China Economic Review 43 (2017) 47–63

Contents lists available at ScienceDirect

China Economic Review

A dilemma of fertility and female labor supply: Identification using Taiwanese twins

Junchao Zhang

Yokohama National University, 79-3 Tokiwadai, Hodogaya ward, Yokohama 240-8501, Japan

ARTICLE INFO

Article history:
Received 28 August 2016
Received in revised form 31 December 2016
Accepted 31 December 2016
Available online 4 January 2017

JEL classification:
J13
J21

Keywords:
Instrumental variable
Twin
Fertility
Female employment
Taiwan

ABSTRACT

Using micro-data from the Taiwan Population and Housing Census, this study estimates the causal effect of fertility on the labor supply of married women. To address the endogeneity of fertility, we exploit exogenous variations in the number of children caused by twin births, which can be considered a natural experiment. The instrumental variable estimates indicate that an additional child reduces female employment by 10.5 percentage points for those who have at least one delivery, and the effects gradually decline for females who have two or more deliveries, with the effects vanishing when females have three or more deliveries. Also, the effect of fertility varies substantially with the time elapsed since the last childbirth, which has a consequence for differences in estimates across different samples in the literature.

1. Introduction

In the decades since family planning policy was introduced in the 1960s, the growth rate of the Taiwanese population has decreased. In 1951, each Taiwanese woman had approximately seven children over her lifetime. However, in the 21st century, Taiwan's fertility rate has become one of the lowest in the world, falling below one child per woman in 2010. The total fertility rate (TFR) of Taiwan was 1.12 in 2015, only slightly higher than Macau and Singapore[1], placing the country third from the bottom among 224 countries (Central Intelligence Agency, 2015). In addition to the development of women's human capital, infant mortality has decreased substantially (Chou, Liu, Grossman, & Joyce, 2010), and fertility behavior is changing within Taiwanese society (Freedman, Fan, Wei, & Weinberger, 1977; Freedman, Hermalin, & Chang, 1975). The extremely low birthrates have become an unprecedentedly challenging social issue in Taiwan (Y.-H. Chen, 2012; Lin & Yang, 2009).

Corresponding to a sharp decline in the fertility rate, the labor force participation (LFP) of married women in Taiwan has experienced a high rate of growth. Fig. 1 shows the long-term time series data of TFR and LFP for married women in Taiwan. On the one hand, the TFR had decreased to 1.165 per woman in 2014 from 2.455 per woman in 1981. On the other hand, the LFP of married women had grown to 49.76% in 2014 from a rate of 33.23% in 1980. The strong negative correlation between the two leads to a policy dilemma regarding whether having more children reduces the likelihood of labor market participation among

E-mail address: ztempest0218@gmail.com.

[1] Singapore 0.81, Macau 0.94, Japan 1.40, China 1.60, U.S. 1.87, UK 1.89.

http://dx.doi.org/10.1016/j.chieco.2016.12.005
1043-951X/

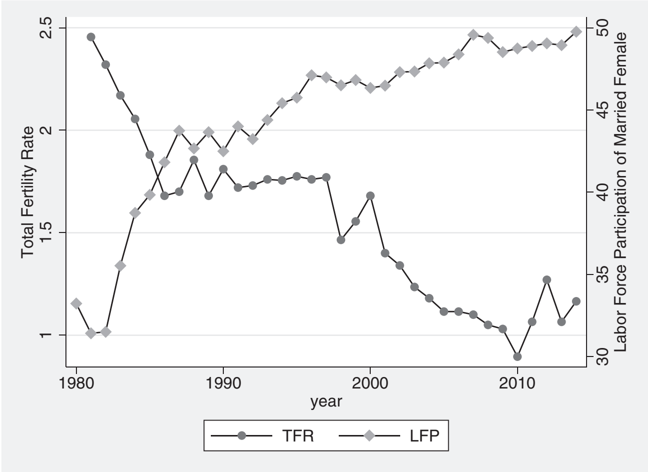

Fig. 1. Time trend of fertility and female labor supply.
Source: (1) Report on the Manpower Utilization Survey and (2) Vital Statistics of Taiwan.

married women. For policy makers, simultaneously increasing the fertility rate and female labor supply is a difficult task, as fertility promotion policy may lead to an increase in the fertility rate but harms female labor supply (Dudel, 2009; Rosenbluth et al., 2002).

Theoretically, economists and demographers have established a variety of models to link fertility and the labor market (Becker, 1985; Gronau, 1973; Mincer, 1962; Rosenzweig & Wolpin, 1980a). They predict a strong negative correlation between the two, as married women could pay more attention to child care and housework. For instance, Mincer (1962) points out that a mother's care of small children is much more difficult to be substituted, compared to food preparation or other houseworks. Labor supply of married women who have small children at present could be affected more than those who have not, or than at other times in the life-cycle. Becker (1985) suggests that child care is more effort intensive than leisure and other household activities, which may further affect women's labor market performance. Empirically, several previous studies show that younger women with lower fertility participate in labor market more frequently and more intensively, which treat fertility decisions as exogenous (Gronau, 1973; Heckman, 1974; Schultz, 1978).

Although the causal link between fertility and labor supply among married women is important, the main difficulty in determining the causality is that fertility decisions and labor market participation are simultaneously determined. Two potential endogeneity issues in this context are as follows. First, unobserved factors such as preferences regarding the number of children and career expectations are heterogeneous across individuals. For example, those who are ambitious at work may have a lower expected number of children while being more likely to participate in the labor market. Any estimates that fail to control for unobserved heterogeneity would be downward biased, as career ambition is negatively correlated with fertility but positively correlated with labor supply, which would overestimate the negative effect of fertility on labor supply. Following Angrist and Evans (1998), several studies have attempted to correct for such bias using mixed sibling-sex composition as an instrumental variable (IV), due to the well-known preference for sons in some Asian countries (Azimi, 2015; Chun & Oh, 2002; Ebenstein, 2009). Following Rosenzweig and Wolpin (1980b), economists also use twin births as a natural experiment for fertility to analyze its effect on labor supply (Bronars & Grogger, 1994; He & Zhu, 2015) or to test the quantity-quality trade-off for children (Angrist et al., 2010; Åslund & Grönqvist, 2010; Black et al., 2005, 2010; Li et al., 2008).

Second, the amount of time passed since the last childbirth is important for the subsequent fertility decision, and the question of whether to have an additional child also affects the labor market activities of married women. The more time that has passed since the last childbirth, the more easily a woman can return to work. Controlling for the youngest child's age as an explanatory variable would attenuate the above-mentioned bias, but the estimates are far from causal because the assumption of a time-invariant preference for the number of children is imposed implicitly. However, the preference for the number of children should be treated as time-variant in the real world when applying either mixed sibling-sex composition or twin births as an instrument. For example, mothers of twins and mothers of non-twins would have different paths (preferences) for having an additional delivery over time, as the burden of caring for children is doubled for mothers of twins. As an alternative approach, this study estimates the effect of fertility using two groups for which the interval since the last childbirth is no more than one year for the first group and no more than three years for the second. The choice of intervals is discussed specifically in the results section.

Other than using exogenous variation in the number of children caused by twin births, this study also effectively compares the difference between mothers of twins and mothers of non-twins using three sub-samples to investigate the causal effect of fertility on female labor supply. Specifically, in each sub-sample, individuals have the same frequency of childbearing experience, and those who have had at least one additional child after twin births are excluded from the sub-samples because these individuals tend to have a higher expected number of children; thus, mothers of twins always have one more child than mothers of non-twins. Each sub-sample allows me to analyze the effect of an additional child caused by twin births on female labor supply, relative to those mothers of non-twins when the preference for the number of children is held constant. In the literature, when a two-stage least squares estimation is applied, preferences for number of children are assumed to be the same among mothers with at least n births. However, mothers with n births and mothers with $n + 1$ births may have different preferences for the number of children. To our knowledge, this research is the first that properly holds constant the preference for number of children for mothers with n births, where $n \in \{1, 2, 3\}$.

The results show that an additional child (caused by twin births) reduces the likelihood of labor market participation by 10.6% for a first-time mother and by 4.7% for a second-time mother, with both estimates accounting for the preference for the number of children and being statistically significant. In contrast to the results for mothers with relatively fewer deliveries, the effect of fertility vanishes for third-time mothers who have a fourth child through twin birth, no longer being significant at the 10% level. This suggests that the effect of fertility on female labor supply is not monotonically decreasing in the number of births. Contrary to previous findings, the IV estimates and sub-sample analysis results are both larger in magnitude than OLS estimates when holding the time since the last childbirth constant, revealing the endogeneity problem caused by the time-variant preference for the number of children suggested in the previous paragraph.

The remainder of this paper is organized as follows. Section 2 describes the data sets. The following section presents the identification strategy. Section 4 compares the estimates among OLS, IV and sub-sample analyses. Section 5 checks the robustness of the estimates. Section 6 discusses the policy implication. Section 7 provides a summary.

2. Data

The data used in this research originate from the 2000 Population and Housing Census, which is a universal sample of all residents of Taiwan. It is the fifth in the series, following the previous four censuses conducted in 1956, 1966, 1980, and 1990. The data set covers 22,300,929 individuals across 6,495,751 households. Individual-level information on demographic characteristics, marital status, education, employment status, ethnicity, etc. is collected.

One advantage of this data set is the huge sample size, which captures the overall characteristics of the Taiwanese population. Such universal census data offer extremely high statistical power in individual-level analyses. Compared with previous studies, we can apply much more strict sample selection rules to obtain a less heterogeneous sample. Moreover, the huge sample size allows me to analyze fertility decisions by delivery order[2] to observe the heterogeneous effect across families with different preferences for the number of children, while it is difficult to do so with survey data. Second, along with the *hukou* system-based population census of Taiwan, our sample has less measurement error caused by migration (because of job allotment, etc.). As the U.S. population census is conducted on the basis of the current residence, previous studies cannot track husbands and wives who are living in two places, as well as their minor children.

We match children to their parents through the relationship identifier within households. First, we identify individuals who are labeled “child” and calculate the fertility information (of their mothers), including the number of children and twin births at the *n*th delivery. Second, we supplement the fertility information with data on mothers who are labeled “household head” or “spouse” in each household. The mothers are the primary observations in this study. For each mother, we also construct the husband's information, including years of schooling and employment status. Individuals who engage in industry, agriculture, and services are defined as employed, as well as those unpaid family workers who work over 15 h every week.

For the analysis of the fertility and labor supply of married women, the sample is restricted as follows: (1) Following Angrist and Evans (1998), we only use children of the household head to construct the fertility information, as we cannot identify the relationship among other relatives within the households. Households with no children are also dropped. (2) We restrict the sample to mothers who are between 16 and 35 years of age and whose eldest child is no more than 20 years of age. In Taiwan, the legal marriage age is 16 for females, and children are not allowed to move out of a household before age 20. We impose such restrictions to ensure that no adult children have already left the household by the time of the survey. (3) Finally, we exclude single mother households because information on fathers cannot be obtained. We also dropped households with fathers who are under 18 or over 50, the latter of which can result in problems of remarriage in a cross-sectional census.

The final sample contains 614,902 females, 9501 of whom have given birth to twins. Because the census does not include an exact identifier for twins, we define twins as children who were born in the same year. Information on the birth month and birthdate is not offered in the census. The potential measurement error in birth year induced by lunar calendar among Chinese societies should be very tiny, as Taiwanese census records the identity card number for each individual (but not available to data user). Among Chinese societies like Mainland China, Taiwan, and Hongkong, females are expected to rest indoors for one full month after giving birth (*zuoyuezi*), sometimes over 40 days, the possibility of giving births to two non-twins in a single year is very low. To avoid other measurement errors, we have dropped households with more than 2 children who were born

[2] It is up to three in this research. Only 2.51% women have more than 3 deliveries in the final sample.

Table 1
Descriptive statistics for married women.

Variables	(1) Overall	Mothers of (2) Twins	Mothers of (3) Non-twins	Difference (4) Unmatched	Difference (5) Matched
Employed	0.596	0.559	0.597	−0.038***	−0.023***
	(0.491)	(0.497)	(0.491)	(0.005)	(0.007)
Number of children	1.949	2.864	1.935	0.929***	0.916***
	(0.787)	(0.828)	(0.778)	(0.008)	(0.012)
Age	30.927	31.281	30.922	0.360***	−0.015
	(3.365)	(3.198)	(3.367)	(0.035)	(0.046)
Age squared/100	9.678	9.888	9.675	0.213***	−0.009
	(1.990)	(1.907)	(1.991)	(0.021)	(0.028)
Years of schooling	11.362	11.153	11.365	−0.212***	−0.002
	(2.627)	(2.614)	(2.627)	(0.021)	(0.038)
Minority	0.022	0.025	0.022	0.003	0.002
	(0.147)	(0.155)	(0.147)	(0.002)	(0.002)
Husband					
Years of schooling	11.627	11.442	11.630	−0.188***	−0.003
	(2.911)	(2.924)	(2.910)	(0.030)	(0.042)
Employed	0.964	0.965	0.964	0.001	−0.002
	(0.187)	(0.185)	(0.187)	(0.002)	(0.003)
Elderly					
Co-resident	0.088	0.093	0.088	0.005	0.003
	(0.284)	(0.291)	(0.284)	(0.003)	(0.004)
Observations	614,902	9501	605,401	614,902	19,002

Notes: Standard errors in parentheses. Column 4 is the raw difference between mothers of twins and mothers of non-twins. Column 5 is the difference after one-to-one matching, which includes 9501 mothers of twins and 9501 mothers of non-twins. The results remain stable when we apply other matching methods such as *k*-nearest neighbors, radius, and kernel. The propensity score used for matching is calculated by logistic regression, which is available upon request. *** p<0.01, ** p<0.05, * p<0.1.

in the same year, which may be caused by adoption. In 2000, there were 3987 adopted children in Taiwan while the number of newborn babies was 305,312[3]. Moreover, approximately 26% of adoptive households abandoned adoptions for unknown reasons in each year during the period from 1993 to 2003[4]. Adoption would not be a severe problem in our data. Also, remarriage rate among divorced and widowed women is only 1.4–1.6 %[5] in late 1990s, which may not cause severe measurement error in defining twin births. The twinning rate in our sample is approximately 0.79%, which is comparable to international literature and Taiwanese statistics (C. J. Chen, Lee, Wang, & Yu, 1992; C. J. Chen, Lin, Chang, & Cheng, 1987; Smits & Monden, 2011).

To investigate the difference between mothers who have and have not given birth to twins, descriptive statistics tabulated by delivery of twins are shown in Table 1. From Column 1 and Column 2, we find that the mothers of twins have a lower probability of being employed than the mothers of non-twins, and the mothers of twins have more children. Although no large differences are observed, the *t*-statistics shown in Column 4 do indicate significant differences in the covariates between the two samples. Specifically, mothers of twins and mothers of non-twins differ with respect to age, years of schooling, and husband's years of schooling, but they do not differ with respect to ethnic minority status, husband's employment status, and co-residence with an elderly parent. The birth of twins is a perfect instrument only when such an event is totally random, and the characteristics of mothers of twins should be the same as those of mothers of non-twins.

To address potential parental selection into twin births, we also construct a matched sample with extremely similar mothers of twins and mothers of non-twins. The *t*-statistics of the matched sample are displayed in Column 5 of Table 1, which are not significant in all covariates. This matched sample will be used in the section on robustness checks, in which the selection issue will be specifically discussed.

3. Identification strategy

3.1. IV Estimation

To estimate the effect of the number of children on married women's labor force participation, the benchmark model is specified as follows:

$$LFP_i = \beta_0 + \beta_1 Children_i + X_i'\delta_1 + Z_i'\delta_2 + \epsilon_i \qquad (1)$$

[3] Source: Deptartment of Household Registration, MOI.
[4] Author's calculation using the data from the *Statistical Year Book of Interior*.
[5] Source: Department of Household Registration, MOI.

where the LFP_i is a binary outcome variable indicating married women's labor force participation, which equals 1 if working and 0 otherwise. β_1 is the coefficient of interest, capturing the effect of the number of children. X_i is a vector of individual characteristics including age, age squared, education, and ethnicity. Z_i is a vector of husband's characteristics and living arrangements, which includes husband's education and labor supply, and a binary variable indicating co-residence with an elder parent.

However, the OLS estimates are consistent only if number of children is not correlated with error term ϵ_i, which is obviously not the case. To address the endogeneity of fertility, we use twin birth as an instrumental variable for the number of children. The first stage of the IV estimation is as follows:

$$Children_i = \gamma_0 + \gamma_1 Twins_i + X_i^{'}\rho_1 + Z_i^{'}\rho_2 + \varepsilon_i \tag{2}$$

where $Children_i$ is the number of children of a married woman. $Twins_i$ is a binary instrumental variable that equals 1 if a woman has given birth to twins at the nth delivery and 0 otherwise. X_i and Z_i are the same vectors of control variables as in Eq. (1).

3.2. Sub-sample OLS estimation using an efficient instrument

In this research, the second identification strategy employs OLS estimation with three sub-samples of our data. As noted by Rosenzweig and Wolpin (1980b), the probability of twin birth increases in the number of deliveries. To control for the preference for the number of children, we follow Li et al. (2008) and restrict the sample to families with at least n births in IV estimations, which assumes that families with n births and $n+1$ births have the same preference for number of children. In this subsection, we will discuss a model that compares mothers of twins and non-twins who have exactly the same frequency of deliveries, and the model is specified as

$$LFP_i = \beta_0 + \beta_1 Twin_i + X_i^{'}\delta_1 + Z_i^{'}\delta_2 + \epsilon_i \tag{3}$$

Table 2 provides graphical descriptions of the sub-samples. Sub-sample A includes mothers of non-twins with a single child and mothers of twins at the first delivery, namely women who only have one delivery. Sub-sample B includes mothers of two non-twins and mothers of twins at the second delivery, namely mothers who have two deliveries. Sub-sample C includes mothers of three non-twins and mothers of twins at the third delivery, namely mothers who have three deliveries. In each sub-sample, mothers of twins have one more child than mothers of non-twins. We construct sub-samples for up to three deliveries because only 2.52% of women have more than 3 deliveries in the final sample, which is also shown in Fig. 2.

In these sub-sample analyses, IV estimations are not necessary as the variation in the number of children is entirely determined by twin birth. In other words, the efficient instrumental variable, $Twin_i$, is equivalent to the number of children. This approach estimates the effect of an additional child (caused by a twin birth) on female labor supply, relative to mothers of non-twins.

4. Results

This section discusses the OLS and IV regression estimates, which were designed to test whether fertility has a negative effect on the labor supply of married women in Taiwan. It is worthwhile to emphasize that we examine some cases that would deliver biased IV estimates, which will be discussed in the subsection on the sub-sample analyses. In all regressions, linear probabilities are used along with robust standard errors. Due to space constraint, only key estimates are reported in all tables, while the full estimates for unconditioned case are attached in the appendix section for reference. Before presenting the results of the OLS and IV estimation, we discuss the validity of twin births as an instrument.

Table 2
Descriptions of sub-samples.

	(1) Sub-sample A	(2) Sub-sample B	(3) Sub-sample C
Mothers of non-twins	①	① ②	① ② ③
vs			
Mothers of twins	❶❶	①❷❷	①②❸❸
Number of children	1 ***vs*** 2	2 ***vs*** 3	3 ***vs*** 4

Notes: ○ indicates non-twin, and ● indicates twins. Numbers in circles show the birth order. Sum of circles shows the total number of children mothers have.

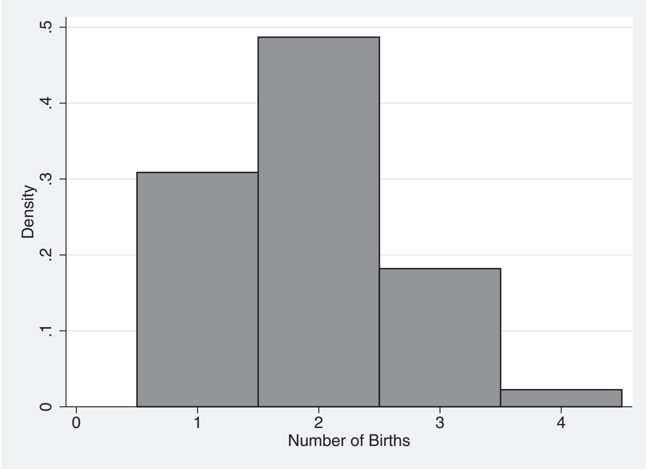

Fig. 2. The distribution of childbirth. Notes: Author's calculation using Population and Housing Census 2000 of Taiwan. Mothers having more than 4 births are dropped, which covers 0.27% of the overall sample as described in Table 1.

4.1. The first stage

Unobserved heterogeneity. A good IV in this case should be highly correlated with number of children but should not affect the labor supply of married women except through the number of children. In other words, a valid IV should not be correlated with unobserved characteristics that are captured by the error term, ϵ_i, in Eq. (1).

One concern is that the occurrence of twin births may not be random and may be correlated with unobserved family characteristics, which by design, cannot be tested. As an alternative approach, we control for a full set of family characteristics including parents' education and living arrangement in regressions. The first stage of full sample is shown in the last four columns of Table A1. The F-statistics of excluded instrument are stably larger than 15,000 across different specifications. Moreover, we examine the difference between families with twins and those without twins as shown in Column 4 of Table 1. The enormous sample size provides the power to show that mothers of twins exhibit significant difference in years of schooling, as do their husbands. To avoid parental selection into twin births, we construct a matched sample with extremely similar families with twins and without twins to check the robustness of the findings, which will be discussed in the section on robustness checks.

Time since the last childbirth. Another concern is that a twin birth may affect female labor supply through the time since the last childbirth. As in a cross-sectional census, women are assigned different durations since the last childbirth because the survey time is fixed. Fig. 3 provides a graphical example. Suppose that mother 1 married at t_0, had her first

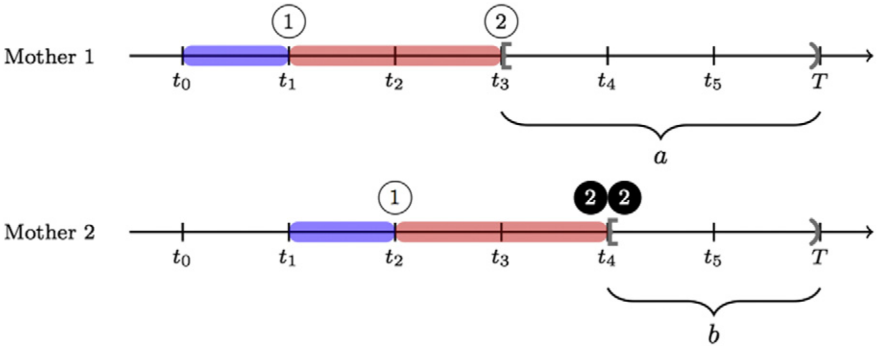

Fig. 3. An example of potential bias. Notes: ○ indicates non-twin, and ● indicates twins. Numbers in circles show the birth order. Braces are durations from last childbirth to the survey time T, where $a = T - t_3$ for mother 1 and $b = T - t_4$ for mother 2.

delivery at t_1 and had her second delivery at t_3; thus, she would have a duration of a from her second delivery until survey time T. Suppose that mother 2 married at t_1, had her first delivery at t_2, and experienced a twin birth at t_4, which is her second delivery. However, mother 2 would have a duration of b until survey time T, which is totally different from mother 1.

There are two possible channels through which twin births might affect female labor supply via the time elapsed since the last childbirth. In both cases, the IV and sub-sample OLS estimates of the effects of fertility could be biased. First, the time since the last childbirth is highly related to presence of a younger baby. For example, mothers of younger babies have to provide more childcare in terms of time and energy, which cannot be completely substituted by market services. Moreover, mothers of twins take on twice the burden of childcare relative to women who are not mothers of twins. Consequently, the former should be more likely to drop out of the labor market relative to mothers of non-twins.

Second, twin births may affect female labor force participation through a time-variant preference for the number of children, which is not observable. Fertility should be a sequential choice, which is not decided at a given time. That is, a woman would decide whether to have another birth after having given birth, and this choice may change over time. As shown in Fig. 3, suppose that both mother 1 and mother 2 expect three children over their lifetime; then, mother 2 (the mother of twins) would likely stop fertility after her twin birth at the second delivery, and the longer the duration b is, the higher the probability that mother 2 will participate in the labor market. Because if duration b is long enough for a child to grow to be school age, mother 2 can devote relatively less attention to her child than when her child is younger. However, mother 1 would prepare to have her next pregnancy, which is censored at survey time T, and this would also affect her labor supply.

To address this possibility, in both the IV estimation and sub-sample analyses, we have regressed unconstrained samples and, separately, those with 3-year and 1-year intervals. To hold the time-variant preference and presence of younger baby constant, a short interval should be given, such as 6 months or 1 year. However, month information is not available in the data, and thus we assign a 1-year interval, along with a 3-year interval as a comparison. Although this approach can measure a precise decline in female labor supply immediately after they give birth, a huge sample size is required. To our knowledge, this paper is the first to analyze the effect of fertility on female labor supply while controlling for the time elapsed since the last childbirth in a cross-sectional census.

4.2. OLS and IV estimations

Table 3 presents the OLS and IV estimates of the effect of the number of children on female labor supply, in addition to the first-stage relationship between the number of children and twins at the nth delivery. The results without controlling for the

Table 3
Estimated coefficients of OLS and IV conditioned on birth order and time since last childbirth.

	Since the last childbirth								
	Unconditioned			No more than 3 years			No more than 1 year		
Variables	(1) OLS	(2) IV	(3) First stage	(4) OLS	(5) IV	(6) First stage	(7) OLS	(8) IV	(9) First stage
*Panel A: Mothers of twins at the first delivery **vs** Mothers of non-twins*									
Number of children	−0.049***	−0.051***		−0.052***	−0.097***		−0.057***	−0.105***	
	(0.001)	(0.008)		(0.001)	(0.010)		(0.002)	(0.014)	
Twins			0.723***			0.771***			0.796***
			(0.009)			(0.012)			(0.017)
Observations	612,504	612,504	612,504	342,877	342,877	342,877	194,748	194,748	194,748
*Panel B: Mothers of twins at the second delivery **vs** Mothers of non-twins with 2 or more births*									
Number of children	−0.047***	−0.028**		−0.039***	−0.041**		−0.038***	−0.065***	
	(0.001)	(0.012)		(0.002)	(0.017)		(0.003)	(0.022)	
Twins			0.884***			0.890***			0.919***
			(0.012)			(0.019)			(0.026)
Observations	421,605	421,605	421,605	213,680	213,680	213,680	114,168	114,168	114,168
*Panel C: Mothers of twins at the third delivery **vs** Mothers of non−twins with 3 or more births*									
Number of children	−0.026***	0.019		−0.017***	−0.011		−0.012**	0.018	
	(0.003)	(0.025)		(0.005)	(0.034)		(0.006)	(0.042)	
Twins			1.027***			1.024***			1.064***
			(0.026)			(0.040)			(0.061)
Observations	125,612	125,612	125,612	62,311	62,311	62,311	31,991	31,991	31,991

Notes: Robust standard errors in parentheses. All specifications control for age, age squared, years of schooling, ethnic minority status, husband's years of schooling, husband's employment status, co-residence with elder parents, and county dummies. In all panels, upper bounds on the number of children are not imposed. *** p<0.01, ** p<0.05, * p<0.1.

time since the last childbirth are reported in the first three columns, the results with the 3-year interval are reported in the middle three columns, and the results with the 1-year interval are reported in the last three columns. From top to bottom, three panels are listed for families with at least *n* births in ascending order of *n* from 1 to 3.

The OLS estimates in Column (1), Column (4), and Column (7) consistently show a significantly negative correlation between the number of children and the mother's labor force participation, regardless of the sample and duration assigned. For example, the OLS coefficient in Panel A (Column 1) suggests that, holding covariates controlled for in the regression constant, an additional child in the family reduces a mother's likelihood of labor force participation by approximately 5 percentage points, when time since the last childbirth is not restricted.

Using twin births as the instrument, the IV estimates in Column (2), Column (5), and Column (7) continue to suggest a negative effect of the number of children on the mother's outcome, except for Panel C of families with at least 3 births, and the coefficient sizes decrease in birth order. In particular, the IV coefficients on number of children are significant at the 1% level for families with at least one birth (Panel A), significant at the 5% level for those with two or more births (Panel B), and not significant, even at the 10% level, for families with three or more births (Panel C). Note that according to the previous discussion, the IV estimates may be subject to upward bias induced by not accounting for the time elapsed since the last childbirth.

As shown in Column (8), when a 1-year interval is assigned, the IV coefficients change dramatically from those in Column (2) except for Panel C, in which time passed since the last childbirth is not held constant. Specifically, for families with one or more births (Panel A), the estimates when the 1-year interval is assigned (Column 8) double to 10.5 percentage points, relative to those when the interval is not assigned (Column 2). Moreover, the estimates in Panel B change to 6.5 percentage points from 2.8 percentage points. However, the IV coefficients for families with 3 or more births (Panel C) remain highly stable even after the 1-year interval is assigned, which indicates no causal effects of fertility on female labor supply.

4.3. Sub-sample analyses

Table 4 presents the estimates of the sub-sample analyses of an additional child (caused by twin births) on female labor supply, where the preference for the number of children is held constant. The results without control variables are reported in odd columns, and the results with control variables are reported in even columns. Furthermore, the estimates are tabulated by duration since the last childbirth, where the first two columns are assigned no intervals, the middle two columns are assigned a 3-year interval, and the last two columns are assigned a 1-year interval. From top to bottom, three panels are listed for families with exactly *n* births in ascending order of *n* from 1 to 3.

The estimates of the sub-sample analyses are very similar to those from the IV analysis in Table 3, indicating negative effects of fertility on female labor supply except for families with more than 3 deliveries. For example, mothers of twins at the first delivery (Panel A) are 5.8 percentage points less likely to participate in the labor market (Column 2), and the results remain stable irrespective of the inclusion of control variables. However, estimates with and without control variables show a different

Table 4
Estimated coefficients from OLS for sub-samples.

	Since the last childbirth					
	Unconditioned		No more than 3 years		No more than 1 year	
	(1) No controls	(2) With controls	(3) No controls	(4) With controls	(5) No controls	(6) With controls
*Panel A: Mothers of twins at the first delivery **vs** Mothers of single child (2 **vs** 1)*						
Twins	−0.057***	−0.058***	−0.097***	−0.095***	−0.107***	−0.106***
	(0.008)	(0.008)	(0.011)	(0.011)	(0.015)	(0.015)
Observations	187,325	187,325	127,309	127,309	79,511	79,511
*Panel B: Mothers of twins at the second delivery **vs** Mothers of two non−twins children (3 **vs** 2)*						
Twins	−0.046***	−0.032**	−0.067***	−0.039**	−0.068***	−0.047**
	(0.013)	(0.012)	(0.018)	(0.017)	(0.025)	(0.023)
Observations	296,315	296,315	151,743	151,743	82,448	82,448
*Panel C: Mothers of twins at the third delivery **vs** Mothers of three non-twins children (4 **vs** 3)*						
Twins	−0.004	−0.001	−0.052	−0.033	−0.016	0.001
	(0.029)	(0.028)	(0.038)	(0.038)	(0.051)	(0.051)
Observations	111,495	111,495	54,528	54,528	27,847	27,847

Notes: Robust standard errors in parentheses. Specifications with controls include controls for age, age squared, years of schooling, ethnic minority status, husband's years of schooling, husband's employment status, co-residence with elder parents, and county dummies. *** p<0.01, ** p<0.05, * p<0.1.

pattern for higher-order births, which are marginally decreasing when controlling for demographic characteristics, husband's characteristics, and living arrangements.

As in the IV estimation, the results of the sub-sample OLS estimation also change dramatically when a 1-year interval is assigned. According to Column (6), an additional child (caused by a twin birth) reduces the labor force participation of a first-time mother by 10.6 percentage points (Panel A), which almost doubles the estimates in Column (2). A similar pattern can be found for a second-time mother (Panel B). However, an additional child no longer decreases female labor supply for a third-time mother (Panel C), and we even observe a positive but insignificant effect when a 1-year interval is assigned (Column 6).

He and Zhu (2015) use census data from Mainland China, which shares a very similar cultural background with Taiwan, however, only find a moderate fertility effect on female labor supply by IV and sub-sample OLS estimation. Because of China's one child policy, they analyze the effect of having a second child through twinning. Using the same instrument, twin births, the causal estimates in their paper consistently indicate a fertility effect by approximately 2 percentage points in 1990, even no significant effect in 2000. As they are focusing on females aged between 20 and 50 (16–35 in this paper), the fertility effect might be attenuated by elder female sample. Those who are over 40 years old would be less likely to have small children in the household, and the childcare could be relatively easy to be substituted by market services. Also, we estimate all models conditioned on time passed since the last childbirth, which measures the labor supply immediately after the fertility. The differences in sample size, women's age, and timing of labor supply might explain the differences in estimates. The estimates from He and Zhu (2015) explain the long-term effects of fertility on female labor force participation, however, this paper mainly explains the short-term effects.

5. Robustness check

In this section, we check the robustness of the results regarding selection and heterogeneity issues. First, we discuss parental selection into twin births, which may render the instrument invalid. Using a matching estimator, we find that the results are robust to IV and sub-sample estimates that have been discussed in the previous section. Second, we investigate the heterogeneous effects of fertility on female labor supply across living arrangement status, regions, and age groups. As the previous results are stable between IV and sub-sample OLS, we use the conventional IV estimates to discuss the heterogeneity issues.

5.1. Selection

To address parental selection into twin births, we construct a matched sample with extremely similar family backgrounds for mothers of twins and those of non-twins. As shown in Column (5) of Table 1, mothers of twins differ only with respect to employment status and the number of children and no longer differ in age, years of schooling, husband's education and employment status, and living arrangements. The test for balanced covariates is displayed in Fig. 4, indicating that the matched sample has much less standardized bias across covariates.

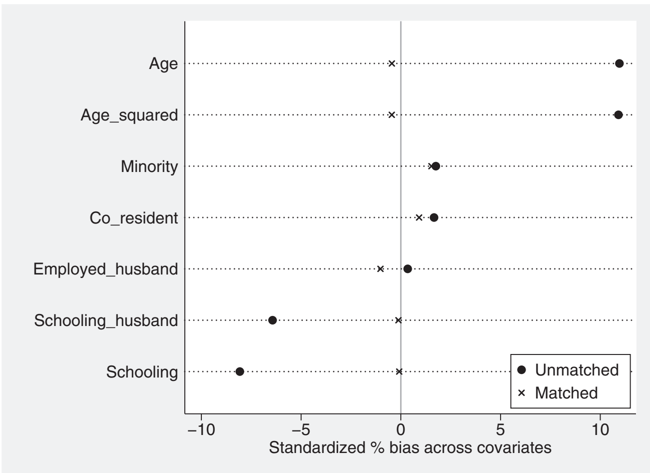

Fig. 4. The bias reduction of covariates after matching. Notes: Y-axis shows all covariates.

Table 5 presents the results of a matching estimation of twin births (treatment) on female labor supply, which is robust to IV estimation and sub-sample OLS estimation. Results from a sample without replacement are shown in odd columns and those from a sample with replacement are shown in even columns.

Compared to IV estimation and sub-sample OLS estimation, the estimates of the matched sample reveal a slightly smaller fertility effect for first-time mothers and a slightly larger fertility effect for second-time mothers when the time elapsed since the last childbirth is held constant. According to Column (6), an additional child (caused by a twin birth) reduces the labor force participation of a first-time mother by 9 percentage points (Panel A), while the estimates from IV and sub-sample OLS are 10.5 and 10.6 percentage points, respectively. For a second-time mother (Panel B), an additional child reduces labor supply by 7 percentage points, while the estimates from IV and sub-sample OLS are 6.5 and 4.7 percentage points, respectively. The fertility effect for a third-time mother remains statistically insignificant across different identification strategies.

5.2. Heterogeneity across living arrangement status

The effect of fertility on female labor supply could be heterogeneous across living arrangement status among Asian countries, where elder parents may help to look after or educate young children. The labor supply of females who are co-resident with elder parents could be less affected by fertility, as co-resident elder parents can partially take responsibilities in housework and child-caring work. Females co-resident with elder parents may have more time constraint to participate in the labor market than those who are living without elder parents. In this subsection, we perform OLS and IV estimations for mothers who are living with and without their elder parents, respectively, to investigate the heterogeneity across living arrangement status.

Table 6 presents the OLS and IV estimates of the effect of the number of children on female labor supply for mothers with and without elder parent co-residence. The results of the first four columns are unconditioned on the time since the last childbirth, while the middle and last four columns are assigned with the 3-year and 1-year interval, respectively. From top to bottom, three panels are listed for families with at least n births in ascending order of n from 1 to 3.

For a first-time mother (Panel A), the short-term fertility effect is quite similar in magnitude across living arrangement status. Holding the time passed since the last childbirth no more than 1 year, females living without elder parents (Column 10) are 11.2 percentage points less likely to participate in the labor market, while females living with elder parents (Column 12) are 11.7 percentage points less likely to do so. On the long-term, the IV estimates for first-time mothers are still significantly negative at 1% level for those who are living without elder parents (Column 2), however, no longer significant for those with elder parent co-residence (Column 4).

For a second-time mother (Panel B), the IV estimates across time since the last childbirth are quite stable for those who are living without elder parents, indicating significantly negative fertility effects by 3.4 –5.7 percentage points (Column 2, Column 6, and Column 10). On the contrary, the corresponding IV estimates for those who are living with elder parents are all not

Table 5
Matching estimator of twin births on female employment.

	Since the last childbirth					
	Unconditioned		No more than 3 years		No more than 1 year	
	(1) Without repl.	(2) With repl.	(3) Without repl.	(4) With repl.	(5) Without repl.	(6) With repl.
Panel A: Mothers of twins at the first delivery ***vs*** *Mothers of single child (2* ***vs*** *1)*						
Twins	−0.042	−0.033***	−0.069***	−0.065***	−0.090***	−0.090***
	(0.012)	(0.012)	(0.016)	(0.016)	(0.021)	(0.021)
Observations	6,984	6,984	3,812	3,812	2,110	2,110
Panel B: Mothers of twins at the second delivery ***vs*** *Mothers of two non-twins children (3* ***vs*** *2)*						
Twins	−0.044**	−0.049***	−0.039	−0.061**	−0.060*	−0.070**
	(0.018)	(0.018)	(0.025)	(0.025)	(0.035)	(0.035)
Observations	3,150	3,150	1,576	1,576	834	834
Panel C: Mothers of twins at the third delivery ***vs*** *Mothers of three non-twins children (4* ***vs*** *3)*						
Twins	−0.010	0.007	−0.012	−0.012	0.086	0.086
	(0.040)	(0.040)	(0.055)	(0.055)	(0.072)	(0.072)
Observations	612	612	330	330	186	186

Notes: Robust standard errors in parentheses. Columns (1), (3), (5) are without replacement, and Columns (2), (4), (6) are with replacement. In all specifications, 1-to-1 caliper matching is applied. The propensity score used for matching is calculated by logistic regression, which is available upon request. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6
Estimated coefficients of OLS and IV, by living arrangement.

	Since the last childbirth											
	Unconditioned				No more than 3 years				No more than 1 year			
	Living without parents		Living with parents		Living without parents		Living with parents		Living without parents		Living with parents	
Variables	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV	(9) OLS	(10) IV	(11) OLS	(12) IV
*Panel A: Mothers of twins at the first delivery **vs** Mothers of non-twins*												
Number of children	−0.049***	−0.067***	−0.042***	−0.026	−0.053***	−0.108***	−0.046***	−0.055*	−0.058***	−0.112***	−0.055***	−0.117***
	(0.001)	(0.007)	(0.003)	(0.023)	(0.001)	(0.009)	(0.004)	(0.029)	(0.002)	(0.012)	(0.005)	(0.038)
Observations	557,669	557,669	53,989	53,989	311,266	311,266	31,169	31,169	176,266	176,266	18,242	18,242
*Panel B: Mothers of twins at the second delivery **vs** Mothers of non-twins with 2 or more births*												
Number of children	−0.048***	−0.034***	−0.040***	−0.008	−0.038***	−0.043***	−0.041***	0.011	−0.037***	−0.057***	−0.045***	0.005
	(0.001)	(0.011)	(0.004)	(0.033)	(0.002)	(0.016)	(0.006)	(0.042)	(0.003)	(0.021)	(0.008)	(0.057)
Observations	382,263	382,263	37,904	37,904	193,101	193,101	19,802	19,802	102,951	102,951	10,780	10,780
*Panel C: Mothers of twins at the third delivery **vs** Mothers of non-twins with 3 or more births*												
Number of children	−0.030***	0.006	−0.007	0.039	−0.017***	−0.000	−0.014	−0.059	−0.013**	0.033	−0.006	−0.002
	(0.004)	(0.022)	(0.010)	(0.065)	(0.005)	(0.030)	(0.014)	(0.084)	(0.007)	(0.038)	(0.018)	(0.094)
Observations	111,712	111,712	13,413	13,413	55,337	55,337	6,714	6,714	28,390	28,390	3,472	3,472

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

significant (Column 4, Column 8, and Column 12), even at the 10% level. Similar with the results of full sample in Table 3, the IV estimates for third-time mothers (Panel C) in Table 6 are insignificant, which is robust across time since the last childbirth and living arrangement status.

5.3. Heterogeneity across regions

The fertility effect could be also heterogeneous between rural and urban areas. The reasons are as follows. First, rural females may have more flexible schedule than urban females if they are engaging in family-oriented enterprise. It is relatively easy for them to balance child caring with household work. Second, in rural areas, elder children are much more likely to help taking care of their younger siblings so that the mothers could do a full-time or part-time job as well. However, it is not possible to distinguish between urban and rural areas in Taiwanese census. As an alternative way, we estimate the OLS and IV models by regions, which are divided into four groups. In 2001, the northern region contributes to 60.0% of the gross domestic product (GDP), followed by southern (21.9%), central (17.1%), and eastern (1.0%) regions[6]. The share of gross product of a particular region in GDP can highly demonstrate its quality of urbanization in the context of Taiwan.

Table 7 presents the OLS and IV estimates of the effect of the number of children on female labor supply by regions. The results of the first two columns are unconditioned on the time since the last childbirth, while the middle and last two columns are assigned with the 3-year and 1-year intervals, respectively. From top to bottom, three panels are listed for families with at least n births in ascending order of n from 1 to 3. In each panel, estimates are listed for the northern, central, southern, and eastern regions.

The IV estimates show significant negative fertility effect for a first-time mother (Panel A) among the northern, central, and southern regions, but not significant for the eastern region[7]. In particular, holding the time passed since the last childbirth no more than 1 year, the fertility effect for the northern region (13.4%) is the largest in magnitude, followed by the central (9.8%) and southern (8%) regions. However, for a second-time mother (Panel B), the corresponding coefficients are only significant for the northern (5.7%) and southern (6.6%) regions at 10% level. Furthermore, the coefficients from IV estimation are all insignificant for a third-time mother (Panel C), regardless of region.

To further understand the regional difference, we estimate IV model by each county (or municipality) separately, and then show the graphical distribution of the short-term fertility effect for a first-time mother in Fig. 5. The estimates are only significant for relatively developed counties and municipalities. The results among the northern region is relatively homogeneous, however, the results of the central and southern regions are not. In the central region, the coefficients are only significant for Taichung County and Taichung City. In the southern region, the coefficients are only significant for Tainan City, Kaohsiung City, and Kaohsiung County.

5.4. Heterogeneity across age groups

In this subsection, we show the IV estimates by age groups for the following reasons. First, although monozygotic twinning is thought to be relatively constant throughout the world at approximately 4 per 1000 births, dizygotic twinning increases substantially with maternal age (Smits & Monden, 2011). Second, labor supply could be correlated with age by various kinds of mechanisms. Assuming a functional form of age in labor supply equation maybe not enough to control for age effect. Estimating the labor supply equation by age groups allows me to observe the heterogeneous fertility effects, which offers more policy implications on female's response to number of children.

Table 8 shows the IV estimates by three age groups, which are 20–25, 25–30, and 30–35. All estimates are conditioned on time since the last childbirth and parity, sharing a similar structure with previous tables. The results of the 20–25 age group are shown in Columns(1), (4), and (7), while those of 25–30 age group are shown in Columns(2), (5), and (8). The rest of the columns show the results of 30–35 age group.

Regardless of age, the effect of the number of children on female labor supply remains negative for a first-time mother when time since the last childbirth is holding constant (Panel A). For example, an additional child reduces the labor force participation of mothers aged 20–25 by 11.3% (Column 7), aged 25–30 by 12.1% (Column 8), and aged 30–35 by 10.2% (Column 9). However, for a second-time mother (Panel B), the corresponding estimate is only significant for 30–35 age group, no longer significant for 20–25 and 25–30 age groups. Note that age is both related to number of children and labor supply, the insignificant results for 20–25 and 25–30 age groups may suffer from sample selection bias. Given the fact of later

[6] Source: National Development Council.

[7] Northern region contains Taipei County, Yilan County, Taiyuan County, Hsinchu County, Keelung City, Hsinchu City, and Taipei City. Central region contains Miaoli County, Taichung County, Changhua County, Nantou County, Yunlin County, and Taichung City. Southern region contains Chiayi County, Tainan County, Kaohsiung County, Pinging County, Penghu County, Chiayi City, Tainan City, and Kaohsiung City. Eastern region contains Taitung County and Hualien County. Kinmen-Matsu region is not used for sample size issue.

Table 7
Estimated coefficients of OLS and IV, by region.

	Since the last childbirth					
	Unconditioned		No more than 3 years		No more than 1 year	
Variables	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
*Panel A: Mothers of twins at the first delivery **vs** Mothers of non-twins*						
Northern	−0.055***	−0.071***	−0.059***	−0.118***	−0.067***	−0.134***
	(0.001)	(0.010)	(0.002)	(0.012)	(0.002)	(0.016)
Observations	283,151	283,151	164,092	164,092	94,208	94,208
Central	−0.035***	−0.062***	−0.038***	−0.090***	−0.041***	−0.098***
	(0.002)	(0.015)	(0.002)	(0.020)	(0.003)	(0.026)
Observations	144,718	144,718	78,550	78,550	43,891	43,891
Southern	−0.036***	−0.047***	−0.040***	−0.086***	−0.048***	−0.080***
	(0.002)	(0.013)	(0.002)	(0.017)	(0.003)	(0.022)
Observations	171,114	171,114	92,805	92,805	52,587	52,587
Eastern	−0.037***	−0.057	−0.036***	−0.071	−0.030***	−0.145
	(0.006)	(0.053)	(0.007)	(0.080)	(0.010)	(0.103)
Observations	11,640	11,640	6,299	6,299	3,398	3,398
*Panel B: Mothers of twins at the second delivery **vs** Mothers of non-twins with 2 or more births*						
Northern	−0.054***	−0.050***	−0.046***	−0.057**	−0.047***	−0.057*
	(0.002)	(0.016)	(0.003)	(0.022)	(0.004)	(0.030)
Observations	186,118	186,118	97,648	97,648	52,704	52,704
Central	−0.030***	0.007	−0.017***	0.034	−0.014***	0.005
	(0.003)	(0.020)	(0.004)	(0.028)	(0.005)	(0.038)
Observations	105,465	105,465	52,234	52,234	27,500	27,500
Southern	−0.035***	−0.035	−0.029***	−0.069**	−0.033***	−0.066*
	(0.003)	(0.021)	(0.004)	(0.029)	(0.005)	(0.036)
Observations	119,275	119,275	58,209	58,209	31,021	31,021
Eastern	−0.048***	−0.029	−0.033***	−0.011	−0.021	0.001
	(0.008)	(0.068)	(0.011)	(0.116)	(0.014)	(0.089)
Observations	8,535	8,535	4,326	4,326	2,205	2,205
*Panel C: Mothers of twins at the third delivery **vs** Mothers of non-twins with 3 or more births*						
Northern	−0.035***	0.024	−0.028***	0.028	−0.027***	0.019
	(0.006)	(0.034)	(0.008)	(0.042)	(0.010)	(0.051)
Observations	49,738	49,738	25,610	25,610	13,253	13,253
Central	−0.016***	0.033	−0.003	−0.019	0.004	0.031
	(0.006)	(0.039)	(0.008)	(0.056)	(0.011)	(0.078)
Observations	37,231	37,231	17,914	17,914	9,065	9,065
Southern	−0.014**	−0.031	−0.008	−0.067	−0.006	0.017
	(0.007)	(0.040)	(0.009)	(0.055)	(0.013)	(0.068)
Observations	34,669	34,669	16,621	16,621	8,532	8,532
Eastern	−0.031**	−0.035	−0.000	−0.008	0.014	0.085
	(0.016)	(0.132)	(0.019)	(0.148)	(0.023)	(0.204)
Observations	3,116	3,116	1,665	1,665	873	873

Notes: Robust standard errors in parentheses. All specifications control for age, age squared, years of schooling, ethnic minority status, husband's years of schooling, husband's employment status, and county dummies. In all panels, upper bounds on the number of children are not imposed. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

marriage in Taiwan[8], mothers who have already given two or more birth before age 25 or 30 maybe different from their peers in characteristics.

6. Policy implications

Among aging societies, fertility has a critical policy implication. The low fertility rate along with the increasing life expectancy leads to potential shortages of labor force. Asian countries like Taiwan, Japan, and Korea have carried out maternity benefits or childbirth subsidies to promote the fertility rate. A key question is that whether more children will discourage female labor force participation.

This paper indicates a possibility that policy makers can simultaneously increase the fertility rate and female labor supply. According to the estimates in Table 3 and Table 4, the negative effect of fertility on female labor force participation is not

[8] The average age at first marriage is 26.1 for female in 2000. Data source: Department of Household Registration, MOI.

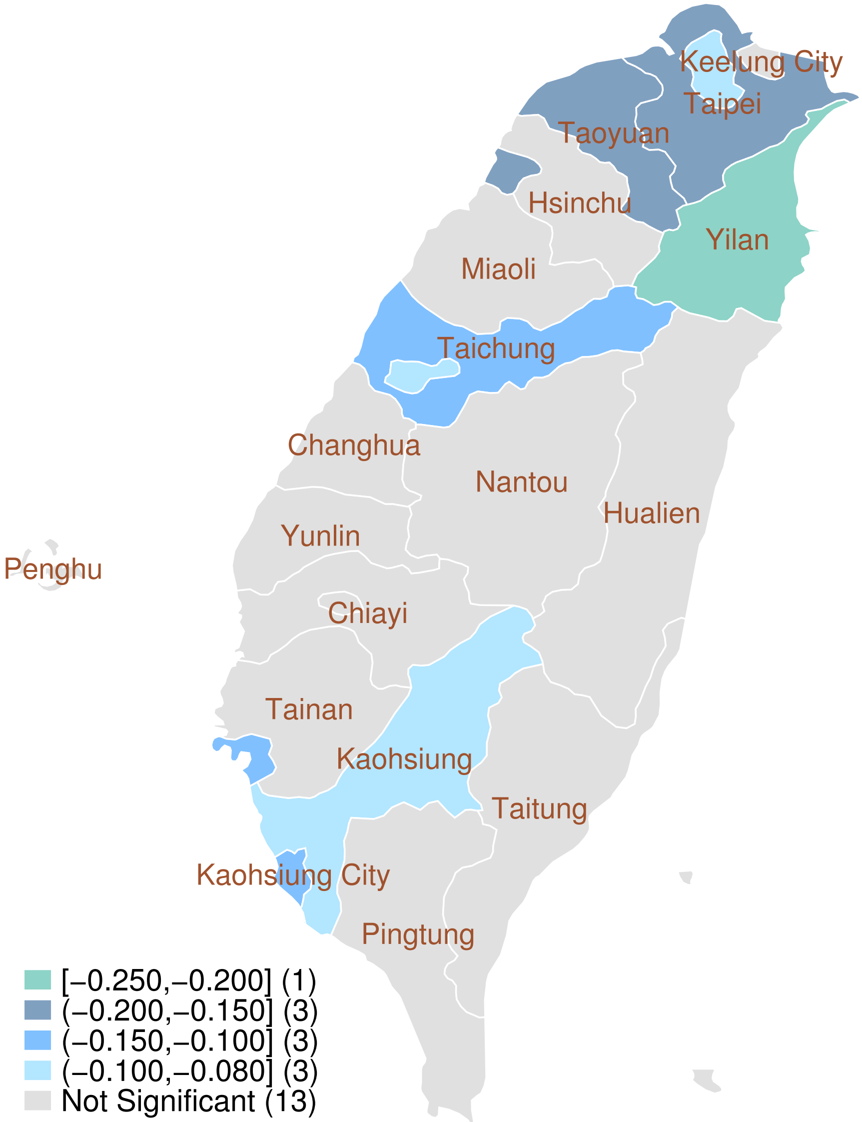

Fig. 5. Graphical distribution of the short-term fertility effect for a first time mother. Notes: We estimate Eq. (1) from IV estimation for 23 counties/municipalities separately. All regressions include age, age squared, years of schooling, ethnic minority status, husband's years of schooling, husband's employment status, and co-residence with elder parents. Insignificant coefficients are displayed by gray color. Due to space constraint, Chiayi City, Taichung City, Hsinchu City, and Taipei City are not marked by text in the figure.

monotonically decreasing in the number of birth. Compared to a first-time mother, the fertility effect gradually declines for those who have two or more deliveries, with the effects vanishing when females have three or more deliveries. Other than offering maternity benefits to a first-time mother, policy makers should encourage females to have multiple births as female labor supply would be less, or not affected. Furthermore, according to Table 6, fertility has no effects on labor supply for females who are living with their elder parents, when they have two or more births. Policy makers should also take living arrangement status into account.

Although this paper focuses on Taiwanese population, the evidence may have implications for Mainland China and Hongkong, where they share a similar cultural background with Taiwan. Especially in Mainland China, the government has relaxed the family planning policy that allows each family to have two children since 2016. According to estimates in this paper, having a second child will significantly decrease the female labor supply on the short run. The Chinese government should additionally conduct policies on female employment or childcare support to avoid the potential decline in labor supply among married women.

Table 8
Estimated coefficients of IV, by age groups.

	Time since the last childbirth								
	Unconditioned			No more than 3 years			No more than 1 year		
Variables	(1) [20, 25]	(2) [25, 30]	(3) [30, 35]	(4) [20, 25]	(5) [25, 30]	(6) [30, 35]	(7) [20, 25]	(8) [25, 30]	(9) [30, 35]
*Panel A: Mothers of twins at the first delivery **vs** Mothers of non-twins*									
Number of children	-0.077^{***}	-0.093^{***}	-0.049^{***}	-0.087^{***}	-0.114^{***}	-0.101^{***}	-0.113^{***}	-0.121^{***}	-0.102^{***}
	(0.027)	(0.013)	(0.009)	(0.028)	(0.015)	(0.013)	(0.032)	(0.018)	(0.018)
Observations	30,691	151,257	342,420	28,274	117,310	171,063	20,757	74,541	87,657
*Panel B: Mothers of twins at the second delivery **vs** Mothers of non-twins with 2 or more births*									
Number of children	−0.073	−0.001	-0.052^{***}	-0.083^{*}	0.007	-0.064^{***}	−0.032	0.011	-0.105^{***}
	(0.048)	(0.023)	(0.014)	(0.049)	(0.027)	(0.020)	(0.057)	(0.034)	(0.028)
Observations	13,786	86,476	249,832	12,770	64,173	116,818	9,154	38,572	57,932
*Panel C: Mothers of twins at the third delivery **vs** Mothers of non-twins with 3 or more births*									
Number of children	0.109	−0.025	0.010	0.109	−0.028	−0.005	0.209	0.016	0.051
	(0.158)	(0.049)	(0.027)	(0.158)	(0.052)	(0.036)	(0.189)	(0.062)	(0.047)
Observations	2,199	20,148	77,277	2,117	15,565	36,829	1,605	9,334	17,759

Notes: Robust standard errors in parentheses. All specifications control for age, age squared, years of schooling, ethnic minority status, husband's years of schooling, husband's employment status, and county dummies. In all panels, upper bounds on the number of children are not imposed. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

7. Concluding remarks

This paper estimates the causal effect of fertility on female labor supply. Using a universal sample of the Taiwanese population, which allows me to analyze the fertility decisions by each order to observe the heterogeneous effect across families with different preferences for the number of children, we exploit the exogenous variation in the number of children caused by twin births to identify the causal effect of fertility on the labor supply of married women.

Four major findings are as follows. First, the effect of fertility on female labor supply is not monotonically decreasing in the number of children. The IV estimates show that fertility reduces the female labor supply by 10.5 percentage points for a first-time mother and 6.5 percentage points for a second-time mother. However, the fertility effect becomes positive and statistically insignificant for a third-time mother. This evidence offers a new policy implication that the government could encourage the fertility of mothers who have given multiple births to simultaneously increase the total fertility rate and female labor force participation rate. The differences in decision-making regarding labor supply are assumed to be associated with unobserved marriage-specific human capital.

Second, this paper shows that the effect of fertility varies substantially with the time elapsed since the last childbirth. We have constructed unconstrained samples and samples with 3-year and 1-year intervals to analyze the heterogeneous effects with respect to the time elapsed since the last childbirth. For a first-time mother, the IV estimate with a 1-year interval assigned is approximately twice the IV estimate with no interval. However, the OLS estimates are stable across samples with different intervals. This evidence indicates that the IV estimation without controlling for birth intervals would depend on the distribution of the time elapsed since the last childbirth, which has consequences for differences in estimates across samples in the literature.

Third, we use a new identification strategy to analyze the effect of one additional child on female labor supply with an efficient instrument, which captures a real sense of the preference for the number of children for mothers with n births. In the literature, when a two-stage least squares estimation is applied, preferences for the number of children are assumed to be the same among mothers who have at least n births. By constructing the three sub-samples described in Table 2, the exogenous variation in the number of children arises solely from twin births among mothers who have the same number of births. Thus, we present a sub-sample OLS estimation to compare with classical IV estimation. The results of the sub-sample OLS are similar to those of IV, which indicates that preferences for the number children may not change substantially between mothers with n and $n+1$ births.

Finally, we investigate the heterogeneous fertility effects across living arrangement status, regions, and age groups. For a first time mother, the short-term fertility effect is quite stable across living arrangement status and age groups, ranging in approximately 10–11 %. The corresponding estimates show some regional differences, ranging in approximately 8–13% among the northern, central, and southern areas, but not significant for the eastern area. For a second-time mother, the short-term estimates are only significant for females who are not living with their elder parents, who are living in well developed areas (northern and southern areas), and who are in the 30–35 age group. The estimates for higher parity consistently remain insignificant, regardless of the subsamples used.

Appendix A

Table A1
Full estimated coefficients of OLS and IV uncontioned on birth order and time since last childbirth.

	OLS				IV				First Stage			
Variables	(1) Employed	(2) Employed	(3) Employed	(4) Employed	(5) Employed	(6) Employed	(7) Employed	(8) Employed	(9) #Children	(10) #Children	(11) #Children	(12) #Children
Number of children	−0.058***	−0.049***	−0.048***	−0.048***	−0.049***	−0.050***	−0.050***	−0.050***				
	(0.001)	(0.001)	(0.001)	(0.001)	(0.005)	(0.005)	(0.005)	(0.005)				
Twins									0.886***	0.848***	0.846***	0.846***
									(0.007)	(0.007)	(0.007)	(0.007)
Age		0.073***	0.071***	0.071***		0.073***	0.071***	0.072***		0.109***	0.117***	0.117***
		(0.003)	(0.003)	(0.003)		(0.003)	(0.003)	(0.003)		(0.003)	(0.003)	(0.003)
Age squared/100		−0.099***	−0.096***	−0.097***		−0.099***	−0.097***	−0.097***		−0.088***	−0.098***	−0.098***
		(0.004)	(0.004)	(0.004)		(0.004)	(0.004)	(0.004)		(0.006)	(0.006)	(0.006)
Years of schooling		0.028***	0.026***	0.026***		0.028***	0.026***	0.026***		−0.082***	−0.057***	−0.056***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Minority		0.017***	0.018***	0.019***		0.017***	0.019***	0.019***		0.166***	0.150***	0.151***
		(0.004)	(0.004)	(0.004)		(0.005)	(0.004)	(0.004)		(0.008)	(0.008)	(0.008)
Husband												
Years of schooling			0.003***	0.003***			0.003***	0.003***			−0.037***	−0.037***
			(0.000)	(0.000)			(0.000)	(0.000)			(0.000)	(0.000)
Employed			0.057***	0.057***			0.057***	0.058***			0.044***	0.044***
			(0.003)	(0.003)			(0.003)	(0.003)			(0.005)	(0.005)
Elderly												
Co-resident				0.038***				0.038***				0.056***
				(0.002)				(0.002)				(0.003)
Constant	0.709***	−0.932***	−0.971***	−0.979***	0.691***	−0.932***	−0.971***	−0.978***	1.930***	0.266***	0.227***	0.216***
	(0.002)	(0.037)	(0.038)	(0.038)	(0.009)	(0.038)	(0.038)	(0.038)	(0.001)	(0.050)	(0.050)	(0.050)
County dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	614,902	614,902	614,902	614,902	614,902	614,902	614,902	614,902	614,902	614,902	614,902	614,902
F of the instrument	-	-	-	-	-	-	-	-	15337.39	15691.32	15825.00	15820.13

Notes: Robust standard errors in parentheses. The observations are slightly different with unconditioned case of Panel A in Table 3, which only uses twins at the first delivery. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

References

Angrist, J. D., & Evans, W. N. (1998). Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review*, *88*(3), 450–477.
Angrist, J. D., Lavy, V., & Schlosser, A. (2010). Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, *28*(4), 773–824.
Åslund, O., & Grönqvist, H. (2010). Family size and child outcomes: Is there really no trade-off? *Labour Economics*, *17*(1), 130-9.
Azimi, E. (2015). The effect of children on female labor force participation in urban Iran. *IZA Journal of Labor and Development*, *4*(1), 1—1 1.
Becker, G. S. (1985). Human capital, effort, and the sexual division of labor. *Journal of Labor Economics*, *3*(1), S33–S58.
Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). The more the merrier? The effect of family size and birth order on children's education. *Quarterly Journal of Economics*, *120*(2), 669–700.
Black, S. E., Devereux, P. J., & Salvanes, K. G. (2010). Small family, smart family? Family size and the IQ scores of young men. *Journal of Human Resources*, *45*(1), 33–58.
Bronars, S. G., & Grogger, J. (1994). The economic consequences of unwed motherhood: Using twin births as a natural experiment. *American Economic Review*, *84*(5), 1141–1156.
Central Intelligence Agency (2015). The World FactBook: Taiwan, https://www.cia.gov/library/publications/the-world-factbook/geos/tw.html.
Chen, C. J., Lee, T. K., Wang, C. J., & Yu, M. W. (1992). Secular trend and associated factors of twinning in Taiwan. *Acta Genet Med Gemellol (Roma)*, *41*(2-3), 205–213.
Chen, C. J., Lin, T. M., Chang, C., & Cheng, Y. J. (1987). Epidemiological characteristics of twinning rates in Taiwan. *Acta Genet Med Gemellol (Roma)*, *36*(3), 335–342.
Chen, Y.-H. (2012). Trends in low fertility and policy responses in Taiwan. *Japanese Journal of Population*, *10*(1), 78–88.
Chou, S.-Y., Liu, J.-T., Grossman, M., & Joyce, T. (2010). Parental education and child health: Evidence from a natural experiment in Taiwan. *American Economic Journal: Applied Economics*, *2*(1), 33–61.
Chun, H., & Oh, J. (2002). An instrumental variable estimate of the effect of fertility on the labour force participation of married women. *Applied Economics Letters*, *9*(10), 631–634.
Dudel, C. (2009). *The demographic dilemma: Fertility, female labor force participation and future growth in Germany 2007–2060*, No. 158. http://ssrn.com/abstract=1357247.
Ebenstein, A. (2009). When is the local average treatment close to the average? Evidence from fertility and labor supply. *Journal of Human Resources*, *44*(4), 955-75.
Freedman, R., Fan, T.-H., Wei, S.-P., & Weinberger, M. B. (1977). Trends in fertility and in the effects of education on fertility in Taiwan, 1961-74. *Studies in Family Planning*, *8*(1), 11–18.
Freedman, R., Hermalin, A. I., & Chang, M.-C. (1975). Do statements about desired family size predict fertility? The case of Taiwan, 1967-1970. *Demography*, *12*(3), 407–416.
Gronau, R. (1973). The effect of children on the housewife's value of time. *Journal of Political Economy*, *81*(2), S168–S199.
He, X., & Zhu, R. (2015). Fertility and female labour force participation: Causal evidence from urban China. *The Manchester School*,
Heckman, J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, *42*(4), 679–694.
Li, H., Zhang, J., & Zhu, Y. (2008). The quantity-quality trade-off of children in a developing country: Identification using Chinese twins. *Demography*, *45*(1), 223–243.
Lin, W.-I., & Yang, S.-Y. (2009). From successful family planning to the lowest of low fertility levels: Taiwan's dilemma. *Asian Social Work and Policy Review*, *3*(2), 95–112.
Mincer, J. (1962). Labor force participation of married women: A study of labor supply. . *Aspects of labor economics*,(pp. 63–105). National Bureau of Economic Research, Inc.
Rosenbluth, F., Light, M., & Schrag, C. (2002). *The politics of low fertility: Global markets, women's employment, and birth rates in four industrialized democracies*,
Rosenzweig, M. R., & Wolpin, K. I. (1980a). Life-cycle labor supply and fertility: Causal inferences from household models. *Journal of Political Economy*, *88*(2), 328–348.
Rosenzweig, M. R., & Wolpin, K. I. (1980b). Testing the quantity-quality fertility model: The use of twins as a natural experiment. *Econometrica*, *48*(1), 227–240.
Schultz, T. P. (1978). The influence of fertility on labor supply of married women : Simultaneous equation estimates. *Research in Labor Economics*, *2*, 273–351.
Smits, J., & Monden, C. (2011). Twinning across the developing world. *PLOS ONE*, *6*(9), 1–5.