WORKING PAPER

Fertility Cost, Intergenerational Labor Division, and Female Employment

NORGES BANK RESEARCH

3 | 2019

HAIYUE YU,
JIN CAO AND
SHULONG KANG

NORGES BANK

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:
FacilityServices@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail
FacilityServices@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-819-0 (online)
ISBN 978-82-8379-069-6 (online)

Fertility Cost, Intergenerational Labor Division, and Female Employment

Abstract

China has set to increase the minimum retirement age, to ease the pressure from pension expenditure and the falling labor supply caused by the aging population. However, policy debates have so far neglected the crucial fact that families in China largely rely on retired grandparents for childcare. Using novel and high-quality survey data, we demonstrate that intra-family downward labor transfer towards childcare significantly increases young females' labor force participation rate and their labor income, and such effects do not exist for males. Furthermore, we show that the positive effects from grandparental childcare are higher for better-educated, urban females with younger children. This paper thus reveals a large, hidden cost in the new retirement policy — the reduced feasibility of grandparental support, due to postponed retirements, may crowd out productive labor of young females, — and rationalizes a series of social protection policies to accompany the phase-in of the new retirement scheme.

JEL-Codes: C240, J130, J220.

Keywords: intergenerational labor division, grandparental childcare, female employment, human capital accumulation, minimum retirement age.

Haiyue Yu
Dongbei University of Finance and Economics, Interdisciplinary Center for Social & Behavioral Studies
217 Jianshan Street
China – 116023 Shahekou District, Dalian
dufebruce@163.com

Jin Cao
Research Department
Norges Bank
Bankplassen 2, PB 1179 Sentrum
Norway – 0107 Oslo
jin.cao@norges-bank.no

*Shulong Kang**
Dongbei University of Finance and Economics
School of Finance
217 Jianshan Street
China – 116023 Shahekou District, Dalian
shulonkang@dufe.edu.cn

*corresponding author

 The authors thank participants of 2018 CESifo Area Conference on Employment and Social Protection, Munich and 2018 Asian Meeting of the Econometric Society, Seoul for helpful comments.

1 Introduction

China has become one of the first major emerging market economies to start to suffer from an ageing population. Its working population has been steadily shrinking since 2012 (World Bank, 2017) when the baby boomers from the early 1960s began to reach retirement age; and gone are the days when manufacturing firms could easily access abundant low-cost labor. As a consequence of the shrinking labor force and ageing population, the acutely underfunded social welfare system is under mounting pressure: The age dependency ratio – defined by the ratio of working employees with pension plans to aged pension-dependent retired employees – has fallen to 2.75:1 in 2016 from 5:1 or more in late 1990s, and it is expected to fall further to 1.3:1 in 2050, while the annual growth rate of the state pension fund's gross income has fallen short of the annual growth rate of pension payments for the past six years (MHRSS, 2017). To ease the increasing shortages in labor supply and pressure on the social security system, the state has determined to delay retirement gradually over the next decades: The statutory retirement age (currently 60 for men, 55 for women in the public sector and 50 for women who work in factories) will be raised gradually until it reaches the level of advanced economies (65 or more) by 2040s or so.

Policy debates and current research about the retirement age have so far much focused on whether increasing the old generation's labor supply has a substitution effect on the young generation's labor supply. In general, the old generation and the young generation do not compete directly in the labor market, thus the rising employment of the old does not necessarily crowd out the employment of the young (Gruber et al. 2010). However, such statement is rather questionable in China, as the labor supply of the old is closely connected with the labor supply of the young through the intra-family labor division. In fact, young children in China have been widely looked after by their retired grandparents, besides the parents. Such intra-family, downward intergenerational labor transfer effectively reduces the childbearing-induced interruption in the careers of working young females as well as their opportunity cost of working, leading to a high female labor force participation rate along with the low birth-related wage penalty in China. Given such a fact, raising the retirement age may reduce grandparents' labor input to childcare for their grandchildren and shift more of the burden of childcare onto young females, hence deterring young females' willingness to work or forcing them to take a longer

break in their career paths. To make matters worse, such an increased interruption in females' careers implies a larger interruption in young females' human capital accumulation, and this effectively lowers the paths of their lifetime labor income flows. Overall, the potential adverse impacts on young females' post childbirth employment and lifetime income may even undo the policy that aims to increase China's aggregate labor supply and the gross income of its pension fund via raising the statuary retirement age.

This paper is one of the first to investigate the impact of grandparent-provided childcare on young females' employment and labor income in China. Using a novel, high-quality dataset, we show that the intra-family downward labor transfer from the older generation significantly increases young females' employment rate and their labor income. This study reveals the hidden costs of raising the statuary retirement age to the female labor supply and rationalizes policies that subsidize childbearing and increase the supply of qualified, affordable childcare services in the market.

1.1 Background and literature review

The current statuary retirement age in China was set in 1978 and has been little changed since then. Men retire at 60 and women at 50 (55 if they work in the public sector), and a large share of retired people commit to childcare within families. In a 2007 survey conducted by the Shanghai Population and Family Planning Commission, 88.7% of the grandparents were involved in caring for their own grandchildren, and 53.3% of the grandparents took major responsibilities in the childcare on a regular basis. The China Health and Retirement Longitudinal Study (CHARLS) shows that 50% of grandparents regularly take on major responsibilities for caring for their grandchildren — much higher than many countries. In contrast, in the US and Europe, grandparent-provided childcare is in general not very common (except in 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 Vermeer 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] Even though in China, childcare services are available in the market, few households actually rely on them. 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.

Such high involvement of the older generation in childcare has a big impact on the labor market for the young generation, especially for young females. It reduces the opportunity cost of childbearing along young females' career paths and leads to two distinguishing features of China's labor market. One is extremely high female labor-force participation. China's labor-force participation rate for females is currently at 61.5% (percentage of the female population aged 15 or over, computed from International Labor Organization, ILOSTAT database. Data retrieved in March 2017), which is among the world's highest (excluding the least developed countries) — not only significantly higher than the US (55.7%) and the euro area (50%), but also higher than Scandinavian countries that are traditionally well known for gender equality in the labor market, such as Denmark (59.2%), Norway (60.8%) and Sweden (60.8%). The other distinguishing feature is that childbearing-induced penalty in females' careers is surprisingly low in China, even though social protection for childbearing female employees is poor. The childbearing-induced penalty for female employment — computed as the childbirth-induced drop in the employment rate — is only 5.6% (for the urban population) or 3.0% (for the rural population) in China (Zhang 2011), much lower than the level in the US (26.3%; see Cristia 2008), while the penalty in labor income — computed as the childbirth-induced drop in wages — is merely 7% for China (Yu and Xie 2014), compared with 18% for Germany, 10% for the UK, 9% for Spain (Gangl and Ziefle 2009), and 16% for the US (Anderson et al. 2002).

It has been well established in the literature that childbearing crowds out young females' working hours, thus significantly reducing their' employment rate (for example, Cristia 2008, Zhou 2008, Xiong and Li 2016) and the intensity in labor supply (for example, Angrist and

[3] It is worth noting 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 Vermeer 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.

Evans 1998, Entwisle and Chen 2002, Zhang 2011). Moreover, childbearing significantly reduces females' wage rate, and results in a wage penalty (Anderson et al. 2002, Yu and Xie 2014) because of reduced productivity (Budig and England 2001), depreciation in human capital caused by the interruption in careers (Polachek 1981) and so on. Therefore, for those young females who can have an older generation to share part of the burden of childcare, it may be naturally expected that they will experience less of a drop in their post-childbirth employment rate and labor income than those females with no such support. In the literature, studies on the impacts of grandparental childcare on female employment mainly follow two strands:

The first strand directly investigates the impact of the older generation's downward labor transfer on females' employment. 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 the 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 strong for ethnic minorities and single mothers. Arpino et al. (2014) document similar positive effect using Italian data with instrumental variables, and this effect is stronger for females with less education and younger children.

The second strand focuses on quantifying the impact using proxies of grandparental childcare, when direct measurements are not available. It has been documented (for example, García-Morán and Kuehn 2012, Compton and Pollak 2014) that the labor-force participation rate is higher for young females living closer to their parents so that grandparental childcare is more likely. Based on a natural experiment from the 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.

Compared with the US and Europe, the intra-family downward labor transfer for childcare in China is much more common and intensive, however, studies of its impact on young females'

employment are still scarce. Although Shen et al. (2012) find that living with the older generation increases overall females' labor-force participation rate and working hours, which may be explained by the higher feasibility of grandparental support for housework, the study itself does not focus on females with young children; therefore, it lacks direct evidence how intergenerational labor transfer affects childbearing females' employment through grandparental childcare. Besides employment, there is also scant literature on how grandparental childcare affects females' labor income, i.e., how grandparental support reduces interruptions to childbearing females' career path, or, path of human capital accumulation, as well as the negative shocks to their after-birth income flow.

Our paper contributes to the literature in three ways: First, this paper is the first to investigate the direct impacts of grandparent-provided childcare on both young females' employment and labor income in China, based on a novel, high-quality dataset. Using Probit and Tobit models with instrumental variables we are able to identify the heterogeneities in the affected females with respect to education level, residential area, and age of children, and provide economic explanations. Second, through the lens of the intra-family intergenerational labor division, we provide convincing explanations of China's extremely high female labor force participation rate along with a very low childbearing-induced penalty for female employment and wages. Third, our findings reveal the hidden cost of young females' employment and human capital development associated with raising statuary retirement age. Our research suggests that policy makers face a clear tradeoff between gaining additional labor supply from the end of old employees' career via postponing their retirement and crowding out the labor supply provided by young females which may have persistent negative effects on their career path due to losing grandparental support for childcare. This rationalizes the need for policies supporting childbearing and female employment, to accompany the phasing-in of new retirement policies.

1.2 Structure of the paper

Section 2 presents the stylized facts on grandparental childcare and briefly describes the data, then section 3 constructs the econometric models for our analysis. Section 4 reports the results from estimating the models, followed by the interpretations and discussions on the results. In section 5 we conduct several robustness checks using various model specifications; finally section 6 outlines policy implications and concludes.

2 Childcare and female employment: Data description and stylized facts

2.1 Data description

To analyze the impact of grandparental childcare on young females' labor-force participation and labor income, we need to combine the information of three generations within the same families: Children, children's parents and grandparents. We extract such information and construct our cross-sectional dataset from China Family Panel Studies (CFPS), 2014. 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 provide a better understanding of the economic, as well as the non-economic, well-being of the 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. In the 2014 survey, the CFPS successfully interviewed almost 15,000 families across China and almost 30,000 individuals within these families, with an approximately response rate of 79%.

Matching the information of children with adults in the same families, we get a sample of 4,277 females living with young children from CFPS 2014.[4] Key variables that we extract from CFPS for our empirical analysis include those on females' employment and labor income, grandparental childcare, females' characteristics, children's characteristics, and so on, such as:

Grandparental childcare (GPC): Dummy variable. For a childbearing female i, if at least one of her children below age 11 is mainly cared for by grandparent(s) during the day time[5] 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 $WORK_i = 0$;

[4] Children in our sample are restricted to be under age 12, as in the CFPS children above 12 are supposed to take care of themselves in the daytime, and the questionnaire for them is different. Under such criteria, there are 4,251 females living with 5,658 small children in the survey; however, considering females currently retire at 50 or so such that there isn't a conflict between childcare and employment anymore once they retire, we dropped 24 females above 50 in our sample.

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

Log female's labor income ($\ln Y_i$*)*: The logarithm of female i's annual labor income in renminbi (CNY), including wage income and other income from agricultural and non-agricultural activities;[6]

Number of young children ($NCHILD_i$*)*: The number of female i's young children at age 0-2;

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

Urban residency ($URBAN_i$*)*: 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;

Log household's annual total net income, excluding the female's ($\ln Y\_FAMILY_i$*)*: The logarithm of female i's household's annual total net income in CNY (including the male's income and transfers, but excluding the female's income).

Table 3 presents the summary statistics of the variables.

Table 3: Summary statistics

Variables	Obs.	Mean	Std. dev.	Min	Max
Intergenerational support, GPC	4,227	0.315	0.465	0	1
Employment, $WORK$	4,227	0.718	0.450	0	1
Log female's annual labor income, $\ln Y$	4,227	5.274	4.612	0	12.506
Number of children of or below age 2, $NCHILD$	4,227	0.328	0.512	0	3
Age of youngest child, AGE_CHILD	4,227	4.862	3.310	0	11
Female's age, AGE	4,227	31.96	6.373	17	49
Education level of the female, EDU	4,216	0.103	0.305	0	1
Years of education for the female, EDU_YEAR	4,218	7.967	4.405	0	19
Urban residency (female),	4,201	0.424	0.494	0	1

[6] The value is set to 0 if her net income is 0 or negative (a few observations with negative net income in our sample are mostly self-employed females who incur losses in their businesses). As our income model (see section 3.2) is based on the Tobit model with a left truncation at 0, such setting does not affect our estimation.

URBAN					
Log household's total income* (excluding female's), $\ln Y\_FAMILY$	4,227	9.145	3.445	0	15.220

* The value is set to 0 if the household's total annual income (excluding the female's) is non-positive.

2.2 Stylized facts of grandparental childcare and female employment

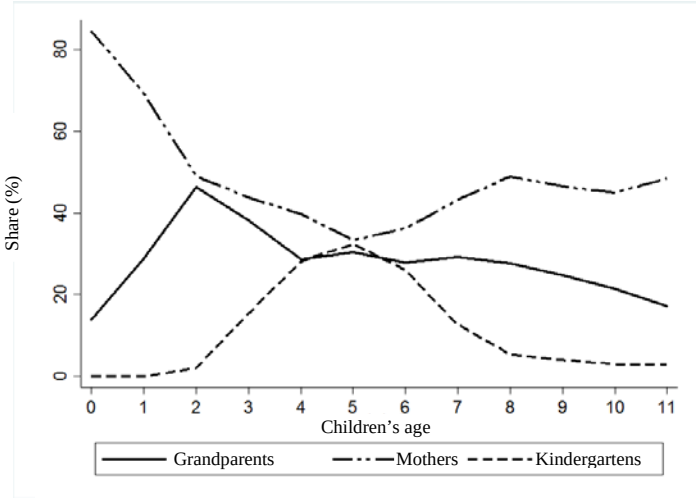


Figure 1: Providers of childcare in the day time for children of different ages

Using the dataset that we construct from CFPS, we characterize three stylized facts on grandparental childcare and its impact on female employment. The first stylized fact is that grandparental childcare is indispensable complement to parental care, which can be directly seen from the CFPS survey. In the questionnaire, it is asked for every child in the sample, "By whom the child is mainly looked after during the daytime?" Among all the options in the answer, over 90% of the respondents choose one of the following four: (1) mother, (2) paternal grandparents, (3) maternal grandparents, and (4) kindergarten. These four options correspond exactly to three major types of childcare in China: Maternal care, non-maternal "non-institutional" care (i.e., provided by grandparents), and "institutional" care (i.e., provided by kindergartens). In Figure 1, we combine childcare provided by paternal grandparents and maternal grandparents as "grandparents" and present the share of each type of childcare that is used across families. It can be seen that at ages 0-1 most children are cared for by their mothers, because maternal care is essential in this stage and females may take maternity leave (often 3-5 months). On the other

hand, grandparental care is almost the only complement available during this period, especially when females start to return to work after maternity leave, grandparents almost entirely fill in the gap: The sum of maternal and grandparental care accounts for more than 98% of total childcare supply at this stage. At ages 2-3, grandparental care peaks, as children at this stage are still below the minimum admission age (which is 3, stipulated by the *Regulation on the Administration of Kindergartens*, Ministry of Education of the People's Republic of China, September 11, 1989) for kindergartens while their mothers' maternity leave is already exhausted. The use of grandparental care is almost as much as maternal, and the sum of these two exceeds 90% in this stage. At ages 4-6, mothers, grandparents and kindergartens take almost equal shares in childcare, as many children are admitted to kindergartens; [7] later, at ages 7-11, kindergartens' share falls to nearly zero after the children are admitted to primary schools, the share of grandparents starts to fall but still remains at a substantially high level. Overall, throughout the entire childhood although maternal care is the most important one, grandparents play a crucial role in childcare. Especially, before children reach kindergarten age, care provided by grandparents is almost the only alternative to mothers'.

Another interesting feature in Figure 1 is that, besides these three major sources of childcare, other sources are almost neglectable. Among the alternatives, daycare service for children aged 0-3, which is widely available in developed countries, is still in very short supply in China. [8] Childcare service purchased from the market is not well used, either: First of all, qualified childcare services so far have been scarce and expensive so that most families can hardly afford the cost of fully relying on employed childcare; at the same time, a credible mechanism is lacking to properly monitor childcare workers, thus even with a full-time employed babysitter at home, monitoring provided by family members (usually by grandparents) is needed mostly for safety reasons. In our sample, among 2,281 children below the age of three, only three are fully

[7] As of 2016, there are more than 240,000 kindergartens across the country and 77.4% of all preschool-aged children attend kindergarten in China. See "Dire shortage of daycare centers may impact China's efforts to increase birth rate", *Global Times*, November 14, 2017, available at http://www.globaltimes.cn/content/1075165.shtml.

[8] The enrolment rate of 0 to 3 year-olds in day-care is only 4.1 percent in China, far lower than in developed countries, where the rate averages 50 percent. Even in the most developed regions such as Shanghai, the enrolment rate is no higher than 14% (*Global Times*, November 14, 2017).

cared for by babysitters hired from the market during the daytime. As a consequence of missing alternatives, grandparents are heavily involved in many families' childcare.[9]

Table 1: Labor market participation rate for females / males with (Yes) or without (No) grandparental support

Labor-force participation rate		Females (%)			Males (%)		
		(1) Yes	(2) No	Difference	(3) Yes	(4) No	Difference
Full sample		82.2	66.9	15.3***	94.0	94.7	-0.7
Children's age	0-2	74.9	39.4	35.5***	92.8	93.5	-0.7
	3-6	86.0	72.8	13.2***	94.6	94.7	-0.1
	7-11	87.4	82.1	5.3**	94.7	95.5	-0.8

Note: (1) *** / ** denote the difference is significant at the 1% / 5% level; (2) source of data: CFPS (2014)

The second stylized fact is that childbearing young females receiving older generation's support are more likely to stay in the labor force than those without, and the difference declines as the children grow up. Table 1 presents the difference in labor force participation rates, between childbearing females benefiting from grandparental childcare and those who do not. Among all 4,227 young females living with young children in the entire sample, 1,333 or 31.54%, get such support. Overall, females with support are 15.3% more likely to work, and the difference is significant at the 1% level. Such differences narrow for females with older children, as the required labor intensity of childcare is lower, as is the crowd-out effect on females' labor supply. As the statuary admission age in China is three for kindergartens and seven for primary schools, respectively, we further divide the whole sample into three groups with respect to the age of each female's youngest child: Ages 0-2, 3-6, and 7-11. As is shown in Figure 1, the use of grandparental care is less common for older children: In our sample, 37.97% of children aged 0-2 are looked after by grandparents, and the share falls to 33.51% and 23.35% for age groups 3-6 and 7-11, respectively. Within-group differences in females' likelihood of working are more striking when the children are younger, as reported in Table 1: Females with support are 35.5%, 13.2%, and 5.3% more likely to work in each of the three groups, respectively. In addition, Table

[9] The strict definition of grandparental childcare in CFPS, that to be eligible as childcare providers grandparents must be the main responsible persons during day time on daily basis, leads to underestimating the role of grandparental childcare, as many non-eligible grandparents are also involved in childcare – just to a lesser degree.

1 also presents the labor force participation rates for males in each of the same groups for comparison. It can be seen that either for the entire sample or for each of the children's age groups, there is no significant difference in employment among males.

Obviously, childcare needs a substantial input of time; in general, as young females in China take the most responsibility in childcare, the burden of childcare may force them to reduce their labor supply, take a leave from their career, or even exit the labor force (Entwisle and Chen 2002). To reduce these adverse impacts, other sources of childcare are needed to share the burden with the young females, especially before their children are eligible to be admitted to kindergartens. Given the lack of alternative childcare providers, grandparental childcare indeed reduces the burdens of maternity. This helps females return to labor force after giving birth, thus reducing the interruption to their career.

In Europe or the US, as grandparental care is not widely available childbearing females mainly face a tradeoff between providing their own childcare via reducing their own labor supply and buying childcare services from the market; therefore, the market price of childcare largely reflects their opportunity cost of working. The higher the price is, the higher the wage threshold females need to earn from work, therefore, the more it deters females from working. For example, Connelly (1992) shows that the low labor force participation rate of the US females living with pre-school children can be much explained by the high prices in the childcare market. In China, with a malfunctioning market for childcare, the price (implicit and explicit) for such services is prohibitively high for most families. Should grandparental support be absent, most childbearing females would have chosen not to work as the opportunity cost is too high; therefore, the widely used grandparent-provided (usually almost free) care in China effectively fills the gap in the market and significantly reduces the opportunity cost of working for young females receiving such support. This is very much reflected in China's high female labor force participation rate.

The third stylized fact is that childbearing females receiving grandparental support earn higher labor income, and the positive effect of grandparental childcare on their labor income persists even as the children grow up. As is shown in Table 2, we compare the average labor income for childbearing females with / without support. In our sample, 2,465 young females involved in childcare receive non-zero labor income at the time of survey, and 36.51% of them receive support from the older generation. As is shown in Panel A, those receiving support earn on

average 45.83% higher annual income, and the difference is significant at the 1% level. If we include observations with zero labor income, the difference is more striking as childbearing females without grandparental support are more likely to drop out of the labor force and earn zero income. In this case, as is shown in Panel B, females with grandparental support on average receive 81.09% higher annual income.

Table 2: Average annual income for childbearing females / males with (Yes) or without (No) intergenerational support

Panel A: Observations with non-zero labor income

		Females (thousand CNY)			Males (thousand CNY)		
		(1) Yes	(2) No	Difference (%)	(3) Yes	(4) No	Difference (%)
Full sample		22.56	15.55	45.08***	31.87	29.78	7.02
Children's age	0-2	21.24	12.13	75.10***	31.64	32.21	-1.77
	3-6	22.47	16.69	34.63***	32.38	30.29	6.90
	7-11	24.35	16.01	52.09***	31.43	28.91	8.72

Panel B: Full sample, including observations with zero labor income

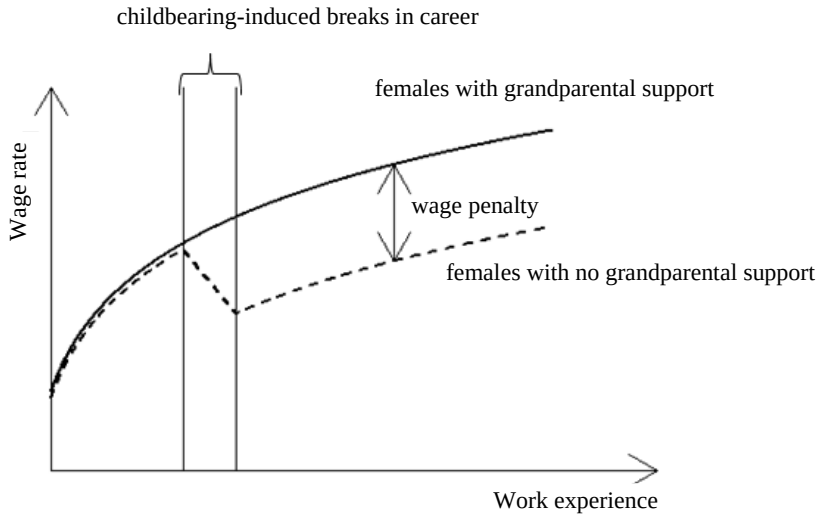
		Females (thousand CNY)			Males (thousand CNY)		
		(1) Yes	(2) No	Difference (%)	(3) Yes	(4) No	Difference (%)
Full sample		15.23	8.41	81.09***	23.50	22.70	3.52
Children's age	0-2	13.16	4.34	203.2***	23.45	24.31	-3.54
	3-6	15.43	9.62	59.70***	23.84	22.95	3.88
	7-11	18.06	10.32	74.92***	23.06	22.34	3.22

Note: (1) *** denotes the difference is significant at the 1% level; (2) source of data: CFPS (2014)

Again, the impact of grandparental support on female labor income varies with children's age. The impact is strongest when the children are youngest: For females with non-zero labor income, as in Panel A, the mean annual labor income for those with support is about CNY 9,110, or 75.10% higher than those without in the age group 0-2, and such difference becomes 5,780 (34.63%) and 8,340 (52.09%) for age groups 3-6 and 7-11, respectively – persistently high even for females with older children. All these differences widen after we include observations with zero income. However, the data do not exhibit any significant difference among the males in the same sample.

It has been widely known that fertility and childcare not only discourage females' labor force participation, but also reduce their labor income in both short run and long run — known as the childbearing-induced "wage penalty". Typically, as work experience increases, a worker's human capital accumulates, and her labor income rises, as Mincer (1974) shows. In the short run, fertility and childcare activities often lead female workers to (temporarily or permanently) drop out of the labor force, interrupting the accumulation of their work experience and human capital. Even for females that remain in the labor force, the need for childcare may force them to switch to less demanding, more flexible, part-time but lower-paid jobs in order to reconcile their role as mothers in the families. Either way, the labor income of childbearing females drops in the short run, as the dashed line shows in Figure 2. Furthermore, the short-run adverse shock to human capital accumulation generates a downward shift in the wage growth path, ensuring that the wage penalty persists in the long run.

Figure 2: Intergenerational childcare support, fertility-induced breaks in career, and wage penalty (inspired by Yu and Xie 2014)



When grandparental childcare support is feasible, young females' burden of childbearing is then shared with the older generation. This reduces childbearing-induced breaks in young females' career paths as well as interruptions to their work experience accumulation and human capital, alleviating the adverse shock to their labor income flow. Such a mechanism is captured by the solid curve in Figure 2: Grandparental support helps reduce or even eliminate females' absence

from the labor force, hence the depreciation in their human capital caused by the absence; this increases their feasible working hours, hence their labor supply intensity, and raises their wage rate as well as labor income, reducing the childbearing-induced wage penalty both in the short run and in the long run.

3 Model specification

The stylized facts in section 2.2 show that childbearing females with grandparental childcare support tend to be more likely to stay in the labor force and receive higher labor income, compared with those without support. In this section, we construct econometric models to further investigate how grandparental support affects females' employment and labor income.

3.1 Model explaining employment

First, we establish an econometric model to explain female employment. As the dependent variable $WORK_i$ is a binary variable, following the specification from Dimova and Wolff (2011) and Aassve et al. (2012), we take a Probit model to analyze the impact of grandparental support on female employment:

$$Pr(WORK_i|GPC_i, X_i) = \Phi(\alpha + \tau GPC_i + X_i\beta)$$

in which $WORK_i$ denotes whether the female i is employed at the time of survey, independent variable GPC_i denotes whether she receives support from the older generation, and vector X_i contains other control variables. $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

Following the literature (Dimova and Wolff 2011, Aassve et al. 2012, Posadas and Vidal-Fernández 2013, Arpino et al. 2014), vector X_i in our model contains three types of control variables:

(1) Characteristics of the female i, including age (AGE_i), squared age (AGE_i^2), level of education (or, whether she has at least a college degree, EDU_i), and urban residency ($URBAN_i$);

(2) Characteristics of her children, including the number of children aged 2 or younger ($NCHILD_i$), the age of the youngest child (AGE_CHILD_i), and

(3) Household characteristics, including the logarithm of the household's total net income, excluding the female's (Y_FAMILY_i).

3.2 Model explaining labor income

Although current literature mostly focuses on the effects of grandparental childcare on females' employment choice, our data allows us to further investigate the impact on childbearing females' labor income. Our labor income model is an augmented version of Mincer earnings function (Mincer 1974), with including the variable for grandparental childcare (GPC_i):

$$\ln Y_i = \begin{cases} \alpha + \gamma_1 AGE_i + \gamma_2 AGE_i^2 + \gamma_3 EDU\_YEAR_i + \tau GPC_i + X_i\beta + \epsilon_i & if \quad \ln Y_i > 0 \\ 0 & otherwise \end{cases}$$

in which $\ln Y_i$ is the logarithm of female i's annual income, AGE_i, AGE_i^2 and EDU_YEAR_i are her age, squared age and years of schooling, respectively.

In the classical Mincer model, the logarithm of wage income is a function of work experience, squared work experience, and years of schooling; while without having information on a female's work experience in our data, we use her age instead. In previous studies, the work experience variable has been indirectly measured in different ways: For example, to take the difference between current age and the age when the first job starts after education (such as Xie and Hannum 1996), or, to take the age minus years of schooling minus 6 (such as Luo 2007) — assuming that everyone in the sample starts primary school at 6 and joins the labor force right after education. Both measurements assume no break in one's career prior to childbearing, which is often not the case in reality and leads to measurement errors. Without being able to get a precise measurement of work experience either, we use age as a proxy instead in the model, based on the regularity that has been documented by Mincer (1974) such that labor income fits well a quadratic function of age. Similar approach has been also adopted by Zhao (2006) as well as Yin and Gan (2010).

Since labor income is usually only observable when an individual's income is positive, the variable Y_i is truncated so that applying linear models such as OLS would lead to a bias in estimation. In principle, labor income is an outcome of employment selection: $Y_i = 0$ when the individual is unemployed and $Y_i > 0$ otherwise, so theoretically the Heckman selection model would be a good candidate for analyzing such questions featured by selections. However, in our

sample, employment and labor income are not perfectly matched: Some females are self-employed in agricultural or non-agricultural sectors and receive negative net income (their $\ln Y_i$ is thus set to 0) for the year of the survey, so that a Heckman selection model would mistakenly interpret these individuals as unemployed. Besides that, in our sample, being employed does not necessarily mean receiving positive income for the year of the survey, either — for instance, some employed individuals receive zero income during their maternity leave, and for the year of survey some unemployed still receive positive income before they drop out of the labor force, etc.; as a result, the selection mechanism of the Heckman model does not work for these observations, either. For these reasons, we use a Tobit model instead for labor income analysis, since in a Tobit model, whether one's labor income is zero or not does not need to correlate with whether she is employed or not.

In addition, for comparison we also investigate the impact of grandparental childcare on employment and labor income for males, or, the spouses of the females in the same families, using the same model setup.

3.3 Exploring heterogeneities

Since the stylized facts in the previous section suggest micro-level heterogeneities in the impacts of grandparental childcare on employment and labor income, we further use models with interacting terms to quantify such heterogeneities. First, it has been seen that heterogeneities may relate to children's age, therefore, in the baseline employment and labor income models we add the interaction between grandparental childcare GPC_i and the age of the youngest child AGE_CHILD_i. If the coefficient of this interaction term in the estimated model is significantly different from zero, we are able to say that the impact of grandparental childcare on female employment and labor income varies with the age of smallest child.

Second, we take into account the heterogeneities with respect to childbearing females' residential areas, the underlying fact that females from urban and rural areas may take very different types of jobs. Compared with females from rural areas, urban females are more likely to be engaged in non-agricultural jobs and their working hours are less flexible; therefore, the conflicts between job and childcare may be more severe so that grandparental support may have a stronger positive impact on female employment and labor income in urban areas. To examine this issue, we augment the baseline employment and income models with the interaction between

grandparental support GPC_i and urban residency $URBAN_i$. The coefficient of this interaction term reflects the heterogeneities coming from urban / rural residency.

Finally, childcare support from the older generation may affect females with different education levels differently. Females with higher education are more likely to take full-time jobs, or, jobs with less flexibility in working hours. For them, grandparental childcare may help ease the conflict between job and childcare, generating a higher positive impact on employment and labor income. Previous studies, such as Anderson et al. (2003), Yu and Xie (2014), indeed find that a break in career leads to higher human capital depreciation for females with higher education, therefore, they suffer more from a childbearing-induced wage penalty. For this reason, we also introduce the interaction between grandparental support GPC_i and female's education level EDU_i in the augmented models. As the baseline income model already includes years of schooling (EDU_YEAR_i) as a control variable, to avoid multicollinearity, here we only take the interaction term using grandparental support and "yeas of schooling" instead of "level of education" (EDU_i).

3.4 Endogeneity and instrumental variable

A potential problem in our model specification is endogeneity. An endogeneity problem may arise from reverse causality, i.e., that the choice of young females to stay in the labor force may force grandparents to provide childcare support; if this is the case, the impact of grandparental childcare on female employment may be overestimated (for example, Posadas and Vidal-Fernández 2013). Endogeneity may also come from missing variables, for example, Aassve et al. (2012) suggest that unobservable household preferences, – such as the work preferences of young and old generations, the strength of family bonds and the household's preference for childcare labor input, – may simultaneously affect grandparents' decision regarding providing childcare *and* childbearing females' decisions regarding employment.

We use an instrumental variable to correct the potential bias caused by endogeneity. In the literature instrumental variables for grandparental childcare are often constructed from family structure such as information on the childbearing female's siblings, characteristics of grandparents, etc. In this paper, we use "paternal grandmother still living" (PGM_i) as an instrumental variable. Although in literature, instrumental variables have been constructed using the information of both paternal and maternal grandparents (Arpino et al. 2014), or the

information of maternal grandmothers only (Aassve et al. 2012, Posadas and Vidal-Fernández 2013), here we choose paternal grandmothers' characteristics instead for the following reasons: First, CFPS 2014 shows that paternal grandparents are much more involved in day time childcare than maternal grandparents; second, paternal grandmothers in general claim a far larger share of childcare than paternal grandfathers, especially when the grandchildren are young. Therefore, the variable "paternal grandmother still living" has a strong correlation with the instrumented variable "grandparental childcare"; on the other hand, whether paternal grandmother is alive or not is normally orthogonal and exogenous to female's employment status and income (Chen 2012) so that the reverse causality from dependent variables is avoided, which justifies PGM_i as a valid instrumental variable.

However, the exogeneity of the instrument may be weakened by the fact that it is likely to be correlated with the other variables that also affect females' employment and income. For example, whether the paternal grandmother is alive or not is correlated with the female's age, an independent variable in our models, — the higher the age is, the less likely the paternal grandmother is still living, — and the age is also correlated with her employment status and income. Furthermore, "paternal grandmother still living" may not always positively correlate with grandparental childcare; for instance, a paternal grandmother in poor health may not contribute to grandparental childcare and on the contrary require care input from the childbearing female, generating an adverse effect on the female's employment and income. To take such concerns into account, in the robustness check we include households' input on old-age support in the regression and see whether the upward labor transfer towards older generation affects the validity of the instrumental variable.

4 Results

4.1 Results from the employment model

The results from the employment model are presented in Table 4. Columns (1) to (3) are the estimates from a series of Probit regressions with various control variables. It can be seen that grandparental childcare significantly boosts the likelihood of working of childbearing females. Columns (4) to (6) report the estimates from the employment model using an IV-Probit estimator, with the instrumental variable "paternal grandmother still living". Column (4) shows the results

using a maximum likelihood estimation, and the results reach qualitatively similar conclusion as an OLS estimation. Columns (5) and (6) report the first- and second-stage results from a Newey two-step estimator of the IV-Probit regression, and the results confirm that grandparental childcare significantly increases the likelihood of working of childbearing females. Additionally, a Wald endogeneity test suggests that the null hypothesis on the non-existence of endogeneities cannot be rejected, so that it is proper to adopt the Probit model, rather than the IV-Probit model.

Table 4: Results from the employment model

	Employment: Female						Male
	Probit			IV-Probit			Probit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Grandparental childcare	0.485*** (0.050)	0.629*** (0.054)	0.674*** (0.055)	0.824** (0.307)		0.826** (0.322)	-0.044 (0.068)
Number of children below age 2		-0.342*** (0.066)	-0.356*** (0.067)	-0.353*** (0.067)	-0.002 (0.022)	-0.354*** (0.064)	-0.082 (0.084)
Age of the youngest child		0.089*** (0.011)	0.052*** (0.013)	0.053*** (0.013)	-0.003 (0.004)	0.053*** (0.012)	0.010 (0.016)
Age of the female / male			0.034*** (0.005)	0.036*** (0.006)	-0.009*** (0.002)	0.036*** (0.006)	4.4e-4 (0.007)
Education level of the female / male			0.559*** (0.084)	0.548*** (0.087)	0.076*** (0.025)	0.550*** (0.088)	0.240* (0.114)
Urban residency (Female / male)			-0.252*** (0.048)	-0.251*** (0.048)	0.015 (0.015)	-0.252*** (0.048)	-0.112* (0.068)
Log household's total income (excluding female's / male's)			-0.066*** (0.015)	-0.067** (0.016)	0.001 (0.004)	-0.067*** (0.015)	-0.003 (0.009)
Paternal grandmother still living					0.154*** (0.015)		
Constant	0.460*** (0.025)	0.140* (0.075)	-0.046 (0.221)	-0.145 (0.301)	0.500*** (0.065)	-0.145 (0.298)	1.619*** (0.241)
Significance test	93.96 (P=0.0000)	390.09 (P=0.0000)	503.20 (P=0.0000)	425.66 (P=0.0000)	34.63 (P=0.0000)	413.70 (P=0.0000)	11.76 (P=0.1087)
Log likelihood	-2187.52	-1997.97	-1930.76	-4272.73			-871.30
Wald endogeneity test				0.24 (P=0.6264)		0.23 (P=0.6328)	
Pseudo R^2	0.0218	0.1066	0.1366				0.0063

Number of obs.	3,795	3,795	3,795	3,795	3,795	3,795	4,089

Note: (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) as for the models' significance tests, F-statistic is reported for the Tobit regression as well as the first-step regression in Newey's two-step estimator, and Wald χ^2-statistic is reported for the other regressions; (4) Wald endogeneity test is based on χ^2-statistic.

As Probit is a non-linear model, we cannot implement a test for weak instrumental variable directly using an F-statistic from the first-stage regression; instead, we can test weak instrumental variables in an IV-2SLS regression, i.e., the linear probability model – see Table 7 in section 5. As the IV-Probit model has the same first-stage estimation as the linear probability model, the F-statistic of IV-linear probability regression can also test for weak instrumental variables in the IV-Probit regression. The F-statistic in Table 7 implies that the weak instrumental variable test cannot reject the null hypothesis that the instrumental variable is not weak.

However, the results of both Probit and IV-Probit regressions show that grandparental childcare has no significant effect on the likelihood of working of males in the same families. Effects of grandparental childcare on childbearing couples' labor force participation reported in Table 4 are consistent with the stylized facts in Table 1, which provides strong support in our conjecture that grandparental childcare lowers the opportunity cost of young females and increases their labor force participation rate.

In order to quantify the impact of grandparental childcare on females' employment, we take a Logit model and regress females' employment status on grandparental childcare. This leads to an odds ratio of being employed for females with grandparental childcare that is 3.15 times as much as for those without such support.[10]

The estimates for the coefficients of the other control variables also have strong economic implications. First, for a childbearing female, the more young children (below the age of 2) she has, the less likely she has a job – this is consistent with previous studies, such that females with more young children have to invest more time in childcare, and this leads to a fall in their labor

[10] In the Logit model, let the likelihood of being employed for a female i with grandparental childcare be $p = Pr(WORK_i = 1|GPC_i = 1, X_i)$, then $P = \frac{p}{1-p}$ denotes the odds that females with grandparental childcare are employed; let the likelihood of being employed for a female j without grandparental childcare be $q = Pr(WORK_j = 1|GPC_j = 0, X_j)$, then $Q = \frac{q}{1-q}$ denotes the odds that females without grandparental childcare are employed. The odds ratio for females with grandparental childcare versus females without grandparental childcare is defined as $\frac{P}{Q}$.

supply; second, the female's likelihood of being employed increases with the age of her youngest child. This is because older children need less care, so that the crowd-out effect of childcare on female's labor supply declines; third, the likelihood of working is higher for females with better education, as better education usually implies better potential for higher labor income, hence the opportunity cost of exiting the labor force rises; fourth, females living in urban areas (most likely, taking non-agricultural jobs) have lower likelihood of working, as the crowd-out effect of childcare on non-agricultural labor (probably less flexible) is possibly stronger than on agricultural labor (probably more flexible); finally, the higher the household's total income (excluding the female's) is, the lower probability she is employed, implying that the rest of the household's income has a substitution effect on the female's labor force participation, as documented by Ashenfelter and Heckman (1974), Yao and Tan (2005), Blau and Kahn (2007), and Posadas and Vidal-Fernández (2013). When a household's income is high, the young female faces low pressure to work and support the household's expenditure, so that she tends to be more likely to quit the labor force after giving birth and spend full time on childcare.

Furthermore, as a comparison, we also investigate the impact of grandparental childcare on males' employment using a Tobit model. The results are reported in column (7) of Table 4. Quite different from the case of females, males' likelihood of working is affected by neither grandparental childcare nor age / number of young children. This clearly shows the division of labor within typical couples: Females generally take much more responsibility for housework and childcare, especially; they are almost indispensable for looking after young children; instead, males contribute more to households' income, and their jobs are often less flexible regarding working hours to reconcile with childcare. As a result, childcare and intra-family downward labor transfer do not usually affect males' labor supply.

4.2 Results from the income model

We further explore the impact of grandparental childcare on young females' labor income through the income model, and the results are reported in Table 5. Columns (1) and (2) are based on Tobit regressions, with different sets of control variables. In our sample, grandparental childcare significantly improves females' labor income. Column (3) to (5) are based on IV-Tobit estimators: Column (3) are results from a maximum likelihood estimation, column (4) reports the first-step results from a Newey two-step regression, and column (5) reports the results from the

second step. In column (4), the estimated coefficient for the instrumental variable "paternal grandmother still living" is significant at the 1% level, with the correct sign, and free of the weak instrument problem. Wald endogeneity tests for IV-Tobit models cannot reject the null hypothesis that the explanatory variables are exogenous; therefore, Tobit is preferred to the IV-Tobit estimator.

Table 5: Results from the income model

	Log annual income: Female					Male
	Tobit		IV-Tobit			Tobit
	(1)	(2)	(3)	(4)	(5)	(6)
Grandparental childcare	2.560*** (0.246)	3.033*** (0.247)	4.145** (1.623)		4.145** (1.629)	-0.255 (0.200)
Years of schooling, female / male		0.181*** (0.029)	0.174*** (0.031)	0.006*** (0.002)	0.174*** (0.031)	0.092*** (0.024)
Age of the female / male		0.702*** (0.180)	0.653*** (0.194)	0.039*** (0.010)	0.653*** (0.193)	0.289*** (0.113)
Squared age of the female / male, divided by 100		-0.829*** (0.262)	-0.736** (0.294)	-0.070*** (0.016)	-0.736** (0.297)	-0.403*** (0.153)
Number of children below age 2		-1.279*** (0.359)	-1.291*** (0.360)	0.020 (0.020)	-1.291*** (0.343)	0.004 (0.263)
Age of the youngest child		0.182*** (0.058)	0.186*** (0.058)	-0.003 (0.004)	0.186*** (0.059)	0.011 (0.044)
Urban residency (female / male)		-0.265 (0.255)	-0.253 (0.256)	0.003 (0.015)	-0.253 (0.254)	-0.110 (0.197)
Paternal grandmother still living				0.151*** (0.015)		
Constant	2.444*** (0.180)	-13.135*** (3.003)	-12.856*** (3.033)	-0.311* (0.171)	-12.856*** (2.987)	0.950 (2.032)
Significance test for the model	108.40 (P=0.0000)	55.29 (P=0.0000)	241.76 (P=0.0000)	43.86 (P=0.0000)	244.53 (P=0.0000)	3.72 (P=0.0000)
Log likelihood	-9825.00	-9688.43	-12273.29			-10849.47
Wald endogeneity test			0.48 (P=0.4887)		0.48 (P=0.4883)	
Pseudo R^2	0.0050	0.0189				0.0013
Number of obs.	4,192	4,192	4,192	4,192	4,192	4,089

Note: (1) *** / ** / * denotes the result is significant at the 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) as for the models' significance tests, F-statistic is reported for the Tobit regression as well as the first-step regression in Newey's two-step estimator, and a Wald χ^2-statistic is reported for the other regressions; (4) Wald endogeneity test is based on χ^2-statistic.

In Table 5, the results from various specifications consistently show that grandparental care significantly increases young females' labor income. Such positive effect may come through two channels: First, through labor supply intensity. It has been well established that childcare reduces females' labor supply intensity, for example, Zhang (2011) finds that a higher number of children leads to lower working hours for both urban and rural females. Therefore, grandparental care frees up part of females' labor input to childcare, increases their labor supply intensity, hence leading to higher labor income.

Second, the effect may also come through changes in females' wage rate, or, productivity. To reconcile with childcare, childbearing females may switch to more child-friendly jobs that are more flexible, less labor-intensive, but less productive and hence with lower pay (Yu and Xie 2014). Polachek (1981) finds that after a childbearing-induced break in the careers, females become less likely to be involved in typically well-paid professional / managerial jobs and more likely to take jobs that demand low skills and pay lower wages. However, if females are able to receive support for childcare from the older generation, their career paths or paths of human capital accumulation will be less interrupted, and they are less likely forced to switch to less productive jobs. This again alleviates the childbearing-related adverse shocks to wages and reduces the wage penalty.

Throughout the results of the Tobit regression, as reported in column (2) of Table 5, all the signs of estimated coefficients are the same as expected. Females' labor income increases significantly with years of schooling, while the signs for coefficients of females' age and squared age are positive and negative respectively, which are in line with the classical Mincer equation. The more young children (below the age of 2) she has, the lower the female's income, and this is consistent with the literature that the childbearing-related wage penalty increases with the number of young children. Females' income increases with the age of the youngest children, because older children become less demanding for care so that more of females' labor supply is freed up; this is also because many females of older children have already returned to the labor force for a while, therefore, the post-birth human capital accumulation further contributes to the rise in labor income.

Similar as in the employment model, for comparison we also explore the impact of grandparental care on males' labor income, and the results are reported in column 6, Table 5. Although

coefficients of years of schooling, age, squared age are significant, – consistent with the Mincer equation, – grandparental care and number / age of children have no significant impact on males' labor income.

4.3 Heterogeneities in the impacts of grandparental childcare

The descriptive statistics in Tables 1 and 2, as well as the empirical evidence from Tables 4 and 5 suggest that the impact of grandparental care on female employment and labor income may vary vis-à-vis children's age, young females' residential area and education level, etc. In this section, we use interaction terms to further explore such heterogeneities, and the results are reported in Table 6. Columns (1) to (3) contain the results from the employment model, and (4) to (6) contain the results from the income model. As we have shown that Wald endogeneity tests do not reject the null hypothesis that control variables are exogenous in the previous section, here we implement all regressions without using instrumental variables.

From columns (1) and (4), one can see that the coefficient for the interaction term (grandparental childcare * age of the youngest child), i.e., $GPC_i * AGE\_CHILD_i$, is significant and negative, implying that the positive effects of grandparental care on females' employment and income are stronger for females living with younger children. This is due to the fact that younger children require more care, the crowd-out effect of childcare on females' labor supply is stronger, thus grandparental care has higher impact on improving females' employment and income by alleviating the crowd-out.

Columns (2) and (5) suggest that grandparental care has a more positive impact on employment and income for young females living in urban areas, compared with those in rural areas. This may be explained by the difference in their jobs: In our sample, among 1,770 employed females from rural areas, 983 or 55.54% of them are involved in agricultural production; in contrast, among 1,245 employed females from urban areas, 1,044 or 83.86% are involved in non-agricultural activities. Compared with agricultural jobs, non-agricultural jobs tend to be more inflexible in working hours, therefore, urban females are more likely to face conflicts between childcare and working; such conflicts are much eased if grandparental support is available, so that grandparental childcare has more positive impacts on employment and income for urban females.

It can also be seen from column (4) that grandparental childcare has a more positive impact for better-educated females. In our sample, only 38.84% employed females without higher education (college degree or above) are involved in non-agricultural jobs, while this share becomes 85.84% for employed females with higher education. This implies that females with higher education are more likely to take inflexible jobs; therefore, grandparental care reduces their childcare burden and encourages them to remain in the same jobs after giving birth. Furthermore, a childbearing-induced break in the career path has a higher adverse impact on human capital accumulation, or, income flow, for females with higher education; for them, support from the older generation obviously reduces the interruption in their careers, hence the negative impact on their post-child birth employment and life-time income flow. Such supportive effect on the career path is therefore higher from females with higher education, compared with those without.

Such results have important implications for the phasing-in policy of raising the minimum retirement age in China. Given that childcare purchased from the market is hardly yet a substitute, childcare burdens are largely shouldered by parents and grandparents. Postponing retirement is likely to reduce the supply of grandparental care, increase the opportunity cost of employment for young females and discourage them from participating in the labor force. Such a negative impact on young females' employment and labor income is more severe for those with higher education, living in urban areas with younger children. And this is worsened by the fact that the retirement policy affects urban grandparents with non-agricultural jobs the most.

Table 6: Results from regressions with interaction terms

	Employment of the females			Log annual income of the females		
	Probit			Tobit		
	(1)	(2)	(3)	(4)	(5)	(6)
Grandparental childcare	0.910***	0.549***	0.638***	4.263***	2.539***	1.936***
	(0.090)	(0.070)	(0.058)	(0.457)	(0.325)	(0.533)
Grandparental childcare* Age of the youngest child	-0.061***			-0.271***		
	(0.019)			(0.078)		
Grandparental childcare* Urban residency (female)		0.298***			1.128**	
		(0.110)			(0.487)	
Grandparental childcare* Female's education level			0.422**			0.129**
			(0.202)			(0.056)
Number of children below age 2	-0.384***	-0.356***	-0.361***	-1.420***	-1.286***	-1.323**
	(0.068)	(0.067)	(0.067)	(0.363)	(0.358)	(0.359)
Age of the youngest child	0.063***	0.052***	0.053***	0.244***	0.179***	0.180***
	(0.013)	(0.013)	(0.013)	(0.061)	(0.058)	(0.058)
Age of the female	0.033***	0.034***	0.034***	0.739***	0.708***	0.707***
	(0.005)	(0.005)	(0.005)	(0.180)	(0.179)	(0.179)
Squared age of the female, divided by 100				-0.890***	-0.837***	-0.844***
				(0.263)	(0.262)	(0.262)
Female's education level	0.548***	0.560***	0.443***			
	(0.084)	(0.084)	(0.096)			
Female's years of schooling				0.178***	0.182***	0.139***
				(0.029)	(0.029)	(0.035)
Urban residency (female)	-0.246***	-0.330***	-0.249***	-0.240	-0.645**	-0.256
	(0.048)	(0.056)	(0.048)	(0.255)	(0.313)	(0.255)

Log household's total income (excluding the female's)	-0.066*** (0.015)	-0.065*** (0.016)	-0.067*** (0.015)			
Constant	-0.055 (0.221)	-0.029 (0.222)	-0.009 (0.222)	-13.922*** (3.020)	-13.047*** (2.997)	-12.790*** (3.002)
Significance test for the model	556.56 (P=0.0000)	505.04 (P=0.0000)	495.30 (P=0.0000)	47.74 (P=0.0000)	49.26 (P=0.0000)	49.22 (P=0.0000)
Log likelihood	-1924.98	-1926.84	-1928.28	-9682.81	-9685.85	-9685.89
Pseudo R^2	0.1392	0.1384	0.1377	0.0194	0.0191	0.0191
Number of obs.	3,795	3,795	3,795	4,192	4,192	4,192

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

5 Robustness check

The robustness of our empirical results may be affected by a variety of factors such as model specification, selection of instrumental variables, and within-sample heterogeneity. In this section, we conduct a series of robustness checks to ensure that our results are robust to different settings.

First, we check the robustness of results under different model specifications. Columns (1) to (3) in Table 7 report the results from the OLS estimator, using various sets of control variables. The results are qualitatively the same as those in Table 4, that grandparental care significantly increases young females' likelihood of working. Columns (4) and (5) report the results using the IV-2SLS estimator; the first-step results are listed in column (4) and the second-step results are listed in column (5). It can be seen that the estimated coefficient for the instrumental variable "paternal grandmother still living" is significant and positive, as expected, and a Cragg-Donald Wald F-statistic shows that the weak instrumental variable hypothesis is rejected. Columns (6) to (8) report the results from the OLS regressions with interaction terms; again, the impact of grandparental care on female employment is heterogeneous: The impact increases when the age of the youngest child is lower, and it is higher for more educated, urban females than less educated, rural females. The results are fully consistent with those from the Probit model.

Second, in this paper we focus on the downward labor transfer within families, i.e., the transfer from the older to the younger generation. However, upward labor transfer may take place in some of the families within the sample, too, i.e., the younger generation may need to take care of the ailing older generation. The upward transfer crowds out young couples' labor supply and generates negative impacts on their labor income, thus weakens the exogeneity of our instrumental variable. In literature, such a problem is often solved by using the health information on the older generation, for example, in the robustness check Arpino et al. (2014) exclude those families containing at least one grandparent who suffers from chronicle disease and may need upward labor transfer. However, in CFPS, grandparents are not necessarily counted as family members as many of them do not live at the same address as their children and grandchildren; therefore, health information for many of them is missing. Here we use another variable from CFPS to control for the upward labor transfer – a dummy variable "substantial upward transfer", i.e., whether the family members provide substantial labor to look after the

older generation. The results are reported in Table 8: Columns (1) and (2) for the employment model, and columns (5) and (6) for the income model. After taking into account the upward labor transfer, the results are still quite close to those in Table 4 and 5. Furthermore, the estimated coefficients for the additional control variable are not significant, implying that downward labor transfer is the major pattern of intra-family labor transfer in our sample and indeed has positive impact on female employment.

Finally, the childbearing females' marriage status may distort the results, too. It has been found that single mothers (including non-married, divorced, widowed) behave differently in the labor market (Arpino et al. 2014), and the likelihood that they receive grandparental support is much lower, compared with the other females in the sample. To disentangle such within-sample heterogeneity, we exclude all single mothers and redo the regressions. The results are listed in columns (3) and (4), (7) and (8) in Table 8, and they are still consistent with the results in Tables 4 and 5. This confirms that our results are robust to the within-sample heterogeneity caused by single mothers.

Table 7: Robustness check (1) Linear model

	Employment of the females							
	OLS			IV-2SLS		OLS with interaction terms		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grandparental childcare	0.151*** (0.014)	0.189*** (0.014)	0.201*** (0.014)		0.213** (0.090)	0.323*** (0.027)	0.167*** (0.018)	0.193*** (0.015)
Grandparental childcare* Age of the youngest child						-0.028*** (0.005)		
Grandparental childcare* Urban residency (female)							0.081*** (0.028)	
Grandparental childcare* Female's education level								0.067* (0.038)
Number of children below the age of 2		-0.134*** (0.021)	-0.135*** (0.021)	-0.0015 (0.0218)	-0.135*** (0.021)	-0.147*** (0.021)	-0.135*** (0.021)	-0.136*** (0.021)
Age of the youngest child		0.026*** (0.003)	0.014*** (0.003)	-0.0033 (0.0035)	0.014*** (0.003)	0.020*** (0.004)	0.014*** (0.003)	0.014*** (0.003)
Age of the female			0.011*** (0.001)	-0.0089*** (0.0014)	0.011*** (0.002)	0.010*** (0.001)	0.011*** (0.001)	0.010*** (0.001)
Female's education level			0.160*** (0.021)	0.0763*** (0.0263)	0.159*** (0.021)	0.150*** (0.021)	0.159*** (0.021)	0.135*** (0.028)
Urban residency (female)			-0.073*** (0.014)	0.0152 (0.0153)	-0.073*** (0.014)	-0.069*** (0.014)	-0.098*** (0.018)	-0.073*** (0.014)
Log household's total income (excluding the female's)			-0.017*** (0.003)	0.0007 (0.004)	0.017*** (0.003)	-0.017*** (0.003)	-0.016*** (0.003)	-0.017** (0.003)
Paternal grandmother still living				0.1543*** (0.0154)				
Constant	0.677*** (0.009)	0.583*** (0.023)	0.484*** (0.056)	0.500*** (0.0642)	0.476*** (0.083)	0.469*** (0.056)	0.492*** (0.056)	0.492*** (0.056)
Significance test for the model	110.73 (P=0.0000)	167.65 (P=0.0000)	100.80 (P=0.0000)	43.73 (P=0.0000)	533.37 (P=0.0000)	92.15 (P=0.0000)	89.05 (P=0.0000)	88.95 (P=0.0000)
R^2	0.0243	0.1243	0.1559		0.1557	0.1633	0.1576	0.1563
Weak IV test				99.10				
Number of obs.	3,795	3,795	3,795	3,795	3,795	3,795	3,795	3,795

Note: (1) *** / ** / * denotes the result is significant at the 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) as for the models' significance tests, F-statistic is reported for OLS estimator, and Wald χ^2-statistic is reported for IV-2SLS estimator; (4) weak instrumental variable test in IV-2SLS regression is based on Cragg-Donald Wald F-statistic.

Table 8: Robustness check (2)

	Employment of the females				Log annual income of the females			
	With upward transfer		Excluding single mothers		With upward transfer		Excluding single mothers	
	Probit	IV-Probit	Probit	IV-Probit	Tobit	IV-Tobit	Tobit	IV-Tobit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Grandparental childcare	0.674*** (0.055)	0.816** (0.307)	0.668*** (0.056)	0.886*** (0.292)	3.033*** (0.247)	4.103** (1.621)	3.063*** (0.251)	4.199*** (1.565)
Number of children below the age of 2	-0.355*** (0.067)	-0.353*** (0.067)	-0.352*** (0.067)	-0.348*** (0.068)	-1.279*** (0.358)	-1.291*** (0.359)	-1.266*** (0.362)	-1.278*** (0.363)
Age of the youngest child	0.053*** (0.013)	0.053*** (0.013)	0.052*** (0.013)	0.053*** (0.013)	0.185*** (0.058)	0.190*** (0.058)	0.195*** (0.058)	0.201*** (0.059)
Age of the female	0.034*** (0.005)	0.036*** (0.006)	0.035*** (0.005)	0.037*** (0.006)	0.695*** (0.179)	0.647*** (0.193)	0.638*** (0.181)	0.587*** (0.195)
Squared age of the female, divided by 100					-0.819*** (0.262)	-0.730** (0.294)	-0.735*** (0.264)	-0.640** (0.295)
Female's education level	0.556*** (0.084)	0.546*** (0.087)	0.553*** (0.084)	0.537*** (0.088)				
Female's years of schooling					0.183*** (0.029)	0.176*** (0.031)	0.186*** (0.030)	0.178*** (0.032)
Urban residency (female)	-0.252*** (0.048)	-0.251*** (0.048)	-0.247*** (0.049)	-0.247*** (0.049)	-0.281 (0.255)	-0.270 (0.256)	-0.298 (0.257)	-0.287 (0.258)
Log household's total income (excluding the female's)	-0.066*** (0.015)	-0.066*** (0.015)	-0.067*** (0.016)	-0.068*** (0.016)				
Upward transfer for old-age care	0.061 (0.079)	0.061 (0.079)			0.636* (0.381)	0.637* (0.381)		
Constant	-0.056 (0.221)	-0.149 (0.301)	-0.049 (0.227)	-0.192 (0.299)	-13.085*** (2.997)	-12.817*** (3.027)	-12.161*** (3.029)	-11.86*** (3.061)
Significance test for the model	505.25 (P=0.0000)	425.75 (P=0.0000)	488.74 (P=0.0000)	425.41 (P=0.0000)	48.80 (P=0.0000)	244.89 (P=0.0000)	54.57 (P=0.0000)	239.78 (P=0.0000)
Log likelihood	-1930.45	-4272.40	-1905.54	-4184.38	-9687.10	-12271.96	-9486.26	-12001.77
Wald endogeneity test		0.21 (P=0.6448)		0.55 (P=0.4603)		0.44 (P=0.5049)		0.54 (P=0.4631)
Pseudo R^2	0.1368		0.1350		0.0190		0.0190	
Number of obs.	3,795	3,795	3,719	3,719	4,192	4,192	4,109	4,109

Note: (1) *** / ** / * denotes the result is significant at the 1% / 5% / 10% level; (2) values in the parentheses are standard errors, except those specified as P values; (3) as for the models' significance tests, LR χ^2-statistic is reported for Probit and Tobit estimators, and Wald χ^2-statistic is reported for IV-Probit and IV-Tobit estimators; (4) Wald endogeneity test is based on χ^2-statistic.

6 Concluding remarks

There are two distinguishing, albeit seemingly paradoxical features of China's labor market: Females' labor force participation rate is among world's highest, yet the childbearing-induced wage penalty is extremely low. Using data from the 2014 China Family Panel Survey (CFPS), we demonstrate the role of grandparental childcare in explaining this paradox. Through the intra-family downward labor transfer, the older generation shares the burden with childbearing females; this reduces the crowd-out effect of childcare on females' employment, and shortens fertility-induced breaks in females' careers. As a consequence, with such support, young females' human capital accumulation is less interrupted by child birth and childcare, which implies smaller adverse shocks to their labor income flows. The effects of grandparental childcare are especially stronger for better educated, urban females with younger children.

Our results reveal an unpleasant hidden cost of China's recent attempt to gradually raise the minimum retirement age. Such a policy is likely to reduce elderly workers' downward labor transfer towards their grandchildren. Without being able to obtain perfect substitutes from the market for childcare, young females will then have to allocate more of their labor to childcare; this may crowd out their labor supply in the labor market and leave a bigger interruption in their career paths, hence a bigger reduction in their future labor income flows. Furthermore, such negative impact is higher for well-educated, urban females. The social welfare loss from young females' life-time careers may well outweigh the social welfare gain from extending elderly workers' labor supply by just a few years.

Our results call for more social protection policies, along with the phasing-in of this retirement policy. A functional market offering sufficient qualified childcare services is needed, which provides a substitute to falling intra-family downward labor transfer caused by elderly workers' postponed retirements. Policies against discrimination against childbearing females, as well as policies that provide insurance to birth-related career breaks and increase the involvement of males in childcare, may also alleviate the negative shocks to females' human capital accumulation and lifetime income flows.

References

Aassve, A., Arpino, B. and Goisis, A. (2012), "Grandparenting and Mothers' Labor Force Participation: A Comparative Analysis Using the Generations and Gender Survey", *Demographic Research* 27(3), 53-84.

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.

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.

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.

Ashenfelter, O. and Heckman, J. (1974), "The Estimation of Income and Substitution Effects in a Model of Family Labor Supply", *Econometrica* 42(1), 73-85.

Blau, F. D. and Kahn, L. M. (2007), "Changes in the Labor Supply Behavior of Married Women: 1980-2000", *Journal of Labor Economics* 25(3), 393-438.

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

Budig, M. J. and England, P. (2001),"The Wage Penalty for Motherhood", *American Sociological Review* 66(2), 204-225.

Chen, L., Fan, H. L., Zhao, N. and Chu, L. L. (2016), "The Impact of Informal Care on Employment for Women in China", *Economic Research Journal* 2016(3), 176-189 (in Chinese).

Chen, Y. S. (2012), "Logic, Imagination and Interpretation: The Application of Instrumental Variables for Causal Inference in the Social Sciences", *Sociological Studies* 2012(6), 192-216 (in Chinese).

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

Connelly, R. (1992), "The Effect of Childcare Costs on Married Women's Labor Force Participation", *Review of Economics and Statistics* 74(1), 83-90.

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.

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.

Dong, X. and An, X. (2014), "Gender Patterns and Value of Unpaid Care Work: Findings from China's First Large-Scale Time Use Survey", *Review of Income and Wealth* 61(3), 540-560.

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.

Feng, J. and Han X. (2017), "The Impact of Retirement Age Policy on Household Care Provision and Labor Participation in China", *Journal of World Economy* 2017(6), 145-166 (in Chinese).

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.

García-Morán, E. and Kuehn, Z. (2012), "With Strings Attached: Grandparent-provided Childcare, Fertility, and Female Labor Market Outcomes", *EUI Working Papers* MWP 2012/29.

Gruber, J., Milligan, K. and Wise, D. (2010), "Introduction and Summary", in: Gruber, J. and Wise, D. (eds), *Social Security Programs and Retirement around the World: The Relationship to Youth Employment*. Chicago: The University of Chicago Press.

Huang, F. (2012), "On the Relationship between Family Care and Female Employment in Urban China from the Perspective of Population Aging", *Journal of Finance and Economics* 2012(9), 16-26 (in Chinese).

Jiang, C. and Zhao, X. J. (2009), "On the Opportunity Cost of Elder Care in China", *Management World* 2009(10), 80-87 (in Chinese).

Liu, L., Dong, X. Y., Chen, G. and Zheng, X. Y. (2010), "The Impact of Parent-caring on Married Women's Labor Time in Rural China", *World Economic Papers* 2010 (5), 1-15 (in Chinese).

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.

Ma, Y. and Li L. (2014), "The Impact of the Care-giving for Elderly Parents on the Employment of Young and Middle-aged Urban Married Women", *Population & Economics*, 2014(2), 39-47 (in Chinese).

Mincer, J. (1974), *Schooling, Experience, and Earnings*. New York: Columbia University Press.

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

Polachek, S. W. (1981), "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure", *Review of Economics and Statistics* 63(1), 60-69.

Posadas, J. and Vidal-Fernández, M. (2013), "Grandparents' Childcare and Female Labor Force Participation", *IZA Journal of Labor Policy* 2(1), 2-14.

Shen, K., Zhang, Y. and Yan, P. (2012), "Family Structure and Female Labor Force Participation in China", *Population Research* 2012(5), 15-27 (in Chinese).

Waldfogel, J. (1997), “The Effect of Children on Women's Wages”, *American Sociological Review* 62(2), 209-217.

Xie, Y. and Hannum, E. (1996), “Regional Variation in Earnings Inequality in Reform-Era Urban China”, *American Journal of Sociology* 101(4), 950-992.

Xiong, R. X. and Li, H. W. (2016), “Childcare, Public Service and Chinese Rural Married Women’s Non-Farm Labor Force Participation: Evidence from CFPS Data”, *China Economic Quarterly* 16(1), 393-414 (in Chinese).

Yao, X. G. and Tan, L. (2005), “Family Income and Labor Force Participation of Married Women in Unban China”, *Economic Research Journal* 2005(7), 18-27 (in Chinese).

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

Yin, Z. C. and Gan, L. (2010), “Smoking, Drinking, and Earnings in China”, *Economic Research Journal* 2010(10), 90-100 (in Chinese).

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).

Zhao, L. T. (2006), “Rising Returns to Education in Rural China”, *Social Sciences in China* 2006(3), 98-109 (in Chinese).

Zhou, C. F. (2011), “Childcare, Elderly Care and Non-Farm Employment of Rural Married Women”, *Journal of Agrotechnical Economics* 2013(11), 94-102. (in Chinese)