The Role of Social Norms in Old-age Support: Evidence from China

Please click here for the newest version of my job market paper

Xuezhu Shi *

November 18, 2018

Abstract

The norm of families providing support to the elderly is common in developing countries without comprehensive pension coverage and is important; it is usually gender-specific. This paper studies the inter-generational transmission of this social norm in China, focusing on the same-gender transmission channel. The mechanism behind this transmission is that parents, by their provision of support to their own parents, shape their same-gender children's preference for old-age support. Given that the gender ratio of Chinese children is not random, I use an interaction term of the timing of the ban on sex-selective abortions in China and the gender of the first-born child as the instrumental variable for the gender of the children to alleviate the possible endogeneity. The empirical results, using two Chinese datasets, show that parents with more same-gender children provide more support to their ageing parents than parents with cross-gender ones. The father effect is more significant in rural subsamples, and the mother effect is seen mainly in the urban ones. The urban-rural difference in the results may indicate a normative shift accompanying economic and demographic changes.

Keywords— Old-age support, Intergenerational transfers, Social norms, Indirect reciprocity

*Department of Economics, STICERD, London School of Economics, United Kingdom. Email: x.shi2@lse.ac.uk.
I wish to thank Nava Ashraf, Oriana Bandiera, Joan Costa-i-Font, Hanming Fang, Greg Fischer, Maitreesh Ghatak, Xavier Jaravel, Henrik Kleven, Michael Kremer, Camille Landais, Daniel Reck, Johannes Spinnewijn, Dirk Van de gaer and especially Robin Burgess, Frank Cowell, and Rachael Meager for their comments and suggestions for this paper in its early stages. I want also to thank Pedro Alves, Michelle Azulai, Shiyu Bo, Sarah Clifford, Jiajia Gu, Chao He, Tillman Hoenig, Yatang Lin, Panos Mavrokonstantis, Arthur Seibold, Eddy Tam, Celine Zipfel and other participants in the Public Economics, Development, and STICERD WIP Seminar in LSE. Thanks too to Jipeng Zhang, Jiayi Bao, Elena Gentili and other participants in the RES 2017 annual conference, TADC Conference in LBS, EDP Jamboree workshop, 2018 CMES and AMES.

1 Introduction

Family support provided by adult children acts as a major income source for ageing parents in developing countries. This social norm of providing support to the elderly is traditional and common, especially in China.[1] Usually the norm is gender-specific: sons provide more support than daughters (Lee et al., 1994). It helps to offset possible risks and expected income drops for the elderly in countries with underdeveloped public pension systems and incomplete financial markets. As a large developing country with an estimated share of the elderly population due to reach 25% in 2030, China is feeling the weight on its public finances of sustaining, improving, and complementing its current pension schemes.[2] Family old-age support has served as the complement for the incomplete public pension system in sustaining the welfare of the elderly in China. A major topic of debate here, with possibly unsustainable pay-as-you-go pension schemes in the future, has been how the norm of providing old-age support can be transmitted to future generations. Given the decline in population growth and the potential problem of ageing in other developing countries, a study of the transmission of social norms of support for the elderly in China may help many developing countries understand better how to encourage such support in the future.

This paper studies the inter-generational transmission of the social norm of old-age support provision in China, focusing on the same-gender channel. Parents convey the social norm of old-age support provision to their same-gender children, in the way that they provide support to their own parents. The hypothesised mechanism behind this norm transmission is the same-gender "demonstration effect". It is based on the demonstration effect established by Cox and Stark (1996). The demonstration effect means that parents treat their parents well if they have "their own children to whom to demonstrate the appropriate behaviour" (Cox and Stark, 2005). This inter-generational demonstration meets the anthropologists' description of an upward and positive indirect reciprocity (Arrondel and Massaon, 2006). Anthropologists believe the indirect reciprocity is an important channel of cultural norm transmission (Mauss, 1950, 1968). I improve Cox and Stark's demonstration effect by adding the same-gender transmission channel for two reasons. First, it is because sociologists and psychologists believe that children are largely influenced by their same-sex parent in their learning of gender norms in society (Lytton and Romney, 1991; Bussey and Bandura, 1999; McHale et al., 1999). Economists have recently found empirical evidence for same-gender intergenerational transmission in individual preferences and social norms (Alesina et al., 2013; Kleven et al., 2018). The second reason is that the gender difference is prominent in the norm of old-age support provision in China and other developing cultures (Das Gupta et al., 2003). Traditionally, sons are responsible for supporting their

[1] In the Chinese Household Finance Survey, 74% of the respondents believed that their children should be fully or at least partly responsible for their care in old age.

[2] United Nations (2015) estimated that, in 2030, the share of the population in China aged 60 and older will be 25%. The current share of the population aged 60 and older in the U.K. is 23.9% and in China is 16.2% (United Nations, 2017). WSJ coverage: https://blogs.wsj.com/chinarealtime/2015/03/10/china-sets-timeline-for-first-change-to-retirement-agesince-1950s/. In 2017 China raised the retirement age, set in the 1950s, to alleviate pressures on its public finances.

elderly parents in China (Lee et al., 1994; Chan et al., 2002).

In my proposed mechanism, parents provide old-age support to their parents and they expect to be recompensed by their same-gender children. A key assumption in this mechanism is that parents internalise the fact that their behaviours regarding old-age support provision may affect their same-gender children (Eccles et al., 1990; Bussey and Bandura, 1999). Under this mechanism, a parent should provide more old-age support when the household includes more same-gender children. This channel of inter-generational transmission of the norm does not only exist in the theoretical framework created by academic researchers, but there are also real-world examples for it. Public service announcement posters in China in Figure 1 show the same-gender demonstration effect that is described. These posters also show the government's efforts to promote the norm of providing family support in old age, which indicates the importance of this norm in Chinese society. By studying the same-gender inter-generational transmission of the norm in old-age support provision, this paper seeks to demonstrate how changes in economic and demographic conditions affect the norm and its transmission in China, both financially and non-financially.

I provide novel evidence for the same-gender transmission of this social norm of support in old age and show that the decision-making regarding old-age support provision involves three generations. Most of the family old-age support studies assume by default that the children will provide old-age support when their parents retire because of altruism or direct reciprocity (Becker and Lewis, 1973; Guttman, 2003). These channels limit the effect of old-age support to two generations, the parents and the children.[3] However, there is a gap in the literature: only a few writers focus on the way in which the social norm of providing old-age support is transmitted to the next generation. Cox and Stark (1996, 2005) provide a theoretical framework for the inter-generational transmission of the norm of providing support in old age. The only empirical evidence for this inter-generational transmission has been collected by Wolff (2001) and Mitrut and Wolff (2009). The present paper helps to fill this gap by providing empirical evidence for the gender-specific effect demonstrated in support for the elderly in China. The empirical results show the importance of the future generation in the process of transmitting the social norm of old-age support. The paper also contributes to the literature by first documenting a normative shift with economic and demographic changes during China's transformation into a modern nation, thanks to the wide urban-rural differences.

When studying the effects in China of the gender of children on the support for the elderly provided by their parents, an empirical difficulty is that the gender of the children is endogenous. The increasing gender ratio of new-borns in China corresponds to the imbalance in the gender ratio of the children in the datasets. The gender ratio of new-borns has been increasing since 1990 (China Population and Employment Statistics Yearbooks, Figure 2). For this, sex-selective abortion is one of the main reasons (Chen et al., 2013). The non-random gender ratio of the children could positively or negatively

[3]Some of the relevant literature evaluates the "manipulation" of children by their parents to ensure more old-age support in the future (Becker et al., 2016).

affect the support for the elderly provided by parents.[4] To address this problem, I utilise two facts: the gender of the first child in households and the timing of a policy ban on sex-selective abortions.

I use the interaction term of the gender of the first child in a household and whether or not a household is affected by the policy ban as the instrumental variable (IV) for the gender ratio of the children. This IV exploits two facts. First, the gender of the first child is closer to the natural rate than the gender ratio for all new-borns in China (Ebenstein, 2010; Wei and Zhang, 2011). Scholars usually regard the gender of the first child as random (Jayachandran and Pande, 2017; Heath and Tan, 2018). However, given the highly skewed gender ratio of new-borns in China, it is difficult to concede that the gender of first-born children is fully exogenous. Second, a policy was introduced to reduce the gender ratio to its natural level, so the gender of first-born children who were born in or after the year of the policy ban is more random. The policy banned the use of ultrasound for prenatal sex determination, and imposed fines on those who conduct sex-selective abortions. It was initiated by the National Family Planning Commission (NFPC) in 2003 affecting all households that have at least one child born in or after 2003.

The timing of the policy change is plausibly exogenous at household-level.[5] I find that the policy, as intended, negatively affects the household-level gender ratio of children. The compliers are those who have not conducted sex-selective abortions since the policy ban. There are two different types of complier: affected and unaffected. The affected compliers are those who have children of the opposite sex to their wishes. Usually the compliers prefer sons to daughters. They capture the time variation of the policy. For example, after 2003 the affected compliers who would have been willing, had no ban existed, to conduct sex-selective abortions, have daughters and this decreases the gender ratio of their children. Unaffected compliers who have sons after 2003 by natural chance provide no variation. The gender ratio of the children of people who would not conduct a sex-selective abortion in any circumstances cannot be affected by this policy, and the gender ratio in their households should be close to the natural rate. The IV thus captures the differences for the affected compliers before and after the policy ban.

The main empirical findings indicate that parents with same-gender children have a higher probability of providing support in old age to their own parents. These results are consistent with different demographic controls. In the datasets, the father and the mother both show gender-specific demonstration behaviour. The results from the robustness check and the heterogeneity analysis are mostly consistent with the main outcomes. The "father" demonstration effect is larger and more significant in low-income and rural subsamples, and also in households with more than one child. The "mother" effect is larger in low-income and urban subsamples, and also in households with a single child. Both gender-specific effects for non-financial support are more significant in households with younger children than in households with one adult child. The result is consistent with sociologists'

[4] This will be further elaborated in the empirical results section.

[5] The law-making processes of most Chinese policies are quite exogenous, as far as members of the public are concerned (Hu, 1998; Shen, 2008).

gender socialisation theory: parents have greater influences on their children during the rearing stage (Lytton and Romney, 1991). The empirical evidence implies that support for the elderly is closely linked to the composition of the gender of parents and their children, and the norm of providing support for the elderly is likely to be transmitted to offspring of the same gender.

However, the two datasets exhibit different gender-dominated demonstration behaviours. CHARLS (the China Health and Retirement Longitudinal Study) mainly presents the father demonstration effect. The mother effect has a more substantial role in the urban subsample and also in the whole sample of CHFS (the China Household Finance Survey). One explanation for this difference is because CHARLS contains much more rural samples than CHFS. It is consistent with results from the urban-rural heterogeneity analysis and subsample check. The discrepancy between the urban and rural subsample results has implications for the norm-shift of providing support for the elderly together with the development of China. Urban areas in China are more developed than rural areas: they have higher pension/insurance coverage, better public infrastructure, and, in particular, fewer gender inequalities and higher female bargaining powers (Fong, 2002; Lee, 2012). The results may suggest that higher female household bargaining power may lead to more significant mother demonstration effects. The mechanism checks also show that the existence of other possible mechanisms, such as altruism and direct reciprocity, does not affect the demonstration effect in the results.

To illustrate the same-gender demonstration effect, I suggest a simple two-period consumption model describing the three-generation interactions in providing old-age support. The model includes inter-household transfers (Banerjee et al., 2014) and a demonstration effect (Cox and Stark, 1996). It also contains the intra-household bargaining components. The model concludes that the parent who holds higher bargaining power in a household is more likely to demonstrate the norm of old age support to offspring of the same gender, which provides a possible explanation of the different gender-dominated demonstration effect in rural and urban subsamples. I also calculate the correlation between the "missing girl" and the demand for support for the elderly in a patrilineal society, using a method from Oster's 2005 paper. Using this method, I calculate the adjusted sex-ratio based only on the correlation between the unbalanced gender ratio and the demand for support for the elderly from sons. The adjusted sex-ratio accounts for 12-18% of the unbalanced gender ratio in the data.

The paper proceeds as follows. More background information on support for the elderly from children in China is in Section 2. Section 3 provides the theoretical background for the same-gender social norm transmission and the model. This is followed by Section 4, which provides the identification strategy and the empirical findings. Section 4 also provides the robustness check for the key empirical findings. Section 5 offers some concluding thoughts.

2 Old-age support in China

The provision of financial and non-financial support to ageing parents is a pro-social norm in China and other countries that are influenced by Confucianism. This family support for the elderly has been acting as an alternative way of sustaining the welfare of elderly to the incomplete public pension system. Table 1 shows that in 2005 less than 50% of the urban elderly viewed public pensions as their major source of income. In rural areas, the percentage was only around 5%. More than 50% of the rural elderly and around 40% of their urban counterparts believed their major source of income to be family support. Even with the development of the public pension in both urban and rural areas in China, , the percentage of rural elderly choosing pensions as their main income source in 2010 was unchanged, although the percentage of those who chose family support declined to 47%. The pension schemes in urban areas have been improved since 2005: around 70% of the urban elderly in 2010 relied on a public pension while only around 20% of them lived mainly on family support. Inferring from the statistics, the public pension coverage shows a wide urban-rural difference. Rural areas in China do not seem to have had an effective pension scheme before 2011, so the elderly there were still depending on the norm of private support for the elderly.

A large proportion of old people in Chinese lives on support from their family members, especially their adult children. The social norm of providing support for the elderly is then important to those who trying to secure their income after their retirement. First, they have to know which characteristics affect the amount of support that they can depend on in old age. The number and the gender of the adult children are two major aspects studied in the relevant literature on China. In the classic literature, such as Becker and Lewis (1973), people believe that more children in a household will lead to more support for the elderly in the future. Cai et al. (2006) and Oliveira (2016) both verify this common belief among Chinese people. As regards the gender of the children, traditionally, males are responsible for providing support to their parents in their old age. Hence the early literature assumed that males provide more than females due to cultural and labour market restrictions (Lee et al., 1994; Chan et al., 2002). The value of male offspring in providing support for the elderly is one of the reasons behind the persistent preference for sons in China and other developing countries (Das Gupta et al., 2003). It was common in China for households to have at least one son, right up to the implementation of the "One-Child" Policy (OCP) (Milwertz, 1997; Ebenstein and Leung, 2010). However, in the recent literature, Xie and Zhu (2009) find that females were providing more support to elderly parents and Oliveira (2016) finds no gender differences in the provision of support in old age. But given the rising gender ratio for new-borns in China, especially in rural areas, it is reasonable to assume that this gender difference still exists, though it may vary between rural and urban areas.

Once those who rely on family support for income in old age know the factors affecting their future income, it is highly likely that they will try to manipulate these characteristics. For most families in

China the number of children is difficult to manipulate. With the strict implementation and high fines of the One Child Policy (OCP), Ebenstein (2010) has found that the policy reduced fertility. Gender, however, was a characteristic that was easier for people to manipulate, with the help of advanced technologies. Chen et al. (2013) has inferred that the increasing gender ratio could be attributed to increased gender selection before birth, thanks to gender-selection technology. For example, B-mode ultrasound allowed people to know the sex of a foetus and was in common use all over the world after 1980 (White, 2001). Qian (2008) has discovered that an increased future income for females also improved the female survival rate. In addition, Ebenstein and Leung (2010) have studied the effects of having a regulated pension system on the sex ratio at birth in China. They find that when a region is covered by a public pension scheme, its gender ratio is more balanced than it is in regions without such coverage. From the literature, it seems that in China, gender is a key factor in the norm of family support in old age. Support for the elderly is also, important enough to affect fertility decisions, such as the number and gender of people's future children. Parents internalise the future support that they will receive from their children when they are old and try to alter the characteristics that affect their own future support.

3 Theoretical framework

3.1 Indirect reciprocity

It is important to learn how best to support the elderly, given their situation. First, we should understand the possible mechanisms for doing so. Altruism and exchange are the two main motives in the standard theoretical models analysing intergenerational transfer. Altruism in the context of support for the elderly means that people are generally willing to support their ageing and retired parents. The theoretical framework for altruistic individuals is developed by Barro (1974) and Becker (1976, 1981). The exchange mechanism is also referred to as (direct) reciprocity. It describes support for the elderly as reciprocal payments for the financial and/or non-financial investment made in the donors' childhood (Cox, 1987). However, the existing empirical results are not robust enough to support these two motives in theoretical models (Arrondel and Masson, 2006). The theory of indirect reciprocity may serve to reconcile the motives of altruism and exchange. Indirect reciprocity is also the theoretical support for the inter-generational transmission of the norm of giving support to the elderly.

The concept of indirect reciprocity is usually attributed to Mauss (1950, 1968), a French anthropologist. He expands the common "gift-return" reciprocity relationship between two parties, the giver and the beneficiary, to three parties. He states that indirect reciprocities involving three successive generations will lead to infinite chains of transfers. He observes that the givers do not get direct payback from the beneficiary but receive it from a third person (Arrondel and Masson, 2001). The channel works for any type of transfer: upward, downward, positive or negative. Cox and Stark (1996) provide a model to describe similar behaviours in the provision of support in old age, which coincides with the upward and positive indirect reciprocity channel. In the context of support for the elderly, the interaction between three parties is that parents educate their children by providing support for the elderly to their own parents, so that the parents when elderly will receive support from their children. It is usually referred to as the "demonstration effect". The model predicts that transfers from individuals to their parents are positively affected by the presence of their children. Cox and Stark (2005) test the prediction using U.S. data. Wolff (2001) and Mitrut and Wolff (2009) also find that the existence of granddaughters increases the visits paid to the grandparents; Becker et al. (2016) believe that parents can "manipulate" the preferences of children, an assumption underlying the demonstration effect.

Except for Mitrut and Wolff (2009), the relevant literature considers only the role of the children in the transmission of the norm of old-age support, without any consideration of the role of gender. Given the gender difference regarding support for the elderly and preference in China for sons, the demonstration effect may also be linked with the gender of the third generation. Godelier (1982) describes indirect reciprocity as gender-specific when it functions as a channel for the transmission of cultural traits and norms. If there is a gender-specific social norm, then it is also a channel for passing on gender norms in the society. Mitrut and Wolff (2009) find that parents' visits to their own parents are largely affected by the presence of daughters rather than sons in their households. This empirical finding is consistent with common beliefs about the role of gender: parents of girls are the more likely ones to pay visits and care for the elderly (Lee et al., 1993). Bau (2016) also studies the connection between the cultural norm and support for the elderly in Ghana and suggests that support for the elderly is a product of cultural norms.

If providing support for the elderly links with gender norms, one vital assumption is that parents should be able to influence their same-gender children more effectively than cross-gender children. Children would also mimic the behaviour of the same-gender parent in the future, a phenomenon which is known in psychology and sociology as "gender socialisation/specification". Many sociologists and psychologists believe that the same-sex parent is the main source for ensuring that children to learn the corresponding gender role that fits social expectations and that the children will perform gender-related behaviours when they become adults (Lytton and Romney, 1991; Bussey and Bandura, 1999; McHale et al., 1999). In the recent economics literature, several papers focus on same-gender intergenerational transmission. Jayachandran and her colleagues show that the effects of same-sex parent on gender attitudes are greater than the peer effects (Dhar et al., 2015). Kleven et al. (2018) reveal that in Denmark preferences over family and career for females are largely influenced by the mother's preference observed during childhood. Alesina et al. (2013) also find that paternal ancestors affect the perspectives of males on the gender role and whether females should enter the labour market.

Parents should also internalise the fact their children's future behaviours will be affected by theirs. This internalisation means that parents will begin to influence their offspring in order to form their children's preferences. Becker (1996), Bisin and Verdier (2000), Guttman (2001), Bronnenberg et al., (2012) and Becker et al. (2016) study whether parents show certain behaviours to or spend more resources on their children in order to formalise their children's preferences. After listing the relevant evidence supporting the demonstration effect and same-gender intergenerational norm transmission, it is reasonable to assume that the demonstration effect works in a more gender-specific way when there is a wide gender difference in the planned support for the elderly. People will demonstrate the norm of support in old age to their same-gender offspring by providing support for the elderly to their own parents.

3.2 The baseline model

The model describing the same-gender demonstration effect in the paper is based on the demonstration effect model by Cox and Stark (1996, 2005), combined with a definition of intergenerational transfers taken from a model by Banerjee et al. (2014). It is a simple inter-temporal two-period consumption model. Cox and Stark (1996, 2005) suggest that "...childhood experience affects behaviour in adulthood". Parents who value support for the elderly will demonstrate the norm of providing support for the elderly to their children by providing support to their own elderly parents. Based on the demonstration effect, the model assumes that parents know that their support to their own elderly parents will affect the future support behaviour of their same-gender children. Another assumption noted above is that children will be affected by the behaviour of their same-gender parents. Given differences in anticipation of the future and same-gender intergenerational transmission, parents will provide support to their own parents, according to the gender of their children. this explains the relationship between parents' support for the elderly and the gender ratio of children.

There are three generations in the model: the mid-age generation (P), the parents; the older generation (O), parents of P, and the younger generation (K), children of P. They correspond to the second generation, the first generation and the third generation respectively only in this subsection and the next (Section 3.2 and 3.3). There are two periods in the model: the first period, $t = 1$, and the second period, $t = 2$. The model uses the notation in Banerjee et al. (2014) and requires a few additional assumptions:

- (i) each household in P has a father and a mother;
- (ii) the father transfers a fraction τ_1^F of his income and the mother transfers a fraction τ_1^M of hers. Both of them have income Y_1 and Y_1 is exogenous;
- (iii) the number of K in each household, n, is exogenous;

- (iv) people value their own consumption and their parents' welfare, so they derive utilities from transfers to their parents. However, there is also a discount factor, $0 < \delta < 1$, for the utility derived from the provision of old-age support, since it does not represent direct consumption for the individuals;

- (v) τ_1^F and τ_1^M are endogenous and change over time. The transfer from the sons and daughters of the father and the mother in the second period will be affected by their same-gender parents' transfer in the first period.[6] In equations, this assumption is expressed as

$$\tau_2^F = \mathcal{T}^F(\tau_1^F) \quad \text{and} \quad \tau_2^M = \mathcal{T}^M(\tau_1^M); \tag{1}$$

  Both functions are strictly concave.

- (vi) the father and the mother in a household make unitary household-level decisions. The male-to-female gender ratio of children in a household is ϕ. The household consumption is c in each time period;

- (vii) for simplicity, I assume that there is no saving in the baseline model;[7]

- (viii) $u(\cdot)$ is a strictly concave function when $u(\cdot) > 0$.

In this model, P is the generation solving the optimization problem in the first period. O passively receives support from P in the first period and dies in the second period. Members of K observe their parents' τ_1 in the first period and provide their parents with τ_2 in the second period. With the assumptions above, a typical household in the second generation solves the following problem:

$$\max_{\tau_1^F, \tau_1^M} \quad U = u(c_1) + \delta u(e_1) + \beta u(c_2) \tag{2}$$

$$s.t.$$

$$c_1 + c_2 \leq Y_1(2 - \tau_1^F - \tau_1^M) + Y_2(\mathcal{T}^F(\tau_1^F)\phi n + \mathcal{T}^M(\tau_1^M)(1-\phi)n);$$

$$e_1 = Y_1(\tau_1^F + \tau_1^M).$$

The father and the mother in generation P make unitary household-level decisions, so that the household consumption is given by:

$$c_1 = Y_1(2 - \tau_1^F - \tau_1^M); \quad c_2 = Y_2(\mathcal{T}^F(\tau_1^F)\phi n + \mathcal{T}^M(\tau_1^M)(1-\phi)n).$$

e_1 is the old-age support provided by the whole household. δ is the discount factor for the utility generated from altruism, and β is the time discount factor. If $u(c)$ is specified as a log or a CRRA

[6] This same-gender demonstration assumption is later relaxed (See Appendix A.2.1).
[7] Saving is included in the basic model in Appendix A.3.

function and τ_2 is a concave function of τ_1 for the fathers and mothers, the FOCs with respect to τ_1^F and τ_1^M are:

$$U^1 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^F} = u'(c_1)(-Y_1) + \delta u'(Y_1(\tau_1^F + \tau_1^M))Y_1 + \beta u'(c_2)Y_2\tau_2^{F'}\phi n = 0; \tag{3}$$

$$U^2 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^M} = u'(c_1)(-Y_1) + \delta u'(Y_1(\tau_1^F + \tau_1^M))Y_1 + \beta u'(c_2)Y_2\tau_2^{M'}(1-\phi)n = 0. \tag{4}$$

The SOCs from these two FOCs are:

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F2}} < 0; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\phi} > 0; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M2}} < 0; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}\phi} < 0; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M} < 0.$$

The detailed expressions and the proof of signs for SOCs are presented in Appendix A.1.1, Equations (9) and (10). Let τ_1^{F*} and τ_1^{M*} be the optimal solution from the FOCs, then

Lemma 1: *In the baseline model, τ_1^{F*} is increasing in ϕ and the optimal τ_1^{M*} is decreasing in ϕ, the gender ratio of K. $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$.*

To prove *Lemma 1* requires me to totally differentiate Equations (3) and (4):

$$\begin{aligned} &U^{11}\mathrm{d}\tau_1^{F*} + U^{12}\mathrm{d}\tau_1^{M*} + U^{13}\mathrm{d}\phi = 0; \\ &U^{21}\mathrm{d}\tau_1^{F*} + U^{22}\mathrm{d}\tau_1^{M*} + U^{23}\mathrm{d}\phi = 0, \end{aligned} \tag{5}$$

where

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*2}} = U^{11}; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi} = U^{13}; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*2}} = U^{22}; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*}\mathrm{d}\phi} = U^{23}; \quad \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\tau_1^{M*}} = U^{12} = U^{21}.$$

and the asterisks denote optimal values. These U^{ij}s are the SOCs when τ_1^F and τ_1^M at their optimal values. $i = \{1, 2\}$ and $j = \{1, 2, 3\}$.

The comparative statics from these conditions from Equation (5) are:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23} - U^{13}U^{22}}{U^{11}U^{22} - U^{12}U^{21}}; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23} - U^{13}U^{21}}{U^{12}U^{21} - U^{11}U^{22}};$$

From the SOCs in Equations (9) and (10) in Appendix A.1.1, I obtain $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$. The proof is in Appendix A.1.1. The first interpretation of these equations is that the fraction of the father's transfer to his parents increases with the gender ratio of his children. It also means, he will provide more old-age support to his parents with more sons in his household. This effect is the father-dominated demonstration effect. The mother will transfer more to her own parents if she has more daughters, regardless of whether τ_1^F is larger or smaller than τ_1^M. This is the female/mother-dominated demonstration effect. As noted above, it is usual in China for males to support their parents more than females do. $\tau_1^F > \tau_1^M$ indicates that the father transfers more than the mother does, as a general social norm. However, the condition $\tau_1^F > \tau_1^M$ does not affect the conclusion of the model.

One key assumption for the interpretations is that ϕ should be exogenous. To make sure that ϕ, the gender ratio of the generation Y, is exogenous at the household-level in the empirical part, I use the policy change which started in 2003. The policy bans the selection of unborn children by sex in China. I give a more detailed explanation in the empirical section. The regulation will bring the gender ratio of newborns after 2003 close to the natural rate compared to the gender ratio before the policy changes.

3.3 Collective household model: intra-household bargaining

One of the assumptions in the baseline model is that households in the generation P make unitary household-level decisions and the utility generated by providing old-age transfer counts as a utility of the household. To relax this assumption, assume a collective model for the household-level decisions, which involves intra-household resource allocation. According to Browning and Chiappori (1998), the genders hold different bargaining powers or "distributions of powers" in the households. This can be translated into different weights attached to the father's and mother's utility in the household-level utility function. The additional assumptions on the intra-household bargaining are as follows:

- (i) The father earns Y_1^F and the mother earns Y_1^M. The weight for the father's utility function is ρ_1 when $t = 1$. $\rho_t = \mathcal{P}(\frac{Y_t^F}{Y_t^M})$ for $t = \{1, 2\}$. The mother's weight is $1 - \rho_1$. ρ_t is increasing in $\frac{Y_t^F}{Y_t^M}$ and $0 \leq \rho_t \leq 1$. When $Y_t^F = Y_t^M$, $\rho_t = 0.5$.

- (ii) The father and the mother each have an individual-level consumption, c_1^F and c_1^M, respectively

- (iii) η_t is a result from the intra-household resource allocation between Y_t^F and Y_t^M. $\eta_t = \mathcal{H}(\frac{Y_t^F}{Y_t^M})$ when $t = \{1, 2\}$. It is increasing in $\frac{Y_t^F}{Y_t^M}$ and $0 \leq \eta_t \leq 2$. When $Y_t^F \geq Y_t^M$, $1 \leq \eta_t \leq 2$, and when $Y_t^F < Y_t^M$, $0 \leq \eta_t < 1$.

- (iv) The results of the intra-household resource allocation are that the father provides a proportion, η_1, of his original fraction of provision, τ_1^F, while the mother provides $2 - \eta_1$ of her original fraction of provision, τ_1^M in the first period.

- (v) In the second period, when neither the father nor the mother earns income, the previous parameters cannot apply to them. So I assume that they share the transfer that they received from the next generation when $t = 2$. The consumption in the second period (c_2) is also at the household level.

The additional assumptions describe how the fraction of the household resources, $\frac{Y_1^F}{Y_1^M}$, affects the allocation of household resources in the first peiod in terms of providing old-age support to the father's ageing parents (τ_1^F) and those of the mother (τ_1^M). The father provides $\eta_1 Y_1^F \tau_1^F$ and the mother provides $(1-\eta_1)Y_1^M \tau_1^M$ to his/her own parents. The optimisation problem becomes:

$$
\begin{aligned}
\max_{\tau_1^F, \tau_1^M} \quad U = \rho_1 u(c_1^F) + (1-\rho_1)u(c_1^M) + \delta\rho_1 u(e_1^F) \\
+ \delta(1-\rho_1)u(e_1^M) + \beta u(c_2);
\end{aligned}
\tag{6}
$$

$$
\begin{gathered}
s.t. \\
c_1^F + c_1^M + c_2 \leq Y_1^F(1-\eta_1\tau_1^F) + Y_1^M(1-(2-\eta_1)\tau_1^M) \\
+ Y_2^F \eta_2 \mathcal{T}^F(\eta_1\tau_1^F)\phi n + Y_2^M(2-\eta_2)\mathcal{T}^M((2-\eta_1)\tau_1^M)(1-\phi)n; \\
e_1^F = \eta_1 Y_1^F \tau_1^F; \\
e_1^M = (2-\eta_1)Y_1^M \tau_1^M.
\end{gathered}
$$

In this section, I discuss only two extreme cases of the model: when $Y_1^F \geq Y_1^M$, then $\eta_1 = 2$ and when $Y_1^F < Y_1^M$, then $\eta_1 = 0$. When $Y_1^F \geq Y_1^M$ and $\eta_1 = 2$, the father provides τ_1^F to his parents and the mother provides no support to her parents accordingly. When $Y_1^F < Y_1^M$ and $\eta_1 = 0$, vice versa.[8] When $\eta_1 = 2$ or $\eta_1 = 0$, the $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\tau_1^M} = 0$

$$
c_1^F + c_1^M = Y_1^F(1-\eta_1\tau_1^F) + Y_1^M \quad \text{or} \quad c_1^F + c_1^M = Y_1^F + Y_1^M(1-(1-\eta_1)\tau_1^M),
$$

and

$$
c_2 = Y_2^F \mathcal{T}^F(\eta_1\tau_1^F)\phi n \quad \text{or} \quad c_2 = Y_2^M \mathcal{T}^M((2-\eta_1)\tau_1^M)(1-\phi)n
$$

depending on the value of η_2.

Lemma 2: *In the model, when* $Y_t^F \geq Y_t^M$ *and* $\eta_t = 2$ *for all* t, $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ *and* $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = 0$. *When* $Y_t^F < Y_t^M$ *and* $\eta_t = 0$ *for all* t, $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = 0$ *and* $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$.

With all the previous assumptions, and also the assumption $u(0) = 0$, the proof of *Lemma 2* is: If $Y_t^F \geq Y_t^M$ and $\eta_t = 2$ for all t, the FOC with respect to τ_1^F:

$$
U^1 = \rho_1 u'(Y_1^F(1-2\tau_1^F))(-Y_1^F) + \rho_1\delta 2Y_1^F u'(Y_1^F\tau_1^F) + \beta u'(c_2)2Y_2^F \tau_2^{F'}\phi n = 0;
$$

τ_1^{F*} is the optimal solution of τ_1^F from $U^1 = 0$. The total differentiation of U^1 is

$$
U^{11}\mathrm{d}\tau_1^{F*} + U^{13}\mathrm{d}\phi = 0.
$$

[8] The general case is discussed in Appendix A.2.2.

Given the conditions $U^{11} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*2}} < 0$ and $U^{13} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi} > 0$,[9] the conclusion from these conditions is:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = -\frac{U^{13}}{U^{11}} > 0.$$

When $Y^F \geq Y^M$ and $\eta_t = 2$ for all t, the gender ratio of the generation K does not affect the mother's transfers to her parents, so

$$\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = 0.$$

In another extreme case, if $Y^F < Y^M$ and $\eta_t = 0$ for all t, then the FOC with respect to τ_1^M:

$$U^2 = (1-\rho_1)u'(Y_1^M(1-2\tau_1^M))(-Y_1^M) + (1-\rho_1)\delta 2Y_1^M u'(Y_1^M\tau_1^M) + \beta u'(c_2)2Y_2^M\tau_2^{M'}(1-\phi)n = 0;$$

τ_1^{M*} is the optimal solution of τ_1^M from $U^2 = 0$. The total differentiation of U^2 is:

$$U^{22}\mathrm{d}\tau_1^{M*} + U^{23}\mathrm{d}\phi = 0;$$

so

$$\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = -\frac{U^{23}}{U^{22}}.$$

The equations of the SOCs in Equation (13) in Appendix A.1.2 show:

$$U^{22} < 0; \quad U^{23} < 0;$$

and as in the case when $Y_t^F < Y_t^M$ and $\eta_t = 0$ for all t, the comparative statics from these conditions are:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = 0; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0.$$

The model conclusion fits what is described in *Lemma 2*.

There are two parameters representing collective households in this model: ρ_t and η_t. As stated in the results, the parameter ρ_t does not affect the results during the process of deriving the comparative statics for $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}$. This parameter does not represent the process of bargaining in terms of old-age support provided by the father or the mother. It simply represents the different weightings attached to different members of the households. η_t is the component that is linked to the old-age support provided by P and the same-gender members of K.

The conclusions from the intra-household bargaining model are as follows: when the father in a household with more sons has a higher bargaining power than the mother, he provides more old-age

[9] Equations for the SOCs and the corresponding signs are in Appendix A.1.2 Equation (12)

support to his own parents. When the mother with more daughters has a higher bargaining power than the father, she provides more. Different bargaining powers possessed by fathers and mothers lead to different demonstration effects from these fathers and mothers. Again, the key assumption is that ϕ is exogenous. The identification strategy in the next section tries to make the corresponding ϕ in the regressions exogenous.

4 Empirical results

4.1 Data

Two datasets are used to assess the gender effects of children on the norm transmission of old-age support, more specifically, how the gender of children affects the support for the elderly provided by their parents. The first dataset is the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is a longitudinal survey of 28 out of the 34 provinces of the country for three waves in the years 2011, 2013 and 2015 up to the present day.[10] It collects a representative sample of residents aged 45 or above. The main wave used in this paper is the 2011 wave. The data set contains information on each respondent's family, work, retirement, wealth, health and income. The main demographic group in the survey is people aged 45 or above. In the 2011 sample, this covered about 17,708 individuals in 10,257 households from 28 provinces. The sample was randomly selected from four samplings at different levels: county-level, neighbourhood-level, household-level and respondent-level.[11]

CHARLS provides detailed information on inter-generational and inter-household transfers. One advantage of this dataset is that it clearly distinguishes between the transfers from each child of the respondents. The survey also identifies different types of support, whether regular or non-regular. The regular support acts as income received from the children of the respondents at fixed times. Regular support is similar to the support for the elderly as defined: a certain amount of income paid repetitively to the elderly at a fixed time. Non-regular support is the support provided at different times of the year, and is not necessarily repetitive, whereas the regular one is.[12] Given the high average age of the respondents, the sample size for the available observations in terms of the transfer provided by the respondents to their parents is small. But most of the respondents have children of working age, so most of them receive support from their children. I regard support for the elderly provided by the children of the respondents as the support from parents to their elderly parents discussed in the previous section. The respondents in the survey are the passive aspect of support for elderly recipients, namely, the elderly, in the model and the main regressions. The grandchildren of the survey's respondents are the children of the respondents' children.

To fit the original dataset into my setting, I construct a new sample that covers the adult children

[10]The detailed distribution in provinces and counties is presented in Figure A.1.

[11]The detailed sampling method at each level can be accessed at: http://charls.pku.edu.cn/en/page/about-sample-2011.

[12]In CHARLS, non-regular support is defined as "support at Spring Festival or/and Mid-Autumn Festival or/and birthday or/and wedding or/and funerals or/and others".

of the survey respondents, namely, the parents.[13] In the newly constructed sample, the sample size decreases to about 14,000 observations. In the reconstructed 2011 wave, around 65% of people come from rural areas, and more than 75% of the samples has rural *hukou* ("household registration"). However, due to the questionnaire design of CHARLS, the demographic information on the parents and their children is not as detailed as the information on their elderly parents. The available demographic variables in the 2011 wave about the children are only the gender and the number of them. In the 2013 and 2015 wave, the only available demographic variable is the number of the children. This is the reason why I can conduct only cross-sectional analyses with CHARLS.

I used a second dataset to verify the generalisation of the results from CHARLS and also to provide supplementary evidence for the demonstration effect. The dataset is the China Household Finance Survey (CHFS). CHFS is a panel dataset covering 25 provinces in China, by Southwestern University's Department of Finance and Economics and Research Institute of Economics and Management. This survey focuses on household-level financial behaviours. It currently has three waves: for the years 2011, 2013, and 2015. The survey does not have the same age limitation on the survey respondents as CHARLS does; hence, there is no need to reconstruct the dataset. In this paper I treated the parents as the main respondents of the survey. The sample in the 2011 wave includes only 8,438 households, and its questionnaire includes only the gender of the children who are living together with the respondents. In the 2013 wave, the number of observations increased significantly: 28,142 households and 97,916 individuals. Accordingly, I used the 2013 wave in CHFS for more observations and more precise information on the gender ratio of the children.

I include only the main respondent for each household in my CHFS sample for regression. The main respondents know the household financial situation best. They are responsible for answering the household-level financial questions, which includes the questions regarding inter-household transfers. If I included only the main respondents, there would be a selection bias. In this sample the parents are in charge of the household finances. So, one possible effect would from females who were in charge of the household finances, who may have higher power in their household than is held by females who are not in financial charge. A possible result of this selection would be that the females in my CHFS sample transferred more to their parents. Regarding the households' support for the elderly, the main respondents may know only the exact amount of their own transfers, and not that of their partner. Their partner may hide the information from them (Ashraf, 2009). Moreover only the main respondents have information about their own parents. The 2013 wave also includes the gender of all the children of the respondents. One limitation of CHFS is that the information about the intergenerational and inter-household transfer collected in the survey is not as detailed as the information available in CHARLS. Due to the sample selection problem, also, I use the CHARLS dataset as my main dataset and for the main results. Thus the results from CHFS act as supporting evidence for the CHARLS results.

[13]A detailed discussion of the dataset reconstruction is in Appendix 8.

4.2 Main regression

The paper sets out to examine the gender effects of the children on the support for the elderly provided by their same-sex parent. The main regression includes the gender of the parents, the gender ratio of their children in their household and the interaction term for these two variables. The main regression is:

$$y_i = \alpha + \beta genderK_i + \gamma maleP_i + \delta(maleP_i \times genderK_i) + \mathbf{X_i'}\boldsymbol{\theta} + \phi_c + \epsilon_i. \tag{7}$$

In the equations, i stands for a parent i. y_i represents the outcome variables testing various aspects of old-age support. The error term is ϵ_i and ϕ_c is the province fixed effects. For the main regressors, I use the three-generation setting specified in the model: P is the parents, K represents the children of P, and O is the parents of P. $maleP_i$ is the gender of a parent i in the P generation. It equals 1 if the parent is male and 0 otherwise. The regressor $genderK_i$ is the gender ratio of K in P's household. It varies between 0 and 1. The gender ratio of K equals the number of sons for a parent i divided by the total number of K in the household if i has more than one child. For i with one child, if the only child is a boy, then $genderK_i = 1$. If the only child is a girl, then $genderK_i = 0$. $genderK_i \times maleP_i$ is the interaction term, and $\mathbf{X_i}$ is the set of demographic variables for P and O to be controlled for in the regression.[14] I run separated regressions for CHARLS and CHFS, since the difference between the two datasets is quite large. The error term is clustered at household-level in generation O in CHARLS because under the data reconstruction some P are from the same family in O. It is a common robust standard errors in CHFS. Using this regression equation, I manage to calculate the within parental gender differences in terms of providing support for the elderly caused by the gender ratio of their children, while controlling for the parents' own gender.

There are six different outcome variables in CHARLS. Five of them are financial old-age support: dummies equal 1 if the parents provide any, regular and non-regular transfer, and natural logarithms of the amount of regular and non-regular transfer provided. The regular transfer is the transfer that parents make to their elderly parents at fixed times. The non-regular transfer represents transfers provided by the parents at festivals, birthdays, weddings, funerals, and for medical treatments, and also for other non-regular but important social events. The dummy for any transfer equals 1 if the parents provide regular and/or non-regular transfers. For the amount of the transfer, I unify it to the annual amount. The non-financial outcome is the natural logarithm of the days per year spent on visits. For the outcome variables in CHFS are the dummy indicating whether parents provide any financial transfer to their elderly parents, the natural logarithms of the amount of any transfer provided, the number of days spent on visits paid to their elderly parents per year and the natural logarithm of this number. The summary statistics for the outcome variables, key regressors and key control variables in different datasets are shown in Table 2.[15]

[14] The controls are different in CHARLS and CHFS. I try to make the controls consistent between the two datasets. The control variables for O are more in CHARLS than in CHFS, but information on P and K is more precise in CHFS.

[15] The full summary statistics for all the controls and the summary statistics by gender of the adult children are in

The OLS results from the regression Equation (7) for CHARLS and CHFS are shown separately in Table 3 and 4.Before analysing the gender effects of children, I first want to verify whether there are gender differences in the provision of support for the elderly in CHARLS and CHFS. In the recent literature, it seems that males are no longer provide more old-age support than females. Oliveira (2016) shows that there is no gender difference in the support for the elderly provided by parents. Xie and Zhu (2009) show that females in urban areas provide more to their parents than males do. I want to use the OLS results from two regressions to check whether the male parents provide more in the datasets used.

The results imply that there are gender differences of P in old-age support. In Table 3, the OLS results from CHARLS show that there is no gender difference between the parents in the probability of providing any kinds of transfer. But if I look into types of transfer in detail, gender differences appear. For regular transfers, male P have higher probabilities of provision; while females are more likely to provide non-regular transfers. Males transfer more in the amounts of their regular transfers. The gender difference is not significant for non-regular transfers. The gender difference for non-pecuniary support is also significant: males visit their elderly parents more than females do. To sum up, males still provide more support than females, especially when it comes to regular support and visits paid to elderly parents recorded in CHARLS.

However, the OLS results in CHFS seem to show the opposite situation. The coefficients of $maleP$ or the probability of providing any kind of transfer and for the total amount of the transfer are both significant and negative. The coefficient for the log of the visit days is insignificant. The only positive and significant coefficient for $maleP$ is the one for the days spent visiting their ageing parents. The results suggest that in CHFS females on average support their parents more. From the CHFS results, it seems that at least regarding pecuniary transfer, female parents provide more than males do. The greatest difference between the two datasets arise from the composition of samples living in urban and rural areas, as shown in the summary statistics (see Table 2).[16] The discrepancy between the OLS results from CHARLS and CHFS for $maleP$ may suggest that there is a difference in the gender norm for providing support for the elderly in urban and rural areas in China. Combining the results in CHARLS and CHFS, it is reasonable to assume that males still provide more in the rural areas and urban females may have more important roles in terms of providing old-age support, supported by the empirical finding in Xie and Zhu (2009).

Regarding the gender of children, the datasets suggest that most of the effects are insignificant. From now on I refer to females in the P generation as mothers, and their male counterparts as fathers, instead of male P. The gender effects of K work through the same-gender P. In Equation 7, $-\beta$ indicates the differences between mothers with all daughters and mothers with all sons, which I use to indicate the working of the mother demonstration effect on daughters. $\beta + \delta$ shows the same

the Appendix, Tables A.2, A.3, A.4, A.5, A.6.

[16] I discuss more discrepancies between CHARLS and CHFS as possible reasons for the discrepancies later in this section.

differences between fathers, and I define this difference as the father demonstration effect. If the same-gender channel works, the expected coefficients of β should be negative and significant for the mother demonstration effect. The coefficients of $\beta + \delta$ should be positive and significant to show the father effect. For six outcome variables in CHARLS, only $genderK$ for $log(visit\ days)$ is negative and significant, the mother and father demonstration effect for different pecuniary support outcomes are all insignificant. The coefficients for β and δ are also insignificant in CHFS. In spite of the most insignificant effects, the significant mother demonstration effect for non-pecuniary support for the elderly appears to fit the common gender norm support in old age for females: providing the parents with more support in terms of time or care rather than monetary support (Broday, 1985; Silverstein et al., 1995).

4.3 Identification strategy

The OLS results in both datasets do not appear to support the proposed demonstration effect. It may be that the results under the OLS model suffer from biases caused by potential endogenous problems. One main endogeneity problem comes from the gender selection issue affecting the gender ratio of the children, $genderK$. According to the China Population and Employment Statistics Yearbooks, the yearly national-level gender ratio of new-borns has been increasing since the late 1980s.[17] The national gender in 2011 shows the ratio of boys to girls to be as high as 1.25 to 1, revealing the gender selection problem as quite severe. Households with a preference for a son would be likely to conduct selective abortions, and these are usually the households holding the traditional stereotypes of daughters. Households with modern views on gender equality are less likely to select their children's gender. In my sample, the gender ratio of the parents is almost free from this problem. In CHARLS the average age of the parents in the sample is 40 and in CHFS it is 48. It is around 0.51 in both datasets. When they were born, gender selection technology was not yet available in China (Chen et al., 2013). The endogeneity problem of $genderK$ may affect the OLS outcomes in two opposite ways as illustrated by males with a preference for sons. First, if a male is eager to have a boy only to secure his own future support, then gender-selection will lead to an upward bias for the father demonstration effect. Second, if, alternatively, a male wants to have a boy to enhance the household's prosperity, he will invest more family resources in a son's upbringing. So the father effect is biased downwardly.

To alleviate the bias, I use the timing of a regulation announced in late 2002 by the Ministry of Health, State Food and Drug Administration (SFDA) together with the National Family Planning Commission (NFPC), together with the timing of its implementation. The regulation bans the use of B-scan ultrasonography and other technologies for determining foetal sex from January 1st 2003.[18] It states that all methods of gender selection should be banned and imposes fines for different levels of violation of the regulation. Fines are imposed on individuals who choose the sex of a foetus allowed

[17] The yearly national-level gender ratio of new-borns is shown in Figure 2.
[18] Website: http://www.gov.cn/banshi/2005-10/24/content_82759.htm. Last accessed: September 2018.

to survive and on the hospitals that conduct scans and abortions. The policy was intended to make the gender of the children born in or after 2003 relatively random, unlike that of the children born before. The policy was designed to reduce the gender ratio of new-born males to females, so it would be relevant to the average gender ratio of children in households, which is the variable $genderK$ in the main regression equation. Figure 3 shows the estimated yearly gender ratios of new-borns using the 2011 wave in CHARLS and the estimated yearly gender ratios of the first-born children using the 2013 wave in CHFS. This graph shows that both gender ratios fall after the year 2003.

I use mainly the timing of the policy change to construct the first part of the instrumental variable employed in the paper. The policy covers most of the provinces, and the provincial congresses passed the policy at much the same time,[19] with no great time difference between them. I assign the value of the policy timing variable to 1 for households with at least one child born in or after 2003, and 0 otherwise. The increasing gender ratio of male to female new-borns is a heated social issue that usually attracts public attention. So public discussion may accompany the agenda-setting process of the policy. However, Hu (1998) and Shen (2008) declare that detailed information and plans are rarely revealed to the Chinese public in the policy planning stage. Thus, the timing of the policy implementation is exogenous to the general public. Regarding this policy, in particular, most of the news about it on Baidu.com or Google.com appears after the provincial governments or the central government passed the associated regulation. In this setting, the exogeneity assumption of the policy timing is reasonable.

Although Figure 3 shows the gender ratio in CHARLS and CHFS to be decreasing, the situation is not quite the same as in Figure 2. The national gender ratio has been stagnating at a high level since 2003, although has not increased since then. The figure implies a slight chance that the policy does not ban sex-selective abortions outright.[20] To address this concern, I combined the dummy indicating the timing of the policy together with the gender of the first-born child in the households surveyed. The gender ratio of the oldest child in a family is relatively balanced in China. The One Child Policy (OCP) does not strictly require all households to have only "one child", especially in rural areas, so the first child's gender is relatively close to the natural ratio of new-borns (Ebenstein, 2010). In Figure 4, the graphs show the ratio of new-born boys who are not the eldest to their girl counterparts are all larger than the proportions among first-born babies.[21] For the relevance condition for this variable, the gender of the oldest child is usually correlated with the gender ratio of children in households (Angrist and Evans, 1998; Heath and Tan, 2018). Together with the timing of the policy, they form a dummy instrumental variable equal to 1 for households with at least one child born on or after 2003 where the oldest child is a son. The dummy is 0 otherwise.

This instrumental variable borrows the concept of the instrumented difference-in-differences design

[19] The provincial congresses all passed the policy at some time between November 2002 and January 2003. The information was collected from the provincial government websites.

[20] Because the policy did not make the gender ratio of new-borns completely random, I cannot use the subsample of households with new babies in or after 2003 to test the demonstration effect.

[21] The data comes from the National Population Census.

(DDIV) (Dulfo, 2001; Hudson et al., 2017).[22] The key variation comes from the policy compliers: those who were not allowed to conduct sex-selective abortions after the policy implementation, and had babies of the undesired gender. Take, for example, people with only one child and a preference for sons; assume they want to have a son but cannot conduct sex-selective abortions due to the policy ban. If they happen to have a son by natural chance, they are not the compliers that I expect under this policy. The compliers are people of the same type who have a daughter eventually. The gender ratio of the first child in the compliers' households will decrease after the policy implementation beyond the gender ratio of the first child before the policy change. The constructed instrumental variable is used for two datasets. As noted above, CHARLS gives limited information on the children of the parents that it surveys. Hence, constructing the gender of the first child in a household using CHARLS entails a few assumptions, which are included in Appendix 8.

One additional assumption that should be stated is that the support for the elderly provided by the parents does not change over time after controlling for the demographic variables, because the outcome variables in the DDIV are usually also time-varying. Due to the data limitation, I manage to get only cross-sectional datasets, so I use the CHFS dataset to compute the average probability that the group of parents who have their last child in the same year will provide support for the elderly. If there is no increasing trend in these averages in the different years of the last childbirth, this shows that the DDIV assumption is likely to be satisfied in the datasets. The graphs for plotting the "time-trend" are shown in Figure A.2 in the Appendix. They show that for the parents' generation, there is no significant decrease in the trend in the year of birth of the last child in households until the last 2 years before 2013.

I also construct another instrumental variable to proxy for the household-level gender ratio for CHARLS only. It is the prefectural-level compliance index of the policy implementation. Only CHARLS has the detailed names of the different prefectural-level cities. One of the components included in the index is the time when the provinces included the policy change in their provincial-level Regulation on Population and Family Planning. The policy change was announced in late 2002, and the actual implementation date was in early 2003. The time when the policy was introduced in the provincial regulation may indicate the level of compliance in the different provinces. Another component included in the index concerns a campaign in early 2005 initiated by the Ministry of Health with NFPC targeting illegal clinics and under-qualified doctors in several prefectures.[23] The illegal clinics are usually ones which illegal conduct sex-selective abortions. The policy acts to complement the policy ban of 2003. Either the central or the provincial governments decide to use this top-down approach because the local governments may have better control of the actual implementation of the campaign. Different prefectural-level cities have different ways of executing this campaign. Some

[22] The use of the interaction term of the gender of the first child and whether a household is affected by the policy is necessary. I cannot use only the subsample of households that is affected by the policy ban r use the gender of the first children as IV. This is because, even with the policy ban, the gender ratios in some provinces are still higher than the natural rate. A more detailed explanation and the subsample regression results are shown in Appendix 8.

[23] Website: http://www.gov.cn/zwgk/2006-08/02/content_352694.htm. The regulation date was in 2006, but in the content, it states that the campaign started early in 2005.

cities have mounted this campaign every year since the campaign started. Others may have implemented the campaign in 2005 for only one year or may even have started the campaign later than the NFPC requirement. The number of years that a city has enforced the campaign and also the year each started to do so are indicators of the strictness with which the regulation was implemented at prefectural level. I take the relevant information from the various prefectural government websites and also from newspapers and generate an index showing the various compliance levels of the listed prefectural cities regarding this policy and this campaign. The constructed compliance index varies from 0 to 2, where 2 is the highest level of allegiance to the aims of the campaign.

To summarise, the instrumental variables used in the paper are the gender of the first child for households having at least one child in or after 2003 and the prefectural-level compliance index. The IV method exploits three facts: first, that the gender of the first child is closer to the natural rate than the total gender ratio for all new-borns; second, that the amongst the first-born children, the gender of those who were born in or after the year of the policy ban is more random; third, that the prefecture-level policy compliance level is higher when the gender ratio of the children in general is lower. The results from the IV regressions are shown in Tables 5 and 6. The first stage results are shown in Table A.10 in the Appendix. I first discuss the results in CHARLS and follow it by discussing the CHFS results. Again, $-\beta$ indicates the mother demonstration effect, and $\beta + \delta$ shows the father effect. The probit regression marginal effect results for the dummy dependent variables are shown in Table A.7. The probit results are consistent with the IV results.

In CHARLS, the coefficients of $maleP$ are consistent with the OLS results for the probability of providing regular support and the amount of regular support provided. For $any\text{-}transfer$, $nonregular$ and $log(nonregular)$, the coefficients for $maleP$ are negatively significant and have larger magnitudes than the OLS coefficients. $log(visitdays)$ also loses the significance of the $maleP$ coefficient, but it is still positive and marginally insignificant. The coefficients indicating the father demonstration are all significant and positive. Compared to fathers with daughters only, fathers with sons only are 7.9%, 2.6% and 5.6% are, correspondingly, more likely to provide support of any kind, regular support and non-regular support to their own parents. Concerning the absolute amount of different kinds of transfer, they provide 16.4% more as a regular transfer and 24.6% more as a non-regular transfer than do fathers with daughters only. They also pay 30% more annual visits to their own parents.

Negative and significant $maleP$ coefficients mean that mothers are actually providing more than fathers in the probability of providing any and non-regular transfers, and also the log of the amount of the non-regular transfer. For these outcomes, the coefficients indicate possibly the mother demonstration effect: $genderK$ are insignificant but all negative. For the probability of providing non-regular support, in particular, the mother demonstration effect is negative and marginally insignificant ($\beta = -0.065$, $p = 0.104$). It implies that mothers may also try to demonstrate filial piety to their daughters, as the fathers in CHARLS do, especially in the case of non-regular transfers. The results suit the traditional norm of old-age support as provided by daughters: they are not mainly responsible for the living expense of their parents.

The demonstration effect in CHFS is different from the father effect in CHARLS. As with the CHFS OLS results, fathers provide less than mothers. The maternal demonstration effect is stronger and more significant than the paternal counterpart. The coefficients for $genderK$ are negative and significant for the probability of providing any support, the total number spent in visiting and its logarithm. Mothers with only daughters are 5.5% more likely to provide any support to their own parents than mothers with sons only. They will also devote 20 (56.1%) more days per year to visiting their own parents. In CHFS, it is difficult to draw any conclusion about the father effect. The coefficients for $genderK + maleP \times genderK$ are insignificant, and the signs of these coefficients are also inconsistent.

The gender ratio of the third generation varies between 0 and 1. When interpreting the regression results, I assume a linear correlation between the increment in $genderK$ and its effects on the outcome variables. I use a dummy variable for the gender ratio of the young children in the households to check the robustness of the results. The dummy is $moresons$. It equals 1 if $genderK$ is larger or equal to 0.5 and 0 otherwise. The results are presented in Tables A.8 and A.9 in the Appendix. The coefficients are similar to and consistent with the ones in Tables 5 and 6.

One explanation may reconcile these different effects from different datasets. If a certain gender is on average providing more than the opposite gender provides for certain types of support, the people with this gender are more likely (albeit insignificantly) to demonstrate to their same-gender children. In CHFS, mothers on average provide more than fathers, and it is the other way around in CHARLS. The results from the mothers in CHARLS and CHFS fit the prediction. Yet for fathers in CHARLS, it seems that they are trying to demonstrate all types of old-age support to their sons, not only those types that fathers on average provide more of than mothers do.

The IV results from CHARLS and CHFS, they show a very interesting phenomenon. The fathers in CHARLS and the mothers in CHFS both demonstrate to their same-gender children. One possible explanation may be that CHARLS and CHFS focus on different samples. As shown in the summary statistics, one major difference between CHARLS and CHFS is the proportion of urban samples in each dataset. CHFS has a sample of which 65.2% lives in an urban area, while the sample in CHARLS contains 33.2% urban dwellers. In CHARLS OLS results, fathers in general support their own parents more than mothers do. This result is consistent with the hypothesis that sons in rural areas are still preferred for their propensity to provide old-age support. In China's rural areas, a higher proportion of people accept the traditional gender discrimination/stereotype, and females have less bargaining power in their households than males (Wang and Zhang, 2018). Urban areas contain more households with a single child than rural areas do as a result of the "1.5" Child Policy implemented in China (Rosenzweig and Zhang, 2009; Wang and Zhang, 2018).[24] If a singleton household has a daughter, both parents are more likely to demonstrate to this daughter so that they can look forward to receiving support

[24] The gender preference in CHFS is in Table A.11 in the Appendix.

when they grow old. Urban areas in China also have more opportunities for female labour market participants and more gender equality compared to rural areas. My predictions for the discrepancies in CHARLS and CHFS are an urban-rural difference and a singleton/non-singleton household difference. The significance of the female or male demonstration effect is driven by the corresponding subsamples with more observations. The results of a subsample check and heterogeneity analysis provide more empirical findings on these two conjectures - see below.

4.4 Subsample analysis and heterogeneity check

To verify the effect of the gender composition of K working mostly through the demonstration mechanism, I use results from the subsample analysis and the heterogeneity check to verify whether , in different circumstances, the results are still consistent with the predicted results from this mechanism. The analyses are conducted for both or only one of the datasets, depending on the available information. I mainly describe the subsample analysis results and then mention the consistency of the results with the corresponding heterogeneity checks. Since the CHARLS data mainly exhibits the father demonstration effect and CHFS mainly shows the mother effect, I focus only on the father effect in different groups from CHARLS and the mother effect in different groups from CHFS. Six categories are used for the analysis: income-level, singleton or non-singleton households regarding the children, urban or rural residence, the number of brothers of the parents, the pension coverage of the parents, and membership of the *Han*/non-*Han* ethnic group. The category for the singleton or non-singleton households and the urban-rural residence are the two categories that may explain the discrepancies between the result from CHARLS and from CHFS.

4.4.1 Income-level difference

As the future support for the elderly received from the offspring acts as an economic incentive to have children (Banerjee and Duflo, 2011; Alfano, 2018), households at different income levels will have different patterns for the demonstration effect. People in the high-income group will have enough savings, investments, and pension income to support their consumption after retirement. So, their incentive to demonstrate to their children by pecuniary support for the elderly is not as large as those who in the low-income group. For the financial old-age support, I expect more demonstration effects for people in the lower-income group than those with higher income. Regarding the non-pecuniary support, the high-income group may demand more than the other group, so more significant father and mother demonstration effects are expected for $log(visitdays)$ in the high-income group.

The subsample IV regression results for CHARLS are shown in Table 7. CHARLS only have a categorical variable of household income level of the parents. To get a balanced subsample in CHARLS, I classify those whose household income-level above the 20,000 RMB per year category as the high-income group. The father demonstration effect for pecuniary outcomes is more significant in the low-income group than in the high-income group. The father effects in the low-income group are significant for most of the pecuniary outcomes; while for the high-income, only the probability of providing any and non-regular transfer have significant coefficients for the father demonstration effect. For the nonpecuniary outcome, the father demonstration effect is also significant in both high and low-income group, but the magnitude of the effect is greater in the high-income group. The coefficient for $genderK$ for $log(visitdays)$ is significant for the high-income group, which means that more mothers in relatively rich households demonstrate to their daughters than in poorer households. The mothers in households with a low income do not show mother-dominated demonstration effects but rather visit their parents more than do the mothers of families with more sons. Apart from these results, other coefficients seem to be consistent with the prediction.

With the detailed income information in the CHFS data, I classify those who have above the average income in the high-income group and the rest of the sample in the low-income group. Table 8 hows that in the low-income group mothers demonstrate filial piety to their daughters only with regard to non-pecuniary outcomes; while in the high-income group, mothers with more daughters are more likely to provide support for the elderly and more likely to visit the mother's parents. This appears to contradict expectations. However, in the relevant literature, the parents are more likely to demonstrate to their daughters that the elderly should be supported (Wolff, 2001). This may be because females care more about family relationships than males do.

The heterogeneity check provides similar results to those of the subsample analysis. It can also check whether there are differences in the demonstration effect between the high and low-income groups. The results of the heterogeneity check for the income-level are shown in Tables A.13 and A.14 in the Appendix. The CHARLS results in Table A.13 show that the father demonstration effect is more common in the low-income group than in the high-income group. The difference in the father demonstration effect between these two subgroups is significant. Table A.14 shows that in CHFS the differences between these two income groups, though not significant, are mostly negative. The results are consistent with the subsample analysis.

4.4.2 The number of the children

This paper mainly focuses on the way that the gender of the children affects the support for the elderly provided by their parents. While the paper does not on the whole discuss other characteristics of the children, I can use the number of members of the third generation to conduct a subsample analysis and heterogeneity check. Most of the households with only one child ('singleton households') are the households that strictly comply with the OCP, even when they have an only daughter. These households may hold modern views of gender roles; hence, females in these households may be able to enjoy higher bargaining powers. A preference for sons is a good indicator of whether a household has more traditional views on gender roles. Such households are more likely to violate the OCP (or be allowed by "1.5" Child Policy) to have a second child if their first child is a girl. I expect more mother demonstration effects in singleton households. Tables 9 and 10 display the results for CHARLS and CHFS. The number of children may also correlate with the gender of the first child in the households, which may lead to biased results in my IV regressions. The subsample analysis of the number of the children also helps to get rid of the possible bias arising from selection in the number of children.

Table 9 shows a clear contrariety between single and non-single child families in CHARLS. The father effects show their significance only for non-singleton households, especially regarding the financial support in old age. As discussed in the previous section, given the OCP, households with more than one child are usually rural households or urban households with a relatively strong son preference. In rural areas, the OCP allows households in which the first child is a daughter to have a second child, and the law and the enforcement of fines in rural areas are not as strict as they are in urban areas (Ebenstein, 2010). According to Ebenstein (2010), if the OCP is violated in an urban area the fine is quite high. Non-singleton households usually possess a stronger preference for sons than singleton ones do; hence, females may have less bargaining power in the non-singleton households. The singleton households in CHARLS also show a significant mother demonstration effect for the probability of providing non-regular old-age support. The coefficient of $genderK$ for $log(nonregular)$ is also negative and marginally insignificant.

The CHFS results show more significant mother demonstration effects in singleton households. The mother demonstration effect in the full sample of CHFS is driven by the mother effect from this subsample. The father demonstration effects in the non-singleton households are insignificant, but they are more positive than the coefficients in the subsample of singleton households. The results of the heterogeneity check for the singleton and non-singleton households are shown in Tables A.15 and A.16 in the Appendix. The CHARLS results show that the father demonstration effect is on average greater in non-singleton households than in singleton households. Table A.16 shows that in CHFS the mother demonstration is greater in the singleton households. The heterogeneity analysis results are in general consistent with the subsample analysis. That females in the singleton households have higher bargaining power is one of my conjectures for explaining the difference between the CHARLS and the CHFS results. But I may need to interpret this together with the results from the urban-rural analysis.

4.4.3 Urban-rural differences

Another conjecture in explaining the discrepancies between the CHARLS and the CHFS results is the urban-rural difference. Urban areas in China enjoy more developed public pension systems, more opportunities for female to be employed and more gender equality. I expect less significant father demonstration effects and more significant mother ones in the urban subsample.

Table 11 presents the regression results for the urban and rural subsamples in CHARLS using the IV regressions. In the urban and rural areas, most of the gender effects of the children are insignificant for pecuniary and non-pecuniary outcomes of the mother demonstration effect. Another interesting finding in the urban subsample is that the coefficients for $maleP$ are all insignificant; while in the rural subsample they are all significant. The empirical findings may indicate that the gender norms as regards support for the elderly are not strong in urban areas compared to rural areas. The subsample results from CHARLS seem to support my prediction.

The difference in the gender effects of the children between rural and urban areas in CHFS also corresponds to my prediction. Table 12 shows that the mother demonstration effect is significant for most of the outcome variables in the urban subsample. In the rural subsample, there is no demonstration effect for mothers to their daughters nor fathers to their sons. The coefficients of $genderK + maleP \times genderK$ larger and more positive than the corresponding coefficients in the urban subsample. Moreover, in the rural subsamples, $maleP$ has two positive coefficients for the days spent in visits. This indicates that the gender norm of providing support for the elderly may be different between rural and urban areas. Although these coefficients are insignificant in rural areas, there may still be gender-role differences concerning the demonstration effects. The heterogeneity check results from two datasets in this part are also consistent with the subsample analysis results.[25]

The urban-rural subsample analysis supports my prediction of more mother demonstration effects and fewer father effects in urban areas. Scholars believe that females have higher bargaining power in urban areas in China (Fong, 2002). However, certain urban households where the first-born is a girl would pay the high fine to have a son (Ebenstein, 2010). Lee (2012) and Hu and Shi (2018) find that the human capital investment for boys and girls is not significantly different in singleton households, but the gap is still wide in multiple-child households. Fong (2002) also limits the rising female empowerment in urban China only to daughters in singleton households. I run a simple urban-singleton and other types of household subsample in CHARLS. I find there is no significant father demonstration effect in the urban-singleton subsample, except with non-pecuniary outcomes. The signs of the male effect are inconsistent between different outcome variables. In the CHARLS subsample of non-urban-singleton households, the father demonstration effects are again significant for all of the outcomes. The CHFS subsample results also show significant mother demonstration effects in the urban-singleton group for all outcomes. The mother effects in urban-singleton households are larger than the corresponding effects in urban households. The results for this simple subsample are shown in Tables A.20 and A.21 in the Appendix. The discrepancies between main results from using CHFS and CHARLS can be explained by the results for the urban households, especially the urban-singleton households: these households drive the significant mother or insignificant father demonstration effects.

4.4.4 Siblings of the parents

Supporting ageing parents is crucial for most males in China owing to the enduring cultural impact of Confucianism. Some parents have to support their own parents, regardless of the gender of their children. This is especially the case for males who provide regular support to their own parents. It

[25] Tables A.17 and A.18 in the Appendix.

may also be the case for some females in the second generation if they are the oldest child in their own family or have no brothers. If people are not fully responsible for the support of their elderly parents and they want only to demonstrate the norm of providing support for the elderly to their children, the results may show greater effects from the gender ratio of the children. I use the same regression equations and the identification methods that I used in the previous section, and I run regressions separately on those who are with and se who are without older brothers. The results are shown in Table 13. CHFS provides only the number of siblings for the main respondents in households, but no information on his or her rank in the siblings. So this subsample analysis is conducted only using the CHARLS data set.

The results indicate that, for the probability of providing regular support and also the days spent in visiting their parents, the father demonstration effects are more significant or greater for those with older brothers than for those without. Yet, with regard to non-regular support, the father effect is more likely to exist among those who have no brothers. The type of difference in the support for the elderly given by people without older brothers can explain why the results do not conform to the prediction. It is because the oldest son in the family is responsible for providing regular support but not non-regular support. So people without older brothers are more likely to show the demonstration effect in their non-regular support, but not in regular support. For mothers, the results still imply that for those with or without older brothers, the gender of their offspring does not have any significant effect on either pecuniary or non-pecuniary outcomes. From the results, I may also infer that the burden on females of supporting elderly parents is not as heavy as it is on males. The heterogeneity results are shown in Table A.19, and the results are more significant than the subsample analysis. The father demonstration effect is significant only in the group without brothers. But the differences between the group with and the group without older brothers are insignificant for non-regular outcomes, which are consistent with the subsample analysis.

4.4.5 Pension coverage

In the introduction, I treat family support for the elderly as a complement of the public pension scheme. As Table 1 shows, in rural areas where the public pension coverage is low family-provided support for the elderly is the primary source of support of China's elderly according to 48% of the NBS survey respondents. So I expect that parents without proper pension coverage of their own are more likely to provide more support to their own elderly parents if they have more same-gender children. The demonstration effect will be higher or more significant for parents without good pension coverage.

To check this hypothesis, I run a heterogeneity analysis on parents with and without a pension scheme. In CHARLS, due to the data reconstruction, I have no information on the parents' pension coverage. However, I can use the occupation of the parents as a proxy for their pension status. CHARLS provide six categories of occupation for the parents, namely, managers; professionals and technicians; clerks, commercial and service workers; agricultural, forestry, husbandry, and fishery producers; and production and transportation workers. Of these six categories, the agricultural, forestry, husbandry, and fishery producers are less likely to be covered by a pension scheme. I create a dummy that equals 1 if a parent is not classified as an agricultural, forestry, husbandry, or fishery producer, and 0 otherwise. The results from this heterogeneity analysis are shown in Table 14 and they make it clear that the father demonstration effect is more significant for parents if they are less likely to be covered by a pension system. Yet the difference between the father demonstration effects in the group with pension coverage and without is insignificant. The empirical results from CHARLS fit the description of the relationship between a public pension scheme and family old-age support.

In CHFS, the information is available for defining the exact pension status of the parents. I create a dummy which equals 1 if a parent is covered by a pension scheme, and 0 otherwise. The results are shown in Table 15. They are also consistent with the prediction: the mother demonstration effect is more significant for parents without good pension coverage. Yet mothers both with and without pension coverage have negative coefficients corresponding to the mother demonstration effect. The difference in terms of visits paid to the grandparents is also significant for mothers with pension coverage and mothers without. From the heterogeneity analysis in Tables 14 and 15, the results provide a piece of suggestive evidence on the relationship between public pension schemes and family support for the elderly suggested previously in the paper.

4.4.6 *Han* culture and norm

As discussed in the background section, the norm of providing support for the elderly is closely linked with Confucianism and filial piety. This raises a possible concern: because the culture of Confucianism is well-known in Chinese society, not only do parents teach their children to provide support for the elderly in the future through the demonstration effect, but also the surrounding community, in schools, the neighbourhood, or the media (i.e. television programmes), could shape young children's predilection to provide support to their parents in their old age. If other channels affect children's preferences regarding old-age support, the demonstration effect from the parents will be smaller or less significant in a *Han*-ethnic dominated community or an exclusively *Han*-ethnic group.

There is no detailed ethnic information about parents in CHARLS. However, in the community survey questionnaire in CHARLS, there is information on whether minority ethnic groups are living in the same community that the parents live in. I generate a dummy that equals 1 if there are minority ethnic groups living in the community, and 0 otherwise. From the results in Table A.22 in the Appendix, the father demonstration effect for most of the pecuniary outcomes in communities with people from minority ethnic groups is significant and greater than the father effect in communities without. However, when it comes to non-pecuniary outcomes, communities without people from minority ethnic groups have a significant father demonstration effect.

The detailed information on individuals' ethnic groups can be obtained in CHFS. I create a *Han* dummy that equals 1 for members of the *Han* ethnic group, and 0 otherwise. In the heterogeneity analysis results in Table A.23 hown in the Appendix, the mother demonstration effects are insignificant for pecuniary support for the elderly in all ethnic groups, whether *Han* or *non-Han*. But the mother demonstration effect in the *non-Han* group is greater than the effect in the *Han* group. For the non-pecuniary outcome variables, it is the *Han* group that has a greater and significant mother demonstration effect. Combining the heterogeneity check results from CHARLS and CHFS, it seems that there still are possible mother and father demonstration effects in *non-Han* ethnic families and communities with *Han* households only.

5 Robustness check

5.1 Mechanism check

There are other different channels that may explain the effects of children on the support for the elderly provided by their parents. For example, the education investment in the children could be one possible explanation for the mother demonstration effect. The significant mother demonstration effect in CHFS may result from the fact that mothers with daughters are less likely to have a human-capital investment in their daughters so that they can provide more for their parents. However, this argument does not work for the father demonstration effects because the human capital investment in sons in China is on average higher than the investment in daughters, except for urban singleton households (Fong, 2002). In addition, the evidence from CHFS shown in Table A.24, suggests that mothers do not invest less in their daughters, as the argument claimed. Testing the subsamples may show the effects of children's gender on the parents' support for the elderly provision; hence, I check other mechanisms discussed in the literature review section and examine which mechanism is working as the main one in providing support for the elderly in the results. I first discuss the channels of altruism and direct reciprocity that may affect my empirical results, and go on to discuss the effectiveness of the demonstration effect.

5.1.1 Altruism and Direct reciprocity channel

The first possible mechanism is altruism. If the main mechanism is pure altruism, the only reason behind the parents providing support to their own elderly parents is that these parents are poor and in need of help. There should not be any significant coefficients for the gender of the adult children, the gender ratio of the children or their interaction term after controlling for the income of the elderly parents in the regression. I run heterogeneity checks on the elderly parents' income-level as shown in CHARLS. In the sample, most of the elderly parents observed have no income, so I create a dummy $income\ of\ O$ which equals 1 if the elderly parents have some income, and 0 otherwise. The results are shown in Table 16, and and they reveal that for most of the outcome variables, the father effect is significant even for elderly parents without any income, whereas for the high-income group, the effects are positive but insignificant. The difference between these two groups is also insignificant. I may draw the conclusion that there is a certain degree of altruism among the motives of providing support to one's elderly parents, but the demonstration effect also exists at the same time.

Another mechanism discussed in the previous section is direct reciprocity. One kind of direct reciprocity in the context of support in old age is the parents' desire to support ageing parents to repay the investment in their childhood. It is sequential direct reciprocity. It may also explain why females provide less support to the elderly to their parents because, according to CHARLS, they did not get enough financial or non-financial investment from their parents during their childhood. Only CHARLS includes this type of information, so I use only this data set to check this mechanism.

If sequential direct reciprocity is the only channel for old-age support to flow along, then controlling in the regression for the same financial and non-financial investment received by the parents in their childhood should confirm that males and females in the parents' generation should provide the same amount of old-age support. Moreover, the gender of the children should not have different effects on the transfers provided by the parents. I control for different variables indicating in the main regression the financial investment and nonfinancial investment that the parents received during their childhood and the results of doing so are shown in Table 17. *awaytime* is the variable representing how long a parent has been away from his or her parents in childhood and *awayage* indicates the age when the parent left her/his parents. These two variables represent the time investment (non-financial support) during the parents' childhood.

I also show the coefficients for *edu* in the table, which is the education level controlled in the main regression. It is another indicator of the size of the financial investment. Table 17 shows that after controlling for the non-financial and financial investment, the results are still similar to the results in Table 5. The significance level and the magnitude of the coefficients have minor decreases, which may imply that there are some effects from the direct reciprocity channel. But, with most of the coefficients representing the father demonstration effect being still significant, it also suggests that the same-gender demonstration effect also works as one of the potential channels. In addition to the results appearing in Table 17, the CHFS main results may also demonstrate that this direct reciprocity channel is not the main mechanism. Mothers in general provide more to their own parents in CHFS, given the fact that females on average have a lower education level than males.

Another direct reciprocity channel is the transfers from the elderly parents to the parents in the same period. In the previous regressions in CHARLS and CHFS, I control the transfer from the elderly parents to the parents. This variable would, in theory, have positive effects on the outcome variable, and vice versa. I also control for the time that the elderly parents spend on taking care of the children of the parents and also the transfer to the children in the regressions in CHARLS. For the robustness check, I show the regression results without these controls in Tables A.26 and A.27, also their corresponding coefficients in Tables A.28 and A.29. The key results are similar to the main results. The coefficients for these transfers exhibit positive and negative effects on the transfers provided.

The rationale behind this direct reciprocity is that if the parents with the same-gender children also receive more from their elderly parents than those without receive, then they also provide more old-age support. However, when I run the same regression on the transfer received by the parents from their elderly parents, the results appearing in the second and the fifth column of Table 18 show that people who provide more to their elderly parents, namely fathers with more sons, receive less. The results may fit the explanation by Li et al. (2010): the elderly parents show more altruism toward the parents who do not provide more transfer than others, rather than expecting commensurate paybacks from the parents who receive their support. To conclude, direct reciprocity may exist, but there is still room for the proposed mechanism: the demonstration effect.

The results from Tables A.28 and A.29 are also interesting. The coefficients for both time and financial transfer from elderly parents to their grandchildren are positive for most of the outcome variables, which may suggest another form of indirect reciprocity. The ageing parents can transfer to their favourite grandchildren. If the favourite grandchildren receive more, their parents (the parent generation) are more likely to provide support to the ageing parents in return. This type of indirect reciprocity has no time lag for the payback, unlike the demonstration effect studied in the present paper. Usually, the preferred grandchildren are grandsons. This could be one explanation of the father demonstration effect in CHARLS. If the indirect reciprocity works in this way, I would expect male parents with more male children to receive more transfers from their elderly parents. However, the third column in Table 18 shows that they do not statistically receive more than males with more daughters. These grandchildren gender effects are not significant for transfers from elderly parents.

5.1.2 Effectiveness of the demonstration effect

Apart from verifying the possible channels, I also have to test for the actual demonstration effect in the datasets. The parents expect their children to provide support for the elderly in the future. The previous results imply only that the parents demonstrate filial piety to their children, but they do not show whether the children actually go on in the future to provide old-age support to their parents. Using the CHARLS dataset only, I obtain the information on support in old age that was provided by the elderly generation to their own parents, who are the grandparents of the parent generation. I run a simple OLS regression to regress the upward-transfers of males and females among the elderly parents to their own parents on the outcome variables in the main regression. I run the regression separately for male and female parents. The types of transfer provided by the elderly parents to their own parents on the left-hand side of the equation also match the corresponding dependent variables. Take, for example, the regressions for $log(regular)$, two key regressors, $father's\ transfer$ and $mother's\ transfer$, these are the logarithm amount of the regular transfer provided by the father and mother of the present parents' generation to their own parents in an earlier sequence. The control variables are the same as the controls in Table 2. One extra control that I have for the particular regressions is the average self-reported health of the grandparents of the parents. The health problems of the elderly may affect the support provided, so I also add to the controls one for the health of the grandparents of the present generation of parents.

To simplify the description of the results, I continue to use the O, P, and K setting in this subsection. The results are combined in Table 19. The key regressors for male and female P panels are $father's\ transfer$ and $mother's\ transfer$. For male and female P, the demonstration effects seem to take into account the effects from the same gender channel: males are more affected by the support for the elderly provided by their fathers than their mothers' and the converse is true for females. The same-gender demonstration effect is more significant for female members of P. The magnitude and also the significance level for $father's\ transfer$ are much smaller than the $mother's\ transfer$ for females, while for males, the difference is not very great. The results show that if the members of O provide more to their parents, they are more likely to receive more from P.

5.2 Panel results: Event study

The main regression results mainly show the cross-sectional empirical evidence of the demonstration effect. The conclusion will be more convincing if there is empirical evidence from a panel dataset. Both CHARLS and CHFS are longitudinal datasets, but CHARLS does not provide information on the gender of the children for the whole sample in the 2013 and 2015 wave. CHFS contains this necessary information in the 2013 and 2015 wave. Although the dataset has only two waves, the results may have some implications for the way in which the behaviours of the parents change after the advent of a new child in their household. Because of the limited number of waves, I use only the panel result as a robustness check for the main results.

To examine the yearly effect of having a son or a daughter on old-age support, I use a quasi-experimental event studies approach. The event is the birth of the first child. The event usually causes sharp changes in several outcomes for the parents, especially labour market outcomes (Kleven et al., 2018). I use a similar event study approach to that used by Kleven et al. (2018) and aim to show possible causal results in the event study approach. The dataset that I used for the event study is still the CHFS dataset. However, in this section, I use the panel dataset including the waves in 2011, 2013, and 2015. The reason for using this three-wave dataset is to gain more yearly data before and after the event. The drawback of using this data is that I can test the demonstration effect on only one consistent outcome variable - the probability of providing old-age support - for three different waves.

In the three-wave panel dataset, the sample is still limited to household respondents. Given the event study approach setting and the limited number of waves for the data, the panel sample includes only those respondents whose first child was born between 2011 and 2015. For each household respondent, I set the event time $e = 0$ for the year in which the respondent has his or her first child.

The value of other years is set relative to the $e = 0$ year. Using the specification in Kleven et al. (2018), the regression is:

$$y_{ite} = \sum_j \alpha_j \times \mathbf{I}[j = e] + \sum_k \beta_k \times \mathbf{I}[k = age_{it}] + \sum_l \gamma_l \times \mathbf{I}[l = t] + \epsilon_{ite}, \tag{8}$$

where i stands for individual i, t for wave t, and e for the event time e. y_{ite} is the probability of providing support to elderly parents. $\mathbf{I}[j = e]$ represents the event time dummies, $\mathbf{I}[k = age_{it}]$ is for the age dummies, and $\mathbf{I}[l = t]$ is the wave fixed effects. By controlling the age dummies, I can control the non-parametrical underlying life-cycle trend (Kleven et al., 2018). I run this regression separately for four different groups: fathers with a first son (father-son), fathers with a first daughter (father-daughter), mothers with a first son (mother-son), and mothers with a first daughter (mother-daughter). Then I compare the results within the same-gender parents and observe that the difference between having a first son and having a first daughter is significant. The reason why the results may be causal is that I examine the variation in the results caused by the gender of the first child. As noted in the previous section, the gender of the first child is almost exogenous. In addition, the timing of the birth for the first child is after 2003, which is after the ban on the use of ultrasonography techniques for sex-detective abortions.

The graph for the plot of the event time dummies coefficients is shown in Figure 5. The graph on the left shows the difference between fathers with a first son and fathers with a first daughter. The right graph is the difference between mothers. After the birth of a first child the mothers with a first daughter provide more than those with sons, whereas the difference between fathers is almost zero. For the pre-trend of the event study, I can only observe one period before the birth of the first child in the panel dataset due to the limitations of the data. But from this one-period pre-trend result, it seems that for mothers, the pre-trend difference is smaller than the differences after the event. For fathers, the pre-trend difference is greater but insignificant. Lack of the pre-trend time period will affect the validity of the inference and the causality of the event study results. But the results may provide some insights into the effects of the gender of the children on the same individual. The regression results are shown in Table 21. The sample size for each group is around 800 observations, which also indicates that the gender of the first child in the event study sample is satisfactorily balanced.

5.3 Other robustness checks

Different outcome variables: In the previous regression equations, the outcome variables regarding the amount of the transfer are the logarithm amounts. The logarithms help to reduce the sensitivity of the results caused by the outliers, which are common in survey datasets. But they also add to the difficulty of interpreting the exact effects of the key regressors. For both datasets, I run Equation (7) on the new outcome variables for the amount of the transfer: the absolute amount of the transfer and the amount of the transfer as a percentage of total income. The results are shown in Table A.25 in the Appendix. For the CHARLS results, the father demonstration effect for the percentage of income appears to be consistent with the results in Table 5, although with an 88% significance level. The absolute amount of the non-regular transfer has the significant mother demonstration effect that is consistent with the main results using CHARLS data. The father demonstration effects for the two absolute values of the transfer in CHARLS are both positive and insignificant. With CHFS, the results in Table A.25 show the insignificant but negative mother demonstration effect for the percentage outcome of any transfer provided by the parents.

Furthermore, the transfers from the elderly are not included in the construction of the outcome variables used in the main regressions. I change the transfer outcome variables to net transfer variables. If $any\ transfer$, *regular* or *nonregular* equals 1 and the parents receive the transfers from or are living together with their elderly parents, I change the corresponding value to 0. For the amount of monetary transfers, I use the net transfer provided by the parents, which is the amount of transfer provided to the parents minus the amount of the transfer received by them from their elderly parents. The change is made for both datasets. The results for the net transfers are shown in Table 22. They are consistent with the main results, except for the negative father demonstration effect for $any\text{-}transfer$ in CHFS. The magnitudes of the demonstration effect also increase beyond the main results.

Cohabitation with the ageing parents One important way of supporting ageing parents is to live together with them. Although this may count as mutual care of the family members, it seems that parents are more likely to be taking care of their ageing parents with respect to income earning and the ability to offer a home. In the literature, cohabitation with one's ageing parents is generally used as an outcome variable. In my main results, the probability of providing monetary support and the outcome variable *visit days* partially captures cohabitation. I use cohabitation as an outcome variable in the robustness check. Similar regression equations are applied but without controlling for parents as members of the sae household member as their elderly parents. The prediction of the results would be similar: the same-gender demonstration effects of cohabitation. The results are shown in Table 20. Both mothers and fathers are more likely to cohabit with their own parents to demonstrate filial piety to their same-gender children. The same-gender demonstration effect is more significant for this outcome variable than the main results.

6 Welfare analysis

6.1 "Missing women" and the old-age support

As pointed out by Qian (2008), the future income of children will affect their gender ratio and parents expect in their old age to receive support from their children . My empirical results show the causality of the support for the elderly provided by the parents and the gender of their children. It may be possible to draw some inferences on the correlations between the support for the elderly and the gender of children from the literature and my results. The results in the paper show that, in rural areas, males provide more in general, and the effects are more persistent for households with consistent male heirs over several generations. I may be able to argue that support for the elderly is at least correlated with the gender ratio, or the "missing women" in China may be correlated with the demand for support in old age. I follow Oster's method in 2005. She calculates the number of "missing women" in China due to hepatitis B infection. She estimates the effect of prevalent hepatitis B on the male-female gender ratio and calculates the hepatitis B-adjusted gender ratio using the percentage of the population infected with hepatitis B. She draws the conclusion that hepatitis B accounts for 75% of the "missing women" in China. Her results are not entirely accurate due to the data that she collected, but her method provides a reasonable estimation strategy to evaluate how much the unbalanced gender ratio can be correlated with hepatitis B infection. I use her method of estimation to measure the possible correlation between the "missing women" and the need for support in old age.

I use CHFS to conduct the estimation.[26] One of the advantages of using CHFS is that the dataset provides people's attitudes on family, children and support for the elderly. There are two relevant questions in the survey: "Do you prefer daughters or sons?" and "Who do you think is responsible for your care in old age?". I created a dummy that equals 1 if people prefer sons and believe that their children should be responsible for supporting their elderly parents and 0 otherwise. I obtain the mean of this dummy in the full sample in the main regressions, the urban subsample and rural subsample. Running the dummy on the gender ratio of the children, the coefficient for the dummy, which is the prevalence of old-age support, is shown in the second column of Table 23. Then I calculate the estimated gender ratio on the sole basis of the prevalence of support for the elderly shown in the third column. Given that the natural gender ratio is 1.049 and the percentage of people who prefer sons and who believe that their children should be responsible for their old-age support, I calculate the adjusted gender ratio in column 5.[27] The equation for the adjusted gender ratio is $GR_{adjust} = GR_{old-age} \times percentage_{old-age} + GR_{nonold-age} \times (1 - prevalence_{old-age})$ (Oster, 2005). The percentage of the gender ratio correlated with the needs of the future support for the elderly is listed in the last column. From the estimations, around 12%-18% of the unbalanced gender ratio is correlated with the needs for support in old age. 17.23 million Chinese babies were born in 2017, and the medium gender ratio for 2015-2017 is 1.150 according to the world population prospects in 2017.[28] If people conduct gender selections to secure future support in their old age, the number of "missing girls" due to this would have been around 93,000 in 2017.

[26] CHARLS does not contain many questions on people's ideology, so I cannot distinguish those who have a preference for sons from whose who do not.

[27] The percentage of people who prefer daughters or who have no preference, and who believe that their children should be responsible for their support in old age is 34.38% for the full sample, 47.48% for the rural subsample and 18.39% for the urban subsample.

[28] The source for these data can be accessed through the following websites: https://www.statista.com/statistics/250650/numberof-births-in-china and http://data.un.org/Data.aspx?d=PopDiv&f=variableID%3A52.

7 Conclusion

The existence of a younger generation plays an essential role in parents' decisions on the support that they provide for the elderly. This paper finds that the gender of the children in China affects the support for the elderly provided by their parents. The parents are more likely to provide more financial and non-financial support to their ageing parents when they themselves have more same-gender offspring, which is the demonstration effect. However, the demonstration effects by mothers and fathers are exhibited in different kinds of area in China. Rural areas show the father demonstration effects while mother demonstration effects appear in urban areas. The urban-rural difference may be due to female empowerment in urban areas, but this needs to be verified by future studies. The demonstration effect is a way for the norm of providing support in old age to be conveyed to future generations. The intergenerational transmission of norms is also gender-specific.

This paper also provides a model with which to describe the same-gender demonstration effect. The model is based on a simple inter-temporal consumption model with additional assumptions on same-gender intergenerational transmissions and intra-household resource allocations. The paper predicts that support for the elderly provided by a father increases when more sons are added to his family and when he has greater bargaining power than his wife. However, the support for the elderly provided by mothers increases with the advent of more daughters and when mothers earn more income. The empirical results of the gender ratio for the household's children match the prediction s of the model. In China, urban females have more bargaining power in their households than females in rural areas have. The findings indicate that the mother demonstration effect mainly shows up in the dataset which has more urban samples. The heterogeneity analysis for the single-child households further suggests that the assumption of intra-household bargaining is valid.

The empirical evidence shows that the gender of the parents and their children in China jointly affect the likelihood and the amount of old-age support, both financial and non-financial, that they provide. The story behind this is more complicated than any pure gender effect from the children. The proposed mechanism, with same-gender intergenerational transmission, is indirect reciprocity, or the demonstration effect. It carries the social norm of providing private support for the elderly across the generations. Given the heavy financial burden of the public pension system facing the central government in China, the government has realised that private support for the elderly is a crucial complement to the public pension. In 2017, the central government started a pilot implementation of "homebased old-age care services". One of the expected goals of this pilot implementation is to collect information on the demographics of all households with ageing parents and use the information to set future policies or incentives for completing the home-based system of care services for old people.[29] The empirical results in the present paper can offer some insights into the demographics of those who provide or do not provide support to their ageing parents: policy-makers could introduce diverse

[29]Website: http://xinhuanet.com/gongyi/yanglao/2017-04/17/c_129543350.htm

incentives in order to target different groups. The rural-urban discrepancies in the results will also help the government to set targeted policies in rural and urban areas.

Although the Chinese government has become aware of the importance of private support for the elderly and has started to promote "filial piety", there may be a hidden hazard behind this action. As this paper shows, sons in the rural areas in China provide more support for the elderly than daughters do. The previous literature also states that economic incentives, especially old-age support, provide one reason for sex selection before birth (Qian, 2008; Ebenstein and Leung, 2010). The gender ratio might stagnate at a high level, to create a damaging equilibrium,. The government needs to promote gender equality by legislating to protect the right of females to inherit, own property and compete in the labour market, especially in rural areas. In urban areas, there is already a healthier balance in the gender ratio of new-borns. Mother demonstration effects showing in urban areas alone may also be due to female empowerment and higher bargaining powers in the household for females. More research is needed to confirm this possible mechanism.

References

[1] Alberto Alesina, Paola Giuliano, and Nathan Nunn. On the origins of gender roles: Women and the plough. *The Quarterly Journal of Economics*, 128(2):469–530, 2013.

[2] Alberto Alesina and Paola Giuliano. Chapter 4 - Family ties. In Philippe Aghion and Steven N. Durlauf, editors, *Handbook of Economic Growth*, volume 2, pages 177–215. Elsevier, 2014.

[3] Macro Alfano. Daughters, dowries, deliveries: The effect of marital payments on fertility choices in India. *Journal of Development Economics*, 125:89–104, 2017.

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

[5] Luc Arrondel and André Masson. Family transfers involving three generations. *The Scandinavian Journal of Economics*, 103(3):415–443, 2001.

[6] Luc Arrondel and André Masson. Altruism, exchange or indirect reciprocity: what do the data on family transfers show. In Serge Christophe Kolm and Jean Mercier Ythier, editors, *Handbook of the Economics of Giving, Altruism and Reciprocity*, pages 971–1053. North Holland., 2006.

[7] Nava Ashraf. Spousal control and intra-household decision making: An experimental study in the Philippines. *American Economic Review*, 99(4):1245–77, 2009.

[8] Joan Baker. The history of sonographers. *Journal of Ultrasound in Medicine*, 24(1):1–14, 2005.

[9] Abhijit Banerjee and Esther Duflo. *Poor economics: a radical rethinking of the way to fight global poverty.* New York: Public Affairs, 2011.

[10] Abhijit Banerjee, Xin Meng, and Nancy Qian. The life cycle model and household savings: Micro evidence from Urban China. Technical report, Yale University, Working Paper, 2010.

[11] Abhijit Banerjee, Xin Meng, Tommaso Porzio, and Nancy Qian. Aggregate fertility and household savings: A general equilibrium analysis using micro data. Technical report, NBER Working Paper No. 20050, 2014.

[12] Robert J. Barro. Are government bonds net wealth? *Journal of Political Economy*, 82(6):1095–1117, 1974.

[13] Robert J. Barro and Gary S. Becker. A reformulation of the economic theory of fertility. *The Quarterly Journal of Economics*, 103(1):1–25, 1988.

[14] Natalie Bau. Can policy crowd out culture? Technical report, Harvard University, mimeo, 2016.

[15] Gary S. Becker. Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology. *Journal of Economic Literature*, 14(3):817–826, September 1976.

[16] Gary S. Becker. Altruism in the Family and Selfishness in the Market Place. *Economica*, 48(189):1–15, February 1981.

[17] Gary S. Becker and H. Gregg Lewis. On the interaction between the quantity and quality of children. *Journal of Political Economy*, 81(2):S279–S288, 1973.

[18] Gary S. Becker, Kevin M. Murphy, and Robert Tamura. Human capital, fertility, and economic growth. *Journal of Political Economy*, 98(5):S12–S37, 1990.

[19] Gary S. Becker. *Accounting for Tastes*. Cambridge, Mass.: Harvard University Press, 1996.

[20] Gary S. Becker, Kevin M. Murphy, and Jörg L. Spenkuch. The manipulation of children's preferences, old-age support, and investment in children's human capital. *Journal of Labour Economics*, 34(S2):S3–S30, 2016.

[21] Alberto Bisin and Thierry Verdier. "Beyond the melting pot": Cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988, 2000.

[22] David E. Bloom, David Canning, and Günther FinkJocelyn E. Finlay. Fertility, female labor force participation, and the demographic dividend. *Journal of Economic Growth*, 14(2):79–101, 2009.

[23] K. Bolin, B. Lindgren, and P. Lundborg. Informal and formal care among single-living elderly in Europe. *Health Economics*, 17(3):393–409, 2007.

[24] Bart J. Bronnenberg, Jean-Pierre H. Dubé, and Matthew Gentzkow. The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*, 102(6):2472–2508, May 2012.

[25] Kay Bussey and Albert Bandura. Social cognitive theory of gender development and differentiation. *Psychological Review*, 106(4):616–713, 1999.

[26] Fang Cai, John Giles, and Xin Meng. How well do children insure parents against low retirement income? an analysis using survey data from urban China. *Journal of Public Economics*, 90(9):2229–2255, 2006.

[27] Cecilia Chan, P.C. Ho Paul Yip Ernest Ng, Celia Chan, and Jade Au. Gender selection in China: Its meanings and implications. *Journal of Assisted Reproduction and Genetics*, 19(9):426–430, 2002.

[28] Yuyu Chen, Hongbin Li, and Lingsheng Meng. Prenatal sex selection and missing girls in China: Evidence from the diffusion of diagnostic ultrasound. *The Journal of Human Resources*, 48(1):36–70, 2013.

[29] Taha Choukhmane, Nicolas Coeurdacier, and Keyu Jin. The One Child Policy and household saving, 2015.

[30] Nicolas Coeurdacier, Stéphane Guibaud, and Keyu Jin. Credit constraints and growth in a global economy. *American Economic Review*, 105(9):2838–2881, 2015.

[31] Donald Cox. Motives for private income transfers. *Journal of Political Economy*, 95(3):508–546, 1987.

[32] Donald Cox and Oded Stark. Intergenerational transfers and the demonstration effect. Technical report, Boston College Working Papers in Economics, 329, 1996.

[33] Donald Cox and Oded Stark. On the demand for grandchildren: Tied transfers and the demonstration effect. *Journal of Public Economics*, 89(9):1665–1697, 2005.

[34] Monica Das Gupta, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok. Why is son preference so persistent in East and South Asia? A cross-country study of China, India and the Republic of Korea. *The Journal of Development Studies*, 40(2):153–187, 2003.

[35] Francine M. Deutsch. Filial piety, patrilineality, and China's One-Child Policy. *Journal of Family Issues*, 27(3):366–389, 2006.

[36] Diva Dhar, Tarun Jain, and Seema Jayachandran. Intergenerational transmission of gender attitudes: Evidence from India. Technical report, NBER Working Paper No. 21429, 2015.

[37] Esther Duflo. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813, September 2001.

[38] Avraham Ebenstein. The missing girls of China and the unintended consequences of the One Child Policy. *The Journal of Human Resources*, 45(1):87–115, 2010.

[39] Avraham Ebenstein and Steven Leung. Son preference and access to social insurance: Evidence from China's Rural Pension Program. *Population and Development Review*, 36(1):47–70, 2010.

[40] Jacquelynne S. Eccles, Janis E. Jacobs, and Rena D. Harold. Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issue*, 46(2):183–201, 1990.

[41] Vanessa L. Fong. China's One-Child Policy and the empowerment of urban daughters. *American Anthropologist*, 104(4):1098–1109, 2002.

[42] W. Hu. *Government Processes.* Hangzhou, Zhejiang Province: Zhejiang People's Press, 1998.

[43] Yang Hu and Xuezhu Shi. The impact of China's One-Child Policy on intergenerational and gender relations. *Contemporary Social Science*, 0(0):1–18, 2018.

[44] Sally Hudson, Peter Hull, and Jack Liebersohn. Interpreting instrumented difference-in-differences. Technical report, Massachusetts Institute of Technology, mimeo, 2017.

[45] Li Gan, Zhichao Yin, Nan Jia, Shu Xu, Shuang Ma, and Lu Zheng. *Data you need to know about China: Research Report of China Household Finance Survey 2012.* Springer Heidelberg New York Dordrecht London, 2014.

[46] Henrik Kleven, Camille Landais, and Jakob Egholt Søgaard. Children and gender inequality: Evidence from Denmark. Technical report, NBER Working Paper No. 24219, 2018.

[47] Gary R. Lee, Jeffrey W. Dwyer, and Raymond T. Coward. Gender differences in parent care: Demographic factors and same-gender preferences. *Journal of Gerontology*, 48(1):S9–S16, 1993.

[48] Yean Ju Lee, William L. Parish, and Robert J. Willis. Sons, daughters, and intergenerational support in Taiwan. *American Journal of Sociology*, 99(4):1010–1041, 1994.

[49] Ming-Hsuan Lee. The One-Child Policy and gender equality in education in China: Evidence from household data. *Journal of Family and Economic Issues*, 33(1):41–52, Mar 2012.

[50] Ronald Lee and Andrew Mason. Cost of aging. *Finance and Development*, 54(1):7–9, 2017.

[51] Hongbin Li, Mark Rosenzweig, and Junsen Zhang. Altruism, favoritism, and guilt in the allocation of family resources: Sophie's Choice in Mao's mass send-down movement. *Journal of Political Economy*, 118(1):1–38, 2010.

[52] Hugh Lytton and David Romney. Parents' differential socialization of boys and girls: A meta-analysis. *Psychological Bulletin*, 109(2):267–296, 1991.

[53] Maurice Godelier. *The making of great men: male domination and power among the New Guinea Baruya.* Cambridge: Cambridge University Press, 1982.

[54] Joel M. Guttman. Self-enforcing reciprocity norms and intergenerational transfers: Theory and evidence. *Journal of Public Economics*, 81(1):117–151, 2001.

[55] Marcel Mauss. *Sociologie et anthropologie.* PUF, Paris., 1950.

[56] Marcel Mauss. *Essais de sociologie.* Editions de Minuit, Paris., 1968.

[57] Susan McHale, Ann Crouter, and Corinna Tucker. Family context and gender role socialization in middle childhood: Comparing girls to boys and sisters to brothers. *Child Development*, 70(4):990–1004, 1999.

[58] Cecilia Nathansen Milwertz. *Accepting Population Control: Urban Chinese Women and the One-Child Family Policy*, chapter Demographic Change in the People's Republic of China, pages 35–64. Richmond Surrey: Curzon Press., 1997.

[59] Andreea Mitrut and François Charles Wolff. A causal test of the demonstration effect theory. *Economics Letters*, 103(1):52–54, 2009.

[60] Jaqueline Oliveira. The value of children: Inter-generational support, fertility, and human capital. *Journal of Development Economics*, 120:1–16, 2016.

[61] Emily Oster. Hepatitis B and the case of the missing women. *Journal of Political Economy*, 113(6):1163–1216, 2005.

[62] Nancy Qian. Missing women and the price of tea in China: The effect of sex-specific earnings on sex imbalance. *The Quarterly Journal of Economics*, 123(3):1251–1285, 2008.

[63] X. Shen. *Contemporary Chinese Government and Processes.* Tianjin, China: Nankai University Press, 2008.

[64] Merril Silverstein, Tonya M. Parrott, and Vern L. Bengtson. Factors that predispose middle-aged sons and daughters to provide social support to older parents. *Journal of Marriage and Family*, 57(2):465–475, 1995.

[65] Michéle Tertilt. Polygyny, fertility, and savings. *Journal of Political Economy*, 113(6):1341–1371, 2005.

[66] Shang-Jin Wei and Xiaobo Zhang. Sex ratios, entrepreneurship, and economic growth in the People's Republic of China. Technical report, NBER Working Paper No. 16800, 2011.

[67] Tyrene White. Domination, resistance, and accommodation in China's One-Child Campaign. In Elizabeth J. Perry and Mark Selden, editors, *Chinese Society: Change, Conflict and Resistance (Asia's Transformations)*, pages 171–196. Routledge, 2001.

[68] United Nations. Share of population aged 60 and older in china from 1950 to 2100. Technical report, United Nations, 2017.

[69] Department of Economic and Social Affairs. World population ageing 2017: Highlights. Technical report, United Nations, (ST/ESA/SER.A/397), 2017.

[70] François Charles Wolff. Private intergenerational contact in France and the demonstration effect. *Applied Economics*, 33:143–153, 2001.

[71] Yu Xie and Haiyan Zhu. Do sons or daughters give more money to parents in urban China? *Journal of Marriage and Family*, 71(1):174–186, 2009.

8 Figures and Tables

Figure 1: Public service announcement posters in China

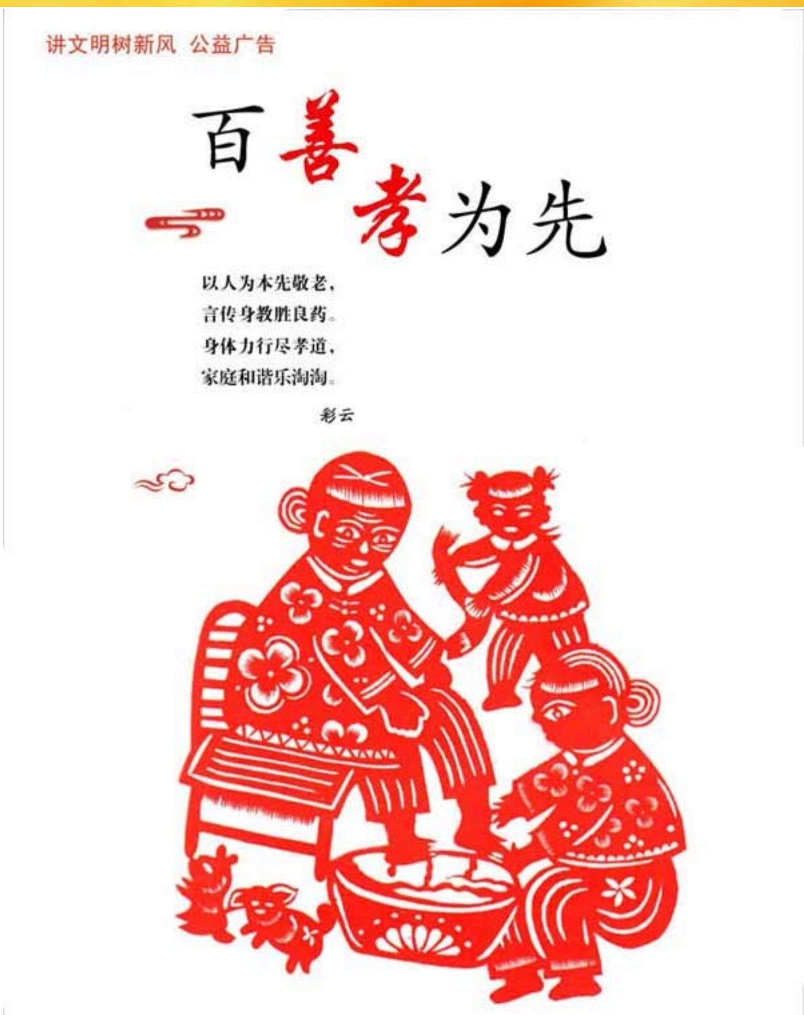
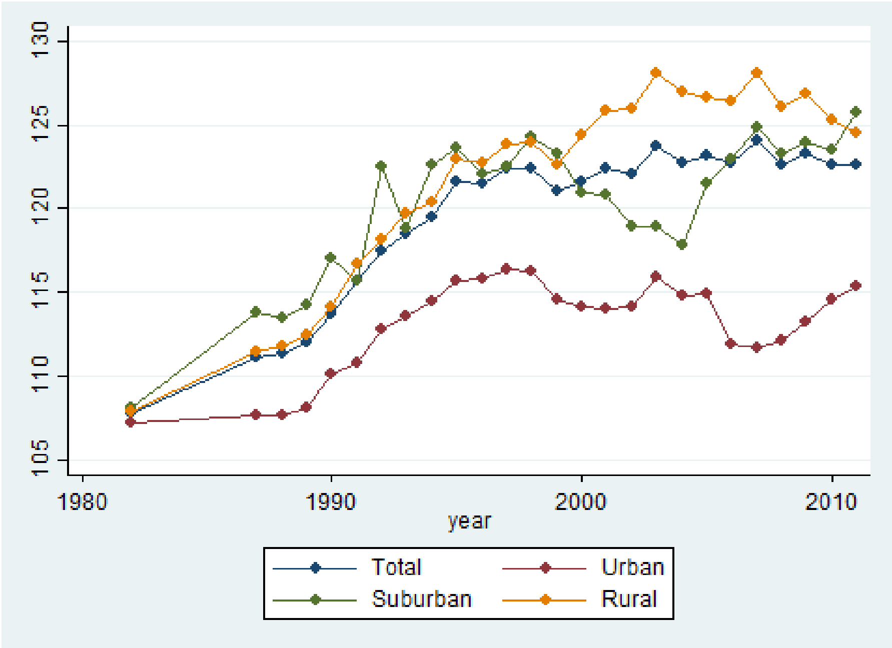


Figure 2: Gender ratios for the newborns in China: yearly trend



Data source: China Population and Employment Statistics Yearbook. 1982-2011.

Table 1: Primary Source of Support of China's Elderly, 2005 and 2010

2005

	Urban			*Rural*		
Source of support	Average	Male	Female	Average	Male	Female
Labour income	13.0	18.4	7.9	37.9	48.5	27.5
Pensions	45.4	56.9	34.6	4.60	8.1	1.3
Dibao	2.4	1.8	2.9	1.3	1.8	0.9
Insurance and subsidy	0.3	0.3	0.2	0.1	0.2	0.0
Property income	0.5	0.5	0.5	0.2	0.2	0.1
Family support	37.0	20.7	52.3	54.1	39.3	68.5
Other	1.5	1.4	1.6	1.8	2.0	1.7

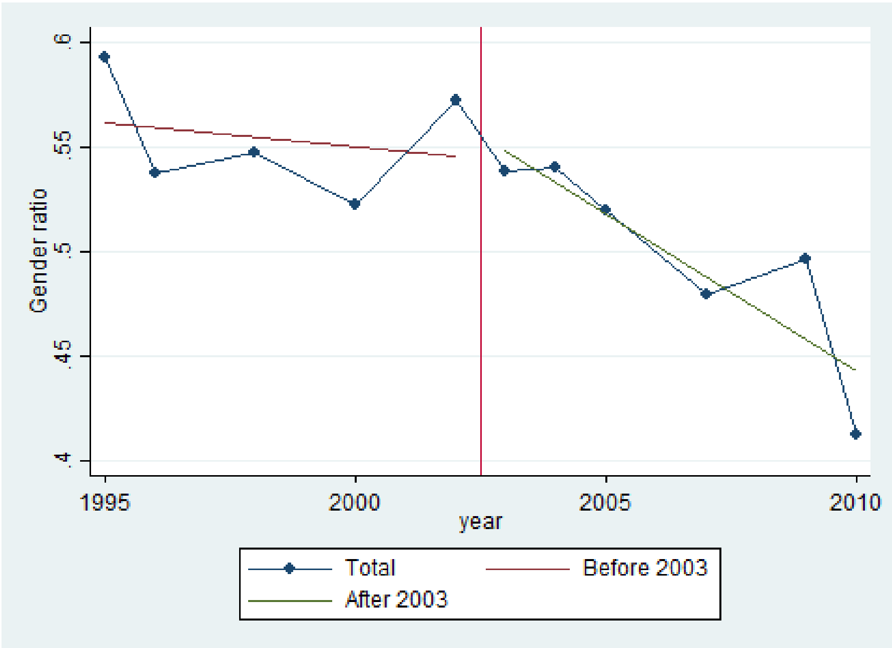
Source: NBS, 2006. Most significant share of support reported.

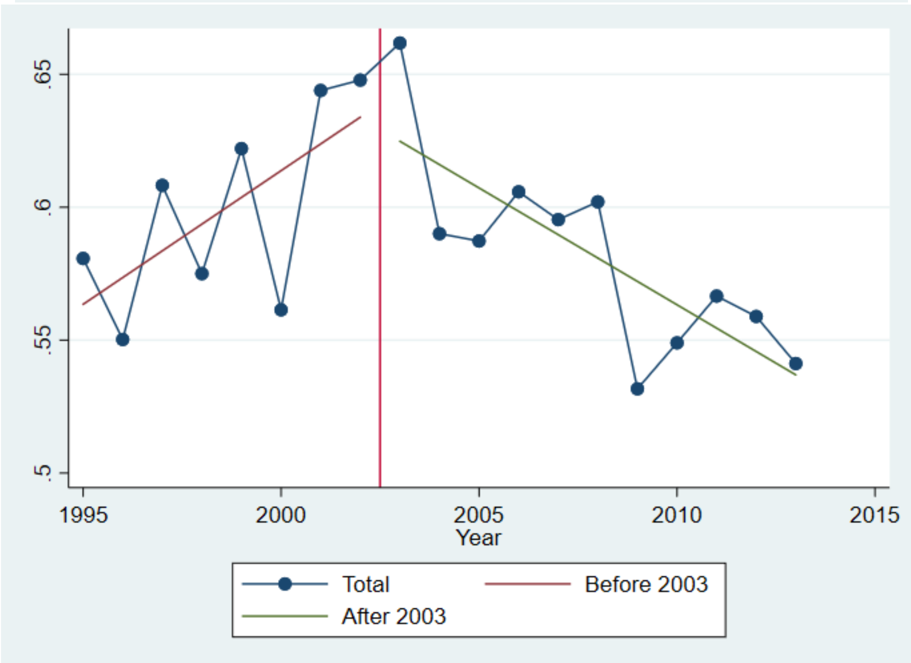
2010

	Urban			*Rural*		
Source of support	Average	Male	Female	Average	Male	Female
Labour income	6.16	9.72	3.75	41.18	50.53	32.14
Pensions	66.30	74.21	58.99	4.60	7.19	2.09
Dibao	2.33	1.76	2.87	4.48	5.14	3.85
Insurance and subsidy	-	-	-	-	-	-
Property income	0.68	0.75	0.62	0.19	0.21	0.16
Family support	22.43	12.13	31.95	47.74	35.13	59.93
Other	1.64	1.44	1.83	1.81	1.79	1.83

Source: NBS, 2011. Most significant share of support reported.

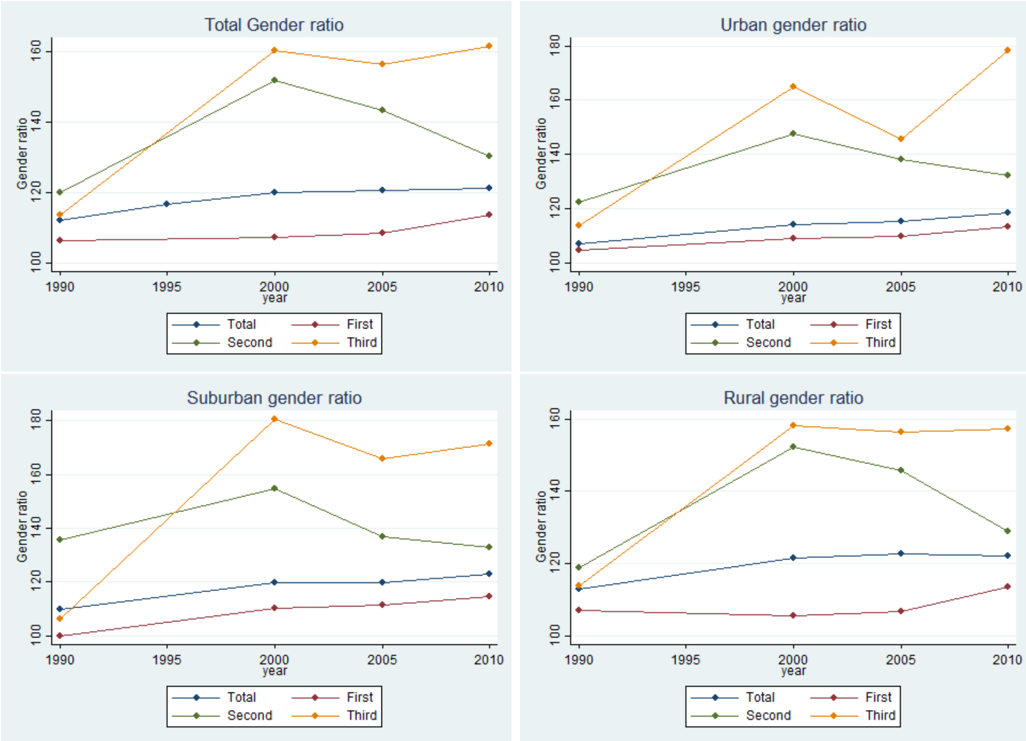
Figure 3: Estimated yearly gender ratios of the newborns using the CHARLS 2011 wave and the CHFS 2013





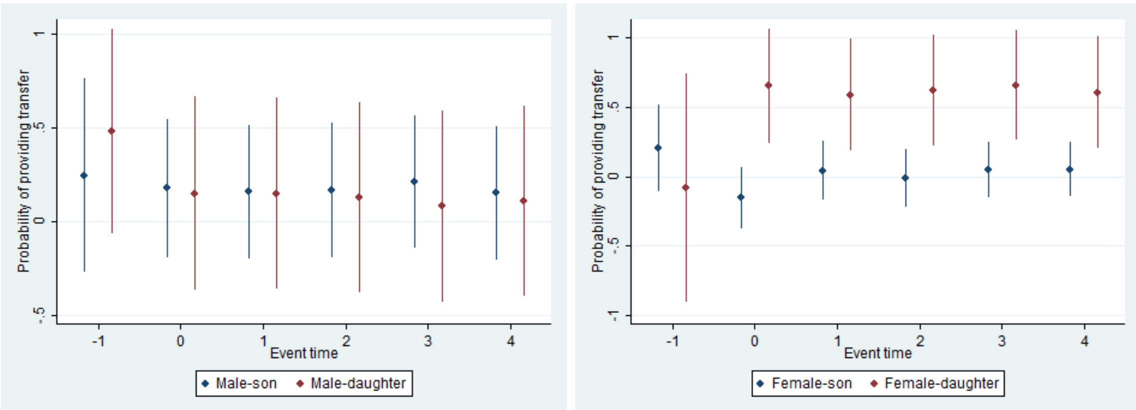
Note: The first graph is estimated using the 2011 wave CHARLS and the second graph uses the 2013 wave CHFS

Figure 4: Gender ratio by birth order for the newborns in China



Data source: National Population Census. 1990, 1995, 2000, 2005 and 2010.

Figure 5: Impact of the gender of the first child on the probability of provide any old-age support



Note: Event time is the time of individuals having their first child. Due to data limitation, I can only get one period before the event in the panel dataset.

Table 2: Summary Statistics: Key variables

	CHARLS				CHFS			
VARIABLES	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
whether P provides								
any transfers	0.284	0.306	0	1	0.265	0.441	0	1
regular transfer	0.105	0.272	0	1	-	-	-	-
non-regular transfer	0.243	0.308	0	1	-	-	-	-
log amount of								
total transfer	-	-	-	-	1.831	3.200	0	11.92
regular transfer	0.582	1.970	0	17.73	-	-	-	-
non-regular transfer	2.393	3.183	0	18.42	-	-	-	-
log visit days	3.234	2.374	0	5.900	1.847	2.256	0	5.900
gender of P	0.513	0.500	0	1	0.499	0.500	0	1
gender ratio of K	0.562	0.405	0	1	0.567	0.416	0	1
No. of Y	1.643	0.774	1	8	1.662	0.889	1	9
age of P	39.73	9.287	21	65	48.17	10.71	21	65
income level of P	5.078	1.420	1	11	-	-	-	-
income of P	-	-	-	-	21779	43639	0	1649439
education of P	0.892	0.496	0	2	0.832	0.646	0	2
whether P has a rural *hukou*	0.680	0.466	0	1	0.546	0.498	0	1
P living in rural areas	0.652	0.476	0	1	0.332	0.471	0	1
No. of siblings of P	3.758	1.612	1	10	3.218	1.856	0	16
any transfers from O	0.037	0.190	0	1	0.144	0.351	0	1

Table 3: Gender effect of K and P on provision of old-age support in CHARLS

	OLS: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
$maleP$	0.003	0.037***	-0.034**	0.275***	-0.088	0.172***
	(0.015)	(0.008)	(0.015)	(0.059)	(0.099)	(0.057)
$genderK$	0.003	0.001	-0.001	0.005	-0.009	-0.132**
	(0.016)	(0.007)	(0.016)	(0.052)	(0.102)	(0.052)
$maleP \times genderK$	-0.008	-0.009	0.004	-0.055	0.035	0.178**
	(0.020)	(0.010)	(0.020)	(0.075)	(0.132)	(0.072)
$genderK+$	-0.005	-0.008	0.003	-0.051	0.026	0.046
$maleP \times genderK$	(0.013)	(0.007)	(0.013)	(0.054)	(0.086)	(0.050)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.206	0.076	0.142	0.072	0.141	0.533

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 4: Gender effect of K and P on provision of old-age support in CHFS

	OLS: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
$maleP$	-0.056***	-0.384***	0.011	20.41***
	(0.010)	(0.075)	(0.051)	(2.933)
$genderK$	-0.008	-0.033	0.020	2.563
	(0.010)	(0.076)	(0.048)	(2.645)
$maleP \times genderK$	0.010	0.031	0.027	0.251
	(0.014)	(0.105)	(0.072)	(4.077)
$genderK+$	0.002	-0.002	0.048	2.814
$maleP \times genderK$	(0.010)	(0.075)	(0.054)	(3.150)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.227	0.230	0.162	0.112

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 5: Gender effect of K and P on provision of old-age support in CHARLS (IV results)

	IV:CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
$maleP$	-0.066**	0.028*	-0.102***	0.252**	-0.382**	0.134
	(0.028)	(0.015)	(0.027)	(0.106)	(0.178)	(0.097)
$genderK$	-0.035	0.020	-0.065	0.178	-0.314	0.052
	(0.040)	(0.022)	(0.040)	(0.154)	(0.262)	(0.136)
$maleP \times genderK$	0.114***	0.006	0.124***	-0.014	0.560*	0.247
	(0.046)	(0.024)	(0.045)	(0.170)	(0.293)	(0.153)
$genderK+$	0.079***	0.026**	0.059***	0.164**	0.246*	0.299***
$maleP \times genderK$	(0.022)	(0.010)	(0.019)	(0.070)	(0.126)	(0.068)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.203	0.074	0.140	0.070	0.140	0.532

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 6: Gender effect of K and P on provision of old-age support in CHFS (IV results)

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
$maleP$	-0.086***	-0.476***	-0.199	11.08
	(0.030)	(0.075)	(0.151)	(9.092)
$genderK$	-0.055*	-0.187	-0.561***	-20.01**
	(0.031)	(0.248)	(0.145)	(8.758)
$maleP \times genderK$	0.063	0.268	0.402	17.23
	(0.054)	(0.406)	(0.263)	(15.99)
$genderK+$	0.009	0.081	-0.160	-2.775
$maleP \times genderK$	(0.037)	(0.283)	(0.197)	(11.98)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.226	0.216	0.156	0.110

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 7: Gender effect of K and P on provision of old-age support in CHARLS: Income-level

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
Low income group						
maleP	-0.0750**	0.0225	-0.102***	0.177	-0.316	0.535***
	(0.0377)	(0.0182)	(0.0359)	(0.120)	(0.222)	(0.133)
genderK	-0.0532	0.0114	-0.0664	0.0884	-0.236	0.462**
	(0.0546)	(0.0260)	(0.0526)	(0.170)	(0.328)	(0.187)
$maleP \times genderK$	0.131**	0.0166	0.122**	0.102	0.426	-0.288
	(0.0614)	(0.0296)	(0.0581)	(0.192)	(0.362)	(0.207)
$genderK+$	0.078***	0.028**	0.057**	0.190**	0.190	0.174**
$maleP \times genderK$	(0.026)	(0.012)	(0.023)	(0.081)	(0.140)	(0.089)
Observations	7,048	7,048	7,048	7,048	7,048	7,048
R-squared	0.178	0.081	0.133	0.075	0.126	0.513
High income group						
maleP	-0.0595	0.0390	-0.110***	0.361**	-0.535*	-0.428***
	(0.0424)	(0.0241)	(0.0418)	(0.180)	(0.287)	(0.148)
genderK	-0.0140	0.0323	-0.0626	0.290	-0.376	-0.421**
	(0.0579)	(0.0348)	(0.0584)	(0.261)	(0.400)	(0.198)
$maleP \times genderK$	0.0900	-0.0138	0.132*	-0.183	0.775*	1.017***
	(0.0709)	(0.0395)	(0.0682)	(0.295)	(0.470)	(0.242)
$genderK+$	0.076*	0.019	0.069*	0.101	0.399	0.597***
$maleP \times genderK$	(0.041)	(0.021)	(0.036)	(0.157)	(0.250)	(0.141)
Observations	5,184	5,184	5,184	5,184	5,184	5,184
R-squared	0.238	0.080	0.160	0.077	0.154	0.555
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 8: Gender effect of K and P on provision of old-age support in CHFS: Income-level

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
Low income group				
$maleP$	-0.0519	-0.197	-0.290	-5.133
	(0.0452)	(0.328)	(0.239)	(14.44)
$genderK$	-0.0197	0.0467	-0.630***	-41.15***
	(0.0491)	(0.351)	(0.218)	(13.26)
$maleP \times genderK$	0.00912	-0.163	0.662	56.87**
	(0.0805)	(0.586)	(0.421)	(25.70)
$genderK+$	-0.011	-0.117	0.032	15.71
$maleP \times genderK$	(0.056)	(0.413)	(0.325)	(19.84)
Observations	13,237	13,237	13,237	13,237
R-squared	0.213	0.202	0.150	0.105
High income group				
$maleP$	-0.118***	-0.769**	-0.240	11.28
	(0.0411)	(0.318)	(0.187)	(10.88)
$genderK$	-0.0731*	-0.384	-0.445**	-0.999
	(0.0441)	(0.342)	(0.187)	(11.13)
$maleP \times genderK$	0.106	0.671	0.221	-9.858
	(0.0702)	(0.550)	(0.311)	(18.49)
$genderK+$	0.032	0.287	-0.224	-10.86
$maleP \times genderK$	(0.039)	(0.380)	(0.221)	(13.02)
Observations	6,287	6,287	6,287	6,287
R-squared	0.216	0.214	0.164	0.141
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 9: Gender effect of K and P on provision of old-age support in CHARLS: Single-K family

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log regular*	*log nonregular*	*log (visit days)*
Single K HH						
$maleP$	-0.0460	0.0109	-0.0723**	0.104	-0.318	-0.0300
	(0.0341)	(0.0171)	(0.0331)	(0.125)	(0.223)	(0.117)
$genderK$	-0.0459	0.0165	-0.0754*	0.135	-0.452	0.0325
	(0.0430)	(0.0214)	(0.0420)	(0.159)	(0.283)	(0.147)
$maleP \times genderK$	0.0773	-0.00226	0.103**	-0.0495	0.573*	0.274*
	(0.0501)	(0.0243)	(0.0478)	(0.179)	(0.324)	(0.166)
$genderK+$	0.031	0.014	0.028	0.085	0.121	0.307***
$maleP \times genderK$	(0.022)	(0.010)	(0.019)	(0.072)	(0.133)	(0.073)
Observations	5,909	5,909	5,909	5,909	5,909	5,909
R-squared	0.211	0.066	0.152	0.064	0.153	0.572
Non-single K HH						
$maleP$	-0.0802	0.0257	-0.111**	0.284	-0.270	0.153
	(0.0584)	(0.0316)	(0.0561)	(0.214)	(0.354)	(0.198)
$genderK$	0.00891	0.0172	-0.0217	0.201	0.118	0.0754
	(0.0874)	(0.0468)	(0.0870)	(0.313)	(0.543)	(0.299)
$maleP \times genderK$	0.149	0.0575	0.109	0.288	0.194	0.481
	(0.108)	(0.0585)	(0.103)	(0.390)	(0.649)	(0.356)
$genderK+$	0.158***	0.065**	0.087*	0.489**	0.312	0.556***
$maleP \times genderK$	(0.060)	(0.032)	(0.053)	(0.214)	(0.345)	(0.189)
Observations	6,323	6,323	6,323	6,323	6,323	6,323
R-squared	0.199	0.097	0.141	0.094	0.142	0.503
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 10: Gender effect of K and P on provision of old-age support in CHFS: Single-K family

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
Single K HH				
$maleP$	-0.137***	-0.750***	-0.446***	1.018
	(0.035)	(0.270)	(0.166)	(9.902)
$genderK$	-0.110***	-0.538**	-0.671***	-15.71*
	(0.035)	(0.270)	(0.152)	(9.404)
$maleP \times genderK$	0.126**	0.563	0.499*	11.68
	(0.057)	(0.446)	(0.263)	(16.15)
$genderK+$	0.016	0.025	-0.172	-4.031
$maleP \times genderK$	(0.039)	(0.306)	(0.187)	(11.31)
Observations	10,426	10,426	10,426	10,426
R-squared	0.222	0.218	0.149	0.119
Non-single K HH				
$maleP$	0.021	0.000	0.010	6.836
	(0.067)	(0.476)	(0.363)	(21.82)
$genderK$	0.107	0.677	-0.243	-35.65*
	(0.080)	(0.554)	(0.344)	(20.18)
$maleP \times genderK$	-0.106	-0.344	0.391	53.88
	(0.129)	(0.919)	(0.696)	(42.06)
$genderK+$	0.000	0.332	0.149	18.23
$maleP \times genderK$	(0.090)	(0.648)	(0.548)	(33.40)
Observations	9,098	9,098	9,098	9,098
R-squared	0.200	0.190	0.159	0.115
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 11: Gender effect of K and P on provision of old-age support in CHARLS: Urban-rural differences

	IV: CHARLS (rural)					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
Urban						
$maleP$	-0.0478	0.0206	-0.0906*	0.225	-0.476	-0.109
	(0.0550)	(0.0288)	(0.0512)	(0.229)	(0.358)	(0.189)
$genderK$	-0.0121	0.0549	-0.102	0.500	-0.715	-0.00895
	(0.0712)	(0.0435)	(0.0681)	(0.352)	(0.472)	(0.265)
$maleP \times genderK$	0.0734	-0.0326	0.144*	-0.341	0.853	0.136
	(0.0929)	(0.0495)	(0.0850)	(0.405)	(0.595)	(0.314)
$genderK+$	0.061	0.022	0.042	0.160	0.138	0.127
$maleP \times genderK$	(0.057)	(0.024)	(0.048)	(0.194)	(0.345)	(0.169)
Observations	2,582	2,582	2,582	2,582	2,582	2,582
R-squared	0.233	0.095	0.161	0.090	0.159	0.560
Rural						
$maleP$	-0.0856***	0.0385**	-0.128***	0.321***	-0.469**	0.321***
	(0.0331)	(0.0173)	(0.0324)	(0.119)	(0.208)	(0.112)
$genderK$	-0.0545	0.0134	-0.0723	0.112	-0.289	0.106
	(0.0477)	(0.0246)	(0.0475)	(0.167)	(0.306)	(0.157)
$maleP \times genderK$	0.158***	0.00905	0.158***	0.0216	0.658*	0.101
	(0.0539)	(0.0273)	(0.0523)	(0.183)	(0.337)	(0.173)
$genderK+$	0.104***	0.023**	0.086***	0.133*	0.368***	0.207***
$maleP \times genderK$	(0.024)	(0.011)	(0.021)	(0.078)	(0.138)	(0.072)
Observations	9,650	9,650	9,649	9,650	9,648	9,651
R-squared	0.201	0.085	0.141	0.082	0.143	0.534
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 12: Gender effect of K and P on provision of old-age support in CHFS: Urban-rural differences

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
Urban				
$maleP$	-0.091***	-0.509**	-0.351**	-2.238
	(0.031)	(0.238)	(0.144)	(8.422)
$genderK$	-0.060*	-0.225	-0.581***	-20.56**
	(0.035)	(0.267)	(0.146)	(8.766)
$maleP \times genderK$	0.063	0.264	0.384	13.25
	(0.054)	(0.429)	(0.251)	(14.92)
$genderK+$	0.003	0.039	-0.196	-7.307
$maleP \times genderK$	(0.039)	(0.304)	(0.184)	(10.92)
Observations	12,990	12,990	12,990	12,990
R-squared	0.226	0.220	0.164	0.125
Rural				
$maleP$	-0.054	-0.440	0.030	10.62
	(0.090)	(0.606)	(0.514)	(31.64)
$genderK$	-0.002	-0.124	-0.347	-27.01
	(0.031)	(0.626)	(0.461)	(28.77)
$maleP \times genderK$	0.032	0.424	0.496	66.32
	(0.153)	(1.032)	(0.874)	(54.05)
$genderK+$	0.029	0.300	0.149	39.31
$maleP \times genderK$	(0.98)	(0.674)	(0.642)	(39.28)
Observations	6,534	6,534	6,534	6,534
R-squared	0.181	0.157	0.168	0.132
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 13: Gender effect of K and P on provision of old-age support in CHARLS: P with or without brothers

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log regular*	*log nonregular*	*log (visit days)*
Without older bro.						
maleP	-0.086**	-0.008	-0.088**	0.005	-0.261	0.257**
	(0.035)	(0.017)	(0.034)	(0.121)	(0.227)	(0.122)
genderK	-0.033	-0.008	-0.034	-0.018	-0.086	0.108
	(0.048)	(0.025)	(0.048)	(0.178)	(0.320)	(0.170)
$maleP \times genderK$	0.113**	0.024	0.102*	0.112	0.370	0.107
	(0.056)	(0.028)	(0.054)	(0.194)	(0.362)	(0.190)
$genderK+$	0.080***	0.016*	0.068***	0.094	0.284*	0.215***
$maleP \times genderK$	(0.025)	(0.010)	(0.022)	(0.071)	(0.146)	(0.080)
Observations	6,912	6,912	6,912	6,912	6,912	6,912
R-squared	0.209	0.071	0.153	0.069	0.153	0.550
With older bro.						
maleP	-0.040	0.078***	-0.118**	0.598***	-0.537*	0.0443
	(0.049)	(0.027)	(0.046)	(0.189)	(0.296)	(0.164)
genderK	-0.029	0.062*	-0.100	0.469*	-0.618	0.007
	(0.0674)	(0.037)	(0.066)	(0.265)	(0.420)	(0.222)
$maleP \times genderK$	0.115	-0.018	0.145*	-0.192	0.778	0.361
	(0.082)	(0.044)	(0.076)	(0.306)	(0.487)	(0.261)
$genderK+$	0.087*	0.043*	0.045	0.277*	0.160	0.368***
$maleP \times genderK$	(0.44)	(0.023)	(0.036)	(0.155)	(0.243)	(0.136)
Observations	5,320	5,320	5,320	5,320	5,320	5,320
R-squared	0.202	0.095	0.135	0.090	0.138	0.513
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 14: Heterogeneity Check: the parents' pension coverage

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0786**	0.00621	-0.106***	0.147	-0.393*	-0.482***
	(0.0361)	(0.0199)	(0.0350)	(0.144)	(0.234)	(0.126)
genderK×	-0.0880	0.0324	-0.0374	0.243	-0.104	0.999***
P pension	(0.0978)	(0.0418)	(0.0965)	(0.288)	(0.610)	(0.323)
P pension	0.00556	-0.0531**	0.0106	-0.370**	-0.125	-0.148
	(0.0578)	(0.0257)	(0.0576)	(0.176)	(0.360)	(0.195)
genderK	-0.0135	0.00606	-0.0550	0.0832	-0.294	-0.324*
	(0.0502)	(0.0287)	(0.0498)	(0.207)	(0.336)	(0.166)
maleP × *genderK*	0.110*	0.0246	0.127**	0.0525	0.480	1.217***
	(0.0599)	(0.0326)	(0.0573)	(0.233)	(0.385)	(0.205)
maleP×	0.0341	0.0485	0.0305	0.181	0.0227	2.400***
P pension	(0.0768)	(0.0372)	(0.0747)	(0.252)	(0.474)	(0.283)
maleP × *genderK*	-0.00700	-0.0452	-0.0488	-0.112	0.123	-3.355***
×*P pension*	(0.125)	(0.0576)	(0.120)	(0.393)	(0.771)	(0.448)
High P pension father	0.001	0.018**	-0.014	0.267	0.205	-1.463***
demonstrate effects	(0.062)	(0.032)	(0.057)	(0.211)	(0.384)	(0.247)
Low P pension father	0.096***	0.0371**	0.072***	0.136	0.186	0.893***
demonstrate effects	(0.030)	(0.014)	(0.026)	(0.097)	(0.170)	(0.113)
Differences in father	-0.095	-0.013	-0.086	0.131	0.018	-2.356***
demonstrate effects	(0.078)	(0.040)	(0.071)	(0.264)	(0.475)	(0.303)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.203	0.075	0.140	0.071	0.140	0.531

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table 15: Heterogeneity Check: the parents' pension coverage

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	0.000663	-0.111	-0.0725	4.655
	(0.0632)	(0.448)	(0.324)	(20.07)
genderK	-0.108*	-0.864*	-1.119***	-61.60***
(Without pension mother demonstration effects)	(0.0634)	(0.454)	(0.284)	(16.61)
With pension	-0.00885	-0.183	-0.183	-21.37*
	(0.0462)	(0.336)	(0.208)	(12.27)
maleP × *genderK*	-0.0390	-0.0925	0.724	50.57
	(0.110)	(0.781)	(0.553)	(34.48)
maleP × *pension*	-0.0579	-0.347	0.111	18.32
	(0.0757)	(0.550)	(0.378)	(23.45)
genderK × *pension*	0.0716	0.766	0.639*	44.55**
(With or without pension mother effect differences)	(0.0822)	(0.599)	(0.360)	(21.30)
genderK × *pension*	0.0562	0.251	-0.470	-43.17
×*maleP*	(0.133)	(0.966)	(0.648)	(40.53)
With pension mother demonstration effects	-0.036	-0.098	-0.479**	-17.05
	(0.047)	(0.355)	(0.200)	(12.29)
With or without pension father effect differences	0.127	1.016	0.169	1.381
	(0.103)	(0.752)	(0.538)	(34.36)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.303	0.228	0.330	0.362

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table 16: Income of the first generation and the demonstration effect

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0691	0.0375	-0.116***	0.300*	-0.438*	0.249*
	(0.0425)	(0.0236)	(0.0414)	(0.164)	(0.266)	(0.145)
genderK× *high income O*	0.00321	0.0119	0.0149	0.199	-0.0561	-0.134
	(0.0920)	(0.0465)	(0.0912)	(0.336)	(0.599)	(0.300)
high income O	-0.00941	-0.0100	-0.0143	-0.138	0.0760	0.309*
	(0.0562)	(0.0294)	(0.0564)	(0.207)	(0.364)	(0.181)
genderK	-0.0393	0.0150	-0.0732	0.0954	-0.300	0.125
	(0.0576)	(0.0304)	(0.0574)	(0.206)	(0.371)	(0.192)
maleP × *genderK*	0.124*	0.00500	0.138**	0.0156	0.652	0.204
	(0.0690)	(0.0364)	(0.0665)	(0.248)	(0.429)	(0.226)
maleP× *high income O*	0.00804	-0.0195	0.0310	-0.0920	0.131	-0.238
	(0.0695)	(0.0372)	(0.0684)	(0.262)	(0.444)	(0.237)
maleP × *genderK* ×*high income O*	-0.0235	-0.00326	-0.0284	-0.120	-0.218	0.0399
	(0.115)	(0.0577)	(0.112)	(0.412)	(0.731)	(0.380)
High income O male demonstrate effects	0.064	0.029	0.052	0.190	0.079	0.235
	(0.048)	(0.024)	(0.044)	(0.167)	(0.288)	(0.150)
Low income O male demonstrate effects	0.084**	0.020	0.065**	0.111	0.353*	0.329***
	(0.035)	(0.017)	(0.032)	(0.117)	(0.210)	(0.111)
Differences in males demonstrate effects	-0.020	0.009	-0.013	0.079	-0.274	-0.094
	(0.070)	(0.036)	(0.066)	(0.246)	(0.428)	(0.222)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.203	0.075	0.140	0.071	0.140	0.532

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 17: Effects of education and time investment on provision of old-age support

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0650**	0.0283**	-0.101***	0.256***	-0.398**	0.130
	(0.0263)	(0.0137)	(0.0252)	(0.0981)	(0.166)	(0.0924)
genderK	-0.0300	0.0236	-0.0656*	0.200	-0.356	0.0505
	(0.0395)	(0.0208)	(0.0386)	(0.149)	(0.255)	(0.134)
$maleP \times genderK$	0.112***	0.00519	0.123***	-0.0251	0.588**	0.251*
	(0.0429)	(0.0219)	(0.0404)	(0.155)	(0.268)	(0.143)
awayage	0.0158	0.00985	0.0144	0.0521	0.115	-0.0584
	(0.0184)	(0.0107)	(0.0177)	(0.0726)	(0.126)	(0.0612)
awaytime	-0.00790	0.00109	-0.0129***	0.00412	-0.0820**	-0.0385*
	(0.00598)	(0.00281)	(0.00498)	(0.0170)	(0.0381)	(0.0220)
log edu expense	-0.00245	-0.000891	-0.00101	-0.00941	-0.00177	-0.0219***
	(0.00255)	(0.00127)	(0.00252)	(0.00972)	(0.0183)	(0.00846)
edu level	0.0135	0.00263	0.0103	0.0556	0.146*	0.145***
	(0.0134)	(0.00761)	(0.0134)	(0.0494)	(0.0832)	(0.0445)
$genderK+$	0.082***	0.029**	0.057**	0.175*	0.232	0.301***
$maleP \times genderK$	(0.026)	(0.013)	(0.022)	(0.091)	(0.151)	(0.086)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *genderK* represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table 18: The third generation effect on inter-household transfer

	IV: CHARLS			IV: CHFS	
VARIABLES	*any transfer*	*any receipt by P*	*any receipt by K*	*any transfer*	*any receipt by P*
maleP	-0.066**	0.026**	0.089***	-0.0527	0.013
	(0.028)	(0.011)	(0.030)	(0.0340)	(0.032)
genderK	-0.035	-0.043***	0.033	-0.0674**	0.170***
	(0.040)	(0.016)	(0.029)	(0.0364)	(0.034)
$maleP \times genderK$	0.114***	0.008	-0.088*	0.0196	-0.022
	(0.046)	(0.016)	(0.052)	(0.0602)	(0.057)
any receipt by P	-0.019	-	0.170***	0.313***	-
	0.028	-	(0.026)	(0.010)	-
any transfer	-	-0.004	0.090***	-	0.213***
	-	(0.006)	(0.009)	-	(0.007)
$genderK+$	0.079***	-0.035***	-0.055	-0.048	0.148***
$maleP \times genderK$	(0.022)	(0.007)	(0.042)	(0.042)	(0.034)
P demographics	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	19,524	19,524
R-squared	0.203	0.040	0.087	0.303	0.250

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *genderK* represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I. The additional information for this table is that new outcomes *anyreceiptbyM* and *anyreceiptbyY* are the transfer from O to P's household and P's children Y.

Table 19: The demonstration effect in the first generation

VARIABLES	OLS: CHARLS				
	any-transfer	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*
Male P					
$father's\ transfer$	0.064**	0.103***	0.102***	0.114***	0.102***
	(0.027)	(0.030)	(0.029)	(0.037)	(0.035)
$mother's\ transfer$	0.048**	0.067**	0.109***	0.111**	0.116***
	(0.021)	(0.028)	(0.023)	(0.045)	(0.027)
Observations	6,688	6,688	6,688	6,688	6,688
Female P					
$father's\ transfer$	0.056	0.031	0.112***	0.058*	0.113**
	(0.035)	(0.025)	(0.039)	(0.030)	(0.045)
$mother's\ transfer$	0.108***	0.075**	0.185***	0.171***	0.206***
	(0.048)	(0.031)	(0.030)	(0.054)	(0.034)
Observations	5,540	5,540	5,540	5,540	5,540
P demographics	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The $father's\ transfer$ and $mother's\ transfer$ are the transfer provided by O to P's paternal and maternal grandparents. The controlling variables for P are age, marriage status, rural *hukou*, provinces, education, professional title, income level, whether P lives with parents and the distant to parents place, visit frequency to O, the number and rank of siblings and the number of children. The detail of the controlling variables are explained in the Appendix I.

Table 21: Impact of the gender of the first child on the probability of provide any old-age support

VARIABLES	*any-transfer*			
	father-son	father-daughter	mother-son	mother-daughter
Event time				
-1	0.244	0.479*	0.207	-0.0824
	(0.264)	(0.278)	(0.160)	(0.418)
0	0.175	0.148	-0.155	0.655***
	(0.186)	(0.262)	(0.114)	(0.211)
1	0.157	0.148	0.0436	0.588***
	(0.181)	(0.258)	(0.108)	(0.206)
2	0.163	0.125	-0.0116	0.618***
	(0.183)	(0.258)	(0.105)	(0.204)
3	0.208	0.0787	0.0499	0.660***
	(0.180)	(0.259)	(0.102)	(0.201)
4	0.150	0.105	0.0507	0.607***
	(0.182)	(0.258)	(0.0991)	(0.204)
Age fixed-effect	Yes	Yes	Yes	Yes
Wave fixed-effect	Yes	Yes	Yes	Yes
Observations	809	771	811	765
R-squared	0.140	0.142	0.093	0.064

Notes: Robust standard errors in parentheses. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The event is having the first child in the respondents' household. The event time equal 0 at the year of having the first child. All the other event times adjusted accordingly. male-son is the male group with the first child as son, male-daughter is the male group with the first daughter. female-son and female-daughter are the female group accordingly. The outcome variable is the probability of providing any transfer to the elderly parents.

Table 22: The demonstration effect on the net-transfer provided by the second generation

	IV: CHARLS					IV: CHFS	
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*any transfer*	*log(transfer) P*
$maleP$	-0.0781***	0.0265*	-0.107***	0.196*	-0.619***	-0.0112	-0.605*
	(0.0278)	(0.0148)	(0.0269)	(0.107)	(0.198)	(0.0316)	(0.346)
$genderK$	-0.0239	0.0211	-0.0501	0.177	-0.206	-0.0949***	-0.782**
	(0.0402)	(0.0215)	(0.0396)	(0.156)	(0.291)	(0.0328)	(0.383)
$maleP \times genderK$	0.117**	0.00661	0.117***	0.00888	0.647**	-0.0454	0.610
	(0.0462)	(0.0240)	(0.0442)	(0.173)	(0.322)	(0.0554)	(0.625)
$genderK+$	0.093***	0.027***	0.067***	0.186**	0.441***	-0.141***	-0.173
$maleP \times genderK$	(0.022)	(0.010)	(0.019)	(0.072)	(0.140)	(0.039)	(0.421)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,229	12,232	12,227	19,524	19,524
R-squared	0.199	0.073	0.140	0.064	0.117	0.057	0.055

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I, except the outcome variables are all net transfers, which count the transfer from O to P in during the variable construction.

Table 20: The demonstration effect on cohabitation

VARIABLES	IV: CHARLS	IV: CHFS
	Ageing parents cohabitation	
$maleP$	-0.409***	0.041*
	(0.023)	(0.021)
$genderK$	-0.055***	-0.034*
	(0.017)	(0.018)
$maleP \times genderK$	0.763***	0.083
	(0.039)	(0.037)
$maleP \times genderK$ $+genderK$	0.708***	0.048*
	(0.034)	(0.029)
P demographics	Yes	Yes
O demographics	Yes	Yes
Observations	12,232	19,524
R-squared	0.371	0.141

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I, except the outcome variables are now the cohabitation with ageing parents.

Table 23: The "missing girls" and the needs of future old-age support

Sample	(1) Estimated Old-age support prevalence	(2) Predicted sex ratio under old-age support prevalence	(3) Sample percentage value old-age support and sons	(4) Old-age support prevalence adjusted sex ratio	(5) Acutal sex ratio	(6) Percentage explained by old-age support
Full sample	0.0317***	1.1913	34.38%	1.0958	1.3070	18.13%
Urban	0.0376***	1.2200	18.39%	1.0784	1.2127	17.96%
Rural	0.0210***	1.1412	47.48%	1.0918	1.3912	12.57%

Notes: The estimations are fully based on Oster's method (2005). Column (1) shows the coefficients of the old-age support dummy on the gender ratio of the third generation. This is the prevalence of the needs of future old-age support on the gender ratio. The natural male-to-female gender ratio is 1.049. Column (2) is the estimated gender ratio for people only care about the future old-age support from the next generation. Column (3) is the percentage of the sample who value both sons in the household and old-age support provided by their children. Column (4) is the weighted gender ratio of the whole population using Column (2) and (3) and the natural gender ratio. The equation for the adjusted gender ratio is $GR_{adjust} = GR_{old-age} \times percentage_{old-age} + GR_{nonold-age} \times (1 - percentage_{old-age})$ (Oster, 2005). The actual gender ratio in the data is in Column (5) and Column (6) is the percentage of the unbalanced gender ratio correlates with the needs of old-age support.

Appendix I

Data and IV construction in CHARLS: I have had to make certain assumptions when constructing the gender of the first child IV in CHARLS. As discussed above, I have restructured the original dataset from a dataset where the main respondents came from the first generation in my setting to a dataset where the main observations are made by the second generation. In the model , the children of the respondents are the second generation and the regression results and the notation for this generation is P. The original dataset gives no information on the birth year, but gives the gender and number of the third generation. The year of birth is available only if grandchildren are living with the first generation.

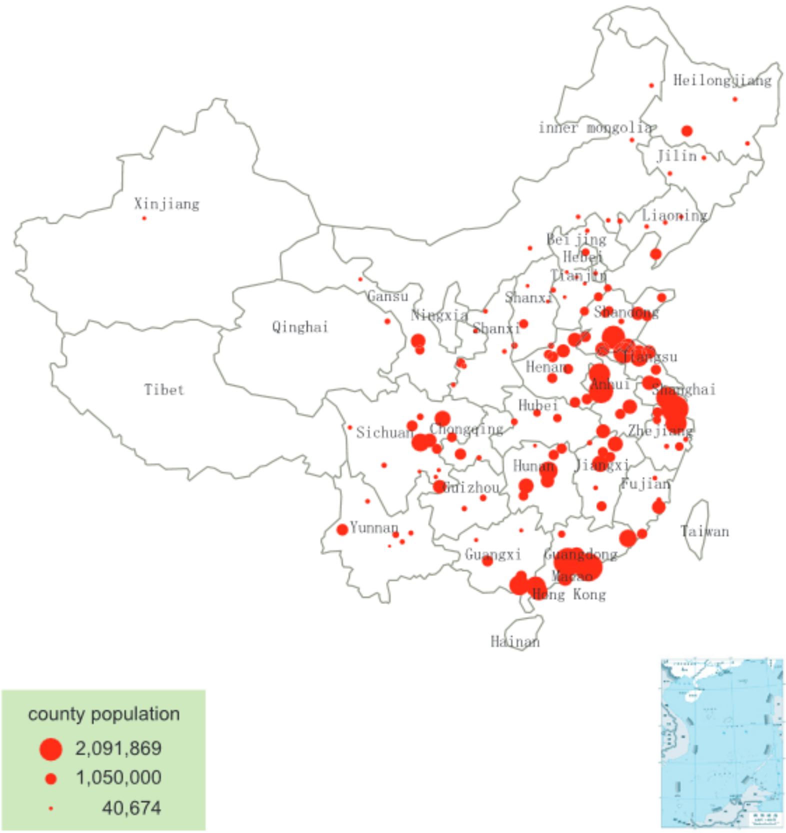
Moreover, many observations are missing for families that are living together with the first generation. Apart from this information, the dataset does provide information on the gender and number of the third generation if s/he is above the age of 16. For most households I can use this information to work out the gender of the first child. But some estimations are still needed in this process; they are based on the parents' age and the average age of females when their children are born, in order of birth, in both urban and rural areas.

Notes for the CHARLS regression results: $Notes$: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. P stands for parents and O represents the ageing parents. K is the third generation, namely the children of P. The coefficient for $genderK$ shows the difference between female P with sons and female P with daughters, which represents the negative female-dominated demonstration effect. The coefficient for $genderK + maleP \times genderK$ shows the difference between male P with sons and male P with daughters, which is the male demonstration effect. These are not the regression coefficients, but a linear combination of the coefficients for $genderK$ and $maleP \times genderK$. The controlling variables for P are age, marriage status, rural *hukou*, provinces, education, professional title, income level, how far away the parents live from P, the frequency of visit frequency to O, the number and rank of siblings and the number of children. The controlling variables for O are age, education level, rural *hukou*, pension status, household wage, the income of O and whether the members of O provide any transfer to P. The dependent variable $any\text{-}transfer$, $regular$ and $nonregular$ equals 1 if P provides any, or regular, or non-regular transfer to O. $log(regular)$ and $log(nonregular)$ are the natural log of the amount of regular and non-regular transfers. $log\ (visitdays)$ is the the natural log of days spent on visits in the previous year. The instrumental variables used in the regressions are the gender of the first child in a household with at least one child born after 2003 and the prefectural-level compliance index.

Notes for the CHFS regression results: $Notes$: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. P stands for parents and O represents the ageing parents. K is the third generation, namely the children of P. The coefficient for $genderK$ shows the difference between female P with sons and female P with daughters, which represents the negative female-dominated demonstration effect. The coefficient for $genderK + maleP \times genderK$ shows the difference between male P with sons and male P with daughters, which is the male demonstration effect. These are not the regression coefficients, but a linear combination of the coefficients for $genderK$ and $maleP \times genderK$. The controlling variables for P are age, marriage status, rural *hukou*, provinces, education, professional title, income level, the frequency of visit frequency to O, the number and rank of siblings and the number of children. The controlling variables for O are age, education level, rural *hukou*, the income of O and whether the members of O provide any transfer to P. The dependent variable $any\text{-}transfer$ equals 1 if P provides any transfer to O correspondingly. $log(transfer)$ is natural log of the amount of any transfers. *visit days* and $log\ (visitdays)$ are the number of days spent on visiting and its natural log in the previous year. The instrumental variables used in the regressions are the gender of the first child in a household with at least one child born after 2003.

For households affected by the policy ban after 2003 As discussed above, using a subsample includes only households affected by the policy ban after 2003 and does not provide well-identified results when the gender of the first child is kept as the instrumental variable. This is because, even with the policy ban, the gender ratio in some provinces is still high. I use a subsample check to provide relevant evidence. I divide the sample that includes only households affected by the policy ban after 2003 into two subsample, one showing a high gender-ratio and the other showing a low gender-ratio. A province is classified as a high gender-ratio province (1) if in the 2010 Population Census gender ratio there is above the national gender ratio, and 0 otherwise. Table A.30 shows the results of this simple subsample check. The father demonstration effects are mostly positive (predicting regular support) and some of them are marginally significant also in the low gender-ratio province subsample. The results from CHFS are shown in Table A.31, which also shows that the mother demonstration effect is greater in absolute value and has a lower standard error. The results from this simple sample check add a piece of suggestive evidence, since it is not sufficient to use only the subsample containing households affected by the policy ban.

Figure A.1: Distribution of CHARLS sample counties and districts



Data source: Official report by CCER. Website: http://charls.pku.edu.cn/uploads/document/public_documents/application/Challenges-of-Population-Aging-in-China-final.pdf

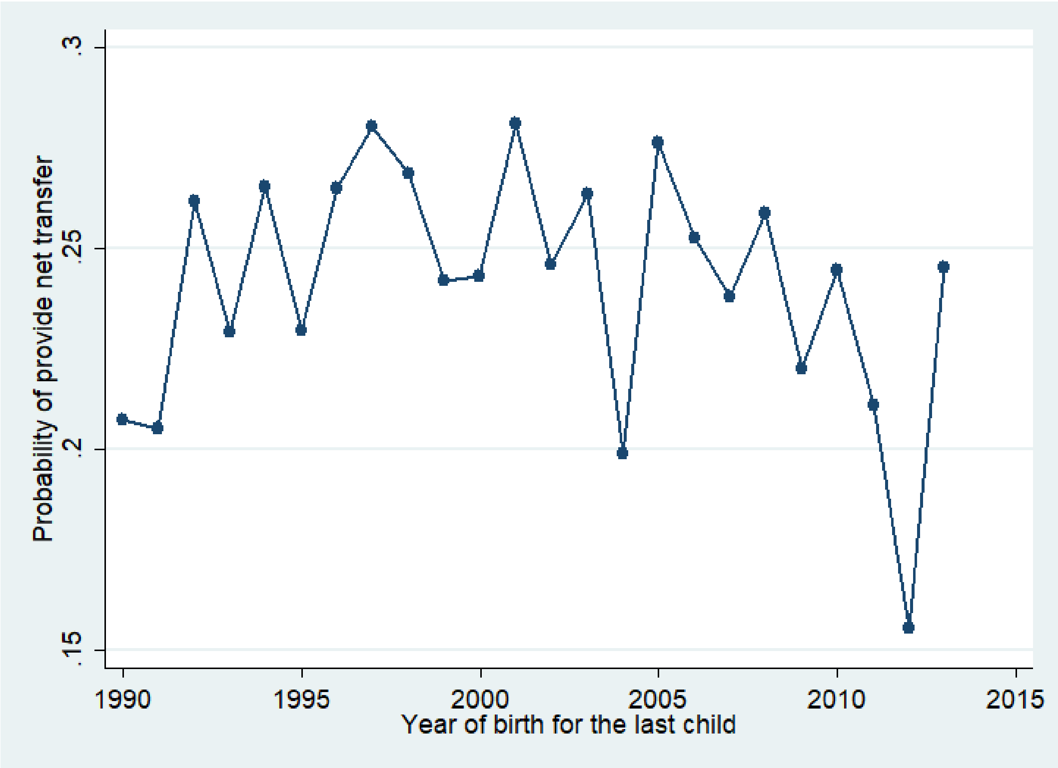
Table A.2: Summary Statistics: key variables females and males

VARIABLES	CHARLS				CHFS			
	Females		Males		Females		Males	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
whether P provides								
any transfers	0.254	0.264	0.314	0.341	0.301	0.459	0.228	0.420
regular transfer	0.045	0.166	0.164	0.336	-	-	-	-
non-regular transfer	0.222	0.262	0.265	0.346	-	-	-	-
log amount of								
total transfer	-	-	-	-	2.089	3.318	1.573	3.056
regular transfer	0.411	1.648	0.771	2.257	-	-	-	-
non-regular transfer	2.597	3.157	2.117	3.199	-	-	-	-
log visit days	2.608	2.115	3.806	2.454	1.898	2.135	1.795	2.371
gender ratio of K	0.567	0.406	0.558	0.405	0.559	0.426	0.575	0.407
No. of Y	1.648	0.781	1.637	0.766	1.585	0.833	1.740	0.936
age of P	38.11	8.956	38.81	8.737	46.91	10.35	49.44	9.822
income level of P	5.085	1.417	5.076	1.419	-	-	-	-
income of P	-	-	-	-	22510	43919	21049	43347
education of P	0.814	0.531	0.960	0.444	0.801	0.652	0.864	0.638
whether P has a rural *hukou*	0.766	0.423	0.767	0.423	0.493	0.500	0.597	0.491
P living in rural areas	0.649	0.477	0.655	0.476	0.268	0.443	0.395	489
No. of siblings of P	3.875	1.598	3.645	1.617	3.189	1.821	3.248	1.890
any transfers from O	0.034	0.182	0.041	0.197	0.169	0.375	0.118	0.323

Table A.3: Summary Statistics: Full variables list CHARLS

	CHARLS			
VARIABLES	Mean	Std. Dev.	Min	Max
whether P provides				
any transfers	0.284	0.306	0	1
regular transfer	0.105	0.272	0	1
non-regular transfer	0.243	0.308	0	1
log amount of				
regular transfer	0.582	1.970	0	17.73
non-regular transfer	2.393	3.183	0	18.42
log visit days	3.234	2.374	0	5.900
gender of P	0.513	0.500	0	1
gender ratio of K	0.562	0.405	0	1
No. of Y	1.643	0.774	1	8
age of P	39.73	9.287	21	65
income level of P	5.078	1.420	1	11
education of P	0.892	0.496	0	2
whether P has a rural *hukou*	0.680	0.466	0	1
whether P is married	0.998	0.040	0	1
P living in rural areas	0.348	0.476	0	1
No. of siblings of P	3.758	1.612	1	10
P's ranking in siblings	2.391	1.396	1	10
professional title of P	0.105	0.547	0	4
distance from O	3.265	1.837	0	7
household head of O	0.439	0.496	0	1
average age of O	63.94	10.441	42	101
average working status of O	0.568	0.453	0	1
average pension of O	0.185	0.388	0	1
average education level of O	2.898	1.665	1	9.5
who should support O	1.626	1.042	1	5
have O retired	1.875	0.301	1	2
whether O have deposit	0.137	0.347	0	1
household income of O	157661	4336359	0	2.00e+8
hours of O taking care of grandchildren	530.901	1816.5	0	17136
any transfers from O	0.037	0.190	0	1

Figure A.2: Trend assumption for the instrumental variable (DDIV)



Note: X-axis is the year of birth for the last child in households and y-axis shows the average probability of providing old-age support for people who have their last child born in the same year

Table A.4: Summary Statistics: Full variables list CHFS

	CHFS			
VARIABLES	Mean	Std. Dev.	Min	Max
whether P provides any transfers	0.265	0.441	0	1
log amount of total transfer	1.831	3.200	0	11.92
log visit days	1.847	2.256	0	5.900
gender of P	0.499	0.500	0	1
gender ratio of K	0.567	0.416	0	1
No. of K	1.662	0.889	1	9
age of P	48.17	10.71	21	65
income of P	21779	43639	0	1649439
education of P	0.832	0.646	0	2
whether P has a rural *hukou*	0.546	0.498	0	1
marital status of P	1.929	0.264	0	2
P living in rural areas	0.332	0.471	0	1
No. of siblings of P	3.218	1.856	0	16
whether P is working	0.688	0.463	0	1
occupation of P	0.902	1.717	0	8
whether P has loan	0.095	0.293	0	1
annual expenditure of P household	20201	42379	0	1607725
No. of O alive	1.230	0.929	0	2
average education level of O	1.894	1.104	0	7
whether O are party members	2.729	0.551	0	3
hukou status of O	2.086	0.9291	0	3
any transfers from O	0.144	0.351	0	1

Table A.5: Summary Statistics for CHARLS: females and males subsamples

	CHARLS			
	Females		Males	
VARIABLES	Mean	Std. Dev.	Mean	Std. Dev.
whether P provides				
any transfers	0.254	0.264	0.314	0.341
regular transfer	0.045	0.166	0.164	0.336
non-regular transfer	0.222	0.262	0.265	0.346
log amount of				
regular transfer	0.411	1.648	0.771	2.257
non-regular transfer	2.597	3.157	2.117	3.199
log visit days	2.608	2.115	3.806	2.454
more sons in K	0.679	0.467	0.688	0.464
No. of Y	1.648	0.781	1.637	0.766
age of P	38.11	8.956	38.81	8.737
income level of P	5.085	1.417	5.076	1.419
education of P	0.814	0.531	0.960	0.444
whether P has a rural *hukou*	0.766	0.423	0.767	0.423
whether P is married	0.999	0.031	0.998	0.0462
P living in rural areas	0.351	0.477	0.345	0.476
No. of siblings of P	3.875	1.598	3.645	1.617
P's ranking in siblings	2.827	1.445	1.978	1.210
professional title of P	0.077	0.481	0.130	0.600
distance from O	3.874	1.332	2.703	2.048
household head of O	0.433	0.496	0.431	0.495
average age of O	65.25	9.622	66.04	9.552
average working status of O	0.550	0.455	0.536	0.456
average pension of O	0.180	0.384	0.182	0.385
average education level of O	2.735	1.564	2.690	1.556
who should support O	1.592	1.024	1.567	1.003
have O retired	1.874	0.302	1.870	0.305
whether O have deposit	0.124	0.330	0.129	0.336
household income of O	103669	3454041	129728	3796947
hours of O taking care of grandchildren	217.61	1124	827.9	2248
any transfers from O	0.034	0.182	0.041	0.197

Table A.6: Summary Statistics for CHFS: females and males subsamples

	CHFS			
	Females		Males	
VARIABLES	Mean	Std. Dev.	Mean	Std. Dev.
whether P provides any transfers	0.301	0.459	0.228	0.420
log amount of total transfer	2.089	3.318	1.573	3.056
log visit days	1.898	2.135	1.795	2.371
gender ratio of K	0.559	0.426	0.575	0.407
No. of K	1.585	0.833	1.740	0.936
age of P	46.91	10.35	49.44	9.822
income of P	22510	43919	21049	43347
education of P	0.801	0.652	0.864	0.638
whether P has a rural *hukou*	0.493	0.500	0.597	0.491
P living in rural areas	0.268	0.443	0.395	489
No. of siblings of P	3.189	1.821	3.248	1.890
whether P is working	0.576	0.494	0.801	0.400
occupation of P	0.789	1.597	1.014	1.822
whether P has loan	0.096	0.295	0.934	0.291
annual expenditure of P household	20658	43260	19746	41482
No. of O alive	1.279	0.948	1.181	0.904
average education level of O	1.974	1.137	1.813	1.064
whether O are party members	2.722	0.546	2.736	0.555
hukou status of O	1.372	0.504	1.283	0.904
any transfers from O	0.144	0.351	0.118	0.323

Table A.7: Gender effect of K on provision support for different gender of P: IV results

	Probit IV: CHARLS			Probit IV: CHFS
VARIABLES	*any-transfer*	*regular*	*nonregular*	*any-transfer*
Male P				
genderK	0.098***	0.065***	0.038**	-0.103
	(0.021)	(0.013)	(0.018)	(0.114)
Observations	6,690	6,690	6,690	9,628
Female P				
genderK	-0.131	-0.011	-0.045	-0.162*
	(0.104)	(0.019)	(0.036)	(0.095)
Observations	5,542	5,542	5,542	9,632
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I, except the regressions are run separately for the females and males and the regression model is the probit IV model.

Table A.8: Gender effect of K and P on provision of old-age support in CHARLS

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0791**	0.0268	-0.119***	0.251**	-0.467**	0.0856
	(0.0332)	(0.0172)	(0.0320)	(0.123)	(0.210)	(0.113)
moresons	-0.0312	0.0202	-0.0622*	0.178	-0.309	0.0497
	(0.0382)	(0.0203)	(0.0378)	(0.145)	(0.249)	(0.129)
$maleP \times moresons$	0.113**	0.00707	0.127***	-0.00840	0.583**	0.274*
	(0.0463)	(0.0237)	(0.0441)	(0.168)	(0.291)	(0.152)
$moresons+$	0.082***	0.027**	0.064***	0.169**	0.274**	0.324***
$maleP \times moresons$	(0.024)	(0.011)	(0.020)	(0.076)	(0.136)	(0.074)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.202	0.074	0.137	0.071	0.138	0.531

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *moresons* represents the female dominated demonstration effect and $moresons + maleP \times moresons$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.9: Gender effect of K and P on provision of old-age support in CHFS

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	-0.092***	-0.524**	-0.237	9.183
	(0.035)	(0.266)	(0.177)	(10.61)
moresons	-0.054	-0.213	-0.557***	-20.45**
	(0.033)	(0.244)	(0.143)	(8.658)
$maleP \times moresons$	0.063	0.307	0.404	17.83
	(0.054)	(0.404)	(0.263)	(16.00)
$moresons+$	0.009	0.094	-0.153	-2.611
$maleP \times moresons$	(0.037)	(0.283)	(0.199)	(12.05)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.226	0.223	0.154	0.110

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *moresons* represents the female dominated demonstration effect and $moresons + maleP \times moresons$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.10: First stage for two constructed instrumental variables

VARIABLES	*genderK*	
	CHARLS	CHFS
$genderK\_1\_2003$	0.263***	0.483***
	(0.007)	(0.007)
$prefectural\_index$	-0.033***	-
	(0.009)	-
P demographics	Yes	Yes
O demographics	Yes	Yes
Observations	12,232	19,524
F-test	289.96	551.02
Under-identification test		
Kleibergen-Paap rk LM statistic	422.4	1637.99
Weak identification test		
Cragg-Donald Wald F-stat.	161.45	2561.16
Anderson-Rubin Wald F test	7.08	4554.97
Over-identification test		
Hansen J statistic	3.616	-

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The coefficient presented here for first stage coefficients for the IV regression. $genderK\_1\_2003$ is the gender of the first-born child in the family after 2003 together and $prefectural\_index$ is the index that indicating how strict the cities on the gender selection behaviours at prefectural-level. The controls are the same as the previous regressions.

Table A.11: Do you think is better to have sons than daughters?

	Urban		Rural	
CHFS	No.	Percentage	No.	Percentage
Prefer sons	1,159	8.43%	621	9.25%
Prefer daughters	2,904	21.12%	672	10.01%
Same	9,685	70.45%	5,423	80.75%

Table A.12: Predictor means for the synthetic control results

	Male Complier		Female Complier	
Predictor	Real	Synthetic	Real	Synthetic
age	44.41	44.64	43.94	43.63
rural *hukou*	0.709	0.560	0.687	0.673
whether working	0.875	0.868	0.828	0.793
occupation	0.874	1.273	1.159	0.948
total kids	1.991	1.697	1.922	2.081
O living together	0.228	0.241	0.083	0.157
No. of O alive	2.335	2.063	2.061	2.040
No. of siblings	3.066	3.054	3.262	3.252
education	0.798	0.939	0.790	0.704
marital status	1.961	1.969	1.892	1.964
household expenditure	25563	23396	21738	23074
han ethic	0.910	0.919	0.869	0.927

Table A.13: Heterogeneity Check: household income level

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0932**	0.00746	-0.103**	0.0427	-0.357	0.593***
	(0.0421)	(0.0213)	(0.0399)	(0.142)	(0.250)	(0.148)
genderK×	0.118*	0.0676*	0.0394	0.620**	0.235	-1.030***
high income	(0.0680)	(0.0378)	(0.0661)	(0.262)	(0.436)	(0.237)
high income	0.0619	0.00887	0.0689	0.00247	0.681**	0.669***
	(0.0534)	(0.0287)	(0.0532)	(0.203)	(0.345)	(0.181)
genderK	-0.00793	0.0282	-0.0329	0.209	0.0421	0.500**
	(0.0581)	(0.0283)	(0.0561)	(0.189)	(0.352)	(0.197)
maleP × *genderK*	0.182***	0.0457	0.141**	0.346	0.676	-0.347
	(0.0695)	(0.0349)	(0.0654)	(0.231)	(0.412)	(0.235)
maleP×	-0.0559	-0.0175	-0.0656	-0.0729	-0.740	-0.949***
high income	(0.0890)	(0.0464)	(0.0875)	(0.329)	(0.577)	(0.301)
maleP × *genderK*	-0.256**	-0.128**	-0.106	-1.085***	-0.801	1.333***
×*high income*	(0.113)	(0.0600)	(0.108)	(0.417)	(0.713)	(0.384)
High income HH male	-0.137***	-0.072***	-0.063	-0.603***	-0.823***	0.537***
demonstrate effects	(0.046)	(0.027)	(0.043)	(0.184)	(0.285)	(0.159)
Low income HH male	0.174***	0.074***	0.108***	0.555***	0.718***	0.153
demonstrate effects	(0.033)	(0.017)	(0.029)	(0.119)	(0.193)	(0.112)
Differences in males	-0.312***	-0.145***	-0.171***	-1.158***	-1.541***	0.384*
demonstrate effects	(0.065)	(0.038)	(0.060)	(0.258)	(0.399)	(0.229)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.197	0.069	0.138	0.063	0.135	0.529

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Appendix: Figures and Tables

Appendix A Appendix: The model

A.1 FOCs and SOCs from the basic models

A.1.1 Basic model without intra-household bargaining

In this section, I need to first prove the signs of SOCs. Recall that function u is strictly concave in c_1. $\mathcal{T}^F$ and $\mathcal{T}^M$ are both strictly concave functions. $c_1 = Y_1(2-\tau_1^F-\tau_1^M)$ and $c_2 = Y_2(\tau_2^F\phi n+\tau_2^M(1-\phi)n)$. From U^1, the SOCs are:

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F2}} = u''(c_1)(Y_1^2) + \delta u''(Y_1(\tau_1^F + \tau_1^M))Y_1^2$$
$$+ \beta u'(c_2)Y_2\tau_2^{F''}\phi n + \beta u''(c_2)(Y_2\tau_2^{F'}\phi n)^2 < 0; \tag{9}$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\phi} = \beta u''(c_2)(Y_2^2\tau_2^{F'}\phi n)(n\tau_2^F - n\tau_2^M) + \beta u'(c_2)Y_2\tau_2^{F'}n.$$

Again $U^{11} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*2}}$ and $U^{13} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi}$, which are the SOCs at the optimal value of τ^F and τ^M. If the function $u(\cdot)$ is specified as a log or a CRRA function, $|u''(c_2)(Y_2^2\tau_2^{F'}\phi n)(n\tau_2^F-n\tau_2^M)|< |u'(c_2)Y_2\tau_2^{F'}n|$, so $U^{13} > 0$.

From U^2, the SOCs are:

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M2}} = u''(c_1)(Y_1^2) + \delta u''(Y_1(\tau_1^F + \tau_1^M))Y_1^2$$
$$+ \beta u'(c_2)Y_2\tau_2^{M''}(1-\phi)n + \beta u''(c_2)(Y_2\tau_2^{M'}(1-\phi)n)^2 < 0; \tag{10}$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}\phi} = \beta u''(c_2)(Y_2^2\tau_2^{M'}\phi n)(n\tau_2^F - n\tau_2^M) - \beta u'(c_2)Y_2\tau_2^{M'}n < 0.$$

The SOC for τ_1^F and τ_1^M is:

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M} = u''(c_1)(Y_1^2) + \delta u''(Y_1(\tau_1^F + \tau_1^M))Y_1^2 + \beta u''(c_2)Y_2^2\tau_2^{F'}\tau_2^{M'}\phi(1-\phi)n^2 < 0. \tag{11}$$

$U^{22} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*2}}$ and $U^{23} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*}\mathrm{d}\phi}$, which are the SOCs at the optimal value of τ^F and τ^M. Also $U^{12} = U^{21} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\tau_1^{M*}}$.

From the basic model, I get the following conditions:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23} - U^{13}U^{22}}{U^{11}U^{22} - U^{12}U^{21}}; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23} - U^{13}U^{21}}{U^{12}U^{21} - U^{11}U^{22}};$$

The signs for SOCs when τ_1^F and τ_1^M at their optimal values are:

$$U^{11} < 0; \quad U^{13} > 0; \quad U^{22} < 0; \quad U^{23} < 0; \quad U^{12} = U^{21} < 0.$$

Recall that the U^{ij}s are the SOCs when τ_1^F and τ_1^M at their optimal values. $i = \{1, 2\}$ and $j = \{1, 2, 3\}$. From the equations for SOCs, I can get the sign of the numerator and denominator:

$$U^{12}U^{23} - U^{13}U^{22} > 0;$$

$$U^{11}U^{23} - U^{13}U^{21} > 0;$$

$$U^{11}U^{22} - U^{12}U^{21} > 0,$$

so the comparative statistics are:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23} - U^{13}U^{22}}{U^{11}U^{22} - U^{12}U^{21}} > 0; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23} - U^{13}U^{21}}{U^{12}U^{21} - U^{11}U^{22}} < 0;$$

Lemma 1 holds under these conditions.

A.1.2 Basic model with intra-household bargaining

In this model, the FOCs and SOCs are different under different circumstances. $u(c)$ is again specified in log utility or CRRA function form if $c > 0$ and I assume $u(0) = 0$. The detailed equations and signs for the SOCs are:

When $Y^F \geq Y^M$ and $c_2 = Y_2^F \tau_2^F \phi n$:

$$\begin{aligned} \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F2}} &= \rho_1 u''(Y_1^F(1-2\tau_1^F))4Y_1^{F2} + \rho_1\delta 4Y_1^{F2}u''(Y_1^F\tau_1^F) \\ &\quad + \beta u'(c_2)4Y_2^F\tau_2^{F''}\phi n + \beta u''(c_2)(2Y_2^F\tau_2^{F'}\phi n)^2 < 0; \end{aligned} \tag{12}$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\phi} = \beta u'(c_2)2Y_2^F\tau_2^{F'}n + \beta u''(c_2)2(Y_2^Fn)^2\tau_2^{F'}\tau_2^F\phi > 0.$$

When $Y^F < Y^M$, $c_2 = Y_2^M\tau_2^M(1-\phi)n$ and $\eta_1 = 0$:

$$\begin{aligned} \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M2}} &= (1-\rho_1)u''(Y_1^M(1-2\tau_1^M))4Y_1^{M2} + (1-\rho_1)\delta 4Y_1^{M2}u''(Y_1^M\tau_1^M) \\ &\quad + \beta u'(c_2)4Y_2^M\tau_2^{M''}(1-\phi)n + \beta u''(c_2)(2Y_2^M\tau_2^{M'}(1-\phi)n)^2 < 0; \end{aligned} \tag{13}$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}\phi} = -\beta u'(c_2)2Y_2^M\tau_2^{M'}n - \beta u''(c_2)2(Y_2^Mn)^2\tau_2^{M'}\tau_2^M(1-\phi) < 0.$$

Again $U^{11} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*2}}$, $U^{13} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi}$, $U^{22} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*2}}$, and $U^{23} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*}\mathrm{d}\phi}$, which are the SOCs at the optimal value of τ^F and τ^M. When $Y^F \geq Y^M$, $U^{11} < 0$ and $U^{13} > 0$. When $Y^F < Y^M$, then $U^{22} < 0$ and $U^{23} < 0$. Then *Lemma 2* holds.

A.2 Extensions of the basic model

A.2.1 The demonstration effect from different gender parents

In the basic model without intra-household resource allocation, I assume the parents transfer would affect the future transfer only from the same-gender next generation. The formulas for these assumptions are $\tau_2^M = \mathcal{T}_2^M(\tau_1^M)$ and $\tau_2^F = \mathcal{T}_2^F(\tau_1^F)$, and they are both strictly concave functions. Now I relax this assumption to both parents would affect the future transfer of the next generation, but the parent would have larger influences to the same-gender next generation than the hetero-sex one. The equations of these assumptions are: $\tau_2^M = \mathcal{T}_2^M(\tau_1^F, \tau_1^M)$ and $\tau_2^F = \mathcal{T}_2^F(\tau_1^F, \tau_1^M)$. The functions are still concave in both τ_1^F and τ_1^M, but I assume

$$\frac{\mathrm{d}\tau_2^F}{\mathrm{d}\tau_1^F} > \frac{\mathrm{d}\tau_2^F}{\mathrm{d}\tau_1^M}; \quad \frac{\mathrm{d}\tau_2^M}{\mathrm{d}\tau_1^M} > \frac{\mathrm{d}\tau_2^M}{\mathrm{d}\tau_1^F}; \quad \frac{\mathrm{d}\tau_2^M}{\mathrm{d}\tau_1^M} > \frac{\mathrm{d}\tau_2^F}{\mathrm{d}\tau_1^M}; \quad \frac{\mathrm{d}\tau_2^F}{\mathrm{d}\tau_1^F} > \frac{\mathrm{d}\tau_2^M}{\mathrm{d}\tau_1^F}.$$

For the second order conditions:

$$|\frac{\mathrm{d}^2\tau_2^F}{\mathrm{d}\tau_1^{F2}}| > |\frac{\mathrm{d}^2\tau_2^F}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M}|; \quad |\frac{\mathrm{d}^2\tau_2^M}{\mathrm{d}\tau_1^{M2}}| > |\frac{\mathrm{d}^2\tau_2^M}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M}|.$$

The interpretation of the SOCs for τ_2^F and τ_2^M is that the effects from the same-gender old-age support on the marginal return of the old-age support are larger than the opposite-gender effects. Under these new assumptions, I re-consider the basic model without the intra-household bargaining. The maximization problem is still the same except for the equations of τ_2^F and τ_2^M. I would still want *Lemma 1* to hold given the new SOCs and FOCs derived from the optimisation problem. The new maximisation problem is:

$$\max_{\tau_1^F, \tau_1^M} U = u(c_1) + \delta u(e_1) + \beta u(c_2) \tag{14}$$

$$s.t.$$

$$c_1 + c_2 \leq Y_1(2 - \tau_1^F - \tau_1^M) + Y_2(\tau_2^F\phi n + \tau_2^M(1-\phi)n);$$

$$e_1 = Y_1(\tau_1^F + \tau_1^M).$$

The expressions for c_2 is $Y_2(\tau_2^F\phi n + \tau_2^M(1-\phi)n)$. For τ_1^F and τ_1^M, the FOCs:

$$\begin{aligned} U^1 = &\frac{\mathrm{d}U}{\mathrm{d}\tau_1^F} = u'(Y_1(2-\tau_1^F-\tau_1^M))(-Y_1) + \delta u'(Y_1(\tau_1^F+\tau_1^M))Y_1 \\ &+ \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^F}^{F'}\phi + \tau_{2,\tau_1^F}^{M'}(1-\phi)) = 0; \\ U^2 = &\frac{\mathrm{d}U}{\mathrm{d}\tau_1^M} = u'(Y_1(2-\tau_1^F-\tau_1^M))(-Y_1) + \delta u'(Y_1(\tau_1^F+\tau_1^M))Y_1 \\ &+ \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^M}^{F'}\phi + \tau_{2,\tau_1^M}^{M'}(1-\phi)) = 0; \end{aligned}$$

The SOCs are:

$$\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{F2}} = u''(Y_1(2-\tau_1^F-\tau_1^M))Y_1^2 + \delta u''(Y_1(\tau_1^F+\tau_1^M))Y_1^2$$

$$+ \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^F}^{F''}\phi + \tau_{2,\tau_1^F}^{M''}(1-\phi)) + \beta u''(c_2)(Y_2 n(\tau_{2,\tau_1^F}^{F'}\phi + \tau_{2,\tau_1^F}^{M'}(1-\phi)))^2 < 0;$$

$$\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\phi} = \beta u''(c_2)Y_2^2 n^2(\tau_{2,\tau_1^F}^{F'}\phi + \tau_{2,\tau_1^F}^{M'}(1-\phi))(\tau_2^F - \tau_2^M) + \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^F}^{F'} - \tau_{2,\tau_1^F}^{M'}).$$

$$\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{M2}} = u''(Y_1(2-\tau_1^F-\tau_1^M))Y_1^2 + \delta u''(Y_1(\tau_1^F+\tau_1^M))Y_1^2$$

$$+ \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^M}^{F''}\phi + \tau_{2,\tau_1^M}^{M''}(1-\phi)) + \beta u''(c_2)(Y_2 n(\tau_{2,\tau_1^M}^{F'}\phi + \tau_{2,\tau_1^M}^{M'}(1-\phi)))^2 < 0;$$

$$\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^M \mathrm{d}\phi} = \beta u''(c_2)Y_2^2 n^2(\tau_{2,\tau_1^M}^{F'}\phi + \tau_{2,\tau_1^M}^{M'}(1-\phi))(\tau_2^F - \tau_2^M) + \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^M}^{F'} - \tau_{2,\tau_1^M}^{M'}).$$

$$\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\tau_1^M} = u''(Y_1(2-\tau_1^F-\tau_1^M))Y_1^2 + \delta u''(Y_1(\tau_1^F+\tau_1^M))Y_1^2$$

$$+ \beta u''(c_2)Y_2^2 n^2(\tau_{2,\tau_1^M}^{F'}\phi + \tau_{2,\tau_1^M}^{M'}(1-\phi))(\tau_{2,\tau_1^F}^{F'}\phi + \tau_{2,\tau_1^F}^{M'}(1-\phi))$$

$$+ \beta u'(c_2)Y_2 n(\tau_{2,\tau_1^F,\tau_1^M}^{F''}\phi + \tau_{2,\tau_1^F,\tau_1^M}^{M''}(1-\phi)) < 0;$$

Depending on the value of τ_1^F, τ_1^M and $\tau_{2,M}^{F'}$ and $\tau_{2,F}^{M'}$, the signs for $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\phi}$ and $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^M \mathrm{d}\phi}$ vary. If the function $u(\cdot)$ is specified as a log or a CRRA function and $\tau_2^F < \tau_2^M$,

1. $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\phi} > 0$

2. If $\tau_{2,M}^{F'}$ is small enough or $\tau_{2,M}^{F'} \to 0$:

   - $\tau_{2,M}^{F'}\tau_2^M - \tau_2^F\tau_{2,M}^{M'} < 0$

   - $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^M \mathrm{d}\phi} < 0$

If the function $u(\cdot)$ is specified as a log or a CRRA function and $\tau_2^F > \tau_2^M$,

1. $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^M \mathrm{d}\phi} < 0$

2. If $\tau_{2,F}^{M'}$ is small enough or $\tau_{2,F}^{M'} \to 0$

   - $\tau_{2,F}^{F'}\tau_2^M - \tau_2^F\tau_{2,F}^{M'} > 0$

   - $\frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^F \mathrm{d}\phi} > 0.$

Recall that for the comparative statics are:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23} - U^{13}U^{22}}{U^{11}U^{22} - U^{12}U^{21}}; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23} - U^{13}U^{21}}{U^{12}U^{21} - U^{11}U^{22}};$$

τ_1^{F*} and τ_1^{M*} is the optimal solution of τ_1^F and τ_1^M from $U^1 = 0$ and $U^2 = 0$. The U^{ij}s are the SOCs listed above when τ_1^F and τ_1^M at their optimal values.[30] From the equations for SOCs, under two sets of conditions: (1) $\tau_2^{F*} < \tau_2^{M*}$ and $\tau_{2,\tau_1^{M*}}^{F*'}\tau_2^{M*} < \tau_2^{F*}\tau_{2,\tau_1^{M*}*}^{M*'}$; or (2) $\tau_2^{F*} > \tau_2^{M*}$ and $\tau_{2,\tau_1^{F*}}^{F*'}\tau_2^{M*} > \tau_2^{F*}\tau_{2,\tau_1^{F*}}^{M*'}$, I can get the sign of the numerator:

$$U^{12}U^{23} - U^{13}U^{22} > 0;$$

$$U^{11}U^{23} - U^{13}U^{21} > 0.$$

For the denominator, if $|\tau_{2,\tau_1^{F*}}^{F*''}| > |\tau_{2,\tau_1^{F*},\tau_1^{M*}}^{F*''}|$, $|\tau_{2,\tau_1^{M*}}^{M*''}| > |\tau_{2,\tau_1^{F*},\tau_1^{M*}}^{M*''}|$, $\tau_{2,\tau_1^{F*},\tau_1^{M*}}^{F*''} \to 0$, and $\tau_{2,\tau_1^{F*},\tau_1^{M*}}^{M*''} \to 0$, then

$$U^{11}U^{22} - U^{12}U^{21} > 0.$$

According to the SOCs, and the assumptions imposed for the sign of these conditions, the comparative statics for $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}$ are consistent with the model under more restrictive assumptions: $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$, and *Lemma 1* holds under these conditions. These conditions are quite strict and less likely to be true in the real world scenario unless the cross-gender effect is really small.. But in the combined model in Section A.2.3, the conditions are more likely to be satisfied.

A.2.2 Relaxed intra-household resource allocation condition

In the basic intra-household bargaining model, I only analyse the cases when $\eta_1 = 1$ or $\eta = 0$. These are extreme cases that the party lose the bargaining provides no transfer to their parents. I relax this assumption and the party lose the bargaining provides a smaller proportion of their transfer to their parents. The winning party provides a larger proportion. The proportion can be represented by η_1 defined in this intra-household bargaining model. Again, $\eta_t = \mathcal{H}(\frac{Y_t^F}{Y_t^M})$ for $t = \{1, 2\}$. $\mathcal{H}(\frac{Y_t^F}{Y_t^M})$ is an increasing function in $\frac{Y_t^F}{Y_t^M}$ and $0 \le \eta_t \le 2$. When $Y_t^F = Y_t^M$, $\eta_t = 1$. In this model, I only assume the same-gender demonstration effect, which $\tau_2^F = \mathcal{T}^F(\tau_1^F)$ and $\tau_2^M = \mathcal{T}^M(\tau_1^M)$. I again assume the new maximization problem for a household in the second generation is:

$$\begin{aligned} \max_{\tau_1^F, \tau_1^M} \quad & U = \rho_1 u(c_1^F) + (1-\rho_1)u(c_1^M) + \delta\rho_1 u(e_1^F) \\ & + \delta(1-\rho_1)u(e_1^M) + \beta u(c_2); \end{aligned} \tag{15}$$

[30] $i = \{1, 2\}$ and $j = \{1, 2, 3\}$. $U^{11} = \frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{F*2}}$, $U^{13} = \frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi}$, $U^{22} = \frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{M*2}}$, $U^{23} = \frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{M*}\mathrm{d}\phi}$, and $U^{12} = U^{21} = \frac{\mathrm{d}^2 U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\tau_1^{M*}}$, which are the SOCs at the optimal value of τ^F and τ^M.

$$s.t.$$
$$c_1^F + c_1^M + c_2 \leq Y_1^F(1-\eta_1\tau_1^F) + Y_1^M(1-(2-\eta_1)\tau_1^M)$$
$$+ Y_2^F\eta_2\mathcal{T}^F(\tau_1^F)\phi n + Y_2^M(2-\eta_2)\mathcal{T}^M(\tau_1^M)(1-\phi)n;$$
$$e_1^F = \eta_1 Y_1^F\tau_1^F;$$
$$e_1^M = (2-\eta_1)Y_1^M\tau_1^M.$$

Lemma 3: *$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$. When $Y_1^F \geq Y_1^M$ for all t, $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$ and when $Y_1^F < Y_1^M$ for all t, $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| < |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$.*

When $c_2 = \eta_2 Y_2^F\tau_2^F\phi n + (2-\eta_2)Y_2^M\tau_2^M(1-\phi)n$, the FOCs for τ_1^F and τ_1^M following the previous setting are:

$$U^1 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^F} = \rho_1\eta_1 u'(Y_1^F(1-\eta_1\tau_1^F))(-Y_1^F) + \rho_1\eta_1\delta Y_1^F u'(\eta_1 Y_1^F\tau_1^F) + \beta u'(c_2)\eta_2 Y_2^F\tau_2^{F'}\phi n = 0$$

$$U^2 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^M} = (1-\rho_1)(2-\eta_1)u'(Y_1^M(1-(2-\eta_1)\tau_1^M))(-Y_1^M) + (1-\rho_1)(2-\eta_1)_1\delta Y_1^M u'((2-\eta_1)Y_1^M\tau_1^M)$$
$$+ \beta u'(c_2)(2-\eta_2)(1-\phi)Y_2^M\tau_2^{M'} n = 0; \tag{16}$$

τ_1^{F*} and τ_1^{M*} is the optimal solution of τ_1^F and τ_1^M from $U^1 = 0$ and $U^2 = 0$. The corresponding SOCs are:

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F2}} = \rho_1\eta_1^2 u''(Y_1^F(1-\eta_1\tau_1^F))Y_1^{F2} + \rho_1\eta_1^2\delta Y_1^{F2}u''(\eta_1 Y_1^F\tau_1^F)$$
$$+ \beta u'(c_2)\eta_2 Y_2^F\tau_2^{F''}\phi n + \beta u''(c_2)(\eta_2 Y_2^F\tau_2^{F'}\phi n)^2 < 0;$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\phi} = \beta u'(c_2)\eta_2 Y_2^F\tau_2^{F'} n + \beta u''(c_2)\eta_2 Y_2^F\tau_2^{F'}\phi n^2(\eta_2 Y_2^F\tau_2^F - (1-\eta_2)Y_2^M\tau_2^M);$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M2}} = (1-\rho_1)(2-\eta_1)^2 u''(Y_1^M(1-(2-\eta_1)\tau_1^M))Y_1^{M2} + (1-\rho_1)(2-\eta_1)_1^2\delta Y_1^{M2}u''((2-\eta_1)Y_1^M\tau_1^M);$$
$$+ \beta u'(c_2)(2-\eta_2)Y_2^M\tau_2^{M''}(1-\phi)n + \beta u''(c_2)((2-\eta_2)Y_2^M\tau_2^{M'}(1-\phi)n)^2 < 0;$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}\phi} = -\beta u'(c_2)(2-\eta_2)Y_2^M\tau_2^{M'} n$$
$$+ \beta u''(c_2)(2-\eta_2)Y_2^M\tau_2^{M'}(1-\phi)n^2(\eta_2 Y_2^F\tau_2^F - (2-\eta_2)Y_2^M\tau_2^M);$$
$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M} = \beta u''(c_2)\eta_2(2-\eta_2)\phi(1-\phi)Y_2^M\tau_2^{M'}Y_2^F\tau_2^{F'}n^2 < 0.$$

Again The U^{ij}s are the SOCs listed above when τ_1^F and τ_1^M at their optimal values.[31] $U^{11} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*2}}$, $U^{13} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\phi}$, $U^{22} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*2}}$, $U^{23} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*}\mathrm{d}\phi}$, and $U^{12} = U^{21} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}\tau_1^{M*}}$. If $u(\cdot)$ is specified as

[31] $i = \{1, 2\}$ and $j = \{1, 2, 3\}$.

a log or a CRRA function, then $U^{13} > 0$ and $U^{23} < 0$. The comparative statistics from the total differentiation equations of FOCs in equation (16) are:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23} - U^{13}U^{22}}{U^{11}U^{22} - U^{12}U^{21}}; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23} - U^{13}U^{21}}{U^{12}U^{21} - U^{11}U^{22}}. \tag{17}$$

From the SOCs, I can infer $U^{11}U^{22} > U^{12}U^{21}$, so:

$$U^{12}U^{23} - U^{13}U^{22} > 0;$$

$$U^{11}U^{23} - U^{13}U^{21} > 0;$$

$$U^{11}U^{22} - U^{12}U^{21} > 0.$$

Under these conditions, $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$, so the first half of *Lemma 3* holds. For the second part of *Lemma 3*, I need to analyse under different circumstances. In *Lemma 3*, I need to prove $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$ when $Y_t^F > Y_t^M$ for $t = \{1, 2\}$. From equation (17), $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$ if $U^{12}U^{23} - U^{13}U^{22} > U^{11}U^{23} - U^{13}U^{21}$, then:

$$U^{12}U^{23} - U^{13}U^{22} > U^{11}U^{23} - U^{13}U^{21};$$

$$U^{23}(U^{12} - U^{11}) + U^{13}(U^{21} - U^{22}) > 0;$$

When $\tau_2^{M*}\tau_2^{F*'} < \tau_2^{F*}\tau_2^{M*'}$, then $|U^{13}| < |U^{23}|$. When $Y_t^F \geq Y_t^M$ for $t = \{1, 2\}$, then $0 < U^{21} - U^{22} < U^{12} - U^{11}$. Under these two sets of conditions, $U^{12}U^{23} - U^{13}U^{22} < U^{11}U^{23} - U^{13}U^{21}$. So $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| < |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$.

When $\tau_2^{M*}\tau_2^{F*'} > \tau_2^{F*}\tau_2^{M*'}$, then $|U^{13}| > |U^{23}|$. When $Y_t^F < Y_t^M$ for $t = \{1, 2\}$, then $U^{21} - U^{22} > U^{12} - U^{11} > 0$. Under these two conditions, $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$. The second half of the *Lemma 3* does not hold.

This is because the diminishing marginal return of the utility function. Take the scenario under $Y_t^F > Y_t^M$ when $t = \{1, 2\}$ for example. Because the fathers in the next generation earns more and the mothers earn less, so the magnitude of the same-gender demonstration effect is smaller for fathers and larger for mothers to obtain the same amount of the transfers from the next generation. Although the second half of *Lemma 3* cannot hold, the same-gender demonstration effect still exists. If I times the corresponding income to $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}$, then I have $|\frac{\mathrm{d}Y_2^F\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}Y_2^M\tau_1^{M*}}{\mathrm{d}\phi}|$ when the difference between Y_t^F and Y_t^M for $t = \{1, 2\}$ is large enough.

A.2.3 Combined model

In the previous setting, I show different models under various relaxed assumptions. I combined all the relaxed assumptions in this model. The first relaxed assumption is the transfer from the next generation is affected by both parents. The second relaxed assumption is that the party earns less then provides a smaller proportion of their transfer to their parents. Under this two new assumptions and the basic bargaining model, the new optimisation problem is

$$\begin{aligned}\max_{\tau_1^F,\tau_1^M} \quad U = & \rho_1 u(c_1^F) + (1-\rho_1)u(c_1^M) + \delta\rho_1 u(e_1^F) \\ & + \delta(1-\rho_1)u(e_1^M) + \beta u(c_2);\end{aligned} \tag{18}$$

$$s.t.$$

$$\begin{aligned} c_1^F + c_1^M + c_2 \leq & Y_1^F(1-\eta_1\tau_1^F) + Y_1^M(1-(2-\eta_1)\tau_1^M) \\ & + Y_2^F\eta_2\mathcal{T}^F(\tau_1^F)\phi n + Y_2^M(2-\eta_2)\mathcal{T}^M(\tau_1^M)(1-\phi)n; \\ e_1^F = & \eta_1 Y_1^F\tau_1^F; \\ e_1^M = & (2-\eta_1)Y_1^M\tau_1^M. \end{aligned}$$

In this combined model, I would like to *Lemma 3* to hold. $c_2 = \eta_2 Y_2^F\tau_2^F\phi n + (2-\eta_2)Y_2^M\tau_2^M(1-\phi)n)$. The FOCs for τ_1^F and τ_1^M following the previous setting are:

$$\begin{aligned} U^1 = & \frac{\mathrm{d}U}{\mathrm{d}\tau_1^F} = \rho_1\eta_1 u'(Y_1^F(1-\eta_1\tau_1^F))(-Y_1^F) + \rho_1\eta_1\delta Y_1^F u'(\eta_1 Y_1^F\tau_1^F) \\ & + \beta u'(c_2)(\eta_2 Y_2^F\tau_{2,\tau_1^F}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M'}(1-\phi)n) = 0; \\ U^2 = & \frac{\mathrm{d}U}{\mathrm{d}\tau_1^M} = (1-\rho_1)(2-\eta_1)u'(Y_1^M(1-(2-\eta_1)\tau_1^M))(-Y_1^M) + (1-\rho_1)(2-\eta_1)\delta Y_1^M u'((2-\eta_1)Y_1^M\tau_1^M) \\ & + \beta u'(c_2)(\eta_2 Y_2^F\tau_{2,\tau_1^M}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M'}(1-\phi)n) = 0. \end{aligned}$$

The corresponding SOCs are:

$$\begin{aligned} \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F2}} = & \rho_1\eta_1^2 u''(Y_1^F(1-\eta_1\tau_1^F))Y_1^{F2} + \rho_1\eta_1^2\delta Y_1^{F2}u''(\eta_1 Y_1^F\tau_1^F) \\ & + \beta u'(c_2)(\eta_2 Y_2^F\tau_{2,\tau_1^F}^{F''}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M''}(1-\phi)n) \\ & + \beta u''(c_2)(\eta_2 Y_2^F\tau_{2,\tau_1^F}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M'}(1-\phi)n)^2 < 0; \\ \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\phi} = & \beta u'(c_2)(\eta_2 Y_2^F\tau_{2,\tau_1^F}^{F'}n - (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M'}n) \\ & + \beta u''(c_2)(\eta_2 Y_2^F\tau_2^F n - (2-\eta_2)Y_2^M\tau_2^M n) \\ & \times(\eta_2 Y_2^F\tau_{2,\tau_1^F}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M'}(1-\phi)n). \end{aligned}$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M2}} = (1-\rho_1)(2-\eta_1)^2u''(Y_1^M(1-(2-\eta_1)\tau_1^M))Y_1^{M2} + (1-\rho_1)(2-\eta_1)_1^2\delta Y_1^{M2}u''((2-\eta_1)Y_1^M\tau_1^M)$$

$$+\beta u'((c_2)(\eta_2Y_2^F\tau_{2,\tau_1^M}^{F''}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M''}(1-\phi)n)$$

$$+\beta u''(c_2)(\eta_2Y_2^F\tau_{2,\tau_1^M}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M'}(1-\phi)n)^2 < 0;$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}\phi} = \beta u'(c_2)(\eta_2Y_2^F\tau_{2,\tau_1^M}^{F'}n - (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M'}n)$$

$$+\beta u''(c_2)(\eta_2Y_2^F\tau_2^Fn - (2-\eta_2)Y_2^M\tau_2^Mn)$$

$$\times(\eta_2Y_2^F\tau_{2,\tau_1^M}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M'}(1-\phi)n);$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}\tau_1^M} = \beta u''(c_2)(\eta_2Y_2^F\tau_{2,\tau_1^F}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F}^{M'}(1-\phi)n)$$

$$\times(\eta_2Y_2^F\tau_{2,\tau_1^M}^{F'}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^M}^{M'}(1-\phi)n)$$

$$+\beta u'(c_2(\eta_2Y_2^F\tau_{2,\tau_1^F,\tau_1^M}^{F''}\phi n + (2-\eta_2)Y_2^M\tau_{2,\tau_1^F,\tau_1^M}^{M''}(1-\phi)n) < 0.$$

If the function $u(\cdot)$ is specified as a log or a CRRA function, and the cross-gender FOCs of τ_1^F and τ_1^M is small enough, $U^{13} > 0$ and $U^{23} < 0$. The proofs are similar to the proofs in the Section A.2.1.[32] These two conditions are especially true when $\frac{\mathrm{d}\tau_2^M}{\mathrm{d}\tau_1^F} \to 0$ and $\frac{\mathrm{d}\tau_2^F}{\mathrm{d}\tau_1^M} \to 0$.

The comparative statics from the total differentiation equations of the FOCs are again:

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{U^{12}U^{23}-U^{13}U^{22}}{U^{11}U^{22}-U^{12}U^{21}}; \quad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{U^{11}U^{23}-U^{13}U^{21}}{U^{12}U^{21}-U^{11}U^{22}}.$$

τ_1^{F*} and τ_1^{M*} is the optimal solution of τ_1^F and τ_1^M from $U^1 = 0$ and $U^2 = 0$. The U^{ij}s are the SOCs listed above when τ_1^F and τ_1^M at their optimal values.[33] From the SOCs, I can infer $U^{11}U^{22} > U^{12}U^{21}$, so:

$$U^{12}U^{23} - U^{13}U^{22} > 0;$$

$$U^{11}U^{23} - U^{13}U^{21} > 0;$$

$$U^{11}U^{22} - U^{12}U^{21} > 0.$$

Under these conditions, $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0$, so the first half of *Lemma 3* holds. For the second part of *Lemma 3*, I need to analyse under different circumstances. In *Lemma 3*, I need to prove $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$ when $Y_t^F > Y_t^M$ for $t = \{1,2\}$. From equation (17), $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$ if

[32] Depending on the value of τ_1^F, τ_1^M and $\tau_{2,M}^{F'}$ and $\tau_{2,F}^{M'}$, the signs for U^{13} and U^{23} vary. If the function $u(\cdot)$ is specified as a log or a CRRA function and $\tau_2^F < \tau_2^M$, then $U^{13} > 0$ and the sign of U^{23} depends on $\tau_{2,M}^{F'}\tau_2^M$ and $\tau_2^F\tau_{2,M}^{M'}$. If $\tau_{2,M}^{F'}$ is small enough or $\tau_{2,M}^{F'} \to 0$, then $U^{23} < 0$. If the function $u(\cdot)$ is specified as a log or a CRRA function and $\tau_2^F > \tau_2^M$, then $U^{23} < 0$ and the sign of U^{13} depends on $\tau_{2,F}^{F'}\tau_2^M$ and $\tau_2^F\tau_{2,F}^{M'}$. If $\tau_{2,F}^{M'}$ is small enough or $\tau_{2,F}^{M'} \to 0$, then $U^{13} > 0$.

[33] $i = 1, 2$ and $j = 1, 2, 3$.

$U^{12}U^{23} - U^{13}U^{22} > U^{11}U^{23} - U^{13}U^{21}$, then:

$$U^{12}U^{23} - U^{13}U^{22} > U^{11}U^{23} - U^{13}U^{21};$$

$$U^{23}(U^{12} - U^{11}) + U^{13}(U^{21} - U^{22}) > 0;$$

When $\tau_2^{M*}\tau_2^{F*'} < \tau_2^{F*}\tau_2^{M*'}$, then $|U^{13}| < |U^{23}|$. When $Y_t^F \geq Y_t^M$ for $t = \{1, 2\}$, then $0 < U^{21} - U^{22} < U^{12} - U^{11}$. Under these two sets of conditions, $U^{12}U^{23} - U^{13}U^{22} < U^{11}U^{23} - U^{13}U^{21}$. So $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| < |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$.

When $\tau_2^{M*}\tau_2^{F*'} > \tau_2^{F*}\tau_2^{M*'}$, then $|U^{13}| > |U^{23}|$. When $Y_t^F < Y_t^M$ for $t = \{1, 2\}$, then $U^{21} - U^{22} > U^{12} - U^{11} > 0$. Under these two conditions, $|\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}|$. The second half of the *Lemma 3* does not hold.

This is because the diminishing marginal return of the utility function. Take the scenario under $Y_t^F > Y_t^M$ for $t = \{1, 2\}$ for example. Because the fathers in the next generation earns more and the mothers earn less, so the magnitude of the same-gender demonstration effect is smaller for fathers and larger for mothers to obtain the same amount of the transfers from the next generation. Although the second half of *Lemma 3* cannot hold, the same-gender demonstration effect still exists. If I times the corresponding income to $\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi}$ and $\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi}$, then I have $|\frac{\mathrm{d}Y_2^F\tau_1^{F*}}{\mathrm{d}\phi}| > |\frac{\mathrm{d}Y_2^M\tau_1^{M*}}{\mathrm{d}\phi}|$ when the difference between Y_t^F and Y_t^M for $t = \{1, 2\}$ is large enough.

A.3 The baseline model with saving

I will only illustrate the model with saving in the baseline model without intra-household bargaining. The optimisation problem with savings is:

$$\max_{\tau_1^F, \tau_1^M, s_1} \quad U = u(c_1) + \delta u(e_1) + \beta u(c_2) \tag{19}$$

$$s.t.$$

$$c_1 + \frac{c_2}{1 + r_2} \leq Y_1(2 - \tau_1^F - \tau_1^M) + \frac{Y_2}{1 + r_2}(\tau_2^F \phi n + \tau_2^M (1 - \phi) n);$$

$$e_1 = Y_1(\tau_1^F + \tau_1^M).$$

Again, the father and the mother in the generation P make unitary household-level decisions. e_1 is old-age support provided by the whole household. δ is the discount factor for the utility generated from altruism. If the function $u(\cdot)$ is specified as a log or a CRRA function, and $\tau_2(\tau_1)$ is a concave function of τ_1 for the fathers and mothers, the FOCs regarding to τ_1^F, τ_1^M, and s are:

$$U^1 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^F} = 0$$

$$U^2 = \frac{\mathrm{d}U}{\mathrm{d}\tau_1^M} = 0$$

$$U_4 = \frac{\mathrm{d}U}{\mathrm{d}s_1} = -u'(c_1) + \beta u''(c_2)(1+r_2) = 0,$$

where $c_1 = Y_1(2 - \tau_1^F - \tau_1^M) - s_1$ and $c_2 = Y_2(\tau_2^F \phi n + \tau_2^M(1-\phi)n) + (1+r_2)s_1$. The expressions for U^{11}, U^{12}, U^{13}, U^{21}/U^{12}, U^{22}, and U^{23} similar to the SOCs listed in equations (9), (10), and (11) the previous model without savings (Appendix A.1.1). With savings, the signs of these SOCs will not change with savings included in the model. The expressions for U^{14}/U^{14}, U^{24}/U^{42}, U^{44}, and U^{34} are:

$$\frac{\mathrm{d}^2U}{\mathrm{d}s_1^2} = u''(Y_1(2-\tau_1^F-\tau_1^M)-s_1) + \beta u''(c_2)(1+r_2)^2 < 0;$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^F\mathrm{d}s_1} = u''(Y_1(2-\tau_1^F-\tau_1^M)-s_1)(Y_1) + \beta u''(c_2)Y_2\tau_2^{F'}\phi n(1+r_2) < 0;$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^M\mathrm{d}s_1} = u''(Y_1(2-\tau_1^F-\tau_1^M)-s_1)(Y_1) + \beta u''(c_2)Y_2\tau_2^{M'}(1-\phi)n(1+r_2) < 0;$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}s_1\mathrm{d}\phi} = \beta u''(c_2)Y_2(n\tau_2^F - n\tau_2^M);$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}s_1\mathrm{d}\phi} > 0 \quad \text{if} \quad \tau_2^F - \tau_2^M < 0$$

$$\frac{\mathrm{d}^2U}{\mathrm{d}s_1\mathrm{d}\phi} < 0 \quad \text{if} \quad \tau_2^F - \tau_2^M > 0$$

τ_1^{F*}, τ_1^{M*}, and s^* are the optimal solution from the FOCs. The U^{ij}s are the SOCs listed above when τ_1^F and τ_1^M at their optimal values.[34] The total differentiation equations of FOCs for τ_1^{F*}, τ_1^{M*}, and s^* are:

$$\begin{aligned} U^{11}\mathrm{d}\tau_1^{F*} + U^{12}\mathrm{d}\tau_1^{M*} + U^{13}\mathrm{d}\phi + U^{14}\mathrm{d}s_1^* &= 0 \\ U^{21}\mathrm{d}\tau_1^{F*} + U^{22}\mathrm{d}\tau_1^{M*} + U^{23}\mathrm{d}\phi + U^{24}\mathrm{d}s_1^* &= 0 \\ U^{41}\mathrm{d}\tau_1^{F*} + U^{42}\mathrm{d}\tau_1^{M*} + U^{34}\mathrm{d}\phi + U^{44}\mathrm{d}s_1^* &= 0 \end{aligned} \tag{20}$$

and

$$U^{44} = \frac{\mathrm{d}^2U}{\mathrm{d}s_1^{*2}}; \quad U^{14} = U^{41} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{F*}\mathrm{d}s_1^*}; \quad U^{24} = U^{42} = \frac{\mathrm{d}^2U}{\mathrm{d}\tau_1^{M*}\mathrm{d}s_1^*}; \quad U^{34} = U^{43} = \frac{\mathrm{d}^2U}{\mathrm{d}s_1^*\mathrm{d}\phi}.$$

From equation (20), I get the expressions

[34] $i = \{1,2,3,4\}$ and $j = \{1,2,3,4\}$.

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} = \frac{bq - dp}{bc - ad},$$

$$\frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} = \frac{aq - cp}{ad - bc};$$

where, when $\tau_2^{F*} - \tau_2^{M*} > 0$, then $U^{34} < 0$,

$$a = U^{11} - \frac{U^{14}U^{41}}{U^{44}} < 0; \qquad b = U^{12} - \frac{U^{24}U^{14}}{U^{44}} < 0;$$

$$c = U^{21} - \frac{U^{14}U^{42}}{U^{44}} < 0; \qquad d = U^{22} - \frac{U^{24}U^{42}}{U^{44}} < 0;$$

$$p = \frac{U^{14}U^{43}}{U^{44}} - U^{13} < 0; \qquad q = \frac{U^{24}U^{43}}{U^{44}} - U^{23}.$$

Given the SOCs and $b = c$, if $\tau_2^{F*'} \approx \tau_2^{M*'}$, then $|d|> |b|$ and $|a|> |c|$. So $bc - ad < 0$. However the sign of the q is undetermined. If $q > 0$, then

$$bq - dp < 0; \quad aq - cp < 0;$$

then

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0; \qquad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0.$$

Recall that $q = \frac{U^{24}U^{43} - U^{23}U^{44}}{U^{44}}$, from the SOCs, if $Y_1(1+r_2) - Y_2\tau_2^{M*'}(1-\phi)n$ is negative, q is positive for sure. If not, q is also highly likely to be positive, especially when c_2 is large enough.

When $\tau_2^{F*} - \tau_2^{M*} < 0$ and $U^{34} > 0$. Then the signs for a, b, c, and d do not change, and $q > 0$ for sure. In this scenario, I need $p < 0$, so that $bq - dp < 0$ and $aq - cp < 0$. If $Y_1(1+r_2) - Y_2\tau_2^{F*'}\phi n < 0$, then $U^{14}U^{34} - U^{13}U^{44} > 0$, so that $p < 0$. The I get the same conclusion as when $\tau_2^{F*} - \tau_2^{M*} > 0$, then

$$\frac{\mathrm{d}\tau_1^{F*}}{\mathrm{d}\phi} > 0; \qquad \frac{\mathrm{d}\tau_1^{M*}}{\mathrm{d}\phi} < 0.$$

The conclusion is consistent with the previous comparative statics in the baseline model.

Table A.14: Heterogeneity Check: household income level

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	-0.058	-0.242	-0.333	-7.470
	(0.045)	(0.329)	(0.240)	(14.54)
genderK	-0.007	0.210	-0.385*	-27.50**
	(0.051)	(0.301)	(0.220)	(13.39)
high income	0.053	1.174	0.354**	4.712
	(0.041)	(0.833)	(0.180)	(10.54)
maleP × *genderK*	0.011	-0.144	0.684	58.91**
	(0.081)	(0.587)	(0.421)	(25.80)
maleP × *high income*	-0.043	-0.427	0.170	22.46
	(0.061)	(0.458)	(0.307)	(18.35)
genderK × *high income*	-0.091	-0.461	-0.418	6.909
(High-low income differences)	(0.071)	(0.528)	(0.306)	(18.35)
genderK × *high income*	0.093	0.818	-0.453	-68.80**
×*maleP*	(0.107)	(0.806)	(0.528)	(31.91)
High income HH female	-0.098**	-0.426	-0.803***	-20.60*
demonstration effects	(0.047)	(0.360)	(0.200)	(11.85)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.227	0.225	0.154	0.109

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.15: Heterogeneity Check: single child family

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0553	0.0150	-0.0784	0.195	-0.285	0.239
	(0.0725)	(0.0409)	(0.0698)	(0.286)	(0.449)	(0.253)
genderK×	-0.0723	-0.00836	-0.0669	-0.132	-0.641	-0.218
singleY	(0.109)	(0.0559)	(0.108)	(0.374)	(0.683)	(0.364)
singleY	0.00230	0.0203	-0.0144	0.188	0.0827	0.189
	(0.0640)	(0.0334)	(0.0635)	(0.226)	(0.400)	(0.217)
genderK	0.0127	0.0228	-0.0209	0.253	0.107	0.171
	(0.0951)	(0.0511)	(0.0947)	(0.338)	(0.595)	(0.320)
maleP × *genderK*	0.0873	0.0688	0.0448	0.360	0.217	0.307
	(0.135)	(0.0767)	(0.130)	(0.529)	(0.837)	(0.460)
maleP×	-0.0210	-0.00686	-0.0224	-0.0927	-0.117	-0.307
singleY	(0.0889)	(0.0487)	(0.0856)	(0.343)	(0.557)	(0.311)
maleP × *genderK*	0.0507	-0.0616	0.108	-0.351	0.511	0.0447
×*singleY*	(0.158)	(0.0869)	(0.151)	(0.604)	(0.983)	(0.539)
Single Y HH male	0.078**	0.022	0.065**	0.130	0.193	0.305***
demonstrate effects	(0.031)	(0.016)	(0.028)	(0.111)	(0.188)	(0.104)
Non-single Y HH male	0.100	0.092*	0.024	0.613*	0.324	0.478*
demonstrate effects	(0.083)	(0.047)	(0.334)	(0.317)	(0.506)	(0.279)
Differences in males	-0.022	-0.070	0.041	-0.483	-0.131	-0.173
demonstrate effects	(0.101)	(0.057)	(0.093)	(0.407)	(0.624)	(0.346)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.202	0.072	0.140	0.069	0.139	0.532

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.16: Heterogeneity Check: single child family

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	0.009	-0.073	-0.139	-0.120
	(0.068)	(0.480)	(0.368)	(22.09)
genderK	0.170**	1.011*	0.116	-23.28
(Non-single Y HH	(0.084)	(0.578)	(0.355)	(20.82)
demonstration effects)				
singleY	0.176***	1.028***	0.656***	7.810
	(0.052)	(0.371)	(0.232)	(13.49)
$maleP \times genderK$	-0.105	-0.364	0.579	63.94
	(0.131)	(0.926)	(0.703)	(42.38)
$maleP \times singleY$	-0.135*	-0.603	-0.230	2.902
	(0.077)	(0.553)	(0.407)	(24.41)
$genderK \times singleY$	-0.306***	-1.664**	-0.975***	-0.625
(Single non-single differences)	(0.094)	(0.663)	(0.400)	(23.59)
$genderK \times singleY$	0.229	0.916	-0.114	-52.49
$\times maleP$	(0.143)	(1.030)	(0.753)	(45.48)
SingleY HH female	-0.136***	-0.653**	-0.860***	-23.90**
demonstration effects	(0.036)	(0.282)	(0.160)	(9.789)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.218	0.218	0.149	0.113

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.17: Heterogeneity Check: urban-rural differences

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0649*	0.0401**	-0.108***	0.322***	-0.389*	0.241**
	(0.0341)	(0.0181)	(0.0334)	(0.124)	(0.215)	(0.115)
$genderK \times urban$	-0.0280	-0.00425	-0.0619	0.0957	-0.858	0.0838
	(0.0936)	(0.0541)	(0.0898)	(0.418)	(0.605)	(0.341)
urban	0.0274	0.0475	0.000931	0.275	0.412	0.252
	(0.0570)	(0.0333)	(0.0559)	(0.250)	(0.374)	(0.199)
genderK	-0.0321	0.0203	-0.0543	0.152	-0.141	0.0748
	(0.0489)	(0.0252)	(0.0485)	(0.171)	(0.314)	(0.160)
$maleP \times genderK$	0.104*	0.00335	0.110**	-0.00348	0.452	0.233
	(0.0558)	(0.0284)	(0.0540)	(0.192)	(0.350)	(0.179)
$maleP \times urban$	-0.0272	-0.0442	-0.00822	-0.239	-0.114	-0.216
	(0.0723)	(0.0407)	(0.0682)	(0.301)	(0.457)	(0.247)
$maleP \times genderK$	0.0860	0.00440	0.120	-0.121	0.702	-0.353
$\times urban$	(0.121)	(0.0660)	(0.111)	(0.498)	(0.744)	(0.413)
Urban male	0.130**	0.024	0.114*	0.124	0.155	0.038
demonstrate effects	(0.067)	(0.033)	(0.060)	(0.236)	(0.398)	(0.215)
Rural male	0.072***	0.024**	056**	0.149*	0.311**	0.307***
demonstrate effects	(0.026)	(0.012)	(0.023)	(0.081)	(0.153)	(0.077)
Differences in males	0.058	0.000	0.058	-0.025	-0.156	-0.269
demonstrate effects	(0.077)	(0.038)	(0.069)	(0.267)	(0.466)	(0.245)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.202	0.075	0.139	0.072	0.138	0.533

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.18: Heterogeneity Check: urban-rural differences

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	-0.067	-0.523	-0.104	4.108
	(0.091)	(0.613)	(0.532)	(33.14)
genderK	0.007	-0.341	0.607	40.45
(Rural female demonstration effects)	(0.102)	(0.671)	(0.481)	(30.23)
urban	0.092	0.331	0.889***	51.19***
	(0.064)	(0.433)	(0.303)	(18.73)
maleP × *genderK*	0.038	0.449	0.595	70.92
	(0.155)	(1.044)	(0.902)	(56.43)
maleP × *urban*	-0.019	0.045	-0.211	-4.688
	(0.096)	(0.656)	(0.551)	(34.18)
genderK × *urban*	-0.072	0.167	-1.466***	-79.89***
(Urban-rural differences)	(0.111)	(0.744)	(0.515)	(32.27)
genderK × *urban*	0.025	-0.192	-0.177	-55.40
×*maleP*	(0.165)	(1.128)	(0.936)	(58.34)
Urban female	-0.067*	-0.174	-0.859***	-39.44***
demonstration effects	(0.036)	(0.273)	(0.153)	(9.164)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.227	0.225	0.144	0.091

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.19: Heterogeneity Check: family compositions of P

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.128***	-0.0340*	-0.105***	-0.169	-0.365	0.333**
	(0.0385)	(0.0190)	(0.0372)	(0.136)	(0.247)	(0.132)
genderK× *older bro*	0.100	0.0992**	-0.00509	0.698**	-0.260	-0.316
	(0.0939)	(0.0496)	(0.0915)	(0.350)	(0.590)	(0.300)
older bro	-0.0443	-0.0534*	0.00992	-0.374*	0.150	0.297
	(0.0570)	(0.0297)	(0.0564)	(0.208)	(0.359)	(0.180)
genderK	-0.0764	-0.0172	-0.0663	-0.0846	-0.237	0.162
	(0.0535)	(0.0270)	(0.0526)	(0.194)	(0.351)	(0.179)
maleP × *genderK*	0.222***	0.102***	0.138**	0.630***	0.538	0.0131
	(0.0629)	(0.0312)	(0.0601)	(0.224)	(0.402)	(0.209)
maleP× *older bro*	0.194***	0.187***	0.0134	1.267***	-0.0182	-0.550**
	(0.0701)	(0.0371)	(0.0667)	(0.259)	(0.438)	(0.236)
maleP × *genderK* ×*older bro*	-0.324***	-0.288***	-0.0449	-1.937***	0.00692	0.707*
	(0.118)	(0.0612)	(0.111)	(0.429)	(0.731)	(0.379)
With older brothers male demonstrate effects	-0.078	-0.104***	0.022	-0.693**	0.048	0.566***
	(0.051)	(0.026)	(0.045)	(0.179)	(0.300)	(0.161)
Without older brothers male demonstrate effects	0.145***	0.084***	0.072***	0.546***	0.301*	0.178*
	(0.029)	(0.014)	(0.025)	(0.102)	(0.170)	(0.096)
Differences in males demonstrate effects	-0.224***	-0.189***	-0.050	-1.239***	-0.253	0.391*
	(0.064)	(0.034)	(0.057)	(0.232)	(0.388)	(0.213)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.199	0.061	0.140	0.059	0.140	0.530

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.20: Gender effect of K on the old-age support provision: urban-singleton households

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log regular*	*log nonregular*	*log (visit days)*
Urban-singleton						
$maleP$	0.0284	0.0354	-0.0161	0.283	0.0237	-0.0454
	(0.0580)	(0.0297)	(0.0550)	(0.243)	(0.387)	(0.197)
$genderK$	0.0138	0.0678	-0.0753	0.597	-0.512	0.129
	(0.0706)	(0.0439)	(0.0677)	(0.368)	(0.476)	(0.266)
$maleP \times genderK$	-0.0277	-0.0651	0.0648	-0.544	0.296	0.00887
	(0.0952)	(0.0513)	(0.0888)	(0.430)	(0.626)	(0.319)
$genderK+$	-0.015	0.003	-0.010	0.053	-0.216	0.241*
$maleP \times genderK$	(0.062)	(0.026)	(0.035)	(0.385)	(0.244)	(0.127)
Observations	2,037	2,037	2,037	2,037	2,037	2,037
R-squared 0.252	0.103	0.181	0.096	0.182	0.574	
Others						
$maleP$	-0.0978***	0.0344**	-0.140***	0.301**	-0.552***	0.232**
	(0.0325)	(0.0172)	(0.0317)	(0.118)	(0.204)	(0.110)
$genderK$	-0.0619	0.0109	-0.0800*	0.0942	-0.344	0.0596
	(0.0474)	(0.0245)	(0.0470)	(0.166)	(0.304)	(0.156)
$maleP \times genderK$	0.170***	0.0134	0.168***	0.0487	0.731**	0.200
	(0.0534)	(0.0273)	(0.0518)	(0.184)	(0.335)	(0.172)
$genderK+$	0.108***	0.024**	0.088***	0.143*	0.387***	0.260***
$maleP \times genderK$	(0.024)	(0.011)	(0.021)	(0.078)	(0.136)	(0.071)
Observations	10,195	10,195	10,194	10,195	10,193	10,196
R-squared	0.199	0.085	0.139	0.081	0.141	0.532
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.21: Gender effect of K on the old-age support provision: urban-singleton households

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
Urban-singleton				
$maleP$	-0.138***	-0.790***	-0.467**	-7.838
	(0.033)	(0.260)	(0.154)	(8.977)
$genderK$	-0.109**	-0.560**	-0.679***	-19.46**
	(0.035)	(0.274)	(0.150)	(9.134)
$maleP \times genderK$	0.137**	0.667	0.439*	12.78
	(0.057)	(0.447)	(0.251)	(15.11)
$genderK+$	0.028	0.107	-0.239	-6.680
$maleP \times genderK$	(0.039)	(0.315)	(0.179)	(10.63)
Observations	8,590	8,590	8,590	8,590
R-squared	0.221	0.214	0.152	0.129
Others				
$maleP$	0.026	0.091	-0.087	5.739
	(0.065)	(0.463)	(0.355)	(21.49)
$genderK$	0.084	0.592	-0.248	-27.14
	(0.073)	(0.506)	(0.319)	(19.20)
$maleP \times genderK$	-0.117	-0.548	0.515	55.54
	(0.117)	(0.835)	(0.635)	(38.70)
$genderK+$	-0.003	0.044	0.268	28.41
$maleP \times genderK$	(0.078)	(0.565)	(0.484)	(29.53)
Observations	10,934	10,934	10,934	10,934
R-squared	0.198	0.186	0.158	0.116
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.22: Heterogeneity Check: living in a community with minority ethnic groups

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0329	0.00959	-0.0537	0.154	0.0596	-0.121
	(0.0469)	(0.0274)	(0.0460)	(0.191)	(0.301)	(0.160)
genderK×	-0.0885	0.0867*	-0.175**	0.581*	-1.321**	0.148
minority community	(0.0904)	(0.0481)	(0.0887)	(0.338)	(0.578)	(0.298)
minority community	-0.0203	-0.0453	0.0268	-0.306	0.277	-0.123
	(0.0558)	(0.0302)	(0.0556)	(0.209)	(0.354)	(0.181)
genderK	0.00245	-0.0200	0.0125	-0.0896	0.283	-0.0349
	(0.0604)	(0.0307)	(0.0612)	(0.214)	(0.403)	(0.200)
maleP × *genderK*	0.0229	0.0586	-0.0115	0.296	-0.516	0.585**
	(0.0773)	(0.0419)	(0.0750)	(0.288)	(0.494)	(0.253)
maleP×	-0.0770	0.0371	-0.108	0.202	-0.942**	0.528**
minority community	(0.0689)	(0.0405)	(0.0661)	(0.281)	(0.431)	(0.241)
maleP × *genderK*	0.204*	-0.106*	0.294***	-0.634	2.270***	-0.712*
×*minority community*	(0.114)	(0.0634)	(0.107)	(0.439)	(0.704)	(0.387)
minority com. father	0.141***	0.019	0.120***	0.153**	0.716***	-0.013
demonstrate effects	(0.036)	(0.021)	(0.032)	(0.143)	(0.209)	(0.131)
non-minority com. father	0.025	0.038	0.001	0.206	-0.233	0.550***
demonstrate effects	(0.043)	(0.024)	(0.039)	(0.169)	(0.261)	(0.138)
Differences in father	0.115*	-0.019**	0.119*	-0.053	0.950**	-0.563**
demonstrate effects	(0.067)	(0.041)	(0.061)	(0.278)	(0.398)	(0.230)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.202	0.067	0.138	0.062	0.137	0.532

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.23: Heterogeneity Check: ethnic group

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	-0.0373	-0.438	0.552	40.03
	(0.105)	(0.750)	(0.575)	(35.58)
genderK	-0.181	-1.181	-0.0492	-4.495
(Non-Han mother demonstration effects)	(0.120)	(0.905)	(0.585)	(36.65)
Han	-0.0466	-0.257	0.468	19.94
	(0.0690)	(0.512)	(0.333)	(20.35)
maleP $\times$ *genderK*	-0.0132	0.222	-0.556	-21.05
	(0.190)	(1.375)	(1.031)	(64.65)
maleP $\times$ *Han*	-0.0143	-0.0113	-0.676	-27.98
	(0.111)	(0.797)	(0.600)	(37.12)
genderK $\times$ *Han*	0.126	0.846	-0.757	-32.42
(Han or non-Han mother effect differences)	(0.129)	(0.967)	(0.618)	(38.57)
genderK $\times$ *Han*	0.0309	-0.000	1.166	51.04
$\times$*maleP*	(0.200)	(1.456)	(1.072)	(67.22)
Han mother	-0.055	-0.336	-0.806***	-36.91
demonstration effects	(0.039)	(0.286)	(0.167)	(10.06)
Han or non-Han	0.157	0.845	0.409	18.63
father effect differences	(0.153)	(1.088)	(0.873)	(54.85)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.301	0.227	0.325	0.359

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. The detailed note is in the Appendix I.

Table A.24: The third generation effect and the education investment

VARIABLES	IV: CHFS any education investment in K	the amount of the education investment	percentage of edu. investment in total expense
$maleP$	-0.0785**	-441.5	-0.0663**
	(0.0338)	(682.8)	(0.0336)
$genderK$	0.00481	-3,042***	0.00976
	(0.0347)	(716.9)	(0.0346)
$maleP \times genderK$	0.129**	468.3	0.108*
	(0.0587)	(1,077)	(0.0584)
$genderK +$	0.133***	-2573***	0.117***
$maleP \times genderK$	(0.042)	(696.7)	(0.041)
(Male with sons-males with daughters)			
$maleP +$	0.050*	26.79	0.041
$maleP \times genderK$	(0.027)	(473.5)	(0.026)
(Male with sons-females with sons)			
P demographics	Yes	Yes	Yes
O demographics	Yes	Yes	Yes
Observations	19,524	19,524	19,524
R-squared	0.158	0.263	0.153

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I, except the outcome variables are the education investment from the second generation on the third generation.

Table A.25: Different ways to present the amount of transfers

VARIABLES	IV: CHARLS Amount of transfer *regular*	 *non-regular*	 Percentage of income *regular*	 *non-regular*	IV: CHFS Amount of transfer	 Percentage of income
$maleP$	514.4	110.6	-0.0272	-0.0116	205.4	-0.008*
	(370.1)	(192.5)	(0.0229)	(0.0154)	(290.3)	(0.005)
$genderK$	269.3	-398.5**	0.000482	-0.000670	570.3	-0.003
	(352.3)	(186.6)	(0.00755)	(0.00694)	(377.1)	(0.006)
$maleP \times genderK$	-78.66	422.6	0.0605	0.0411	-376.3	0.010
	(337.8)	(263.5)	(0.0410)	(0.0277)	(530.9)	(0.009)
$genderK+$	190.8	24.07	$0.061^{\dagger}$	$0.041^{\dagger}$	194.0	0.007
$maleP \times genderK$	(280.0)	(106.0)	(0.039)	(0.025)	(306.9)	(0.006)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
P income level	Yes	Yes	No	No	Yes	No
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	19,524	19,524
R-squared	0.094	0.013	0.616	0.051	0.094	0.087

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I, except the outcome variables are different forms of transfer provided by P.

Table A.26: The demonstration effect without controlling for the transfers from the first generation

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
$maleP$	-0.0727***	0.0251*	-0.105***	0.226**	-0.416**	0.141
	(0.0280)	(0.0148)	(0.0272)	(0.106)	(0.177)	(0.0964)
$genderK$	-0.0477	0.0178	-0.0756*	0.161	-0.398	0.0701
	(0.0406)	(0.0216)	(0.0402)	(0.154)	(0.264)	(0.137)
$maleP \times genderK$	0.143***	0.0146	0.144***	0.0571	0.737**	0.208
	(0.0466)	(0.0240)	(0.0446)	(0.169)	(0.292)	(0.152)
$genderK+$	0.096***	0.032**	0.069***	0.218***	0.339***	0.278***
$maleP \times genderK$	(0.022)	(0.010)	(0.018)	(0.066)	(0.121)	(0.067)
Transfer from O	No	No	No	No	No	No
O taking care for K	No	No	No	No	No	No
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.199	0.072	0.137	0.068	0.137	0.531

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.27: The demonstration effect without controlling for the transfers from the first generation

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
$maleP$	-0.0760**	-0.654**	-0.118	16.89
	(0.0374)	(0.275)	(0.169)	(10.40)
$genderK$	-0.0759*	-0.538*	-0.742***	-28.16***
	(0.0421)	(0.310)	(0.177)	(10.79)
$maleP \times genderK$	0.0759	0.666	0.588**	20.76
	(0.0663)	(0.490)	(0.294)	(18.32)
$genderK+$	0.0000	0.128	-0.153	-7.404
$maleP \times genderK$	(0.0436)	(0.323)	(0.203)	(12.90)
Transfer from O	No	No	No	No
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.243	0.181	0.323	0.362

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.28: Coefficients for the direct transfer from the first generation

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*log(regular)*	*log(nonregular)*	*log (visit days)*
maleP	-0.0799***	0.0285*	-0.117***	0.252**	-0.480***	0.125
	(0.0280)	(0.0148)	(0.0271)	(0.106)	(0.178)	(0.0968)
genderK	-0.0403	0.0202	-0.0701*	0.178	-0.349	0.0490
	(0.0397)	(0.0215)	(0.0392)	(0.154)	(0.258)	(0.136)
$maleP \times genderK$	0.127***	0.00578	0.138***	-0.0136	0.651**	0.255*
	(0.0464)	(0.0240)	(0.0445)	(0.170)	(0.293)	(0.153)
transfer from O to P	-0.0478*	-0.0348***	-0.0132	-0.264***	-0.0629	0.134
	(0.0279)	(0.0102)	(0.0275)	(0.0687)	(0.186)	(0.0814)
O taking care for K	7.44e-06***	2.30e-06*	4.35e-06*	2.12e-05**	4.56e-05***	-1.03e-05
	(2.25e-06)	(1.38e-06)	(2.26e-06)	(1.06e-05)	(1.66e-05)	(8.32e-06)
transfer from O to K	0.173***	-0.00360	0.178***	0.0240	1.152***	0.116**
	(0.0163)	(0.00947)	(0.0170)	(0.0723)	(0.112)	(0.0517)
$genderK+$	0.087***	0.026***	0.068***	0.163**	0.302**	0.304***
$maleP \times genderK$	(0.023)	(0.010)	(0.020)	(0.070)	(0.135)	(0.068)
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,232	12,232	12,232	12,232	12,232	12,232
R-squared	0.215	0.074	0.153	0.070	0.153	0.532

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *genderK* represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I. One new controls is the transfer from O to K and K is the children of P.

Table A.29: Coefficients for the direct transfer from the first generation

	IV: CHFS			
VARIABLES	*any-transfer*	*log(transfer)*	*log (visit days)*	*visit days*
maleP	-0.0527	-0.461*	-0.0651	14.64
	(0.0340)	(0.254)	(0.161)	(9.957)
genderK	-0.0674*	-0.417	-0.741***	-34.13***
	(0.0364)	(0.269)	(0.158)	(9.570)
$maleP \times genderK$	0.0196	0.246	0.501*	24.90
	(0.0602)	(0.451)	(0.280)	(17.52)
transfer from O to P	0.313***	1.950***	0.523***	11.73***
	(0.00991)	(0.0732)	(0.0417)	(2.657)
$genderK+$	-0.048	-0.171	-0.240	-9.230
$maleP \times genderK$	(0.042)	(0.315)	(0.203)	(12.88)
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes
Observations	19,524	19,524	19,524	19,524
R-squared	0.303	0.228	0.328	0.361

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. *genderK* represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.30: Subsample check for high and low gender-ratio provinces

	IV: CHARLS					
VARIABLES	*any-transfer*	*regular*	*nonregular*	*regular amount*	*nonregular amount*	*log (visit days)*
low gender-ratio provinces						
$maleP$	-0.0616*	0.0284**	-0.0850***	-5.030	130.4	-0.317***
	(0.0327)	(0.0116)	(0.0318)	(209.4)	(147.4)	(0.0767)
$genderK$	-0.00933	0.0304**	-0.0267	-121.2	-124.9	0.0460
	(0.0387)	(0.0151)	(0.0381)	(182.0)	(118.2)	(0.0933)
$maleP \times genderK$	0.0428	-0.0338**	0.0602	-38.89	242.8	0.0721
	(0.0459)	(0.0172)	(0.0443)	(293.4)	(214.2)	(0.106)
$genderK+$	0.033	-0.003	0.034	-160.1	117.9	0.118**
$maleP \times genderK$	(0.023)	(0.009)	(0.021)	(202.4)	(155.8)	(0.047)
Observations	3,373	3,373	3,372	3,373	3,373	3,373
R-squared	0.200	0.064	0.162	0.109	0.041	0.854
high gender-ratio provinces						
$maleP$	0.00360	0.0124	-0.0249	774.9***	233.3	-0.566***
	(0.0397)	(0.0210)	(0.0389)	(212.5)	(553.1)	(0.0902)
$genderK$	-0.0414	-0.0238	-0.0474	292.4**	-305.1	-0.0917
	(0.0455)	(0.0253)	(0.0449)	(139.1)	(587.0)	(0.108)
$maleP \times genderK$	0.0238	0.00114	0.0593	-180.5	626.1	0.176
	(0.0529)	(0.0279)	(0.0498)	(184.3)	(817.9)	(0.122)
$genderK+$	-0.018	-0.023*	0.012	321.0	111.9	0.085
$maleP \times genderK$	(0.030)	(0.013)	(0.025)	(398.9)	(128.3)	(0.056)
Observations	2,489	2,489	2,489	2,489	2,489	2,490
R-squared	0.264	0.091	0.192	0.053	0.045	0.878
P demographics	Yes	Yes	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.

Table A.31: Subsample check for high and low gender-ratio provinces

			IV: CHFS	
VARIABLES	*any-transfer*	*amount*	*log (visit days)*	*visit days*
low gender-ratio provinces				
$maleP$	-0.0408	-147.8	32.18***	0.239
	(0.0334)	(151.4)	(10.20)	(0.157)
$genderK$	-0.0315	-221.0	-6.319	-0.193
	(0.0334)	(148.3)	(8.215)	(0.130)
$maleP \times genderK$	0.0378	259.1	-5.102	0.0556
	(0.0497)	(220.0)	(13.33)	(0.207)
$genderK+$	0.006	38.13	-0.134	-11.32
$maleP \times genderK$	(0.037)	(162.4)	(0.163)	(10.63)
Observations	2,672	2,672	2,672	2,672
R-squared	0.147	0.222	0.364	0.285
high gender-ratio provinces				
$maleP$	-0.0612	27.58	55.41***	0.561**
	(0.0496)	(219.5)	(16.39)	(0.246)
$genderK$	0.0440	-62.66	-8.917	-0.104
	(0.0529)	(214.3)	(13.18)	(0.205)
$maleP \times genderK$	0.000331	105.1	2.156	0.0382
	(0.0762)	(315.4)	(21.17)	(0.321)
$genderK+$	0.044	42.45	-0.051	-6.089
$maleP \times genderK$	(0.055)	(226.8)	(0.254)	(17.20)
Observations	1,454	1,454	1,454	1,454
R-squared	0.193	0.289	0.375	0.289
P demographics	Yes	Yes	Yes	Yes
O demographics	Yes	Yes	Yes	Yes

Notes: Robust standard errors in parentheses, clustered at household-level. Stars indicate statistical significance. *** p<0.01, ** p<0.05, * p<0.1. $genderK$ represents the female dominated demonstration effect and $genderK + maleP \times genderK$ shows the male dominated demonstration effect. The detailed note is in the Appendix I.