Who Cares: Deciphering China’s Female Employment Paradox[1]

Haiyue Yu

Interdisciplinary Center for Social and Behavioral Studies, Dongbei University of Finance and Economics, 217 Jianshan Street, CN-116023 Shahekou District, Dalian, China

dufebruce@163.com

Jin Cao

Research Department, Norges Bank, Bankplassen 2, PB 1179 Sentrum, NO-0107 Oslo, Norway and CESifo, Germany

jin.cao@norges-bank.no

Shulong Kang[2]

School of Finance, Dongbei University of Finance and Economics, 217 Jianshan Street, CN-116023 Shahekou District, Dalian, China

shulongkang@dufe.edu.cn

Abstract

China is featured by extraordinarily high female post-childbirth labor market participation rate and labor intensity, given that the public subsidy to childcare is poor and policy support for childbearing female employees is largely absent. Establishing a panel dataset that tracks females’ childbirths and employment, we find that such paradox is well explained by the intra-family grandparental childcare. Correcting the selection bias that stems from females’ fertility choices using the PSM-DID model, we find that females without grandparental support suffer a substantial drop in post-childbirth employment, while the employment of females with grandparental support

[1] This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the author and do not necessarily reflect those of Norges Bank. First version: January 31, 2019. This version: September 12, 2019.

[2] Corresponding author.

even rises after childbirth. It takes females without grandparental support twice as long to recover their employment after childbirths.

Keywords: Grandparental Childcare, Fertility Choice, Female Employment

JEL classification: C24, J13, J22

1 Introduction

Childcare is one of the main reasons that cause interruptions in females' career paths and human capital accumulation. In most of advanced economies in the world, heavy public investment in childcare and supportive policies for female employees have been the keys to maintaining relatively high labor market participation rates for females and low childbearing-induced human capital losses. Compared with these countries, female employment in China seems rather paradoxical: On the one hand, in China the public investment in childcare is rather poor and the protection for childbearing female employees is rather limited; on the other hand, China's female labor market participation rate is not only far above the other emerging market economies, but also above Scandinavian countries that have been for long known for gender equality and high female employment. How females in China reconcile the conflicts between childbearing and jobs, and achieve unusually high employment rate and labor intensity, remain largely unanswered.

In the past decades, China's female labor market participation rate has been always among the world's highest. According to the International Labor Organization (ILOSTAT, 2018), the employment rate for females above age 15 is as high as 61.3% in China --- not only far higher than the US (56%) and EU countries (50%), but also higher than Scandinavian countries (58% for Denmark, 61% for Sweden, 60.2% for Norway) that have been for long known for gender equality and high female employment. Usually, higher female employment reflects better social protection and supportive policies for female employees in one country (Mandel and Semyonov 2006). For instance, longer paid maternity leave and wider kindergarten admission significantly improve the labor market participation rate for females of age 25-55 (Besamusca et al. 2015), and providing more flexible jobs improves female employment (Gomes 2012, Blau and Kahn 2013). However, both supportive policies for childbearing females and public provision of childcare are by no means satisfactory in China, if not poor. For instance, China's public expenditure on kindergartens only accounts for 0.16% of GDP, far below most of OECD countries such as 0.9% for New Zealand, 0.7% for Norway, 0.65% for UK, and 0.45% for Germany. Due to low public investment, the supply of affordable public kindergartens is limited in China, the minimum admission age for kindergartens (3 years old) is too high, and pre-school / day-care services are nearly non-existing. For childbearing females, the statutory maternity leave in China is only 98 days, which is one of the lowest among 43 OECD / emerging market economies. In addition, there is no compulsory paternity leave in China, either; as a result, the burden of childcare in China almost fully falls on the shoulders of young females.

Except for the unusually high labor market participation rate, females' labor intensity in China is also among the world's highest, and their jobs are highly inflexible, too. As of 2017, the average weekly working hours for employed females in China is 45.5, far higher than those in advanced economies such as the US (34), the Netherlands (27), Norway (31) and Denmark (32). From the China Family Panel Survey (CFPS, 2012-2016), we find that the average weekly working hours for pre-childbirth females is 47, while for post-childbirth females is 46.1, which is not much different from the former. This implies that it is hardly possible for childbearing females to switch to jobs of more flexible working hours; the choice for the childbearing females is not how much they shall work, but rather, whether they shall work or not.

Given the low public investment on childcare and poor social protection for childbearing female employees in China, the burden of childcare is largely shouldered by the grandparents. In a 2007 survey conducted by Shanghai Population and Family Planning Commission, 88.7% of the grandparents were involved in taking care of their own grandchildren, and 53.3% of the grandparents took major responsibilities in the childcare on daily basis. The China Health and Retirement Longitudinal Study (CHARLS) shows that 50% grandparents regularly take major responsibilities for taking care of their grandchildren — much higher than many other countries. In contrast, in the US and Europe grandparent-provided childcare is in general not very common (except Mediterranean countries such as Italy), largely due to weaker family bonds and widely available daycare services provided by the market as well as public institutions. In the US, only 16% of grandparents are regularly involved in childcare (Health and Retirement Study, a.k.a. HRS, 2008, see Lumsdaine and Verneer 2015), 15% in Germany / Austria, 30% in Italy / Spain and 2% in Denmark / Sweden (Survey of Health, Ageing and Retirement in Europe, a.k.a. SHARE, 2004).[3] In China, even though childcare service is available in the market, few households actually rely on it. For example, among 2,281 children below the age of three in the whole sample of China Family Panel Studies (CFPS) 2014, only three are fully taken care of by babysitters hired from the market during the daytime. Instead, assistance provided by grandparents is almost always indispensable for a large share of families.

It seems that grandparental childcare is the key to explaining the unusually high employment rate and labor intensity for childbearing females in China. Indeed, using data from European countries (SHARE), Dimova and Wolff (2011) find that grandparental childcare significantly improves young females' labor-force participation rate as well as the intensity in labor supply. In a cross-country study, Aassve et al. (2012) find that such impact is significant and positive in some of the European countries (France, Germany, Bulgaria and Hungary), while not significant in other countries (such as Georgia, the Netherlands and Russia). Using survey data from the US (NLSY79), Posadas and Vidal-Fernández (2013) find that grandparental childcare increases young females' labor-force participation rate by 9%, and the effect is particularly stronger for ethical minorities and single mothers. Arpino et al. (2014) document similar positive effect using Italian data with instrumental variables, and such effect is stronger for females with less education and younger children. García-Morán and Kuehn (2017) and Compton and Pollak (2014) find that labor-force participation rate is higher for young females living closer to their parents so that grandparental childcare is more likely to happen. Based on a natural experiment from Italian pension reform, Bratti et al. (2016) find that grandmothers' retirement increases young females' employment rate by 13% while such effect does not exist for grandfathers, suggesting that childcare is more likely to be provided by grandmothers.

[3] It's worth noted that by definition grandparental childcare in the surveys from China is more intensive: In China regular grandparental childcare is defined as grandparents being the *main* responsible persons on *daily* basis (such as CHARLS and China Family Panel Survey, a.k.a. CPFS), while in HRS for US the threshold for a household's using grandparental childcare on a regular basis is that grandchildren are looked after by grandparents for more than 672 hours in 12 months (Lumsdaine and Verneer 2015), and in SHARE for Europe the threshold is just "at least twice a week" (Arpino et al. 2014). Taken into account the difference in definitions, the contrast between China and US / Europe is even more striking.

However, most of the studies are based on the observations of post-childbirth females, therefore, they are more capable to quantify the impact of grandparental care on post-childbirth females. However, females' fertility choices are endogenous, or, the availability of grandparental care affects females' choice on childbirth. Therefore, the impact of childbirth on female employment under various modes of childcare is better identified only if such endogeneity issue is properly addressed.

Several recent studies attempt to address the endogeneity issue in different ways. Using legislation on abortion as instrumental variable, based on a cross-country panel dataset covering 97 countries, Bloom et al (2009) identify strong significant negative correlation between birth rate and female employment rate. Using children's sex as instrumental variable for the number of children, Angrist and Evans (1998) and Cruses and Galiani (2007) find that having more than 2 children significantly reduces females' labor supply. Using a sample of females treated in fertility clinics who are likely to have similar willingness for childbirth, Crista (2008) finds that the first childbirth causes 26.3% fall in female employment.

Although the endogeneity issue has been addressed in various ways, there is so far little research on how childbirth affects female employment with the availability of grandparental childcare. In this paper, we use a micro-level dataset that contains information on female employment, childbirth, and grandparental childcare, construct records that track females' childbirth and employment, identify the impact of childcare on female employment with / without grandparental childcare, and address the endogeneity of females' fertility choices through propensity score matching difference-in-difference (PSM-DID) model.

Our paper contributes to existing literature in four folds: First, we identify the impact of childbirth on female employment with or without grandparental support, and directly quantify the contribution of grandparental support on reducing childbearing-induced interruption to females' careers; second, by constructing panel data that keep track of females' childbirth and employment, and using PSM-DID method, we provide a novel estimates that address the endogeneity problem stemming from females' self-selection on fertility; third, we provide an explanation on China's female employment paradox, that how females in China maintain high employment rate and labor intensity under poor public support. We show that retired workers take the burden of childcare that much compensates the poor public provision for childcare; finally, our study reveals a hidden cost of new retirement policies that aim to raise retirement age. Retaining old workers in the labor force may shift the burden of childcare to young females, thus crowd out their employment.

Section 2 describes the data and key variables, then section 3 constructs the econometric models for our analysis. First, we use panel regression with fixed effect and random effect to control for the unobserved variables, then we construct PSM-DID model to address the endogeneity in childbirth decisions and quantify the impact of childbirth on employment of females with / without grandparental support. Section 4 provides further discussions such as the persistency of the impact. In the end section 5 provides policy implications and concludes.

2 Data

2.1 Data description

Our dataset is constructed from China Family Panel Studies (CFPS). This is a nationally representative, biannual longitudinal survey of Chinese communities, families, and individuals launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University, China. The project aims to better understanding the economic, as well as the non-economic, well-being of contemporary Chinese population, and it collects individual-, family-, and community-level longitudinal data across the country. The survey contains rich information covering topics such as economic activities, education outcomes, family dynamics and relationships, migration, and health. Using CFPS survey data from 2010, 2012, 2014 and 2016, we established the panel data to keep track of the same females over time, in order to better understand the differences in their employment before and after childbirth, as well as to address to self-selection problem in fertility and employment decisions.

In our sample, we require that the individuals that first appeared in the observations must be females of age 20-49 who never had childbirth before, based on the following reasons: First, the minimum marriage age for females, set by the Marriage Law of China, is 20; given that children born outside marriage are rare in China and females are receiving longer schooling, the likelihood for females below age 20 to give birth is rather low. Second, the statutory minimum retirement age for females in China is 50 (for those from manufacturing firms, 55 for the public employees), females older than 50 would face different employment choices and thus must be excluded from our sample. Finally, those who first appeared in the observations never had childbirth before, while some of them had childbirth during 2010-2016; this allows us to focus on the changes of employment after childbirth.

Since year 2010 was the start of the survey, all observations in the 2010 subsample are our eligible females. 2012 and 2014 subsamples include those who have entered the sample in the previous surveys, plus those first entered the sample for various reasons.[4] Since 2016 was the last year of the survey, all observed females in the 2016 subsample appeared at least once in the previous surveys.[5] Combining the subsamples from the four surveys and deleting the females who appeared only once during 2010-2016, we obtain a sample with 7,551 observations. Those observations correspond to 2,776 females, in which 1,322 are observed twice, 909 are observed three times, and 545 are observed four times. On average, each eligible female is observed 2.72 times in the sample. See more details in Table 1.

Table 1 Tracking of the observed females

Year	**Obs.**	**First-time entries**	**Tracked in 2012**	**Tracked in 2014**	**Tracked in 2016**
2010	1136	1136	912	833	793
2012	1940	1028	--	1537	1409
2014	2302	612	--	--	1835
2016	2173	0	--	--	--

It can be seen that 80% of eligible females in 2010 are still observed in 2012, 73% of them remain in 2014, and 70% in 2016. Among all observed females in 2012, 79% of them are still observable two years later, 73% of them are available four years later. 80% of observed females in 2014 remain in the 2016 survey. This implies that the surveys keep good track of the families, and there is relatively small loss of observed samples because of lost contacts.

2.2 Key variables

We take key variables from CFPS that describe the characteristics of individual females, their households, and their communities.

[4] Including those who became over 20 in the years of surveys, those who joined the family through marriage, etc.
[5] It may happen that some early observed females disappear in the late surveys, because of divorce, lost contact, etc.

Motherhood: Dummy variable, equals to 1 if the observed female has given birth and 0 otherwise;

Grandparental childcare (GPC): Dummy variable. For a childbearing female i, if at least one of her children below age 11 is mainly taken care of by grandparent(s) during the day time[6] on daily basis, she is defined to receive grandparental childcare and her $GPC_i = 1$; otherwise $GPC_i = 0$;

Employment status ($WORK$): Dummy variable. If a young female i is in the labor force (including those on maternity leave) at the time of survey, $WORK_i = 1$; otherwise $WORK_i = 0$;

Weekly working hours: Female's average weekly working hours in the year of survey, including overtime;

College degree (EDU_i): Dummy variable, equals to 1 if female i's highest degree is college degree or above, 0 otherwise;

Female being urban resident ($URBAN_i$): Dummy variable, equals to 1 if the location of female i's residence is classified as "urban" by the National Bureau of Statistics of China;

Studentship: Dummy variable, equals to 1 if the female is going to school (full-time or part-time), 0 otherwise;

Marriage: Dummy variable, equals to 1 if the female is in marriage or cohabitation, 0 otherwise;

Log household's per capita net income, excluding the female's: The logarithm of female i's household's annual per capita net income in CNY (including the male's income and transfers, but excluding the female's income);

Household's total assets: In CNY;

Home ownership: Dummy variable, equals 1 if the household at least partially owns the property, 0 otherwise;

Home size: In square meters;

[6] CFPS asks about childcare providers for both day time and night time. As mostly people work during the day time when there is the most conflict between childcare and jobs, we therefore define GPC as grandparents' providing childcare during the day time.

Hardship in housing: Dummy variable, equals 1 if the household has insufficient home space, 0 otherwise;

Average property price for the community: In CNY per square meter. For urban communities, the price is defined as the average transaction price for the previous month of the survey; for rural communities, the price is defined as the average building cost;

Number of kindergartens in the community: The number of kindergartens within the community;

Number of primary schools in the community: The number of primary schools within the community;

Community's birth rate: The number of newborns per 1,000 inhabitants in the last year.

Table 2 presents the summary statistics of the variables.

Table 2 Summary statistics

Variables	Obs.	Mean	Std dev	Min	Max
Motherhood	7,551	0.238	0.426	0	1
Grandparental childcare	7,551	0.082	0.274	0	1
Employment	7,551	0.614	0.487	0	1
Weekly working hours	7,351	27.1	27.0	0	112
Age	7,551	25.9	5.18	20	56
College degree	7,544	0.305	0.461	0	1
Urban residency	7,153	0.334	0.472	0	1
Studentship	7,551	0.130	0.336	0	1
Marriage	7,549	0.485	0.500	0	1
Household's net per capita income, excluding the female's (CNY)	7,068	11,986	22,205	-100,350	814,600
Household's net asset (CNY)	7,129	459,357	926,195.3	-607,000	1.61e07
Home ownership	7,551	0.832	0.374	0	1
Home size (sqm)	7,087	137.1	106.3	5	2,000

Hardship in housing[7]	5,378	0.156	0.363	0	1
Average property price for the community	7,513	3,619	6,363.7	500	55,000
Number of kindergartens in the community	7,141	1.002	1.506	0	30
Number of primary schools in the community	7,141	0.726	0.713	0	5
Community's birth rate	7074	9.847	10.99	0	200

2.3 Stylized facts

Among 2,776 females in our sample, 1,066 or 38.4% gave births during the sample period. 480 or 45.03% of them received childcare assistance from grandparents after childbirth. We divide all females in our sample into two subsamples: 1,710 females who never had childbirths and 1,066 who had childbirths; for the latter, we can further divide them into two subgroups: Those before giving birth and those after giving birth. In Panel A of Figure 1, we present the employment rates for those who never had childbirths ("No childbirth"), those before giving birth ("Pre-childbirth"), and those after giving birth ("Post-childbirth"). It can be seen that the employment rates for "No childbirth" and "Pre-childbirth" groups are very close, implying that the employment choice for females before childbirths is not much different from those who never had childbirths. Furthermore, the employment rate for the "Post-childbirth" group is slightly lower, implying that there is indeed a childbearing-induced penalty in female employment.

However, if we further divide the "Post-childbirth" group into those who receive grandparental care (480 females, "Post-childbirth with GPC") and those who receive no grandparental care (586 females, "Post-childbirth without GPC"), we find that the drop in the employment of the post-childbirth females is largely caused by those who do not receive grandparental childcare, for example, compared with females without childbirth, the employment rate for post-childbirth females without grandparental childcare almost falls by 50% in 2012. On the contrary, females with grandparental childcare are even more likely to work after childbirth, probably because those females need to work to compensate the losses of grandparents' income due to childcare.

[7] Due to changes in the questionnaire design, this variable is missing for the 2016 survey. However, our econometric models do not need this variable for year 2016.

To better reflect childbearing females' employment decision in the long run, Panel B focuses on those who had childbirth during the sample period. Define the year of a female's first childbirth as "year 0". Before childbirth, the employment rate grows with age, while after childbirth, the patterns of employment are largely driven by the providers of childcare. For those receiving grandparental childcare ("Post-childbirth with GPC"), their employment follows the pre-childbirth trend and remains high, while for those receiving no grandparental childcare ("Post-childbirth without GPC"), their employment rate falls almost 50% in the first two years --- although their employment starts to recover from the third year, the employment gap between these two groups persists.

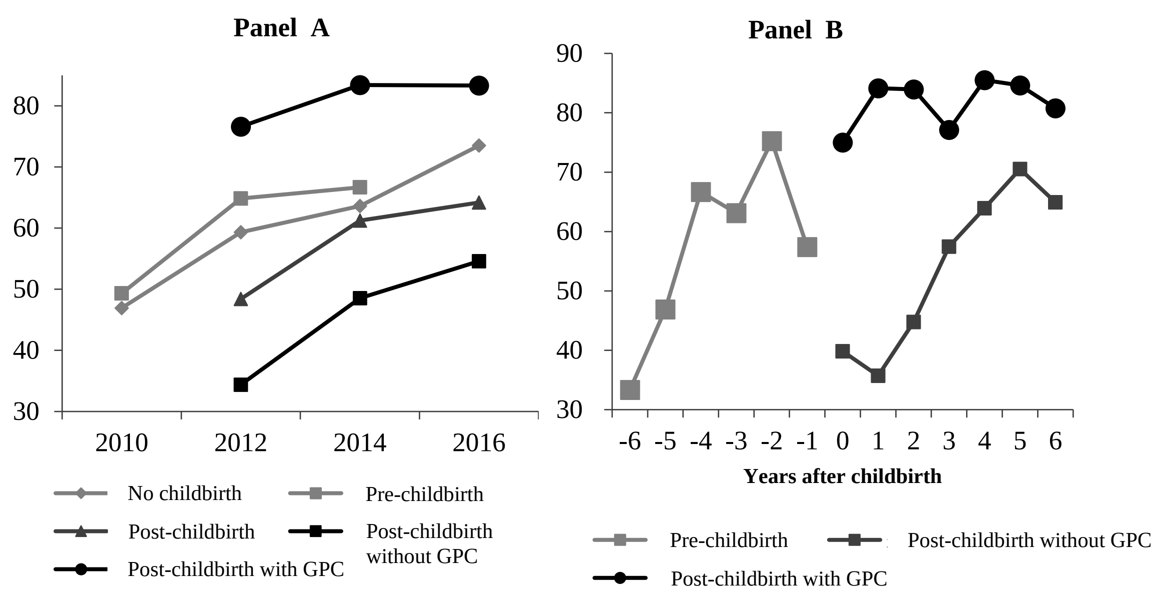


Figure 1 Employment before and after childbirth

On the other hand, grandparental childcare does not seem to much affect females' labor intensity. As is shown in Figure 2, for females who receive post-childbirth grandparental care and those who do not, their pre-childbirth labor intensities are very much similar, and the weekly working hours for both groups are around 45-50 hours. After childbirth, the weekly hours for those without grandparental care is comparatively slightly lower but is still as high as 40-45 hours. This implies that the supply of part-time jobs or jobs with flexible working hours is very limited in China. The dilemma for females without grandparental support is that they either have to stay with intensive jobs or to exit the labor market completely, which explains the sharp drop in their employment during the first two years after childbirths.

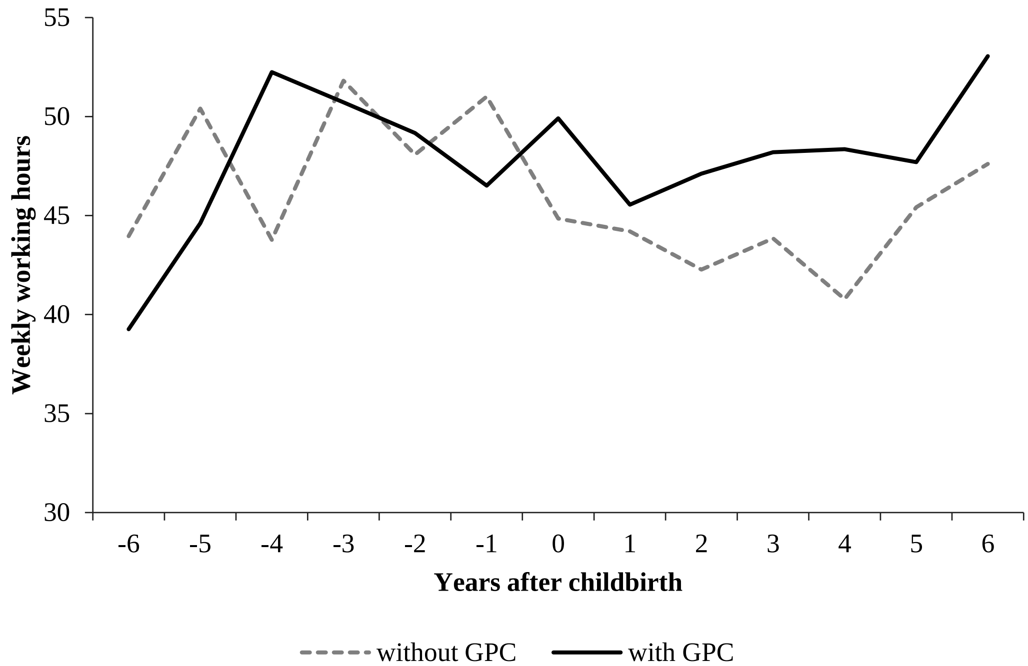


Figure 2 Labor intensity before and after childbirth

3 Empirical analysis and results

3.1 Panel regression with fixed effects

First, we use two-way fixed effect panel regression to analyze the impact of fertility on female employment. The baseline model is defined as

$$Work_{pit} = \alpha + \beta_1 Birth_{pit} + X_{pit}\gamma + \delta_t + \lambda_i + \eta_p + u_{pit} \quad (1)$$

in which subscript p denotes province, i denotes female i, and t denotes year. $Work_{pit}$ denotes whether the female is employed, $Birth_{pit}$ denotes "motherhood" or whether female i has given birth by the year of survey t, the vector X_{pit} contains a group of control variables (including female's age, college degree, urban residency, studentship, marriage, and per capital household's net income --- excluding the female's), δ_t captures year fixed effect, λ_i captures individual fixed effect, and η_p captures province fixed effect.

In order to see the heterogeneities in the impacts of childcare modes on female employment, we add interaction term $Birth_{pit} \cdot GPC_{pit}$ to the baseline model, such as

$$Work_{pit} = \alpha + \beta_1 Birth_{pit} + \beta_2 Birth_{pit} \cdot GPC_{pit} + X_{pit}\gamma + \delta_t + \lambda_i + \eta_p + u_{pit} \tag{2}$$

Together with the variable $Birth_{pit}$, the interaction term divides the females into three subgroups: Females that never gave birth ($Birth_{pit} = 0$, $Birth_{pit} \cdot GPC_{pit} = 0$), females with post-childbirth grandparental care ($Birth_{pit} = 1$, $Birth_{pit} \cdot GPC_{pit} = 1$), and females without post-childbirth grandparental care ($Birth_{pit} = 1$, $Birth_{pit} \cdot GPC_{pit} = 0$).

We may also explain the dummy variable $Work_{pit}$ as a probabilistic outcome, such as

$$Pr\left(Work_{pit} | Birth_{pit}, X_{pit}\right) = f\left(\alpha + \beta_1 Birth_{pit} + X_{pit}\gamma + \delta_t + \lambda_i + \eta_p\right) \tag{3}$$

in which $f(\cdot)$ can be accumulative distribution function $\Phi(\cdot)$ with standard normal distribution (Probit model), or $\Lambda(\cdot)$ with Logistic distribution (Logistic model). And we may also add interaction terms to (3), such as

$$\begin{aligned} &Pr\left(Work_{pit} | Birth_{pit}, GPC_{pit}, X_{pit}\right) \\ &\qquad = f\left(\alpha + \beta_1 Birth_{pit} + \beta_2 Birth_{pit} \cdot GPC_{pit} + X_{pit}\gamma + \delta_t + \lambda_i + \eta_p\right) \end{aligned} \tag{4}$$

Table 3 presents the results from the baseline models (1) and (3), with various specifications. Columns (1) and (2) are estimates from OLS, columns (3) to (7) are estimates from Probit or Logit models. The standard errors of estimated coefficients in columns (1) to (4) are heteroskedasticity robust standard errors, and the standard errors reported in column (7) are Bootstrap standard errors.

Table 3 Results for the baseline panel regressions

	Female's employment						
	(1) FE OLS	(2) FE OLS	(3) RE Probit	(4) RE Logit	(5) FE Logit	(6) FE Logit	(7) FE Logit
Motherhood (β_1)	-0.090*** (0.022)	-0.088*** (0.022)	-0.236*** (0.064)	-0.407*** (0.110)	-0.633*** (0.136)	-0.611*** (0.137)	-0.633*** (0.126)
Age	0.020 (0.027)	0.029 (0.027)	0.015** (0.006)	0.025** (0.011)	0.053 (0.187)	0.700 (0.193)	0.053 (0.168)
College degree	0.082*** (0.028)	0.082*** (0.028)	0.576*** (0.061)	1.003*** (0.107)	0.956*** (0.365)	1.092*** (0.391)	0.956** (0.435)
Urban residency	-0.000 (0.034)	-0.004 (0.034)	0.107* (0.061)	0.182* (0.107)	0.073 (0.250)	0.056 (0.257)	0.073 (0.277)

Studentship	-0.517*** (0.028)	-0.509*** (0.029)	-2.189*** (0.096)	-3.806*** (0.175)	-2.753*** (0.311)	-2.689*** (0.315)	-2.753*** (0.302)
Marriage	-0.199*** (0.023)	-0.199*** (0.023)	-0.537*** (0.062)	-0.941*** (0.108)	-1.334*** (0.145)	-1.339*** (0.148)	-1.334*** (0.152)
Log household's per capital net income	-0.009*** (0.003)	-0.008*** (0.003)	-0.029*** (0.009)	-0.049*** (0.016)	-0.060*** (0.021)	-0.061*** (0.022)	-0.060** (0.024)
Province FE	No	Yes	No	No	No	Yes	No
Individual FE	Yes	Yes	No	No	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.200 (0.595)	0.367 (0.616)	0.074 (0.173)	0.133 0.301			
Significance test	156.38 (P=0.000)		877.51 (P=0.000)	757.84 (P=0.000)	745.27 (P=0.000)	780.32 (P=0.000)	321.24 (P=0.000)
Log likelihood			-3624.49	-3623.11	-850.69	-833.16	-850.69
Within R^2	0.221	0.229					
Hausman test				376.08 (P=0.000)			
Obs.	6,796	6796	6,796	6,796	3,344	3344	3,344

Notes: (1) *** / ** / * denotes the result is significant on 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) heteroskedasticity robust standard errors are reported for FE OLS, RE Probit and RE Logit models, and bootstrap standard errors are reported for the FE Logit model in column (7); (4) In the models' significance tests, FE OLS is based on F-statistic, while RE Probit, RE Logit and FE Logit in column (7) are based on Wald χ^2-statistic; (5) Hausman test is based on χ^2-statistic.

It can be seen from the results that childbirth significantly lowers female employment rate, implying that childbearing-induced penalty in employment does exist in China. Based on the Logit model with fixed effect from column (6), the odds ratio of female employment is 0.54, i.e., *ceteris paribus*, the ratio of likelihood to be employed for post-childbirth females to that for pre-childbirth females is 0.54. Results from the OLS models suggest that childbirth reduces female employment by 9%, much lower than the impacts in most advanced economies.

Table 4 reports the results from models (2) and (4) containing interaction terms with various specifications. Columns (1) and (2) are estimates from OLS, columns (3) to (7) are estimates from Probit or Logit models. The results consistently show that the estimated coefficient of variable "motherhood", β_1, is significantly negative, while the estimated coefficient of the interaction term "motherhood*GPC", β_2, is significantly positive, and $\beta_2 > -\beta_1$. Wald test further rejects the hypothesis that $\beta_1 + \beta_2 = 0$. This implies that for females without grandparental care, their post-childbirth employment rate is significantly reduced, while for females with grandparental care, their post-childbirth employment rate is improved rather than reduced. That is, grandparental childcare reduces young females' burden, eases the conflicts between childcare and employment, thus reduces young females' opportunity cost to work and largely avoids the interruptions in young females' career paths. On the other hand, because of the low retirement age for female workers in China, many females choose to stay in the labor force after retirement.[8] For those females, forcing them to leave the labor force and take care of grandchildren reduces households' total income, as a result, young females receiving grandparental care may have higher incentives to work to compensate for grandparents' income losses. This may explain why females with grandparental support are even more likely to work after childbirth. In addition, the absolute values of estimated β_1 in Table 3 are significantly lower than those in Table 4, suggesting that the rise in post-childbirth employment rate of females with grandparental support largely compensates the fall in the employment of females without grandparental support, which explains why overall childbearing-induced interruption to females' careers in China is relatively low.

Table 4 Results from regressions with interaction terms

	Female's employment						
	(1) FE OLS	(2) FE OLS	(3) RE Probit	(4) RE Logit	(5) FE Logit	(6) FE Logit	(7) FE Logit
Motherhood (β_1)	-0.171*** (0.025)	-0.170*** (0.025)	-0.565*** (0.071)	-0.965*** (0.122)	-1.066*** (0.151)	-1.055*** (0.152)	-1.066*** (0.176)
Motherhood *GPC (β_2)	0.228*** (0.027)	0.231*** (0.027)	1.084*** (0.099)	1.875*** (0.175)	1.431*** (0.190)	1.471*** (0.194)	1.431*** (0.197)
Age	0.016 (0.026)	0.017 (0.026)	0.016** (0.006)	0.026** (0.011)	0.030 (0.190)	0.057 (0.197)	0.030 (0.164)
College degree	0.084*** (0.028)	0.083*** (0.028)	0.562*** (0.061)	0.982*** (0.107)	0.962*** (0.371)	1.102*** (0.399)	0.962** (0.425)
Urban residency	-0.004 (0.034)	-0.008 (0.033)	0.096 (0.061)	0.160 (0.106)	0.060 (0.256)	0.060 (0.263)	0.060 (0.267)

[8] According to China Population and Employment Statistics Yearbook, 2017, 24.7% of female employees are above 50, as of 2016. The China Health and Retirement Longitudinal Study (CHARLS, 2015) shows that the employment rate of urban females of age 50-60 is 45.8%, 74% for rural females.

Studentship	-0.515*** (0.029)	-0.507*** (0.029)	-2.183*** (0.096)	-3.796*** (0.174)	-2.793*** (0.313)	-2.732*** (0.318)	-2.793*** (0.389)
Marriage	-0.198*** (0.023)	-0.198*** (0.023)	-0.537*** (0.062)	-0.941*** (0.108)	-1.308*** (0.147)	-1.308*** (0.149)	-1.308*** (0.152)
Log household's per capital net income	-0.008*** (0.003)	-0.008*** (0.003)	-0.028*** (0.009)	-0.048*** (0.016)	-0.060*** (0.022)	-0.060*** (0.023)	-0.060*** (0.018)
Province FE	No	Yes	No	No	No	Yes	No
Individual FE	Yes	Yes	No	No	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.277 (0.585)	0.436 (0.606)	0.059 (0.173)	0.093 (0.301)			
Significance test	147.2 (P=0.000)		956.56 (P=0.000)	824.26 (P=0.000)	808.89 (P=0.000)	845.22 (P=0.000)	375.22 (P=0.000)
Log likelihood			-3551.57	-3549.91	-818.875	-800.71	-818.875
Within R^2	0.236	0.245					
Hausman test				114.57 (P=0.000)			
Wald test ($\beta_1 + \beta_2 = 0$)	4.35 (P=0.037)	4.89 (P=0.027)	29.25 (P=0.000)	28.24 (P=0.000)	3.49 (P=0.062)	4.42 (P=0.036)	2.19 (P=0.139)
Obs.	6,796	6,796	6,796	6,796	3,344	3344	3,344

Notes: (1) *** / ** / * denotes the result is significant on 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) heteroskedasticity robust standard errors are reported for FE OLS, RE Probit and RE Logit models, and bootstrap standard errors are reported for the FE Logit model in column (7); (4) In the models' significance tests, FE OLS is based on F-statistic, while RE Probit, RE Logit and FE Logit in column (7) are based on Wald χ^2-statistic; (5) Hausman test is based on χ^2-statistic; (6) For the statistic of the Wald test, OLS is based on F-statistic, while Probit and Logit models are based on χ^2-statistic.

3.2 Propensity score matching difference-in-difference model

In order to address the selection bias problem stemming from females' fertility choices, we further use propensity score matching difference-in-difference (PSM-DID) model to better identify the impact of childbearing on female employment. First, we estimate propensity scores through Logit regression for 2010-2012, 2012-2014, 2014-2016 subsamples, and the results are reported in Appendix A. The results suggest that females' fertility choices are largely affected by age, urban residency, studentship, marriage and housing conditions; the signs of estimated coefficients are the same as expected.

Then we match participants to treatment group and control group based on propensity scores. To ensure the robustness of results, we use 4 most common matching methods in the literature: k-nearest neighbor, caliper, k-nearest neighbor with caliper, and kernel matching. We follow Abadie et al (2004) and set $k = 4$, so that the estimates' mean square errors are minimized. We choose matching radius to be 0.05 for both caliper matching and k-nearest neighbor matching with caliper. For kernel matching, the kernel function is the most common quadratic kernel function, and the bandwidth is set to be 0.06.

Appendix B reports the test of balance for the covariates across treatment and control groups. Under all four methods, the post-matching standardized differences of covariates all fall below 5%; The pseudo R^2 of the Logit model falls from 20-30% to 1% after matching; LR test shows that covariates are joint significant before matching, but no longer so afterwards. Figure 8 compares the kernel densities for treatment and control groups. It can be seen that two groups' kernel density curves are significantly different before matching, while they almost coincide after matching.

Figure 8 Kernel densities of treatment and control groups, before and after matching

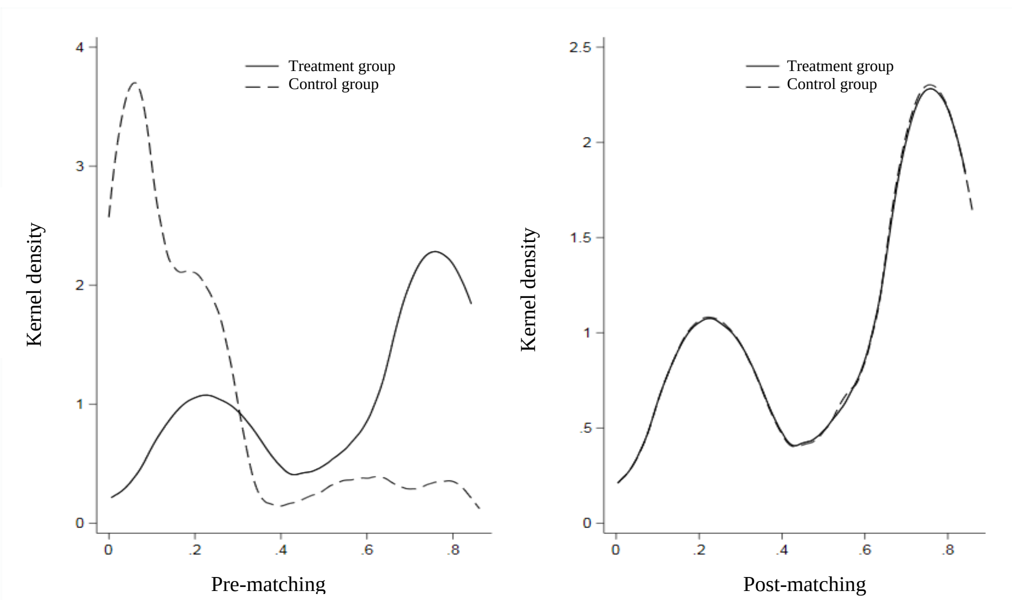


Table 5 reports the overall impact of fertility on female employment. It can be seen that childbirth significantly reduces females' likelihood to work, both before and after matching, and the results are robust for all three subsamples and four matching methods. Take k-nearest neighbor matching as an example, the average fall in employment caused by childbirth is about 17.9%. For every subsample, ATT falls after matching, implying that there is indeed self-selection in fertility decision, or, females who prefer to work reduce or postpone childbirths. The impact of childbirth on female employment would thus be overestimated, if the selection bias were not taken into account.

Table 5 Results from PSM-DID: Overall

Sample	Matching methods	Matching	Mean variation in employment			Std error	t-statistic
			w/ childbirth	w/o childbirth	ATT		
2010-2012 sub-sample	K-nearest neighbor	Before	0.028	0.203	-0.176***	0.044	-4.01
		After	0.028	0.202	-0.174**	0.074	-2.34
	Caliper	Before	0.028	0.203	-0.176***	0.044	-4.01
		After	0.028	0.165	-0.137**	0.068	-2.00
	K-nearest neighbor with caliper	Before	0.028	0.203	-0.176***	0.044	-4.01
		After	0.028	0.202	-0.174**	0.074	-2.34
	Kernel	Before	0.028	0.203	-0.176***	0.044	-4.01
		After	0.028	0.165	-0.138**	0.068	-2.02
2012-2014 sub-sample	K-nearest neighbor	Before	-0.053	0.128	-0.181***	0.040	-4.57
		After	-0.054	0.082	-0.136**	0.060	-2.28
	Caliper	Before	-0.053	0.128	-0.181***	0.040	-4.57
		After	-0.054	0.082	-0.136**	0.054	-2.50
	K-nearest neighbor with caliper	Before	-0.053	0.128	-0.181***	0.040	-4.57
		After	-0.054	0.083	-0.136**	0.060	-2.29
	Kernel	Before	-0.053	0.128	-0.181***	0.040	-4.57
		After	-0.054	0.085	-0.139**	0.055	-2.54
2014-2016 sub-sample	K-nearest neighbor	Before	-0.196	0.139	-0.335***	0.045	-7.46
		After	-0.195	-0.002	-0.194***	0.058	-3.33
	Caliper	Before	-0.196	0.139	-0.335***	0.045	-7.46
		After	-0.195	0.042	-0.237***	0.057	-4.16
	K-nearest neighbor with caliper	Before	-0.196	0.139	-0.335***	0.045	-7.46
		After	-0.195	-0.002	-0.194***	0.058	-3.33
	Kernel	Before	-0.196	0.139	-0.335***	0.045	-7.46
		After	-0.195	0.041	-0.237***	0.057	-4.14

In Table 6, we focus on the females who receive no grandparental support after childbirth. It can be seen that childbirth significantly reduces females’ likelihood to work, both before and after matching. Take k-nearest neighbor matching as an example, the average fall in employment caused by childbirth is about 33.3%, much higher than most advanced economies.

Table 6 Results from PSM-DID: Females without GPC

Sample	Matching methods	Matching	Mean variation in employment			Std error	t-statistic
			w/ childbirth	w/o childbirth	ATT		
2010-2012 sub-sample	K-nearest neighbor	Before	-0.035	0.203	-0.238***	0.050	-4.75
		After	-0.043	0.255	-0.298***	0.085	-3.49
	Caliper	Before	-0.035	0.203	-0.238***	0.050	-4.75
		After	-0.043	0.256	-0.299***	0.079	-3.81
	K-nearest neighbor with caliper	Before	-0.035	0.203	-0.238***	0.050	-4.75
		After	-0.043	0.254	-0.298***	0.085	-3.49
	Kernel	Before	-0.035	0.203	-0.238***	0.050	-4.75
		After	-0.043	0.259	-0.302***	0.079	-3.84
2012-2014 sub-sample	K-nearest neighbor	Before	-0.229	0.128	-0.357***	0.046	-7.69
		After	-0.233	0.074	-0.307***	0.063	-4.86
	Caliper	Before	-0.229	0.128	-0.357***	0.046	-7.69
		After	-0.233	0.075	-0.308***	0.058	-5.26
	K-nearest neighbor with caliper	Before	-0.229	0.128	-0.357***	0.046	-7.69
		After	-0.233	0.074	-0.307***	0.063	-4.86
	Kernel	Before	-0.229	0.128	-0.357***	0.046	-7.69
		After	-0.233	0.074	-0.307***	0.059	-5.24
2014-2016 sub-sample	K-nearest neighbor	Before	-0.342	0.139	-0.481***	0.052	-9.21
		After	-0.345	0.057	-0.402***	0.067	-5.97
	Caliper	Before	-0.342	0.139	-0.481***	0.052	-9.21
		After	-0.345	0.052	-0.398***	0.063	-6.31
	K-nearest neighbor with caliper	Before	-0.342	0.139	-0.481***	0.052	-9.21
		After	-0.345	0.057	-0.402***	0.067	-5.97
	Kernel	Before	-0.342	0.139	-0.481***	0.052	-9.21
		After	-0.345	0.048	-0.393***	0.063	-6.23

However, for females with grandparental support after childbirth, as Table 7 shows, childbirth significantly reduces their employment only in the 2012-2014 subsample before matching; such adverse effect does not exist for other pre-matching subsamples and all post-matching subsamples.

Table 7 Results from PSM-DID: Females with GPC

Sample	Matching methods	Matching	Mean variation in employment			Std error	t-statistic
			w/ childbirth	w/o childbirth	ATT		
2010-2012 sub-sample	K-nearest neighbor	Before	0.155	0.203	-0.048	0.067	-0.73
		After	0.141	0.064	0.077	0.091	0.84
	Caliper	Before	0.155	0.203	-0.048	0.067	-0.73
		After	0.141	0.124	0.017	0.086	0.20
	K-nearest neighbor with caliper	Before	0.155	0.203	-0.048	0.067	-0.73
		After	0.141	0.064	0.077	0.091	0.84
	Kernel	Before	0.155	0.203	-0.048	0.067	-0.73
		After	0.141	0.122	0.019	0.086	0.22

2012-2014 sub-sample	K-nearest neighbor	Before	0.252	0.128	0.124**	0.059	2.11
		After	0.252	0.153	0.100	0.083	1.20
	Caliper	Before	0.252	0.128	0.124**	0.059	2.11
		After	0.252	0.125	0.127	0.076	1.68
	K-nearest neighbor with caliper	Before	0.252	0.128	0.124**	0.059	2.11
		After	0.252	0.143	0.109	0.083	1.32
	Kernel	Before	0.252	0.128	0.124**	0.059	2.11
		After	0.252	0.119	0.133*	0.076	1.76
2014-2016 sub-sample	K-nearest neighbor	Before	0.111	0.139	-0.028	0.072	-0.38
		After	0.111	0.028	0.083	0.084	0.99
	Caliper	Before	0.111	0.139	-0.028	0.072	-0.38
		After	0.113	0.049	0.065	0.082	0.79
	K-nearest neighbor with caliper	Before	0.111	0.139	-0.028	0.072	-0.38
		After	0.113	0.028	0.085	0.085	1.00
	Kernel	Before	0.111	0.139	-0.028	0.072	-0.38
		After	0.111	0.044	0.067	0.082	0.82

4 Discussion

So far we have shown that employment rate for females without grandparental support plummets after childbirth; however, it would be also interesting to know whether such adverse impact is temporary or permanent, and to what extend their employment recovers after childbirth. Answering these questions would help us better understand the impact of childbirth on female employment both in the short run and in the long run. In addition, as both overall and post-childbirth female labor intensities are rather high in China, it is also interesting to see how much grandparental childcare contributes to females' high post-childbirth labor intensity.

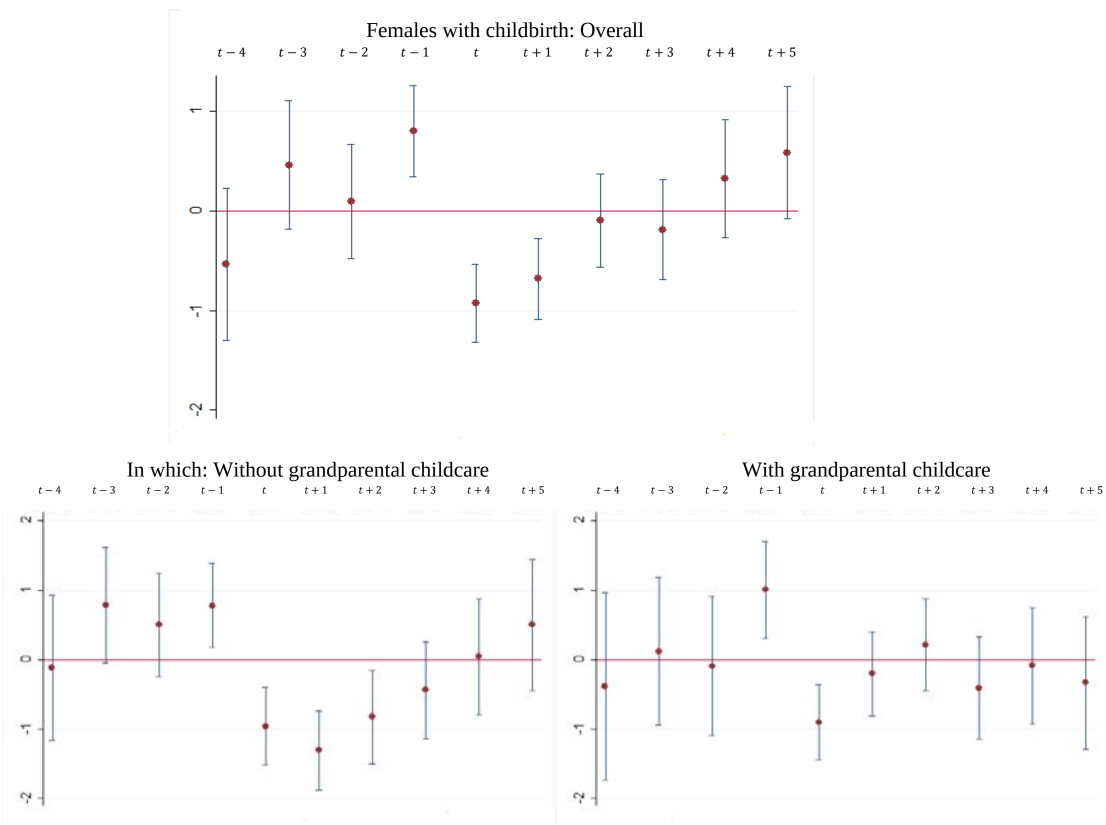
Exiting labor market: Temporarily or permanently?

Childbirth forces some of the females to drop out of the labor market. Drop-outs may be either temporary, --- that females only exit the labor market when children are young, and they return after children grow up, or permanent, --- that females stay out of the labor market for considerably long time, or even life-time. In our paper, although we are not able to tell whether some of the females drop out of the labor force permanently, we can at least see how persistent the adverse impact of childbirth on employment is. First, we separate those females who had childbirth during our sample period, denote the year of childbirth for each one as year t, then we define year dummies $Year_k$ in which k ranges from 4 years before childbirth ($t-4$) to 5 years after childbirth ($t+5$). Using Logit model, we explain females' employment as

$$Pr(Work_{it} = 1|Year_k, X_{it}) = f\left(\alpha + \sum_{k \in \binom{t-4,t-3,t-2,t-1,}{t,t+1,t+2,t+3,t+4,t+5}} \beta_k Year_k + X_{it}\gamma + \delta_t\right) \quad (5)$$

We apply the same model for all females with childbirth, females with no post-childbirth grandparental support, and females with post-childbirth grandparental support, respectively. Results are reported in Figure 3.

Figure 3 Persistency of the impact of childbirth on female employment



Overall, as we can see, the adverse impact of childbirth on employment persists for about 3 years, while in the long run, females tend to return to the labor market. However, the persistency differs significantly for females with / without grandparental support. For females without grandparental support, their employment starts to recover only 4 years after childbirth, while for those with grandparental support, their employment starts to rise only one year after childbirth. Grandparental childcare thus significantly shortens the childbearing-induced interruption to females' careers.

Labor intensity before and after childbirth

Childbirth and childcare may not only force females to drop out of labor force, but may also force them to reduce labor intensity by shifting towards part-time jobs or jobs with flexible hours. To see the impacts on females labor intensity, we separate those females with positive weekly working hours and explain their labor intensity in panel regressions. The results are reported in Table 8.

Table 8 Regression results for labor density

	Weekly working hours			
	(1) RE OLS	(2) FE OLS	(3) RE OLS	(4) FE OLS
Motherhood	-0.148 (0.939)	-0.515 (1.371)	-1.297 (1.136)	-1.836 (1.603)
Motherhood*GPC			2.479** (1.202)	2.650* (1.490)
Controls	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes
Year dummy	Yes	Yes	Yes	Yes
Obs.	3,706	3,706	3,706	3,706

Note: Standard errors in the parentheses are heteroskedasticity robust standard errors.

It can be seen that childbirth does not significantly reduce females' labor intensity, and the females receiving post-childbirth grandparental support work even more than females who had no childbirth. These findings are consistent with the previous results: Because the supply of part-time jobs or flexible jobs is rather limited, for post-childbirth females who return to the labor market, they mostly have to maintain the same labor intensity compared with females who had no childbirth. This deters females' willingness to work after childbirth, especially for those who have no grandparental support. Policies that aim to create part-time jobs are thus helpful to keep childbearing females in the labor force.

Types of jobs taken by childbearing females

We further investigate the types of jobs that are taken by childbearing females, those without grandparental support versus those with. The results are presented in Table 9. It can be seen that almost half of females without grandparental support are either self-employed and / or from agricultural sector, more than double as many as females with grandparental support. Females with grandparental support are twice more likely to receive pension and medical insurance from their employers, and 60% more likely to receive housing fund. It may be that females without grandparental are forced into less formal and secured jobs to reconcile with childcare, or females from these types of jobs are more likely able to reconcile with childcare without grandparental support. We leave the explanation for our future research.

Table 9 Types of jobs: Females without grandparental support versus females with grandparental support

	Agricultural or self-employed	Employed	Managerial	Pension provided by employer	Medical insurance provided by employer	Housing fund provided by employer
w/o GPC	45.26%	54.64%	10.18%	15.12%	14.95%	11.34%
w/ GPC	21.26%	78.74%	12.92%	29.53%	29.72%	18.70%

6 Concluding Remarks

The unusually high female labor market participation rate in China is strongly contrasted by the poor public expenditure in childcare and policy support for female employees. Our paper finds the intrafamily downward labor transfer is the key to understanding such paradox. The intrafamily grandparental childcare largely fills the missing role of public provision of childcare, reduces the opportunity costs for childbearing females to work, improves their labor market participation, and reduces the childbearing-induced interruption to females' careers. Grandparental childcare largely explains why post-childbirth female employment rate and labor intensity are so high, given the public subsidy to childcare is rather limited: With addressing the endogeneity problem of females' fertility decisions using PSM-DID model, we find that the employment rate for females without grandparental support falls substantially after childbirth, and the fall is even significantly larger than that in advanced economies; however, the employment rate for females with grandparental support does not fall after childbirth, instead, it even slightly rises. We further show that the recovery in the employment for females without grandparental support takes twice as long as females with grandparental support.

Our results have strong implication for the recent debate on postponing retirement. People in the debate generally agree that the current statutory retirement age is too low so that it must be raised to relieve the mounting burden of the pension system. However, our research shows that retired workers contribute much to take care of their grandchildren, thus allowing young childbearing females to maintain a high level of employment and labor intensity. Without providing more public support for childcare, postponing the retirement of old workers may shift the burden of childcare towards young females, thus crowd them out from the labor force. Our research therefore calls for more public investment in childcare and social protection policies for childbearing females, in companion with the phasing-in of the new retirement policy.

References

[1] Abadie, A., Drukker, D., Herr, J. L. and Imbens, G. W. (2004), "Implementing Matching Estimators for Average Treatment Effect in Stata", *Stata Journal*, 4(3), 290-311.

[2] Agüero, J. M. and Marks, M. S. (2008), "Motherhood and Female Labor Force Participation: Evidence from Infertility Shocks", *American Economic Review*, 98(2), 500-504.

[3] Agüero, J. M. and Marks, M. S. (2011), "Motherhood and Female Labor Supply in the Developing World Evidence from Infertility Shocks", *Journal of Human Resources*, 46(4), 800-826.

[4] Ahn, N. and Mira, P. (2002), "A Note on the Changing Relationship between Fertility and Female Employment Rates in Developed Countries." *Journal of population Economics*, 15(4), 667-682.

[5] Anderson, D. J., Binder, M. and Krause, K. (2002), "The Motherhood Wage Penalty: Which Mothers Pay It and Why", *American Economic Review*, 92(3), 354-358.

[6] Angrist, J. D. and Evans, W. N. (1998), "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size", *American Economic Review*, 88(3), 450-477.

[7] Arpino, B., Pronzato, C. D. and Tavares, L. P. (2014), "The Effect of Grandparental Support on Mothers' Labour Market Participation: An Instrumental Variable Approach", *European Journal of Population*, 30(4), 369-390.

[8] Beguy, D. (2009), "The Impact of Female Employment on Fertility in Dakar (Senegal) and Lomé (Togo)", *Demographic Research*, 20, 97-128.

[9] Besamusca, J., Tijdens, K., Keune, M. and Steinmetz, S. (2015), "Working Women Worldwide. Age Effects in Female Labor Force Participation in 117 Countries", *World Development*, 74, 123-141.

[10] Blau, F. D. and Kahn, L. M. (2013), "Female Labor Supply: Why is the United States Falling Behind", *American Economic Review: Papers & Proceedings*, 103(3), 251-256.

[11] Bloom, D. E., Canning, D., Fink, G. and Finlay, J. E. (2009), "Fertility, Female Labor Force Participation, and the Demographic Dividend", *Journal of Economic Growth*, 14(2), 79-101.

[12] Bratti, M., Frattini, T. and Scervini, F. (2016), "Grandparental Availability for Child Care and Maternal Employment: Pension Reform Evidence from Italy", *Institute for the Study of Labor (IZA) Working Paper*, No. 9979.

[13] Chen, J., Shao, X., Murtaza, G. and Zhao, Z. (2014), "Factors that Influence Female Labor Force Supply in China", *Economic Modelling*, 37, 485-491.

[14] Compton, J. and Pollak, R. A. (2014), "Family Proximity, Childcare, and Women's Labor Force Attachment", *Journal of Urban Economics*, 79, 72-90.

[15] Cristia, J. P. (2008), “The Effect of a First Child on Female Labor Supply: Evidence from Women Seeking Fertility Services”, *Journal of Human Resources*, 43(3), 487-510.

[16] Cruces, G. and Galiani, S. (2007), “Fertility and Female Labor Supply in Latin America: New Causal Evidence”, *Labour Economics*, 14(3), 565-573.

[17] Dimova, R. and Wolff, F. C. (2011), “Do Downward Private Transfers Enhance Maternal Labor Supply? Evidence from Around Europe”, *Journal of Population Economics*, 24(3), 911-933.

[18] Emara, N. (2016), “Fertility and Female Employment: A Panel Study on Developing Countries”, *Applied Economics and Finance* 3(2), 122-127.

[19] Engelhardt, H. and Prskawetz, A. (2005), “A pooled time-series analysis on the relation between fertility and female employment”, Vienna Institute of Demography, Austrian Academy of Sciences.

[20] Entwisle, B. and Chen, F. (2002), “Work Patterns Following a Birth in Urban and Rural China: A Longitudinal Study”, *European Journal of Population*, 18(2), 99-119.

[21] Fang, H., Eggleston, K. N., Rizzo, J. A. and Zeckhauser, R. J. (2013), “Jobs and Kids: Female Employment and Fertility in China”, *IZA Journal of Labor & Development*, 2(1), 12.

[22] Feng, P. and Ke, S. Z. (2016), “Self-selection and Performance of R&D Input if Heterogeneous Firms: Evidence from China’s Manufacturing Industries”, *China Economic Review*, 41, 181-195.

[23] Gangl M. and Ziefle, A. (2009), “Motherhood, Labor Force Behavior, and Women’s Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States”, *Demography*, 46(2), 341-369.

[24] García-Morán, E. and Kuehn, Z. (2017), “With Strings Attached: Grandparent-provided Child care, Fertility, and Female Labour Market Outcomes”, *Review of Economic Dynamics* 23, 80-98.

[25] Gerster, M. and Lappegård, T. (2010), “Mother’s Employment and Fertility in Norway”, *Statistics Norway Discussion Paper* No. 624.

[26] Gomes, C. S., Oliveira, I. T. D., Pinto, M. L. R. and Cabrita, M. (2012), “Fertility, Full-time and Part-time Female Employment in Europe”, *CIES e-Working Papers* ISSN 1647-0893.

[27] He, X. and Zhu, R. (2016), “Fertility and Female Labour Force Participation: Causal Evidence from Urban China”, *Manchester School*, 84(5), 664-674.

[28] Hwang, J., Park, S. and Shin, D. (2018), “Two Birds with One Stone: Female Labor Supply, Fertility, and Market Childcare”, *Journal of Economic Dynamics & Control*, 90(2), 171-193.

[29] Kögel, T. (2004), “Did the Association between Fertility and Female Employment within OECD Countries Really Change Its Sign”, *Journal of Population Economics*, 17(1), 45-65.

[30] Kögel, T. (2006), “An Explanation of the Positive Correlation between Fertility and Female Employment across Western European Countries”, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=995206.

[31] Longwe, A., Smits, J. and de Jong, E. (2013), “Number and Spacing of Children and Women’s Employment in Africa”, *Nijmegen Cent Econ NiCE Working Paper*, No. 13-103.

[32] Lumsdaine, R. L. and Vermeer, S. J. C. (2015), “Retirement Timing of Women and the Role of Care Responsibilities for Grandchildren”, *Demography*, 52(2), 433-454.

[33] Majbouri, M. (2016), “Fertility and the Puzzle of Female Employment in the Middle East”, *IZA Discussion Paper* No. 11322.

[34] Mandel, H. and Semyonov, M. (2006), “A Welfare State Paradox: State Interventions and Women’s Employment Opportunities in 22 Countries”, *American Journal of Sociology,* 111(6), 1910–1949.

[35] Molina J. A. and Montuenga, V. M. (2009), “The Motherhood Wage Penalty in Spain”, *Journal of Family and Economic Issues*, 30(3), 237-251.

[36] Posadas, J. and Vidal-Fernández, M. (2013), “Grandparents’ Childcare and Female Labor Force Participation”, *IZA Journal of Labor Policy*, 2(1), 2-14.

[37] Rondinelli, C. and Zizza, R. (2011), “(Non-) persistent Effects of Fertility on Female Labour Supply”, *ISER Working Paper,* No. 2011-4.

[38] Sun, W. K. and Wang, Y. J. (2016), “The Effects of Parental Migration on Left-behind Children’s Health in China”, *China Economic Quarterly* 15(3), 963-988 (in Chinese).

[39] Van den Broeck, G. and Maertens, M. (2015), “Female Employment Reduces Fertility in Rural Senegal”, *PLoS ONE*, 10(3), e0122086.

[40] Yu, H. Y., Cao, J. and Kang, S. L. (2018), “Fertility Cost, Intergenerational Labor Division, and Female Employment”, *CESifo Working Paper* No. 7293.

[41] Yu, J. and Xie, Y. (2014), “The Effect of Fertility on Women's Wages in China”, *Population Research* 2014(1), 18-29 (in Chinese).

[42] Zhang, C. C. (2011), “Impacts of Number of Children on the Labor Supply and Wage Earnings of Married Women”, *Population & Economics* 2011(5), 29-35 (in Chinese).

Appendix

A Estimated propensity score using Logit model

	2010-2012 subsample	2012-2014 subsample	2014-2016 subsample
Employment	-0.081 (0.203)	-0.266 (0.186)	-0.209 (0.230)
Age	0.492** (0.249)	0.377** (0.184)	0.572** (0.227)
Age squared, divided by 100	-1.134** (0.451)	-0.847*** (0.322)	-1.105*** (0.397)
College degree	-0.089 (0.281)	-0.325 (0.221)	0.030 (0.235)
Urban residency	-0.739*** (0.274)	-0.137 (0.219)	-0.440* (0.245)
Studentship	-1.736*** (0.433)	-1.929*** (0.405)	-2.200*** (0.554)
Marriage	2.537*** (0.219)	2.432*** (0.186)	1.889*** (0.221)
Log household's per capita net income, excluding the female's	-0.024 (0.044)	-0.004 (0.039)	0.060 (0.040)
Home ownership	-0.280 (0.359)	-0.058 (0.307)	-0.167 (0.279)
Home size	-0.001 (0.001)	-0.001 (0.001)	4.61e-4 (7.05e-4)
Hardship in housing	-0.624** (0.277)	-0.500* (0.277)	-0.618** (0.287)
Household's total assets	-4.36e-8 (1.77e-7)	1.61e-7 (1.57e-7)	-4.56e-8 (1.53e-7)
House price for the community	-1.75e-5 (2.53e-5)	4.46e-5 (2.72e-5)	-4.42e-6 (1.67e-5)
Number of kindergartens in the community	-0.022 (0.095)	-0.015 (0.073)	-0.071 (0.080)
Number of primary schools in the community	-0.022 (0.133)	-0.109 (0.127)	0.175 (0.171)
Community's birth rate	-0.009 (0.014)	-0.016 (0.014)	-0.004 (0.010)
Constant	-5.627* (3.315)	-5.089** (2.564)	-9.189*** (3.144)
Log likelihood	-345.35	-467.80	-332.01
Pseudo R^2	0.3195	0.2839	0.2376
Significance test for the model	324.28 (P=0.000)	370.84 (P=0.000)	206.93 (P=0.000)
Observations	820	1196	911

Notes: (1) *** / ** / * denotes the result is significant on 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) Significant test for the model is based on LR χ^2-statistic.

B Covariates' balance test before and after matching

	Matching	Pseudo R^2	LR statistic	P-value	Standardized error
2010-2012 subsample	Before	0.317	322.09	0.000	26.5

	K-nearest neighbor	0.011	7.87	0.953	4.3
	Caliper	0.011	7.74	0.956	4.2
	K-nearest neighbor with caliper	0.011	7.87	0.953	4.3
	Kernel	0.010	7.33	0.966	4.2
2012-2014 subsample	Before	0.286	373.16	0.000	19.5
	K-nearest neighbor	0.015	11.21	0.796	4.3
	Caliper	0.008	6.36	0.984	3.5
	K-nearest neighbor with caliper	0.015	11.21	0.796	4.3
	Kernel	0.008	6.49	0.982	3.5
2014-2016 subsample	Before	0.239	207.94	0.000	20.0
	K-nearest neighbor	0.011	4.75	0.997	4.3
	Caliper	0.005	2.28	1.000	3.2
	K-nearest neighbor with caliper	0.011	4.75	0.997	4.3
	Kernel	0.005	2.38	1.000	3.2