ELRR

The effect of the universal two-child policy on female labour market outcomes in China

The Economic and Labour Relations Review
2022, Vol. 33(3) 526–546

DOI: 10.1177/10353046221112716
journals.sagepub.com/home/elra
SAGE

Qian Huang and Xiaofei Jin
Nankai University, People's Republic of China

Abstract
To alleviate the imbalance in demographics, the Chinese government initiated the universal two-child policy nationwide in 2016, which has comprehensively impacted society, especially females. Our study investigates whether this policy has negatively affected workforce employment and income among women in urban areas. Based on the DID (difference-in-differences) method and the Heckman Two-Step Estimation, reliable empirical evidence shows that the universal two-child policy has significantly reduced women's employment by 4.06% and decreased their labour income by 10.43%. Surprisingly, this policy has decreased the employment among women under 25 years old by 23.99% and has reduced the income of higher educated females by 29.59%. Furthermore, we find that the influence of the universal two-child policy on female employment has gradually increased from 2016 to 2018, and its impact on income has presented an evident time lag.

JEL Codes: J13, J21, J31

Keywords
The universal two-child policy, fertility, female employment and income, difference-in-differences

Introduction

China has implemented a strict and long-term population control policy resulting in a rapid drop in total fertility rate (TFR) from 5.8 in 1970 to 2.6 in 1981. Since 1990, the

Corresponding author:
Xiaofei Jin, School of Economics, Nankai University, No 94 Weijin Road, Nankai District, Tianjin, 300071, China.
Email: jin_xiaofei@163.com

TFR has fallen below the replacement level (Peng, 2011). The percentage of women in the workforce has remained high because, with fewer children, there is less domestic labour (Maurer-Fazio et al., 2006). However, this demographic predicament has brought China several challenges since 2010, such as a shrinking labour workforce, persistent fertility decline, a severe sex ratio imbalance, and most significantly the deepening of population ageing. The Chinese government has attempted to revise family planning policy to address these challenges. Thus, the universal two-child policy came into force on January 1, 2016, making it legal for all Chinese couples to have two children.

Studies have shown that, as the number of children decreases, women spend more time and energy engaging in work; in addition, a decline in fertility increases women's labour-force participation and income returns (Goldin et al., 2006). For example, from 1963 to 1989, America's fertility experienced an evident decline, and women's employment rate rose from 40.1% to 76.1% (Gash, 2009). In contrast, the increasing number of children has reduced the labour participation of married women (e.g., Angrist and Evans, 1996), and having additional children causes sizable reductions in women's labour supply (Cools et al., 2017). Women sacrifice advantages in the labour market due to motherhood, and organisations penalise those who choose to give birth (e.g., Budig and England, 2001; Doren, 2018). Both women's employment opportunities and earnings decline when they have children (Waldfogel, 1997), referred to as the 'motherhood penalty'.

Women's labour participation in China has been relatively high but has continuously declined from 70.52% to 61.82% during the past twenty years (World Bank, 2020). The gap between men's and women's labour participation rates has widened from 11.6 percentage points to 14.8 percentage points from 1990 to 2019. In contrast, this gap in the United States, the European Union, Japan, and the world, in general, has narrowed during the same period. Furthermore, motherhood leads to declines not only in the women's labour participation rate but also in their earnings throughout their entire career, while fatherhood is always associated with wage premiums (International Labour Organization, 2018). Tang and Scott (2017) estimate that China's gender wage gap has reached 25%.

Under such circumstances, does the universal two-child policy cause women to face a more complicated choice between motherhood and employment? Does it exacerbate gender discrimination in the labour market and harm women's labour rights? Searching for answers to these questions, we may contribute to the existing literature in the following ways.

First, to the best of our knowledge, we are the first to empirically analyse the influence of China's universal two-child policy on women's employment and income using the difference-in-differences method, which has substantial implications for women's labour supply, fertility intentions, and social policymaking. Second, since counter-natal policies rarely exist in developed countries, the present literature mainly concentrates on the 'motherhood penalty,' which refers to the phenomenon under which raising children negatively affects mothers' labour market outcomes. However, China's fertility policy affects all adult women's employment and income. China has converted the mandatory, stable, and long-term one-child policy to a two-child policy, which provides a unique opportunity to demonstrate the systemic influence of fertility policy changes on all women in the workforce. We use China's natural experiment to explore the relationship between fertility and work in this field. Third, we find that the policy's influence plays a different role for

different types of women. This analysis helps further illuminate diverse women's responses to the policy in a transitional economy like China. Finally, we perform the preliminary exploration of whether and how the policy's influence varies over time.

We find that China's universal two-child policy has significantly reduced women's employment by 4.06% and labour income by 10.43%. The policy creates a disadvantage for young and higher-educated women and, those working in private companies. In addition, the policy's influence on women's employment has gradually strengthened from 2016 to 2018, and its effect on income has presented an evident time lag.

Background and literature review

In this section, we introduce the evolution of China's family planning policy, which provides the background for our study. We also review the literature informing our research.

Evolution of family planning policy in China

Chinese society stabilised and experienced sizable population growth after 1949, and the TFR reached more than six births per woman in the early 1960s (Banister, 1987). It directly pushed China's central government to issue an administrative document regarding population control in 1962, which marked the beginning of China's family planning policies.

During the 1970's, in accordance with this document, Chinese local authorities suggested the number of children that each couple could have, be dropped from three to two. Simultaneously, a campaign arose for later marriage, longer intervals between births, and fewer children, known as the 'Later, Longer, and Fewer' policy. This mostly voluntary campaign was successful and led to a massive decrease in the TFR.

However, overpopulation fears persisted as the baby boomers born after the great famine were entering reproductive ages. Then, recognised as the world's strictest family planning policy, the one-child policy was implemented in 1979. It requested that each Han Chinese couple only have one child, and households with 'above-quota births' would be penalised. A second child was permitted only under specific circumstances.

Owing to the traditional preference for boys, the rule was deemed virtually unenforceable. The one-child policy was gradually relaxed to allow couples in rural areas, where around 70% of the total population lived, to have a second birth if the first one was a girl—called the one-and-a-half-child policy. After 1982, more provinces relaxed the conditions for a second birth, such as if both parents were the only children or if the first child was disabled.

China's fertility policy has become relatively stable, and since 1990, the TFR has dropped lower than the replacement level. In recent decades, China has begun to face the mounting pressures associated with continued low fertility, and the one-child policy in such a demographic contest was no longer defensible (Peng, 2011). The Chinese government has begun to relax the population control policy gradually. Around 2014, couples in which one partner was an only child were generally allowed to have a second child, and the universal two-child policy replaced the iconic one-child policy on January 1, 2016.

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Due to the universal two-child policy, the sluggish fertility trend has not been reversed as expected. The number of newborns in China was 17.86 million in 2016, reaching the maximum level since 2000 and decreasing in later years. Considering the fading fertility accumulation effect and women's lukewarm willingness to give birth a second time, the number of newborns has plateaued. Having only one child is endogenously determined within households, not by the policy constraints. Knowing the possible barriers to childbearing and taking feasible action is the top priority for the Chinese government.

Literature review

Previous literature has revealed that having children leads to a reduction in the female labour supply (Angrist and Evans, 1996). This negative effect not only arises among women aged 20-39 but also persists among older women (Bloom et al., 2009). The impact of children on women's labour market outcomes is widespread, affecting in terms of labour force participation, working hours, wages, occupations, sectors, and firms, while men are virtually unaffected (Lokshin, 2004; Kleven et al., 2019). Furthermore, mothers have a lower transition rate to re-employment than childless women (Frodermann et al., 2019).

Several fruitful studies have examined the aforementioned 'motherhood penalty'. Women who take a break from work for childbearing experience downward occupational mobility, long-term wage decline, and a low probability of being re-employed, even for only a few months (Angrist and Evans, 1996). Waldfogel (1997) finds that, compared to unmarried women, hourly wages are 6% lower for women with one child and 13% for women with two children. The annual income gap is even larger because mothers are more inclined to seek part-time jobs or switch their work departments in the process of parenting, and such translations often lead to lower wages (Adair et al., 2002; Juhn and McCue, 2017). Yu and Xie (2014) find that in China having one child reduces women's wages by 7.0%, and two children reduce them by 16.8%, compared to childless women. Budig and England (2001) summarise the four causes of mothers' declining income: childbirth and childcare lead to (1) interruptions in work experience or at least interrupt their full-time work experience; (2) less efficient work; (3) choice of 'mother-friendly' jobs rather than highly-paid jobs; (4) experience discrimination from employers.

Role Conflict Theory asserts that this situation is due to the inherent tension between household labour and career development. Women undertake the roles of mothers and employees, which induces virtually exclusive family-work conflict (Greenhaus and Beutell, 1985). Throughout the world, the general social culture, regardless of Western or Eastern, agrees that women's responsibility for taking care of the family is a greater priority than their career development (Fursman, 2002). In East Asian countries, such as China, Japan, and South Korea, mothers 'opt out' of the labour market because of gender discrimination, limited childcare options, gender-specific role expectations for women, and male-centered employment systems (Kinoshita and Guo, 2015; Rudolf and Kang, 2015). He and Zhu (2015) exploit twin births as a natural experiment and find a negative, but minimal, causal impact of a second child on the female labour force participation in urban China. They infer that relaxing the one-child policy does not significantly affect the urban female labour supply. Kinoshita and Guo (2015) point out that since childcare is generally provided by grandparents in rural areas, reduction in fertility does not

increase parental labour supply. Due to the extended family structure, the researchers suggest a small or even non-existent relationship between fertility and the female labour supply in China, which may otherwise exist in developed countries. They infer that the effect of the two-child policy on the labour supply, if any, tends to be minimal. Agarwal et al. (2019) study the effect of selective two-child policy (if one parent is from an only-child family, the couple is allowed to have a second child) on the female workforce and find a significantly negative influence.

Empirical design

DID method

This paper employs the classic difference-in-differences (DID) method (Angrist and Krueger, 1999), which has been widely utilised in public policy evaluations in recent years, to avoid potential endogenous problems, such as omitted variables or causality, and to improve estimation accuracy.

Considering that the universal two-child policy has been legally effective since 2016, we use the sample of 2014 China Family Panel Studies (CFPS) data as the pre-treatment group, with the 2018 data as the post-treatment group. The greatest challenge in our study is that the treatment group is not distinct. That is, there is no simple distinction between treatment and no treatment, as would be the case in most situations in which a DID method might be applied. This challenge calls for further exploration. Owing to endogenously determined local regulations and variations in one-child policy implementation across provinces (McElroy and Yang, 2000), the universal two-child policy's influence differs by region. If the province had a lower actual fertility rate before the universal two-child policy came into effect, it would probably be more strongly affected by this new policy. If the province's fertility rate was already higher, it could be less influenced, so we need to find a feasible fertility rate as the indicator to divide the sample into the treatment and control groups. Finally, we select the policy fertility rate (PFR). We set aside the commonly used TFR in regions because scholars have debated its accuracy in China (e.g., Chen and Yang, 2014; Guo, 2011; Zhu, 2012). The PFR is defined as the average number of children born if a woman gives birth according to the local policy during her lifetime (Guo et al., 2003). Local authorities' implementation variables, such as local-specific situations, fines for above-quota births, one-child subsidies, and the provision of contraceptives, have often varied across regions; so the PFR differs in each province.

The population-weighted national PFR of 1.465 is used as the cut-off rate based on the calculations in Table 1, following the identification strategy by Duflo (2001). After eliminating Tibet[1], we group 12 provinces as the treatment group (Jiangxi, Shandong, Jilin, Guangdong, Heilongjiang, Liaoning, Chongqing, Sichuan, Tianjin, Beijing, Jiangsu, Shanghai) and 13 provinces as the control group (Hubei, Zhejiang, Hunan, Anhui, Fujian, Shanxi, Henan, Shaanxi, Guangxi, Gansu, Hebei, Inner Mongolia, Guizhou, Yunnan, Qinghai, Ningxia, Hainan, Xinjiang).[2]

According to the Statistical Yearbooks in each province, growths in the birth rate for the treatment group are much higher than those for the control group after the universal two-child policy. Taking 2014 as a reference, growths in the birth rate are 7.38%,

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Table 1. PFR of China's 30 provinces (population-weighted corrected).

Province	PFR	Province	PFR	Province	PFR
Shanghai	1.060	Shandong	1.453	Guangxi	1.527
Jiangsu	1.060	Jiangxi	1.464	Gansu	1.559
Beijing	1.086	Hubei	1.466	Hebei	1.592
Tianjin	1.167	Zhejiang	1.467	Inner Mongolia	1.602
Sichuan	1.188	Hunan	1.479	Guizhou	1.667
Chongqing	1.273	Anhui	1.480	Yunnan	2.006
Liaoning	1.383	Fujian	1.481	Qinghai	2.104
Heilongjiang	1.392	Shanxi	1.487	Ningxia	2.116
Guangdong	1.413	Henan	1.505	Hainan	2.137
Jilin	1.450	Shanxi	1.514	Xinjiang	2.366

Source: Guo, Zhang, et al. (2003).

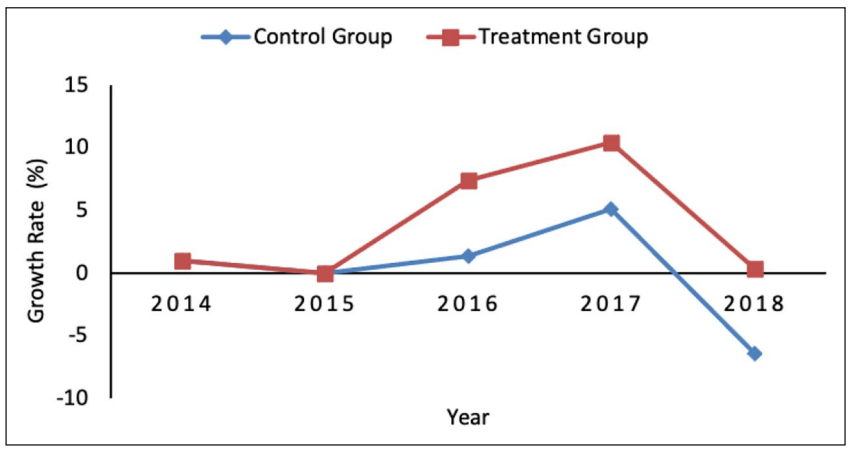

Figure 1. Annual rate of growth of birth rate in China, 2015–2018.
Source: Calculated by the data from China Statistical Yearbook of 2014–2018.

10.45%, and 0.35% for the treatment group, and 1.37%, 5.16%, and −6.44% for the control group in 2016, 2017, and 2018, respectively, as shown in Figure 1. The growths in birth rate differ significantly between the control and treatment groups, and the latter is much more sensitive to the universal two-child policy. The assumption is affirmed to be justified and valid.

In practice, we use the following regression-adjusted DID model to examine the policy impact. Consider

$$Emp_{it} = \alpha_0 + \alpha_1 T_i + \alpha_2 Treat_i + \alpha_3 Treat_i \times T_i + \boldsymbol{\alpha} \boldsymbol{X}_{it} + \theta_j + \varepsilon_{it}, \quad (1)$$

$$lnY_{it} = \alpha_0 + \alpha_1 T_i + \alpha_2 Treat_i + \alpha_3 Treat_i \times T_i + \boldsymbol{\alpha} \boldsymbol{X}_{it} + \theta_j + \varepsilon_{it}, \quad (2)$$

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

where i, t, j denote the individual, year and province, respectively; Emp_{it} denotes the employment dummy, which equals 1 if a woman i in year t is employed, while it equals 0 if she doesn't have a job. ln Y_{it} denotes the logarithm of total labour incomes per month in renminbi (CNY) from her primary job. Price adjustments are made to nominal wages in 2014 and 2018, using the 2010 price index as a benchmark. Also, the original data is two-sided 1% winsorised to eliminate outliers. $Treat_i$ is the treatment group dummy, which equals 1 if the individual comes from the treatment group, with a lower PFR before the universal two-child policy. T_i is the policy timing dummy, which equals 1 if the person is in the post-treatment group, that is, the 2018 CFPS database used in this article. These two variables, $Treat_i$ and T_i, pick up the two main effects of being in a treatment group and after the two-child policy. Essentially, the coefficient of the interaction term is our DID estimator, which measures the employment or income gap attributed to the affirmative universal two-child policy. Because the omitted variables affect the effectiveness of the DID method (Meyer, 1995), we add a vector of variables $\boldsymbol{X}_{it}$ in the equation to control for some demographic characteristics that also contribute to the employment and income, such as age, age square, education level, marital status, health, and enterprise ownership; θ_j is the dummy variable of province j and ε_{it} is the error term. The robust standard errors are clustered at the individual level.

Heckman two-step estimation

Applying linear models such as ordinary least squares (OLS) leads to a bias in the income estimation because the observations' distribution is conducted selectively or not randomly chosen (Heckman, 1976). Specifically, we can only follow women's income when they are hired, so the variable ln Y_{it} does not satisfy the DID method's randomness requirement. A Heckman two-step estimation is strongly recommended to rectify the non-random selection bias and mitigate the effects of heteroscedasticity.

In the selection stage, a Probit model would be conducted. Consider

$$P\left(Emp_{it}=1\right)=\alpha_0+\alpha_1 T_i+\alpha_2 Treat_i+\alpha_3 Treat_i\times T_i+\boldsymbol{\alpha}\boldsymbol{X}_{it}+\theta_j+\varepsilon_{it} \tag{3}$$

where $P(Emp_{it} = 1)$ indicates the probability for a woman to be employed, and other variables are the same as in equation (1).

At the outcome stage, an OLS regression applies to estimate the relationship and the extent of the independent variables constrained by the Probit model results in the first stage. We calculate the Inverse Mills Ratio (IMR) and put it into equation (2) as a correction term. This gives us

$$lnY_{it}=\beta_0+\beta_1 T_i+\beta_2 Treat_i+\beta_3 Treat_i\times T_i+\boldsymbol{\beta}\boldsymbol{Z}_{it}+\lambda IMR_i+\theta_j+\varepsilon_{it} \tag{4}$$

where a vector of variables $\boldsymbol{Z}_{it}$ controls the demographic characteristics that affect women's incomes. Considering that work experience and different industries largely affect earnings, new variables, like work experience, work experience square, and industry, are introduced. Furthermore, to meet the exogeneity requirement of the Heckman

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

two-step estimation, we take the variable 'marital status' as the exclusive variable, which is only in the Probit model of the first stage and is omitted from $\boldsymbol{Z}_{it}$ in the outcome stage. The reasons are as follows.

First, it is an exogenous variable, and it doesn't impair the IMR. Second, it is a conventional independent variable for explaining female labour participation (e.g., Eckstein and Lifshitz, 2011; Juhn and McCue, 2017; Killingsworth and Heckman, 1986). Third, 'marital status' does not directly affect people's income, and has not been considered in many income studies such as the famous Mincer wage equation (Mincer, 1974). Therefore, it is reasonable to omit 'marital status' in the second stage and use the IMR to reflect its possible indirect influence.

Data description

The China Family Panel Studies (CFPS) program was launched in 2000 and is run every 2 years. The surveys exist at three levels: community, family, and individual levels, covering 25 provinces and more than 90% of the population in China.

Prior to the quantitative study, we restrict our participants to women between 16 and 54 years old who are neither full-time students nor incapacitated for work. We set 16 as the lower bound because it is the statutory age for work in China and employed women under 16 are rare. We use 54 as the upper bound according to the current retirement age for women.[3] In addition, since agricultural production is quite different from formal employment in urban areas, those who have rural *hukou* (household registration) are eliminated from the sample. We obtain 3010 and 2095 individuals in 2014 and 2018. Specifically, we consider the following variables:

Employment status (Emp)—Dummy variable. If female *i* has a job, then emp = 1; otherwise, emp = 0.

Log female's labour income (lnY)—The logarithm of female *i*'s monthly wage in renminbi (CNY) from her main job during the past year, including salary, bonuses, cash benefits, and subsidies, but insurances excluded.

Age—Female *i*'s age at the time of the survey.

Age squared—The square of female *i*'s age at the time of the survey.

Marital status—Dummy variable. If female *i* is married, mar = 1; 0 otherwise.

Education level—Categorical variable divided into four groups: junior high school or below; senior high school or a vocational and technical high school; college; and bachelor's degree and above.

Enterprise ownership—Categorical variable divided as follows: government/government agencies/people's organisations; public institutions; state-owned enterprises; private companies/individual businesses; foreign/Hongkong/Macao/Taiwan enterprises; otherwise.

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Health—Categorical variable divided into five groups: very healthy, healthy, relatively healthy, fair, unhealthy.

Work experience—The number of years for work. Data on the actual years of work experience for a large sample of employers are generally unavailable. We have followed Mincer (1970) and defined $A_i - E_i - 6$ as a proxy for actual work experience; where A_i is the age of the individual and E_i is the number of years of schooling completed.

Work experience square—The square of the years of work experience.

Industry—Categorical variable divided as follows: farming, forestry, animal husbandry, and fishery; mining; manufacturing industry; electricity, gas and water production and supply; construction; transportation, storage, and postal industry; information dissemination, computer service, and software industry; wholesale and retail; accommodation and catering industry; finance; real estate; leasing and business services; technical services and geological survey; water, environment, and public facilities management; residence, and other services; education; health, social security, and social welfare industries; sports and entertainments; and public administration and social organisations.

Due to space limitations, the summary statistics of control variables are omitted from this paper and available on request.

Effect of the universal two-child policy on female labour market

We explore whether China's universal two-child policy has affected female labour market outcomes and the magnitude of such an effect; then, we employ the Heckman two-step estimation to rectify the sample selection bias, if any. Also, we test the validity of the DID method's parallel hypothesis and the policy's dynamic effect.

Baseline DID method

In column 1 of Table 2, the coefficient of interaction $\mathrm{Treat_i} \times \mathrm{T_i}$ is negative and significant at the 10% statistical level, demonstrating that the universal two-child policy has significantly decreased female labour participation. The switch from the one-child policy to the universal two-child policy reduces female employment by as much as 4.06%.

Childcare reduces the female labour supply. Under the universal two-child policy, many households decide to have a second child, which can compete with working hours for young women. Furthermore, the statutory admission for kindergarten is children over 3 years old, and few government-funded childcare or affordable childcare services are available for children under 3. Mothers have no choice but to exit the labour market if they have no way to share the burden of childcare with their parents or the parents of their spouses.

The control variables can also be attributed to our model's validity and accuracy, which have substantial economic implications. First, the positive coefficient of age and the negative coefficient of age square indicate that: as women get older, the probability

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Table 2. Effect of the universal two-child policy on female employment and income.

	(1) Employment (OLS)	(2) Income (Heckman)
T_i	0.0072	0.4347***
	(0.0176)	(0.0475)
$Treat_i$	0.0475***	0.2577***
	(0.0159)	(0.0437)
$Treat_i \times T_i$	–0.0406*	–0.1101*
	(0.0226)	(0.0622)
Age	0.0510***	0.0302
	(0.0072)	(0.0295)
Age square	–0.0006***	–0.0004
	(0.0001)	(0.0004)
Marital status	–0.0446***	–
	(0.0168)	
Education level	0.1103***	0.7459***
	(0.0189)	(0.0626)
Enterprise ownership	0.0535*	0.2686***
	(0.0317)	(0.0951)
Health	–0.0371	–0.3327***
	(0.0322)	(0.0800)
Work experience	–	0.0556***
		(0.0106)
Work experience square	–	–0.0014***
		(0.0003)
Industry	–	–0.0089
		(0.1489)
Lambda	–	0.6830***
		(0.2070)
Constant	–0.1470	7.3469***
	(0.1258)	(0.4696)
Observations	3485	2881
Adj-R^2 / Wald chi^2	0.065	1043.92

Note: For the variable education level, the reference group is 'junior high school or below'; for enterprise ownership, the reference is 'other' ownership; for industry, the reference group is 'Public administration and social organisation'; and for health status, the reference is the 'very healthy' group. Due to space limitations, we only report the regression results for employees who have a bachelor's degree or above, are unhealthy, work in foreign/Hongkong/Macao/Taiwan enterprises from the manufacturing industry. The complete results are available by request. Clustered standard errors are in parentheses. *, **, and *** represent significance levels of 10%, 5% and 1%, respectively.

that they stay in the labour market will rise but at a decreasing rate, and once age exceeds a certain threshold, the employment probability declines. Second, marriage has a significantly negative effect on female employment. Third, the coefficient of education level is significantly positive, denoting that more education is beneficial to female employment. A bachelor's degree or above boosts women's employment rate by 11.03%, compared to those who graduate from junior high school or below. Furthermore, among these six

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

different ownership types, foreign enterprises (including Hong Kong, Macao and Taiwan) improve women's employment by 5.35% at the statistical level of 10%, compared to 'others'.

Heckman two-step estimation

Since sample selection bias cannot be ignored in our study, the Heckman two-step estimation is exploited, and the result is presented in column 2 of Table 2.

After controlling for individual characteristic variables, the coefficient of IMR is significant at the 1% statistical level, confirming that selective bias does exist and the result from the simple OLS is unreliable. The coefficient of $Treat_i \times T_i$ means that the universal two-child policy has decreased women's monthly income by 10.43% ($e^{-0.1101} - 1$) at the 10% level. After eliminating bias, we develop the confidence to attribute this decline to the universal two-child policy.

Observing other control variables, the coefficients of female age and age squared indicate that an inverted U-shaped relationship exists between age and income, which is consistent with the classical Mincer wage equation (Mincer, 1974). Work experience and experience square denote an inverted U-shaped relationship as well. Thus, receiving more education could considerably improve women's earnings. Women working in government agencies, public institutions, state-owned enterprises, and foreign enterprises (including Hong Kong, Macao and Taiwan) earn significantly more money than women working for 'other employers'. Regarding health status, self-evaluation of unhealthy status significantly decreases female income by 28.29% ($e^{-0.3327} - 1$) at the 1% level compared to a very healthy status.

Parallel-trend assumption

Utilisation of the DID estimation must satisfy the parallel-trend assumption, which requires the control and treatment group to demonstrate parallel and consistent trends before the exogenous shock occurs.

We use the event study[4] method to test the policy's exogeneity. We take the year 2012 as a reference group and use $Year_i^{\tau}$ is a dummy variable to represent the year before and after the policy. '$Treat_i \times T_i$' is replaced by '$Treat_i \times Year_i^{\tau}$' in equations (1) and (4), and the expanded DID model is shown below:

$$Emp_{it} = \alpha_0 + T_i + Treat_i + \sum_{\tau} \alpha_{\tau} Treat_i \times Year_i^{\tau} + \boldsymbol{\alpha} \boldsymbol{X}_{it} + \theta_j + \varepsilon_{it} \tag{5}$$

$$lnY_{it} = \beta_0 + T_i + Treat_i + \sum_{\tau} \beta_{\tau} Treat_i \times Year_i^{\tau} + \boldsymbol{\beta} \boldsymbol{Z}_{it} + \theta_j + + \lambda IMR_i + \varepsilon_{it} \tag{6}$$

where $Year_i^{\tau}$ (τ =14, 2016, 2018) is the dummy variable and equals 1 if the year is τ and 0, otherwise.

Columns 1 and 2 of Table 3 show that the coefficient of $\mathrm{Treat_i} \times Year_i^{2014}$ is not significant, while $\mathrm{Treat_i} \times Year_i^{2016}$ and $\mathrm{Treat_i} \times Year_i^{2018}$ are significant. That is, the two sets of samples are comparable, and the new fertility policy is exogenous.

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

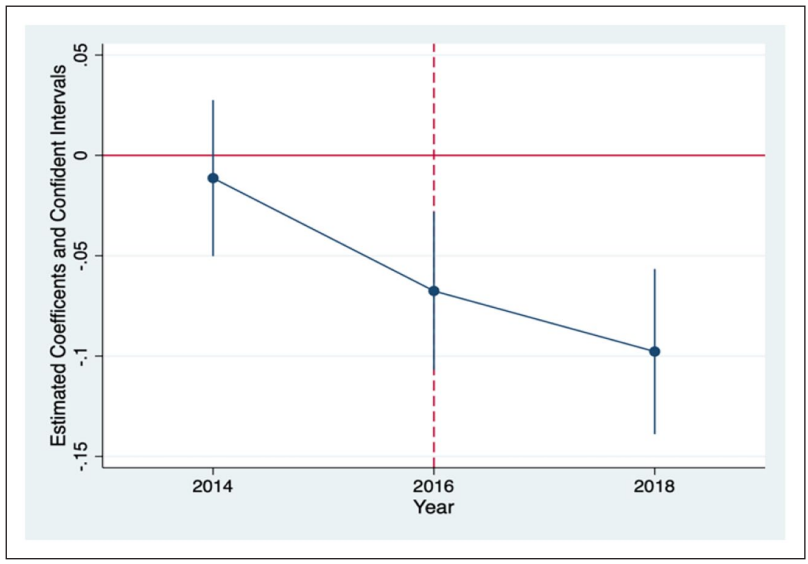

Figure 2. Parallel-trend test for female employment.

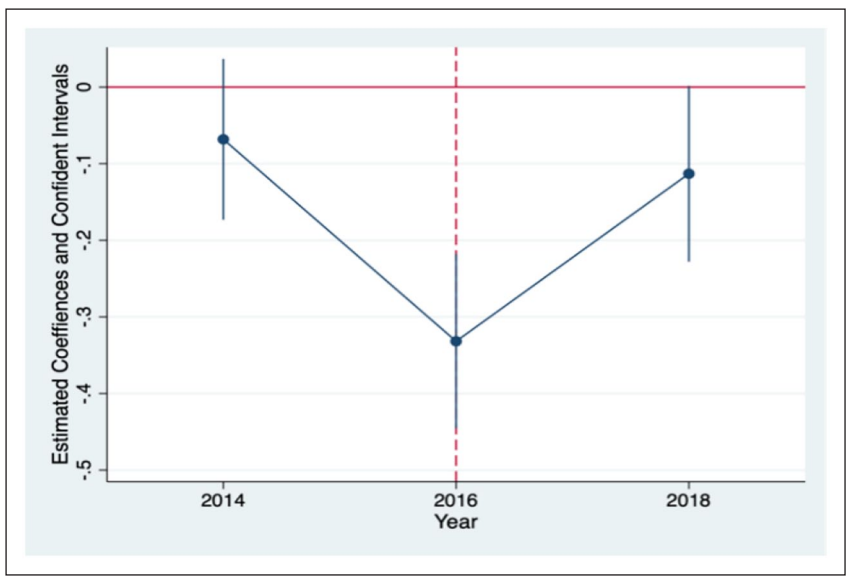

Figure 3. Parallel-trend test for female income.

Figures 2 and 3 show the results of the parallel-trend test more apparently. They reveal that, before the implementation of the universal two-child policy in 2016, the interaction coefficient is not significantly different from 0 (the 95% confidence interval contains a value of 0) but is significantly negative after the policy. This finding indicates that both

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Table 3. Paralled-trend test and dynamic effect.

	Parallel-trend test		Dynamic effect	
	(1) Employment	(2) Income	(3) Employment	(4) Income
$Treat_i \times Year_i^{2014}$	–0.0113	–0.0657	–	–
	(0.0198)	(0.0537)		
$Treat_i \times Year_i^{2016}$	–0.0675***	–0.3286***	–0.0393**	–0.0837
	(0.0201)	(0.0582)	(0.0194)	(0.0590)
$Treat_i \times Year_i^{2018}$	–0.0977***	–0.1110*	–0.0715***	–0.1122*
	(0.0210)	(0.0586)	(0.0206)	(0.0628)
Lambda	–	0.4321***	–	0.4110**
		(0.1621)		(0.1933)
Control variables	Yes	Yes	Yes	Yes
Observations	7564	5750	5627	4269

Note: Due to space limitations, only interaction terms' coefficients are reported, and the control variables are available by request. Clustered standard errors are in parentheses. *, **, and *** represent significance levels of 10%, 5% and 1%.

groups are indeed parallel and consistent before the policy in the female employment or income estimation.

Dynamics of the policy effect

We add the data from 2016 to examine the dynamics of this policy effect. Specifically, we replace t_i in equations (1) and (4) with dummy variables $Year_i^{2016}$ *and* $Year_i^{2018}$ Columns 3 and 4 of Table 3 report the estimations. The universal two-child policy decreases female employment by 3.93% and 7.15% in 2016 and 2018, respectively. We find that the coefficient increases during the two years, indicating that the policy's effect on labour participation remains strong over time.

The policy effect on income is not significant in 2016, whereas it significantly declines by 10.61% ($e^{-0.1122}-1$) in 2018, as shown in column 4. Hence, there is a time lag between policy implementation and income change. For instance, it might take quite a while to decrease or increase employees' salaries due to the rigidity of the salary payments system in a company.

Robustness check

Our model's robustness could be influenced by a variety of factors, such as sample selection, model specification, and within-sample heterogeneity. In this section, a series of tests are conducted to ensure that our results are robust in different settings.

Resample test

A potential problem in our research is whether the designated year impairs our conclusion. Since the universal two-child policy came into force in 2016 and the CFPS database is updated to 2018, we cannot find another post-treatment cohort except, 2018.

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Table 4. Robustness test.

	Resample		Placebo test		PSM-DID		Add macro variable	
	(1) Employment	(2) Income	(3) Employment	(4) Income	(5) Employment	(6) Income	(7) Employment	(8) Income
T_i	0.0567*	0.6322***	–0.0149	0.4235***	0.0070	0.4276***	0.0128	0.4900***
	(0.0322)	(0.0886)	(0.0213)	(0.0525)	(0.0181)	(0.0475)	(0.0186)	(0.0502)
$Treat_i$	0.0922***	0.4552***	0.0372**	0.1476***	0.0514***	0.2559***	0.0548***	0.3241***
	(0.0170)	(0.0486)	(0.0174)	(0.0456)	(0.0164)	(0.0435)	(0.0173)	(0.0482)
$Treat_i \times T_i$	–0.0809***	–0.2892***	–0.0012	–0.0754	–0.0424*	–0.1158*	–0.0415*	–0.1202*
	(0.0240)	(0.0676)	(0.0248)	(0.0627)	(0.0238)	(0.0638)	(0.0227)	(0.0619)
Lambda	–	0.3978*	–	0.3952	–	0.4969**	–	0.4870**
		(0.2157)		(0.2418)		(0.2469)		(0.2319)
Control variable	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3485	2817	3485	2881	3106	2576	3484	2881

Note: Due to space limitations, only the interaction terms' coefficients are reported, and the control variables are omitted but available by request. Clustered standard errors are in parentheses. *, **, and ***represent significance levels of 10%, 5% and 1%.

Therefore, we reselect the individuals in 2012 as the pre-treatment group and conduct the regressions again. Columns 1 and 2 of Table 4 show that the interaction coefficients are still significantly negative, and the absolute values are much higher than the benchmark outcomes. That is probably because, before the universal two-child policy, most provinces in China began to implement the two-child fertility policy for couples in cases in which either the husband or the wife is from a single-child family in 2014. After we choose individuals in 2012, to a large extent, we have picked up the effects of these two fertility policies on female labour market outcomes, which are both negative.

Placebo test

The division of the treatment and control groups plays a crucial role in our study. A placebo test is conducted by constructing a false treatment group and a false control group. In the previous study, the control and treatment groups are divided based on the PFR of 1.465. That is, we include the province with a PFR below 1.465 in the treatment group, and others in the control group. To construct the false groups, we number these 25 provinces and sort them by PFR from smallest to largest. We pick out the odd-numbered provinces as the false treatment group and the even-numbered provinces as the false control group.[5] Therefore, before the universal two-child policy is set, the average PFRs of the two groups are not evidently different from each other.

The results in columns 3 and 4 of Table 4 show that, the coefficients of the false treatment group and the post-treatment period, that is, $\mathrm{Treat}_i \times \mathrm{T}_i$, are insignificant. This finding indicates that, if we change the treatment group, we cannot obtain similar results consistent with Table 2, thus proving that the conclusions of Table 2 are robust and reliable.

PSM-DID test

This subsection selects individual characteristic variables for propensity score matching (PSM) to control the bias induced by choosing the two groups. We employed Logit estimation and 1:4 nearest neighbour matching. The vector of variables $\boldsymbol{X}_{it}$ including age, age squared, marital status, education level, enterprise ownership, and health is used in the employment estimation. 'Industry', 'work experience', and 'work experience square' are added to the vector to assess female income. There is no significant difference after PSM, and the bias of control variables is eliminated, since the *p*-values of the *t*-tests are greater than 10%.[6]

The PSM-DID estimations are shown in Columns 5 and 6 of Table 4, which are consistent with our previous conclusion: the universal two-child policy negatively affects the employment and income of women who live in cities and towns.

Introduction of the macro variable

The control variables in the regressions are microdata at the individual level, but labour market outcomes are also substantially influenced by the macroeconomic environment. We added the GDP growth rate in each province to reduce macroeconomic interference. The results do not evidently change the size and significance of the coefficients in

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

columns 7 and 8 of Table 4, and we believe the benchmark regression model is reliable from this perspective.

How does the policy's effect differ for females?

We wondered whether the policy's influences on different groups were diverse, so further explorations were conducted among groups based on age, education level, and enterprise ownership.

Based on age

We divided the urban female workforce into four groups based on age: under 25 years old, 25 to 29 years old, 30 to 34 years old, and 35 to 39 years old.[7] Then we performed the estimations using the DID method and Heckman two-step estimation within each group.

Column 1 in Table 5 shows that the interaction coefficient is surprisingly low at −23.99%; that is, the universal two-child policy has brought a 23.99% decrease in women's labour participation under 25. The policy has no significant effect on female employment or income in other age groups. Thus, we have the confidence to assert that the universal two-child policy has a rather negative impact on younger females.

Based on education

Education level holds great importance in people's employment and income generation, and we divided education into four groups.

The results in columns 3 to 6 of Table 5 indicate that the universal two-child policy has decreased higher-educated women's income by 29.59% ($e^{-0.3509}-1$), which is significant at the 1% statistical level. Although receiving more education can improve female earnings as discussed in Table 2, the universal two-child policy's implementation leads to a more considerable income decline for women with higher academic qualifications. The 'motherhood penalty' among better educated or high-skilled women is more severe than among the lower-educated (Anderson et al., 2002; Yu and Xie, 2014).

Based on enterprise ownership

After carrying out the regression within each subgroup, it was found that only 'private companies/individual businesses' have a significant adverse relationship with female employment.[8] The interaction coefficient is −6.93% and is significant at the 10% level. After the new policy, women find it more challenging to be hired at a private enterprise, and this is probably because the private companies are more cost-sensitive.

Conclusion

China has been carrying out its family planning policy for more than half a century, and the one-child policy has been controversial ever since its inception in the late 1970s. The dramatic population reduction has brought an economic boost and social development

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

https://doi.org/10.1177/10353046221112716 Published online by Cambridge University Press

Table 5. Heterogeneity analysis.

	Age		Educational Level			
	(1) under 25	(2) under 25	(3) Junior high school	(4) Senior high school	(5) College	(6) Bachelor's degree and above
	Employment	Income	Income	Income	Income	Income
T_i	0.1691**	0.6072**	0.5448***	0.3003***	0.4153***	0.5574***
	(0.0678)	(0.2430)	(0.1446)	(0.1077)	(0.0910)	(0.0732)
$Treat_i$	0.1654***	–0.0147	0.2854***	0.1112	0.3181***	0.4200***
	(0.0579)	(0.2027)	(0.1066)	(0.1046)	(0.0918)	(0.0768)
$Treat_i \times T_i$	–0.2399***	–0.1345	–0.2713	0.1639	–0.0951	–0.3509***
	(0.0883)	(0.3075)	(0.1863)	(0.1473)	(0.1277)	(0.0974)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Lambda	–	–0.4906	1.2246**	–0.5060	0.6190**	–0.1932
		(0.4639)	(0.5666)	(0.4848)	(0.4498)	(0.4097)
Observations	328	260	809	679	660	733

Note: Due to space limitations, only interaction terms' coefficients are reported, and the control variables are omitted but available by request. Clustered standard errors are in parentheses. *, **, and *** represent significance levels of 10%, 5% and 1%.

but also led to a complicated and overwhelming demographic predicament. The universal two-child policy tries to address these challenges; however, it falls far short of expectations. One crucial reason is that children usually induce huge costs to households, especially women, so future research must provide the sorely needed understanding of the policy's influence and trace the root causes of poor policy implementation.

Firstly, based on the CFPS 2014 and 2018 survey, the universal two-child policy has significantly decreased women's labour participation and income. Specifically, from the labour supply perspective, women probably exit the labour workforce, because more domestic labour and more child-rearing responsibilities require vast time and energy spent for their families. From the labour demand perspective, the employer would have to undertake the child-induced costs if they hire female employees. Such costs include, but are not limited to, maternity insurance fees, allowances, pregnancy-test leave, maternity leave, breastfeeding leave, hiring a replacement during her occupational interruption, and the work decrement due to childbearing and childcare, which would probably be doubled after the universal two-child policy. So, employers have strengthened their preference for male workers, which causes more severe gender discrimination in the labour market.

Secondly, the policy has reduced employment opportunities for women under 25 by nearly a quarter; the income of those with a bachelor's degree or above has decreased by almost one-third, attributed to the universal two-child policy. In addition, different enterprises have different responses to this policy, and private companies are more sensitive to it.

Thirdly, based on the data we obtain, the universal two-child policy's influence on female employment is gradually strengthening from 2016 to 2018, and its impact on income has presented an evident time lag. However, the definition of employment is quite broad, which includes formal employment, temporary employment, and other forms, so we cannot claim that the impact of the universal two-child policy on women's formal employment is gradually increasing. Actually, encountering an economic downturn and/or severe employment discrimination, women tend to choose more flexible employment characterised by fixed-term contracts, involuntary part-time, and bogus self-employment (Grimshaw et al., 2017).

Finally, the DID method theoretically requires the control group not to be affected by the policy; that is, there is no policy 'spillover effect'. Since our control group is not utterly free from the policy, our conclusions underestimate this policy's effect. The policy's negative influence on female labour participation and income should be greater than our valuations.

Women shoulder the main proportion of domestic labour and childrearing responsibilities, so after the universal two-child policy, it is more difficult for them to balance family and career constrained by limited time and energy. This policy is not the root of women's dilemma, but it exacerbates the conflict. The Chinese government expects an increase in the number of newborns among its citizens upon relaxing the fertility policy; however, the policy's negative effect on the female labour market presents a disadvantage within the family's reproductive decision. Therefore, the government should take adequate measures to protect women's employment opportunities and wage equality, such as providing more affordable infant and childcare services, supplying employers

with tax subsidies or financial incentives for hiring females, encouraging fathers to participate in parenting duty and housework, and establishing a childbirth-friendly institutional and cultural environment.

From a global perspective, birth rates in many countries have fallen below the replacement rate of 2.1. According to the conclusions of this article, a more relaxed fertility policy or fertility incentives would decrease female employment and income. Thus, if governments place greater priority on guaranteeing women's equal rights in the labour market, they can greatly promote women's fertility willingness and reproductive behaviours.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research on which this article is based was funded by grant 72074128 from the National Natural Science Foundation of China.

ORCID iD

Xiaofei Jin https://orcid.org/0000-0003-4896-8670

Notes

1. Tibet has no PFR.
2. The CFPS database covers 25 provinces in China, so Inner Mongolia, Ningxia, Hainan, Qinghai, and Xinjiang are excluded from our sample.
3. Currently, the female cadres's statutory retirement age is 55, and female workers can retire at the age of 50.
4. The event study is widely used in the literature to test the parallel trend assumption (Gentzkow, 2006; Moser and Voena, 2012).
5. If we choose the even-numbered provinces as the false treatment group and the odd-numbered provinces as the control group, the conclusion still holds.
6. The results of mean-value tests are omitted from this paper and available by request.
7. Generally, the fertility policy has a minor influence on females above 40 years old, so only women under 40 are discussed in this subsection.
8. The regression results are omitted but available by request.

References

Adair L, Guilkey D, Bisgrove E, et al. (2002) Effect of childbearing on Filipino women's work hours and earnings. *Journal of Population Economics* 15(4): 625–645.

Agarwal S, Li K, Qin Y and Wu J (2019) The impact of fertility relaxation on female labor market outcomes. Available at: DOI:10.2139/ssrn.3507515 (accessed 10 June 2022).

Anderson DJ, Binder M and Krause K (2002) The motherhood wage penalty: Which mothers pay it and why? *American Economic Review* 92(2): 354–358.

Angrist JD and Evans W (1996) Children and their parents' labor supply: Evidence from exogenous variation in family size. *American Economic Review* 88(3): 450–477.

Angrist JD and Krueger AB (1999) Empirical strategies in labor economics. In: Ashenfelder O and Card D (eds.) *Handbook of Labor Economics*, vol. 3. Amsterdam: Elsevier, pp. 1277–1366.
Banister J (1987) *China's Changing Population*. Stanford, CA: Stanford University Press.
Bloom D, Canning G, Fink G and Finlay JE (2009) Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth* 14(2): 79–101.
Budig MJ and England P (2001) The wage penalty for motherhood. *American Sociological Review* 66(2): 204–225.
Chen W and Yang S (2014) China's fertility in 2010: An indirect estimation using Brass P/F ratio method. *Population Research* 38(6): 16–24. (in Chinese)
Cools S, Markussen S and Strøm M (2017) Children and careers: How family size affects parents' labor market outcomes in the long run. *Demography* 54(5): 1773–1793.
Doren C (2018) Is two too many? Parity and mothers' labor force exit. J*ournal of Marriage and Family* 81(2): 327–344.
Duflo E (2001) Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91(4): 795–813.
Eckstein Z and Lifshitz O (2011) Dynamic female labor supply. *Econometrica* 79(6): 1675–1726.
Frodermann C and Müller D (2019) Establishment closures in Germany: The motherhood penalty at job search durations. *European Sociological Review* 35(6): 845–859.
Fursman L (2002) *Ideologies of motherhood and experiences of work: Pregnant women in management and professional careers*. Berkeley Collection of Working and Occasional Papers, 18. Berkeley, CA: Center for Working Families, University of California
Gash V (2009) Sacrificing their careers for their families? An analysis of the penalty to motherhood in Europe. *Social Indicators Research* 93(3): 569–586.
Gentzkow M (2006). Television and voter turnout. *The Quarterly Journal of Economics* 121(3): 931–972.
Goldin C (2006) The quiet revolution that transformed women's employment, education, and family. *American Economic Review* 96(2): 1-21.
Greenhaus H and Beutell J (1985) Sources of conflict between work and family roles. *Academy of Management Review* 10(1): 76–88.
Grimshaw D, Fagan C, Hebson G, et al. (2017) *Making Work More Equal: A New Labour Market Segmentation Approach*. Manchester, UK: Manchester University Press.
Guo Z (2011) 2010 Population census data indicates serious miscount in past population estimation and projection. *Chinese Journal of Population Science* 25(6): 2–13.
Guo Z, Zhang E, Gu B, et al. (2003) Diversity of China's fertility policy by policy fertility. *Population Research* 27(5): 1–10. (in Chinese)
He X and Zhu R (2015) Fertility and female labor force participation: New causal evidence from urban China. *The Manchester School* 84(5):664–674.
Heckman J (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement* 5(4): 475–492.
International Labour Organization (2018) *Global Wage Report 2018/19: What Lies behind Gender Pay Gaps*. Geneva: International Labour Organization.
Juhn C and McCue K (2017) Specialization then and now: Marriage, children, and the gender earnings gap across cohorts. *Journal of Economic Perspectives* 31(1): 183–204.
Killingsworth MR and Heckman JJ (1986) Female labor supply: A survey. In Ashenfelter O and Layard R (eds) *Handbook of Labor Economics*, 1st edn, vol. 1. Amsterdam: Elsevier, pp. 103–204.
Kinoshita Y and Guo F (2015) *What Can Boost Female Labor Force Participation in Asia?* Washington DC: International Monetary Fund.

Kleven H, Camille L and Søgaard J (2019) Children and gender inequality: Evidence from Denmark. *American Economic Journal: Applied Economics* 11(4): 181–209.
Lokshin M (2004) Household childcare choices and women's work behaviour in Russia. *Journal of Human Resources* 29(4): 1094–1115.
Maurer-Fazio M, Hughes JW and Zhang D (2006) *Economic reform and changing patterns of labor force participation in urban and rural China*. William Davidson Institute Working Paper No. 787. Ann Arbor, MI: William Davidson Institute, University of Michigan.
McElroy M and Yang DT (2000) Carrots and sticks: Fertility effects of China's population policies. *American Economic Review* 90(2): 389–392.
Meyer B (1995) Natural and quasi-experiments in economics. *Journal of Business and Economic Statistics* 13(2): 151–161.
Mincer J (1970) The distribution of labor incomes: A survey with special reference to the human capital approach. *Journal of Economic Literature* 1970(8): 1–26.
Mincer J (1974) *Schooling, Experience, and Earnings*. New York, NY: Columbia University Press.
Moser P and Voena A (2012) Compulsory licensing: Evidence from the trading with the enemy act. *American Economic Review* 102(1): 396–427.
National Bureau of Statistics of China (2012) *Tabulation on the 2010 Population Census of the People's Republic of China*. Beijing: China Statistical Press.
Peng X (2011) China's demographic history and future challenges. *Science* 333(6042): 581–587.
Rudolf R and Kang SJ (2015) Lags and leads in life satisfaction in Korea: When gender matters. *Feminist Economics* 21(1):136–163.
Tang Y and Scott R (2017) 'Glass ceiling' or 'sticky floor': The evidence from Chinese labor market. *Advances in Economics and Business* 5(10): 531–538.
Waldfogel J (1997) The effect of children on women's wages. *American Sociological Review* 62(2): 209–217.
World Bank (2022) World Development Indicators database. Available at: https://databank.worldbank.org/reports.aspx?source=world-development-indicators (accessed 21 May 2022).
Yu J and Xie Y (2014) The effect of fertility on women's wages in China. *Population Research* 1: 18–29. (in Chinese)
Zhu Q (2012) Estimation of fertility level in China 2000-2010: A preliminary study based on the 6th population census in China. *Chinese Journal of Population Science* 4: 68–77. (in Chinese)

Author biographies

Qian Huang is a distinguished professor of economics at the Institute of Population and Development in Nankai University in China. His research interests and publications are in the areas of labour economics, population economics, and health economics.

Xiaofei Jin holds a PhD from the Institute of Population and Development in Nankai University in China. Her research focuses on labour market dynamics and income distribution.