IV strategies rely not only on random events but also on strong implicit assumptions that are not clearly demonstrated without using an economic model. Then by specifying such a model, RW highlight the several channels through which, in this case, twins, can directly influence maternal labor supply independently of the number of children. The main conclusion is that many natural experiments, including twins based estimators, put strong assumptions on both preferences and costs. We provide empirical evidence that several of those assumptions are unlikely to hold in the Japanese case.

Consider a three period model where a mother either has a singleton birth or twins in period 1, may or may not have a birth in period 2 and then makes a labor supply decision in period 3. Let the number of kids of age a be given by K_a and let the stock of children at time t be given by $N_t = \sum_{i=1}^{t} K_i$. Suppose the mother has utility over consumption, c, the number of children, and leisure. We also allow the utility of leisure to interact with consumption and both the age and the number of children. We focus on the labor supply decision in period 3 where the mother needs to make a binary hours of work decision, $h_3 \in \{0,1\}$. Assume the specification of the mother's utility is given by:

$$\begin{aligned} u_3 =& c_3 + \alpha_1 N_3 - \alpha_2 N_3^2 + \alpha_3(1-h_3) + \alpha_4 c_3(1-h_3) + \\ & (\alpha_5^1 K_1 + \alpha_5^2 K_2 + \alpha_5^3 K_3)(1-h_3) + \alpha_6 N_3(1-h_3) \end{aligned}$$

In particular, we highlight that consumption and leisure may be compliments or substitutes according to the sign of α_4 and that the utility of staying home may depend on the number of children of age a through the α_5^a parameters. The budget constraint is:

$$c_3 = y_3 + w_3 h_3 - p_3 K_3 - p_2 K_2 - p_1 K_1$$

where y_3 is the woman's non-labor income, w_3 is the mother's wage if she works, and p_a is the per child cost for the number of children of age a. The mother's state space at period 3 consists of the age distribution of her children $S_3 = \{K_3, K_2, K_1\}$. Plugging in the budget constraint, the mother's utility if she works is:

$$
\begin{aligned}
u_3(h_3 = 1|S_3) =& c_3 + \alpha_1 N_3 - \alpha_2 N_3^2 \\
=& y_3 + w_3 - p_3 K_3 - p_2 K_2 + \alpha_1 N_3 - \alpha_2 N_3^2
\end{aligned}
$$

and mother's utility if she does not work is:

$$
\begin{aligned}
& u_3(h_3 = 0|S_3) = \\
& c_3 + \alpha_1 N_3 - \alpha_2 N_3^2 + \alpha_3 + \alpha_4 c_3 + \alpha_5^2 K_2 + \alpha_5^3 K_3 + \alpha_6 N_3 = \\
& y_3 - p_3 K_3 - p_2 K_2 - p_1 K_1 + \alpha_1 N_3 - \alpha_2 N_3^2 + \alpha_3 + \alpha_4 (y_3 - p_3 K_3 - p_2 K_2 - p_1 K_1) + \\
& \alpha_5^2 K_2 + \alpha_5^3 K_3 + \alpha_6 N_3
\end{aligned}
$$

Differencing the two utilities, the mother works if the following difference is positive:

$$
\begin{aligned}
& \Delta u_3(S_3) = u_3(h_3 = 1|S_3) - u_3(h_3 = 0|S_3) = \\
& \left[y_3 + w_3 - p_3 K_3 - p_2 K_2 + \alpha_1 N_3 - \alpha_2 N_3^2 \right] - \\
& \Big[y_3 - p_3 K_3 - p_2 K_2 - p_1 K_1 + \alpha_1 N_3 - \alpha_2 N_3^2 + \alpha_3 + \alpha_4 (y_3 - p_3 K_3 - p_2 K_2 - p_1 K_1) + \\
& \alpha_5^2 K_2 + \alpha_5^3 K_3 + \alpha_6 N_3 \Big] = \\
& w_3 - \alpha_3 - \alpha_4 y_3 + \alpha_4 p_3 K_3 + \alpha_4 p_2 K_2 + \alpha_4 p_1 K_1 - \alpha_5^2 K_2 - \alpha_5^3 K_3 - \alpha_6 N_3
\end{aligned}
$$

Now take two identical mothers who have different numbers and ages of children because of the birth of twins. Any difference in the two mothers labor supply can be captured by differencing the work decision equation for different values of S_3, which gives:

$$
\begin{aligned}
& \Delta u_3(\tilde{S}_3) - \Delta u_3(\bar{S}_3) = \\
& \alpha_4 p_3 (\tilde{K}_3 - \bar{K}_3) + \alpha_4 p_2 (\tilde{K}_2 - \bar{K}_2) - \alpha_5^3 (\tilde{K}_3 - \bar{K}_3) - \alpha_5^2 (\tilde{K}_2 - \bar{K}_2) - \alpha_6 (\tilde{N}_3 - \bar{N}_3)
\end{aligned}
$$

This equation illustrates the channels through which differences in fertility can influence differences in labor supply among otherwise identical mothers. Specifically, the costs of age a children, p_a, interact with the leisure-consumption interaction parameter, α_4, through any difference in the number of age a children, $\tilde{K}_a - \bar{K}_a$. Direct preferences to stay home with children of age a enter through α_5^a and the total number of children may also directly affect labor supply through α_6.

Ignoring the birth decision in period 3, consider three birth sequences at time period 3: (1) twins in period 1 and no birth in period 2 ($K_3 = 2, K_2 = 0$) (2) a singleton birth in both period 1 and period 2 ($K_3 = 1, K_2 = 1$) and (3) a singleton birth in period 1 and no birth in period 2 ($K_3 = 1, K_2 = 0$).[6] RW raise the interesting interpretation of twins as a timing failure in fertility. For example, suppose that a family wanted to have two children but had twins at the first birth. Their preferences for total fertility would be satisfied but their preferences for birth spacing may not be satisfied. A different family with twins could have preferred to have only one child in which case twins would be an excess of fertility. Comparing birth sequence (1) to (2) corresponds to the former case and comparing birth sequence (1) to (3) corresponds to the latter case. Of course, unless asked, the distribution of expected fertility is unknown so the twins estimator is a mixture over the two differences.

Now consider differences in the labor decision equation by twins versus non-twins in period 1 for the birth sequence (1) versus (2):

$$\Delta u_3(K_3 = 2, K_2 = 0) - \Delta u_3(K_3 = 1, K_2 = 0) =$$
$$\alpha_4 p_3 - \alpha_5^3 - \alpha_6$$

[6]Our data are conditional on having a birth so all mothers must have at least one child in period 1. Other possibilities are to have twins in period 1 and a singleton baby in period in period 2. In addition, twins and non-twins mothers from period 1 could also have twins in period 2. In total there are 6 cases but the three cases above should comprise the majority of birth sequences given relatively low Japanese fertility. In addition, 95% of Japanese households have at most 2 children.

which indicates that the difference in labor supply between twins and non-twins mothers is determined by the price of children in period 3 interacted with the consumption-leisure interaction term, the utility or disutility of staying home with K_3 children and the utility of staying home with more total number of children. If α_6 is positive, then this answers the classic question of whether more children causes less labor supply. However, notice that the twins based estimators conflates the α_6 parameter with $\alpha_4 p_3$ and α_5^3. If either of those is large and positive, then the occurrence of twins could actually *increase* labor supply if those effects are sufficient to outweigh α_6.

The second birth sequence compares first birth twins with two successive singleton births:

$$
\begin{aligned}
&\Delta u_3(K_3 = 2, K_2 = 0) - \Delta u_3(K_3 = 1, K_2 = 1) = \\
&\alpha_4 p_3 - \alpha_4 p_2 - \alpha_5^3 + \alpha_5^2 \\
&\alpha_4(p_3 - p_2) - \alpha_5^3 + \alpha_5^2
\end{aligned}
$$

in which case the number of children term cancels out because the families are the same size. The difference in child costs over time, $p_3 - p_2$, interacts with α_4 and the utility of staying home α_5^3 for older children lowers the relative labor supply of twins mothers while the α_5^2 increases the relative labor supply of twins mothers. Because the families are the same size, this case stresses the importance of differences in child costs and preferences for leisure. Specifically, if α_4 is positive and child costs increases in children's age so that $p_3 - p_2$ is positive, then twins could increase the likelihood of maternal labor supply. In addition, if period 3 corresponds to the school entry age and mothers have a strong preference to stay home with younger children, then α_5^3 may be small and α_5^2 may be large, which would also generate incentives for twins mothers to work more than non-twins mothers.

The model highlights that even if twins occur randomly with respect to mothers' wage or unobserved tastes for leisure that twins do not correspond solely to an increase in the number of children but also plausibly affect labor supply through other channels that we highlight; namely differences in child costs and maternal preferences that may vary by children's age. In our empirical work, we provide evidence on the role of such channels in affecting labor supply in our data set.

III. Data

In this paper we use data from the *Longitudinal Survey of Newborns in the 21st Century*, an ongoing, annual panel data collected by the Japanese Ministry of Health, Labour and Welfare currently in its 10th wave. The sampling frame targeted all 53,575 Japanese babies born between January 10-17 and July 10-17, 2001.[7] Respondents were primary caregivers, usually mothers, and they answered a battery of questions about the household roster, including previous numbers of children and their ages, paternal and maternal labor supply, household earned and unearned income, child development, time use and questions about expectations for children.[8] A nice feature of our data set is that multiple births, such as twins and triplets, are identified by matching information directly from birth records. This is an improvement on the previous literature using US Census data that infers multiple births indirectly from household rosters and reported information on children's year or quarter of birth. In our data set, due to the small number of triplets, we excluded them from the analysis. Table 1 presents information on the sample size, response rates and attrition rates by survey round. From the initial survey response rate of 75%, the current response rate after 10 waves is still higher than 90% with a total sample attrition rate of 34.3%. The attrition rate is in line with other commonly used panel data sets, such as the Panel Study of Income Dynamics, which had a 10 year attrition rate of 31.8% (Fitzgerald et al., 1998).

[7] A potential concern with such a sampling strategy might be seasonality effects in births. Nakamuro et al. (2013) checked for seasonality effects using monthly Japanese Vital Statistics collected by the Ministry of Health, Labour and Welfare and found no evidence of seasonal patterns of births in Japan.

[8] A drawback of this data is some important questions are missing across rounds. For example, parental employment status was not asked in wave 3.

Rosenzweig and Wolpin (1980a) discuss several issues related to using the twins instrument. First, the probability of ever having twins will statistically increase with the number of children as there are more chances to ever have had a twin. Using the entire set of twins would potentially result in a select sample of mothers with preferences for larger families, which may not be exogenous to unobserved determinants of labor supply. Second, the twinning probability seems to, for biological reasons, increase with with both parity and with the age of mother. They address these issues by conditioning on first born twins and including age of the mother as a control variable. We follow their lead and in our main analyses we condition on the sample of first born twins and include controls for mother's age at first birth.

We present summary statistics from the data in Tables 2a and 2b. The main variables of interest for our study are maternal employment, number of children and the occurrence of twin births. Table 2a shows that maternal employment is near 24.8% in the first year after the child is born in 2001 and gradually climbs to 61.9% by the 10th survey round. In the 10th round mothers have on average 1.97 children and 21% of couples live with one or more their parents in the house. The even lower aggregate fertility rate cited in the introduction can be understood through low Japanese marriage rates and the rarity of out of wedlock births in Japan. For example, our data show that 97.5% of fathers live in the household at the time of birth in 2001.

Table 2b shows several time invariant and baseline characteristics of the sample both on average and decomposed by twins and non-twins mothers. We focus on three estimation samples. First, our main estimation sample is the set of mothers with a first birth in 2001. Among this sample, twins mothers are 2.3 years older on average than non-twins mothers, which is consistent with previous discussion that twinning seems to increase with maternal age and which is why we add maternal age at birth to all our regression equations. The twins and non-twins mothers look statistically identical on educational attainment, % of fathers working and % of fathers in the household in 2001. The survey also asked mothers whether or not they were working one year before the child or twins were born, which we define as ”% mothers working baseline” in Table 2b. In our main estimation sample of mothers with a first birth in 2001, 75% of non-twins mothers report working whereas only 64% of twins mothers report working. This difference is statistically significant. Because labor supply is our main outcome of interest an important concern is obviously whether twins mothers differ importantly from non-twins mothers in our sample for a non-random reason. The main hypothesis would be that perhaps twins mothers may have use assisted reproduction that artificially increased the twinning rate and that these mothers differ importantly in propensity to work.[9] To address this issue, we explore several strategies. First, we present direct evidence on the low incidence of assisted reproduction in Japan. Second, we show that the twinning incidence in our data is line with historical twinning rates in Japan, which suggests that twinning in our data is unlikely to primarily artificial. Third, we explore the sensitivity of our main results to changes in the estimation sample, which we also present in Table 2b. We discuss each below.

A potential concern with use of twins data from more recent years is the increasing use of assisted reproductive technology (ART), which have higher rates of twinning than traditional reproduction.[10] In Figure 1, we plot data released by the Japan Society of Obstetrics and Gynecology that shows the total number of children born in Japan broken down into the three most common methods of ART; in-vitro fertilization (IVF), intracytoplasmic sperm injection (ICSI) and frozen embryo transfer (FET).[11] The graph also shows the percentage of the total births in Japan from 1995 - 2010 that used ART. For example,

[9]Another hypothesis to explain this result is that the first survey was conducted when the children were already 1.5 years old so there is a potential for recall bias for this question especially if twins mothers were encouraged to work less during their pregnancies and mothers conflate their work status the year before birth and during their pregnancies.

[10]An analogous issue to ART for the gender preference instrument is the development of technology to actually choose the gender of the child. Such a technology currently exists and it may render these instrumental variables strategies unusable, unless data are also collected on the use the reproductive technologies. For work analyzing the impact of gender selection technology on fertility see Li and Pantano (2013). Sex selective abortion also may play a role in invalidating gender preference based instruments but such effects are perhaps unlikely to operate in developed countries.

[11]These IVF data come from the Japan Society of Obstetrics and Gynecology (2011) and can be found at http://plaza.umin.ac.jp/ jsog-art/2011data.pdf

in 2001, the year the children in our data were born, the total number of children born from ART is only 1.12% of total births. The graph shows increasing use of ART over time but that the rates of ART use compared to total fertility still remains low.

To understand the potential for ART twins to bias our sample, we also need information on the twinning probability among ART births. The Japan Society of Obstetrics and Gynecology provides an estimate that approximately 10% of assisted reproductive births result in twins, which is obviously higher than the 2% of children who are twins in our data.[12] Therefore, an approximation of the percentage of twins conceived using reproduction assistance is 10% of twins out of 1.12% of ART births, or 0.11% of all births were ART twins. Given that 2% of our estimation sample are twin births, these results suggest that 5.5% of the sample twins in our data, or approximately 20 children or 10 twins pairs/mother observations may have resulted from assisted reproduction, which is likely to be negligible.

A second and more indirect test is to compare the twinning rate of 7.882 twin pairs per 1000 maternities reported in Table 2b in our main estimation sample to estimates of the historical twinning rate from Japan and other countries. Basically the result is in line with previous estimates of twinning rates of around 10 per 1000 maternities (Bortolus et al., 1999) and, interestingly, the Japanese, and Asians more generally, apparently have lower rates of twinning than other ethnic groups, which we also find support for in our data. Imaizumi (2003) provides a time trend of twinning rates in Japan with an increasing trend from 1972 to 1999. The twinning rates fall with the range report in that work. This increase in the twinning rate could obviously be driven by increased use of ART but one counter piece of evidence is that studies of Japanese immigrants to California have found *higher*

[12]Twinning rates are usually expressed as the number of maternities per 1000 in which twins occur, which we used previously to compare the twinning rates in our data to the medical literature. However, in the IVF data, we have the *total* number of reproductive assisted babies born, which means that twins pairs will be counted twice, once for each baby. Therefore, when comparing the twinning rate with the IVF data, we need to use the percentage of newborns that are twins, which is just slightly less than twice the twinning rate because when converting from a twinning rate to percentage of twin the numerator doubles but the dominator also increases additively by the number of twin pairs.

rates of twinning as compared to non-immigrant Japanese (Pollard, 1995). Although it is possible that Japanese immigrants differs innately in their propensity to have twins, genetic researchers favor the hypothesis that environmental factors, such as diet, can influence the twinning rate. To the extent that dietary factors have changed over time this could also have led to the increased in twinning rate observed in the data.[13]

Although we do not find much ancillary support for the hypothesis that ART importantly affects our sample given its limited use in Japan and twinning rates in our data well within normal ranges, we address the concern about the lack of baseline equivalence through a series of robustness checks on our main regression results. The first is to look at only mothers under the age of 35, which we also display in Table 2b. These mothers also look different in terms of baseline labor force participation by twins status even though these mothers are unlikely to have used fertility treatments. The second strategy is to expand the estimation sample to the full sample of twins. Although potentially this results in a biased sample because the probability of ever having twins increases in the number of children, we see that in Table 2b this larger group of mothers actually does not differ in baseline labor force participation by twins; 54% of twins mothers working versus 55% non-twins working. The lower overall labor supply of the full sample captures that some of these women already have children and therefore are less likely to be working in prior to birth in 2001. Finally, we also condition our main empirical results on having worked at baseline. However, as we show in our results section, none of these affect our main findings in any important way.

IV. Econometric Model

[13]Besides ART, another potential concern with using twins data is the use of fertility drugs, which also increase the twinning. In fact, Ooki (2010) provides evidence that 50% of current multiple births in Japan are iatrogenic. This large increase in the twinning rate is primarily due to changes in the twinning rate among women 35 to 39 years of age. In our robustness checks, we estimate the model on exactly this sample of women less than age 35, who are much less likely to have used fertility drugs, to reduce the probability of bias.

We use the following linear probability model for maternal employment:[14]

$$Y_{ia} = C_{ia}'\beta_a + X_i'\beta_1 + W_{ia}'\beta_2 + \varepsilon_{ia}$$

where the labor supply decision, Y_{ia}, for mother i at child's age a is a function of the number of children, C_{ia}, time invariant family characteristics, X_i, and time varying characteristics, W_{ia}, time invariant characteristics, W_{ia}, and a shock, ε_{ia}. The main coefficient of interest is β_a and how this parameter changes over the child's age. A typical concern is that the number of children C_a is correlated with unobserved determinants of maternal labor force participation ε_{ia}. Provided that the occurrence of twins is unrelated to maternal preference or opportunities for employment, this variable can be used as an instrument for number of children in the following first stage regression:

$$C_{ia} = T_0'\alpha_a + X_i'\alpha_1 + W_{ia}'\alpha_2 + u_{ia}$$

where T_0 is an indicator for a twin birth in 2001. To consider the effects over time, we follow Vere (2011) and interact family size and the occurrence of twins with dummy variables for calender time to map out the impact of the effect of an exogenous increase in the number of children over time. This creates an exactly identified model with the number of excluded regressors equal to the number of endogenous regressors, namely number of children at each time period.

V. Results

Our first stage results are presented in Table 3 with the "Twins in 2001" indicator variable as the key excluded instrument. The results show that the birth of twins in 2001 is a strong predictor of subsequent total number of children through round 10 with F-stats

[14]We also tried estimating probit models for maternal labor supply and found the same basic pattern as the linear probability model estimates. We follow the previous literature in using linear probability models.

above 1000 in all rounds.[15] Moving to the OLS and IV estimates in Table 4, we focus on maternal labor supply as an outcome variable, which is coded as 1 if the mother was employed or self-employed and 0 otherwise.[16] The potentially endogenous independent variable of interest is the number of children in the household. Because mothers may change their labor supply behaviors as the children age, we interact the numbers of children and dummies for the survey year, which is the "round" variable in Table 3. The OLS results show a consistent negative association between the number of children and maternal labor supply. In addition, the negative coefficient on having an additional child initially becomes larger in absolute value with years but begins to decrease after round 5. At its peak in round 4, the magnitude of the OLS coefficient implies that maternal labor force participation is approximately 16 percentage points lower per child. By round 10, however, the OLS coefficient shrinks to 7 percentage points lower per child. If the the OLS estimates represented causal effects, summing the OLS coefficients over all rounds suggests that the cumulative impact of having an extra child is 1.07 less years of work over a 10 year time horizon. The impact is statistically significantly different from zero.

Table 4 also presents IV estimates in the second column, where the regression is calculated using a twins birth as a natural experiment to identify the effect of an exogenous increase in the number of children on maternal labor supply. The excluded instrumental variable is defined as equal to 1 if a mother experienced a twin birth during her first birth in 2001. We interact this twins variable with year to identify the impacts over time. The IV column of Table 4 reveals that, overall, the effect on maternal labor supply of the exogenous fertility shock caused by the occurrence of twin birth in 2001 is quite small in

[15] In round 1, "Twins in 2001" exactly predicts having two children because we condition our main analysis on first having a birth in 2001 so that mothers either had 1 or 2 children in 2001. Then this "first-stage" regression for Round 1 is equivalent to including the "Twins in 2001 X Round 1" directly in the second stage regression. Such an effect does not occur when we estimate the model on the expanded sample because some mothers already have children at Round 1 so that the twins instrument does not perfectly predict number of children.

[16] Our data also contain a variety of other information on maternal labor supply including whether in the labor force and hours worked in typical week. We are currently analyzing the robustness of our results to other definitions of maternal labor supply. The results seem to be robust and will be provided upon request.

magnitude and much smaller than the OLS estimates. The more striking result is that, while the impact on labor supply of having an unexpected additional child is negative for the first four years, the impact then actually becomes *positive* and significant. Figure 2 gives a graphical comparison of the OLS and IV estimates over time and illustrates the S-shape of the IV results. One interesting feature of Figure 2 is that the IV estimates show increases in labor supply around the school entry age.

There are two statistical tests of interest for the IV results. The first is whether the IV estimates differ statistically from the OLS estimates. The Durbin-Wu-Hausman test in Table 4 has an extremely low p-value and suggests that the answer is yes. The second is whether the IV estimates differ from a zero impact of number of children on maternal labor supply. Although point-wise the IV estimates are not statistically significantly different from zero at all rounds, when we sum the IV coefficients across rounds, which corresponds to the cumulative effect on labor supply of one additional child over 10 years, we find an overall positive effect of 0.24 additional years of work per child. A χ^2 test suggests this sum is unlikely to be statistically significantly different from zero by chance.

Before proceeding to a discussion of the interpretation of these results, we first investigate the robustness of our main regression results. As discussed previously, a potential concern with our main estimation sample is the lack of equivalence of the mothers in baseline labor supply. Setting aside the reliability of the question, which we have discussed, we also present a series of robustness checks in Table 5 to try to understand whether our main IV results hold using different estimation samples. Although theoretically twinning is random conditional on maternal age, if twinning rates were higher because of fertility treatments this would be unlikely to hold. One strategy, which we pursue in the first columns (1) and (2) of Table 5, is to add additional controls for maternal education, whether the husband lives in the house, whether the grandparents live in the house and husband's income to the main IV regressions.[17] The S-shaped pattern of our results continues to hold

[17]Fathers' income information is missing in rounds 6, 8 and 9 but for the other rounds the point-wise estimates become more precise when we add fathers' income as a control. The same patterns of higher

even adding these controls. In column (3), we further restrict the sample to only mothers who were younger than age 35 in 2001 given the evidence that younger mothers are much less likely to have used any kind of fertility treatment. Again, the results look similar. We also tried in column (4) only estimating the model on mothers who were working in 2000. Interestingly, the positive impacts become larger in magnitude, which suggests that mothers who had some previous connection to the labor market may be particularly influenced by twins to re-enter the labor market as the twins enter school. Finally, we expand the sample to the full set of twins in column (5). This strategy raises the issue that the probability of ever having had twins will increase in the number of children. This concern is balanced, however, by the baseline equivalence in labor supply among this larger set of mothers. Again, these results look quite similar to our main results. The hypothesis of difference from the OLS results holds in every specification and the hypothesis of the cummulative IV effects different from zero holds in every specification except for one. The magnitude of the 10 year positive impacts of fertility on labor supply range from 0.12 to 0.39 years, which is in line with the 0.24 additional years of labor supply result found in Table 4.

As compared with the results from the US, the results from Japan indicate that Japanese twins mothers re-enter the labor market much earlier than their US counterparts. Papers such as Angrist and Evans (1998), Jacobsen et al. (1999), and Vere (2011) find that the effect of an exogenous increase in fertility through twins has a negative impact on maternal labor supply that does not fade out until 13 years after the twins are born. Vere (2011) hypothesizes that mothers may find more time to work after children enter junior high school and the expiration of tax breaks for child care expenses when children turn age 13 may play a role in the timing of the fade out. Jacobsen et al. (1999) find similar results with consistently negative impacts for children less than age 10 but positive impacts when the child is between 11 through 18 year of age. So while there is a consistent cross study evi-

maternal labor supply in the IV estimates continues to hold.

dence that 10-13 years old is a threshold for mothers in the US to change their labor supply behavior, the occurrence of twins in Japan appears to have a completely different pattern with longer term *positive* impacts on maternal labor supply. While the causes of these international differences are beyond the scope of this paper, we attempt to understand this interesting result in Japan by drawing on the intuition of the model developed in section II where we showed that twins may directly impact maternal labor supply through channels other than the number of children. Specifically, in our data, we demonstrate that the two channels identified in the economic model, spacing effects and child rearing costs, seem to differ empirically in important ways between twins and non-twins families in Japan. In addition, the patterns of these preferences and costs also display an S-shape through children's age, which lends additional suggestive evidence to our labor supply findings.

Table 6 shows estimates of labor supply models with both the number of children and the number of children less than age six entering the model. The results show a strong negative impact of the number of children with each additional child associated with 7.3 percentage points lower labor force participation in the full model, which parallels the OLS results previously presented. In addition, holding the number of children and other variables constant, the results in column (3) show that shifting one child to be less than age six increases the negative association of fertility with lower labor supply by approximately 49%. This is in direct violation of one of the identifying assumptions in the Rosenzweig-Wolpin model that spacing of children does not directly affect maternal labor supply. This is consistent with much anecdotal evidence in Japan on pervasive social norms for mothers to stay home with young children and in particular the Japanese "with three-year-old child myth," which is the idea that it is better for childrens development if mothers raise their children at home until age three.[18] Another piece of evidence is that while 60.3% of

[18]These ideas about child raising seem to have become widely accepted during the 1970s (Asahi News Paper, July 17th 2013). According to a survey conducted by the Cabinet Office in 2006, this view is still dominant among parents: approximately 70% of parents with children aged 9-15 years agreed that mothers should focus on raising their children at home until age three. Reflecting exactly these views, Prime Minister Abe requested companies to offer three years of maternity leave in 2013.

mothers with at least one child aged less than six are employed in US, only 42.8% of such mothers are employed in Japan (Bureau of Labor Statistics 2011; Ministry of Internal Affairs and Communications, 2011). This US-Japan difference in maternal labor supply patterns of mothers with young children may partly explain why the spacing of children more strongly affects maternal supply in Japan than in the US and generate different results in the IV estimates. It also highlights an important difference between twin and non-twin families of the same size: twins will enter school simultaneously while for non-twins families with two different aged siblings there will always be a trailing younger sibling that will give non-twins mothers an added incentive to stay home. As the model in section II demonstrated, this channel can actually serve to increase labor supply for twins mothers if mothers have preferences to stay home with younger children, which empirical and anecdotal evidence suggest is common practice in Japan.

In Table 7, we estimate regressions of "per child expenditure" as a dependent variable and the number of children and an indicator for twins as the key independent variables. We also include controls for mother's age at first birth, mother's education, whether the father lives in the house, whether grandparents live in the house and both paternal and maternal income. The coefficient on the number of children is negative and statistically significant and suggests that each additional child is associated with ¥3131 ($\approx$ \$31.31) less spent per month per child. However, the coefficient on whether the child has a twin is positive and statistically significant across all specifications. Given that twinning is plausibly random conditional on both first birth and maternal age at first birth, the result implies that families with twins spends approximately ¥2942 to ¥3512 ($\approx$ \$29.42 - \$35.12) more per child per month than non-twins families. This result shows that child costs differ importantly between twins and non-twins families, which is again in direct violation of one of the identifying assumptions developed in the model section for twins to not affect labor supply directly. The coefficients on the round indicators in Table 7 also show that per child expenditures are higher when children are in school and displays the same S-shape pattern as maternal labor supply. If twins mothers face higher per child expenditures for twins, then the Rosenzweig-Wolpin model suggests that twins mothers may display quite different labor supply patterns independently of the number of children, including the possibility of the S-shaped labor supply patterns that we find in our study.

Whether these costs differences can potentially drive the cross country differences in labor supply patterns of mothers is a much larger research question but some supporting evidence from the OECD is displayed in Table 8. Specifically Table 8 looks at private educational expenditures in the OECD, the US and Japan and shows that Japan is one of the countries where the share of private spending on education is substantial and much higher than the US, especially during primary school. Compared with US and other high-income peer countries, the direct cost to families of raising young children in Japan seems to be extremely high. Competitive entrance examinations into high school and college encourage many parents to purchase supplementary education, such as cram schools, which may substantially increase the burden on parents of educational costs and may provide an additional incentive for mothers in Japan to re-enter the labor market around the time of children's school entry.[19]

VI. Conclusions

The goal of this paper is to study the effects of fertility on maternal labor supply in Japan where issues surrounding fertility and the dynamics of mothers in the work force are currently important policy issues. Using a twins estimation procedure to untangle the joint determination of fertility and labor supply through, we found an interesting and counter-intuitive result: the occurrence of twins in Japan actually *increases* maternal labor supply in the long run. To understand this result we drew on the intuition of a model developed by Rosenzweig and Wolpin (2000) that demonstrates that twins based estimators may pick

[19] According to Japanese Ministry of Education, Culture, Sports, Science and Technology, 41.2%, 70.8% and 64.0% of household expenditures on education are paid for activities outside of school, mainly supplementary education, for public elementary, middle and high school students respectively.

up not only direct effects of additional exogenous fertility on labor supply but also indirect effects through differences in preferences for leisure and child rearing costs that vary by children's ages and may differ importantly between families with twins and non-twins. Our rich panel data provide indirect evidence that those assumptions are likely to violated in Japan and supporting data sets point to important cross country differences in labor supply patterns among mothers with young children. Our main policy conclusion is that fertility and maternal labor supply are not necessarily in conflict in Japan but that labor support policies may face strong preference based resistance, the timing of costs has an important role in maternal labor supply and that policies that assist in child rearing expenses may actually serve to reduce maternal labor supply.

References

Joshua D Angrist and William N Evans. Children and their parents' labor supply: Evidence from exogenous variation in family size. *The American Economic Review*, 88(3): 450–477, 1998.

Yukiko Asai. Parental leave reforms and the employment of new mothers: Quasi-experimental evidence from japan. *Working Paper, Institute of Social Sciences, University of Tokyo*, 2013.

Renata Bortolus, Fabio Parazzini, Liliane Chatenoud, Guido Benzi, Massimiliano Maria Bianchi, and Alberto Marini. The epidemiology of multiple births. *Human Reproduction Update*, 5(2):179–187, 1999.

Stephen G Bronars and Jeff Grogger. The economic consequences of unwed motherhood: Using twin births as a natural experiment. *The American Economic Review*, pages 1141–1156, 1994.

Julio Cáceres-Delpiano. The impacts of family size on investment in child quality. *Journal of Human Resources*, 41(4):738–754, 2006.

Julian P Cristia. The effect of a first child on female labor supply: Evidence from women seeking fertility services. *Journal of Human Resources*, 43(3):487–510, 2008.

John Fitzgerald, Peter Gottschalk, and Robert Moffitt. An analysis of sample attrition in panel data: The michigan panel study of income dynamics. *Journal of Human Resources*, pages 251–299, 1998.

V. Joseph Hotz, Susan Williams McElroy, and Seth G Sanders. Teenage childbearing and its life cycle consequences exploiting a natural experiment. *Journal of Human Resources*, 40(3):683–715, 2005.

Sachiko Imada and Shingo Ikeda. Expand supports for the balance of work and child rearing. *JILPT Discussion Paper Series 04-12*, 2004.

Y Imaizumi. A comparative study of zygotic twinning and triplet rates in eight countries, 1972-1999. *Journal of biosocial science*, 35(2):287–302, 2003.

Joyce P Jacobsen, James Wishart Pearce III, and Joshua L Rosenbloom. The effects of childbearing on married women's labor supply and earnings: using twin births as a natural experiment. *Journal of Human Resources*, pages 449–474, 1999.

Daiji Kawaguchi and Junko Miyazaki. Working mothers and sons preferences regarding female labor supply: direct evidence from stated preferences. *Journal of Population Economics*, 22(1):115–130, 2009.

Wataru Kureishi and Midori Wakabayashi. Son preference in japan. *Journal of Population Economics*, 24(3):873–893, 2011.

Jungmin Lee. Sibling size and investment in childrens education: An asian instrument. *Journal of Population Economics*, 21(4):855–875, 2008.

Qi Li and Juan Pantano. The demographic consequences of gender selection technology. 2013.

Kei Maruyama. The use and determinants of female labor supply after childbirth. *Journal of Population Problems*, 57(2):3–18, 2001.

Jacob Mincer. Labor force participation of married women. In *Aspects of labor economics*, pages 63–106. Princeton University Press, 1962.

Makiko Nakamuro, Ryoji Matsuoka, and Tomohiko Inui. Time trade doesnt happen: The determinants of childrens hours spent studying. Discussion paper, Research Institute of Economy, Trade and Industry (RIETI), May 2013.

Akiko Oishi. The effect of childcare expenses to mothers labor force participation. *The Japanese Journal of Social Security Policy*, 39(1):55–69, 2003.

Syuichi Ooki. Effect of maternal age and fertility treatment on the increase in multiple births in japan: vital statistics, 1974-2009. *Journal of epidemiology/Japan Epidemiological Association*, 21(6):507–511, 2010.

Richard Pollard. Ethnic comparison of twinning rates in california. *Human biology*, pages 921–931, 1995.

Mark R Rosenzweig and Kenneth I Wolpin. Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy*, pages 328–348, 1980a.

Mark R Rosenzweig and Kenneth I Wolpin. Testing the quantity-quality fertility model: the use of twins as a natural experiment. *Econometrica*, 48(1):227, 1980b.

Mark R Rosenzweig and Kenneth I Wolpin. Natural" natural experiments" in economics. *Journal of Economic Literature*, 38(4):827–874, 2000.

Masaru Sasaki. The causal effect of family structure on labor force participation among japanese married women. *Journal of Human Resources*, pages 429–440, 2002.

Sachiko Senda. The relationship between the maternal labor supply and the access to resources for child rearing. *Journal of Population Problems*, 58(2):3–18, 2002.

Yukiko Shigeno and Katsumi Matsuura. The balancing of childbearing, child rearing, and employment: the trade-off between marriage and employment, and the role of maternal leave for working mothers. *The Japanese Journal of Social Security Policy*, 39(1):43–54, 2003.

Yuriko Shintani. The determinants of maternal labor supply around the time of childbirth. *Journal of Population Problems*, 54(4):46–61, 1998.

Shigesato Takahashi, Akira Ishikawa, Hisakazu Kato, Miho Iwasawa, Ryuichi Komatsu, Ryuichi Kaneko, Masako Ikenoue, Mita Fusami, Akiko Tsuji, and Rie Moriizumi. Population projections for japan 2001-2050. *Journal of Population and Social Security*, 1 (1), 2003.

Ryuichi Tanaka. The gender-asymmetric effect of working mothers on children's education: Evidence from japan. *Journal of the Japanese and International Economies*, 22 (4):586–604, 2008.

James P Vere. Fertility and parents labour supply: new evidence from us census data. *Oxford Economic Papers*, 63(2):211–231, 2011.

Table 1
Response and Attrition Rates

Wave	Number Distributed (a)	Number collected (b)	Response Rate (b)/(a)	Attrition Rate (53,575-(b))/53,575
1	53,575	47,015	87.8%	12.2%
2	46,966	43,925	93.5%	18.0%
3	46,897	42,812	91.3%	20.1%
4	44,837	41,559	92.7%	22.4%
5	43,559	39,817	91.4%	25.7%
6	42,187	38,537	91.3%	28.1%
7	40,598	36,785	90.6%	31.3%
8	39,261	36,136	92.0%	32.6%
9	37,932	35,264	93.0%	34.2%
10	36,989	34,124	92.3%	36.3%

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare

Table 2a: Descriptive Statistics

Time varying characteristics	Round 1	2	4	5	6	7	8	9	10
% Mothers employed	24.84	29.21	38.03	43.05	46.52	50.17	55.09	58.39	61.92
	(43.21)	(45.47)	(48.55)	(49.52)	(49.88)	(50)	(49.74)	(49.29)	(48.56)
Average number children	1.01	1.07	1.54	1.67	1.78	1.88	1.92	1.95	1.97
	0.09)	(0.26)	(0.55)	(0.59)	(0.62)	(0.68)	(0.7)	(0.71)	0.73
% live with grandparents	20.25	21.09	19.35	22.03	22.47	20.99	22.86	22.39	21.27
	40.19)	(40.79)	(39.51)	(41.44)	(41.74)	(40.73)	(42)	(41.68)	40.92
Per child expenditures (¥1000 / month)	4.93	3.08	3.54	5.97	5.03	35.5	35.46	36.18	39.47
	(13.02)	(6.04)	(7.79)	(12.97)	(6.3)	(32.66)	(23.69)	(22.56)	(26.45)

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare.
We report means and standard deviations in parentheses.

Table 2b: Descriptive Statistics

	Baseline characteristics			
First birth in 2001	Average	Twins	Non-twins	Difference
Mother's age at 1st birth	27.61	29.89	27.59	2.3‡
	(4.4)	4.17	4.4	
Mother's education	13.18	13.37	13.17	0.2
	(1.65)	(1.44)	(1.66)	
% Mothers working baseline	.75	.64	.75	-0.11‡
	(.44)	(.48)	(.43)	
N	22,583	178	22,405	
Twinning rate per 1000 maternities in 2001	7.882			
Mother age less 35	Average	Twins	Non-twins	Difference
Mother's age at 1st birth	28.14	30.12	28.13	1.99‡
	(3.89)	(3.45)	(3.89)	
Mother's education	13.17	13.25	13.16	0.09
	(1.65)	(1.41)	(1.65)	
% Mothers working baseline	.75	.65	.75	-0.10‡
	(.43)	(.48)	(.43)	
N	21,573	162	21,411	
Twinning rate per 1000 maternities in 2001	7.509			
Full sample	Average	Twins	Non-twins	Difference
Mothers age at 1st birth	27.53	28.81	27.52	1.29‡
	(4.29)	4.33	4.28	
Mothers education	13.07	13.2	13.07	0.13
	(1.63)	1.51	1.63	
% Mothers working baseline	.55	.54	.55	-0.01
	(.5)	(.5)	(.5)	
N	45,503	458	45,045	
Twinning rate per 1000 maternities in 2001	10.065			

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare. We report means and standard deviations in parentheses.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Figure 1: Assisted Fertility Births in Japan 1995-2010

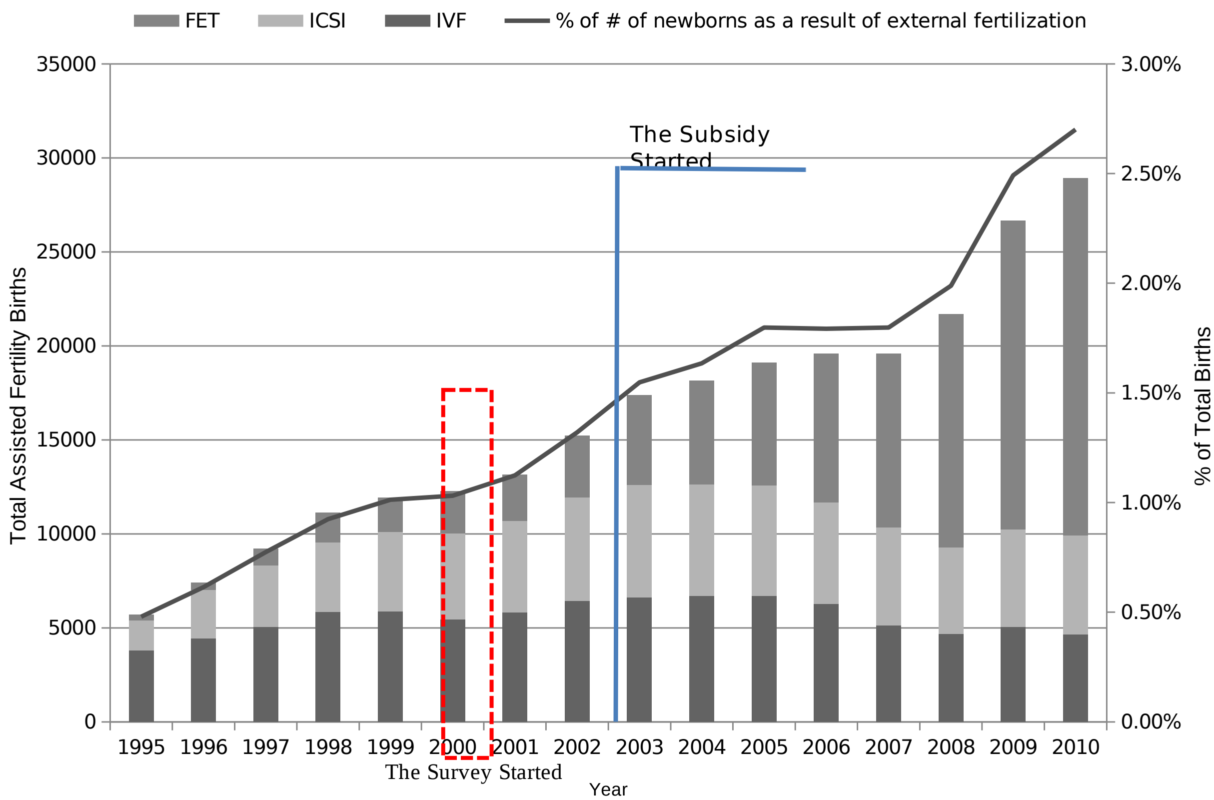

Table 3
First stage regressions
Dependent variable: Number of children

	Round								
	(1)	(2)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Twins in 2001	1.0000	0.9608‡	1.6016‡	1.5154‡	1.4542‡	1.4184‡	1.3822‡	1.3472‡	1.3315‡
	(0.000)	(0.018)	(0.040)	(0.043)	(0.046)	(0.049)	(0.050)	(0.051)	(0.054)
Age Mother First Birth	-0.0000	-0.0051‡	-0.0213‡	-0.0260‡	-0.0311‡	-0.0394‡	-0.0442‡	-0.0482‡	-0.0512‡
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	1.0000	1.2028‡	2.1206‡	2.3879‡	2.6336‡	2.9707‡	3.1485‡	3.2902‡	3.3965‡
	(0.000)	(0.011)	(0.024)	(0.027)	(0.029)	(0.032)	(0.034)	(0.035)	(0.036)
F-test	-	1474.71	1081.58	975.4	926.79	981.68	1035.07	1079.75	1072.52
N	22,583	21,158	20,005	19,178	18,529	17,743	17,439	17,069	16,549
R^2	1.000	0.122	0.098	0.092	0.091	0.100	0.106	0.112	0.115

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Table 4
Dependent Variable: Maternal Labor Supply

	OLS	IV
Number children X Round 1	0.0026	-0.0069
	(0.032)	(0.036)
Number children X Round 2	-0.1202‡	-0.0407
	(0.011)	(0.038)
Number children X Round 4	-0.1599‡	-0.0076
	(0.006)	(0.023)
Number children X Round 5	-0.1569‡	0.0371
	(0.006)	(0.025)
Number children X Round 6	-0.1415‡	0.0756‡
	(0.006)	(0.027)
Number children X Round 7	-0.1065‡	0.0580**
	(0.006)	(0.027)
Number children X Round 8	-0.0950‡	0.0554*
	(0.005)	(0.029)
Number children X Round 9	-0.0838‡	0.0457
	(0.005)	(0.029)
Number children X Round 10	-0.0696‡	0.0493
	(0.005)	(0.032)
Age Mother First Birth	-0.0066‡	-0.0025‡
	(0.001)	(0.000)
Round dummies	Yes	Yes
Sum over 10 rounds	-1.07	0.24
p-value χ^2 test sum = 0	0.00	0.015
Durbin-Wu-Hausman p-value		0.00
N	166,397	166,397
R^2	0.484	0.466

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare.
‡ $p<0.01$, † $p<0.05$, * $p<0.1$

Figure 2. OLS and IV Effects Over Time

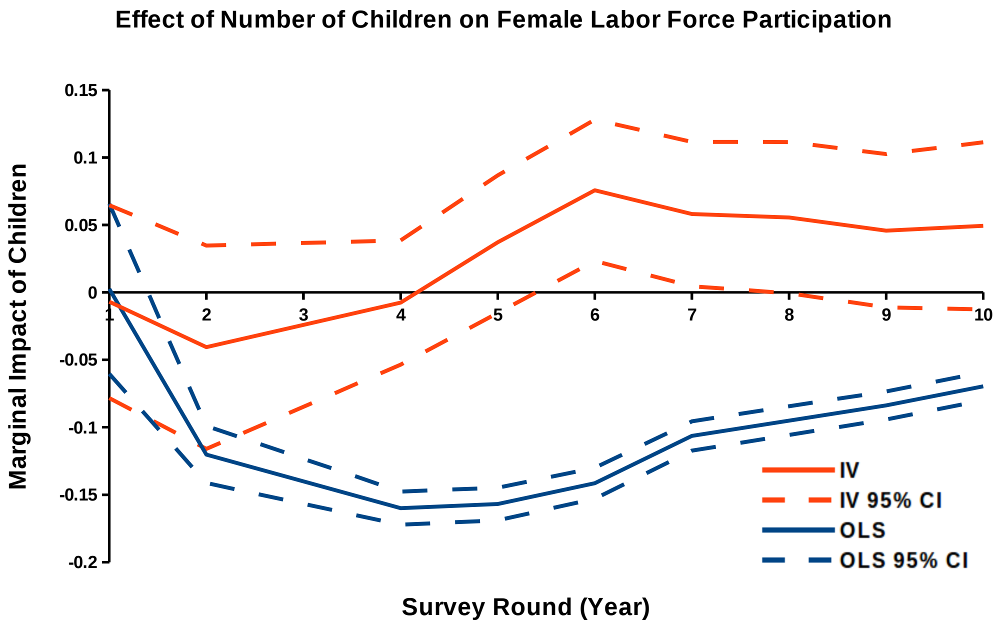

Table 5: Robustness Checks
Dependent Variable: Maternal Labor Supply

	(1)	(2)	(3)	(4)	(5)
Number Children X Round 1	-0.0360	-0.0252	-0.0166	-0.0059	0.0017
	(0.036)	(0.036)	(0.018)	(0.038)	(0.037)
Number Children X Round 2	-0.0592	-0.0455	-0.0444†	-0.0463	-0.0388
	(0.038)	(0.038)	(0.021)	(0.040)	(0.039)
Number Children X Round 4	-0.0126	-0.0005	0.0026	-0.0124	-0.0076
	(0.023)	(0.024)	(0.015)	(0.025)	(0.024)
Number Children X Round 5	0.0301	0.0438*	0.0405†	0.0357	0.0424
	(0.025)	(0.026)	(0.016)	(0.027)	(0.026)
Number Children X Round 6	0.0667†		0.0529‡	0.0773‡	0.0853‡
	(0.026)		(0.017)	(0.028)	(0.027)
Number Children X Round 7	0.0456	0.0575*	0.0407†	0.0536*	0.0674†
	(0.033)	(0.034)	(0.018)	(0.029)	(0.028)
Number Children X Round 8	0.0459		0.0329*	0.0430	0.0646†
	(0.028)		(0.018)	(0.030)	(0.029)
Number Children X Round 9	0.0369		0.0302*	0.0340	0.0543*
	(0.029)		(0.018)	(0.031)	(0.030)
Number Children X Round 10	0.0379	0.0598*	0.0356*	0.0426	0.0599*
	(0.031)	(0.033)	(0.019)	(0.033)	(0.032)
Age Mother First Birth	-0.0017‡	0.0018‡	-0.0049‡	-0.0053‡	-0.0025‡
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
Additional controls	Yes	Yes	No	No	No
Husband's income	No	Yes	No	No	No
Sample					
First birth 2001	Yes	Yes	Yes	Yes	No
Mother's age birth less 35	No	No	Yes	No	No
Only mothers working in 2000	No	No	No	Yes	No
Full sample of twins	No	No	No	No	Yes
Sum over 10 rounds	0.12	-	0.15	0.19	0.39
p-value χ^2 test sum = 0	0.23	-	0.001	0.07	0.002
Durbin-Wu-Hausman p-value	0.00	0.00	0.00	0.00	0.00
N	157,346	96,131	335,486	158,499	159,041
R^2	0.48	0.42	0.52	0.47	0.46

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare. Additional controls include maternal education, whether the husband lives in the house, and whether any grandparents live in the house.
‡ p<0.01, † p<0.05, * p<0.1

Table 6: Children's Age Effects
Dependent Variable: Maternal Labor Supply

	(1)	(2)	(3)
Round 1	0.504‡	0.513‡	0.196‡
	(0.028)	(0.028)	(0.038)
Round 2	0.554‡	0.569‡	0.244‡
	(0.028)	(0.028)	(0.038)
Round 4	0.694‡	0.710‡	0.373‡
	(0.029)	(0.029)	(0.038)
Round 5	0.756‡	0.772‡	0.435‡
	(0.029)	(0.029)	(0.038)
Round 6	0.803‡	0.817‡	
	(0.029)	(0.029)	
Round 7	0.855‡	0.829‡	0.493‡
	(0.030)	(0.030)	(0.037)
Round 8	0.905‡	0.875‡	
	(0.029)	(0.030)	
Round 9	0.941‡	0.903‡	
	(0.030)	(0.030)	
Round 10	0.981‡	0.935‡	0.605‡
	(0.030)	(0.030)	(0.037)
Number children	-0.109‡	-0.087‡	-0.073‡
	(0.004)	(0.004)	(0.005)
Number children less than age 6		-0.038‡	-0.036‡
		(0.004)	(0.004)
Mother's education	0.008‡	0.009‡	0.018‡
	(0.002)	(0.002)	(0.002)
Father lives in house	-0.130‡	-0.126‡	-0.012
	(0.015)	(0.015)	(0.021)
Grandparents live in house	0.158‡	0.158‡	0.136‡
	(0.006)	(0.006)	(0.007)
Age Mother First Birth	-0.006‡	-0.006‡	-0.001
	(0.001)	(0.001)	(0.001)
Husband's income			-0.000‡
			(0.000)
N	157,346	157,346	96,131
R^2	0.492	0.493	0.431

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare
‡ p<0.01, † p<0.05, * p<0.1

Table 7
Dependent variable: per child expenditures / month
Unit ¥1000 (≈ $10)

	(1)	(2)	(3)
Round 1	-7.55‡	-7.36‡	0.47
	(0.803)	(0.802)	(1.324)
Round 2	-9.10‡	-8.90‡	-1.13
	(0.797)	(0.796)	(1.298)
Round 3	-8.86‡	-8.63‡	
	(0.799)	(0.798)	
Round 4	-7.23‡	-6.97‡	0.77
	(0.807)	(0.807)	(1.285)
Round 5	-4.40‡	-4.13‡	3.61‡
	(0.809)	(0.809)	(1.281)
Round 6	-5.07‡	-4.78‡	
	(0.808)	(0.808)	
Round 7	26.15‡	26.45‡	34.01‡
	(0.835)	(0.835)	(1.267)
Round 8	25.76‡	26.06‡	
	(0.809)	(0.809)	
Round 9	26.55‡	26.86‡	
	(0.806)	(0.807)	
Round 10	29.91‡	30.22‡	38.05‡
	(0.808)	(0.809)	(1.254)
Number children	-3.11‡	-3.24‡	-3.25‡
	(0.134)	(0.138)	(0.168)
Twins		3.12‡	2.71‡
		(0.602)	(0.620)
Age Mother First Birth	0.25‡	0.25‡	0.15‡
	(0.016)	(0.016)	(0.029)
Mother's education	0.79‡	0.79‡	0.40‡
	(0.044)	(0.044)	(0.071)
Father lives in house	-1.95‡	-1.89‡	-3.81‡
	(0.499)	(0.500)	(0.866)
Grandparents live in house	-0.12	-0.13	0.23
	(0.149)	(0.149)	(0.225)
Husband's income			0.00†
			(0.002)
N	176,626	176,626	96,370
R^2	0.632	0.632	0.572

Source: Longitudinal Survey of Newborns in the 21st Century, Ministry of Health, Labour and Welfare
‡ p<0.01, † p<0.05, * p<0.1

Table 8
Educational expenditure data

Private spending share of total (public and private) educational expenditures			
	Primary	Secondary	Tertiary
OECD Average	19.3%	8.8%	27.5%
US	22.4%	8.5%	66.0%
Japan	56.6%	10.1%	67.8%

% of household educational expenditure for activities outside of school			
	Elementary	Middle School	High School
Japan	70.8%	64.0%	33.9%

Source: OECD, Ministry of Education, Culture, Sports, Science and Technology