



Contents lists available at ScienceDirect

Journal of Comparative Economics

journal homepage: www.elsevier.com/locate/jce

Fertility, household structure, and parental labor supply: Evidence from China[☆]



Rufei Guo^a, Hongbin Li^b, Junjian Yi^{*,c}, Junsen Zhang^d

^a Economics and Management School, Wuhan University, China

^b Stanford Center on Global Poverty and Development, Stanford Institute for Economic Policy Research, Stanford University, USA

^c Department of Economics, National University of Singapore, Singapore

^d Department of Economics, Chinese University of Hong Kong, Hong Kong

ARTICLE INFO

Keywords:

Fertility
Parental labor supply
Household structure

JEL classification:

J13
J18
J22
O10

ABSTRACT

This study examines the effects of fertility on household structure and parental labor supply in China. To solve the endogeneity problem, we use a unique survey on households with twin children and a comparison group of non-twin households. The ordinary least squares estimates show a negative correlation between fertility and parental labor supply in rural China. Using twinning as a natural experiment, we do not find evidence on the negative effects of fertility on parental labor supply. By contrast, we find that the twinning-induced increase in fertility significantly enhances the coresidence of grandparents in rural China. We suggest that the negative effects of fertility on parental labor supply are mitigated by the childcare provided by grandparents in rural China. We also find that fertility does not induce coresidence of grandparents in urban China. Our results have important implications for population and public childcare policies.

1. Introduction

Standard economic models predict a negative effect of fertility on parental labor supply, especially mothers' (Borjas, 2000), because parents need to allocate time to childcare. Such models assume a nuclear family, in which the only adults who can provide labor are parents. These models fit well for data from developed countries (Killingsworth and Heckman, 1986; Angrist and Evans, 1998),¹ but less so for data from developing countries (Cruces and Galiani, 2007; Agüero and Marks, 2008, 2011; Cáceres-Delpiano, 2012; He and Zhu, 2016).²

[☆] We would like to thank James Heckman and Robert Pollak for their valuable comments and suggestions. Junjian Yi acknowledges financial support from the National University of Singapore HSS Grant No. R-122-000-232-646 and FRC Grant No. R122-000-247-115.

^{*} Corresponding author.

E-mail addresses: rufei_guo@whu.edu.cn (R. Guo), lhongbin@stanford.edu (H. Li), junjian@nus.edu.sg (J. Yi), jszhang@cuhk.edu.hk (J. Zhang).

¹ The empirical literature testing the causal effects of fertility on parental labor supply from developed countries is in abundance (Cain, 1966; Fleisher and Rhodes, 1979; Cramer, 1980; Rosenzweig and Wolpin, 1980; Willis, 1987; Angrist and Evans, 1998; Ebenstein, 2009). See also two survey papers by Nakamura and Nakamura (1992) and Browning (1992).

² The empirical literature from developing countries is relatively small (Cruces and Galiani, 2007; Agüero and Marks, 2008, 2011; Cáceres-Delpiano, 2012). Using twinning as a natural experiment, Cáceres-Delpiano (2012) finds that although fertility decreases labor supply for mothers with more schooling years and living in urban areas, it does not have a significant effect on maternal labor supply for residents from rural areas and Asia. We explain the insignificant effect on maternal labor supply for residents in Asia in Cáceres-Delpiano (2012). Following Angrist and Evans (1998), Cruces and Galiani (2007) uses the sex composition of the first two children as an instrument for fertility. They find that fertility decreases the labor supply of women in Argentina and Mexico. Recently, He and Zhu (2016) find that a very small negative effect of fertility on female labor force participation in urban China.

<http://dx.doi.org/10.1016/j.jce.2017.10.005>

Received 7 September 2016; Received in revised form 4 August 2017; Accepted 29 October 2017

Available online 01 November 2017

0147-5967/© 2017 Association for Comparative Economic Studies. Published by Elsevier Inc. All rights reserved.

Such a negative association between fertility and labor supply may not exist in developing countries, where extended families with three generations living in one household are common (Rosenzweig and Wolpin, 1985; Chu et al., 2011).³ For extended families, the effect of fertility on parental labor supply depends on the substitutability between grandparents' time and parents' time in caring for children. Because the time cost of grandparents is lower than that of parents, in response to the increase in fertility, grandparents may move to live with their children and grandchildren, increasing their time allocated to childcare and household chores. In this scenario, household structure endogenously responds to the change in fertility, and parental labor supply does not necessarily decrease with fertility. Despite the possibility of this phenomenon, no prior study has addressed the issue of mediating role of extended families in examining the effects of fertility on parental labor supply.

This study aims to test the effects of fertility on parental labor supply with particular attention to the mediating role of childcare provided by grandparents in extended families. Drawing on data from rural China, we not only test whether parental labor supply changes with the increase in fertility, but also examine whether household structure responds to the increase in fertility. To our best knowledge, this study is the first to rigorously identify the effect of fertility on household structure.⁴

To address the issue of endogenous fertility decisions, following Rosenzweig and Wolpin (1980), we use twinning as a natural experiment for fertility.⁵ Existing studies on the Chinese economy use the population census to exploit the incidence of twin births as the instrumental variable of fertility. For example, Li et al. (2008) use twin births as the instrumental variable of fertility, and they find that fertility reduces the education attainment of children in rural China.

Of particular relevance to our study is He and Zhu (2016), who exploit twin births to study the effect of fertility on female labor supply in urban China. He and Zhu (2016) find that fertility reduces female labor supply in urban China using the 1990 census, but fertility does not have statistically significant effect on female labor supply in the 2000 census. The magnitude of the negative effect detected in the 1990 census is also very small.

Although census data are appropriate for studying effects of fertility on child quality or parental labor supply, such data are unsuitable for analyzing the effect of fertility on household structure. Households are identified by coresidence of household members in the census data. Coresiding household members only report their relation to the household head. To identify twins and select households in which all children co-reside, the literature (e.g., Li et al., 2008; He and Zhu, 2016) focus on households in which one of the parents is the household head. In reality, grandchildren may also move into households headed by grandparents. Identifying households by coresidence of household members introduces serious sampling bias for studying the effect of fertility on household structure.⁶

We employ the self-collected Chinese child twin survey (CCTS) to study the effects of fertility on parental labor supply and household structure. To the best of our knowledge, CCTS is the first census-type survey on households with twin children. The CCTS is suitable for studying household structure because it collects information on both coresiding and non-coresiding household members. We find that fertility does not have a negative effect on parental labor supply in rural areas. By contrast, we find that the twinning-induced rise in fertility increases the probability of coresidence of grandparents in rural areas. Specifically, an increase of one additional child induced by twinning raises the probability of coresidence of grandparents by seven percentage points, which is an increase of 37%. The empirical evidence in rural areas is consistent with the hypothesis that the childcare provided by grandparents mitigates the negative effect of fertility on parental labor supply in rural China.

Informal institutions, like grandparent-provided childcare, may play distinct roles in different stages of economic development. Urban China is in more advanced stages of economic development compared with rural China. Does informal childcare provided by grandparents explain the lack of a sizable negative effect of fertility on female labor force participation in urban China as documented by He and Zhu (2016)? We use the urban sample of CCTS to provide a comparative study. We find that although fertility does not reduce parental labor supply in urban China, fertility does not induce coresidence of grandparents. The results suggest that in urban China where formal childcare is more accessible, informal childcare provided by grandparents no longer play the mediating role in muting the response of female labor supply to fertility. The comparative analysis between rural and urban China indicates the substitution between informal and formal childcare through different stages of economic development.

Our study not only contributes to the traditional literature on fertility and parental labor supply but also to a strand of more recent literature on the role of grandparental coresidence in family behavior. Zeng and Xie (2014) finds that the educational level of coresiding grandparents directly affects the educational attainment of their grandchildren in China. The effect size is similar to that of

³ Chu and Yu (2009) show that more than 20% of the Chinese rural households live with three generations. Within the first few years after the marriage of young adults, 72% of them live with their parents. In addition, the parents of husbands and the young couples usually live in the same village even if they do not live together. Rosenzweig and Wolpin (1985) find that 62% of all farm households in rural India contain at least two generations of adult kin.

⁴ In a chapter about fertility and family behavior in the Handbook of Development Economics, Schultz (2010) lists the effect of fertility on household living arrangements as an important but under-studied research topic. Many studies have examined the effect of household structure on the labor supply of women (Tienda and Glass, 1985; Wong and Levine, 1992; Gong and Van Soest, 2002; Maurer-Fazio et al., 2011; Posadas and Vidal-Fernández, 2013). They find that the coresidence of grandparents increases the labor supply of women who have children. However, no one has explicitly addressed the mediating role of the childcare provided by grandparents in the relationship between fertility and parental labor supply in extended families.

⁵ To tackle this endogeneity problem, the literature has mainly used two methods based on natural experiments. Rosenzweig and Wolpin (1980) use the natural occurrence of twinning at the first birth parity to identify the effects of fertility on labor supply. They find that, although women who have twins withdraw temporarily from the labor market, their life-cycle labor supply is not affected. See also Bronars and Grogger (1994) for making use of twins for identification. Angrist and Evans (1998) use parental preferences for a mixed sex composition of children to construct instrumental variables estimating the effects of fertility on labor supply. They find fertility has a negative effect on the labor supply of women. Recently, Agüero and Marks (2011) advance a new instrument for family size, infertility. Based on data from 27 mid- and low-income countries, their estimation results show that the presence of children affects neither the likelihood of work nor its intensity for women.

⁶ We thank an anonymous referee for pointing out the sampling problem using census data.

parental education. The result in [Zeng and Xie \(2014\)](#) corroborates our analysis by showing empirical evidence on the importance of grandparental child care in China. [Chu et al. \(2014\)](#) studies the time to first birth by treating coresidence with husbands parents and labor supply as endogenous. After controlling for endogeneity, they find that both coresidence and work are associated with a delay in childbearing. We treat fertility as endogenous. Using twinning as a natural experiment, we find that the increase in fertility enhances significantly the coresidence of grandparents but does not significantly reduce parental labor supply.

Our findings have important policy implications. One major cost of fertility is parental labor supply (or generally, parents' time). Our findings suggest that with the help of grandparents, fertility does not negatively affect parental labor supply, suggesting that as an informal childcare provision mechanism, extended family plays an important mitigating role for the time cost of children. However, with the decline of extended family in China, as well as in other developing countries, formal childcare may substitute for informal childcare. This fact calls for a few policy considerations. First, government-provided formal childcare may supplement the declining informal system. The recent literature from Canada suggests that the highly subsidized and universally accessible childcare significantly increase maternal labor supply ([Baker et al., 2008](#); [Lefebvre and Merrigan, 2008](#)). Such services are often unavailable in developing countries. Second, relaxation of China's one-child policy may not have sizable negative effects on parental labor supply.⁷ The one-child policy was relaxed in 2016 to allow couples to have two children. Compliers of the policy change may either seek help from members of the extended family or leverage on formal childcare.

The rest of the paper is structured as follows. In [Section 2](#), we introduce the CCTS data. In [Sections 3 and 4](#), we present our main estimation results by testing the effects of fertility on parental labor supply and household structure in rural China. In [Section 5](#), we conduct a comparative analysis using data from urban China. [Section 6](#) concludes the study.

2. Data

This section first introduces the CCTS data. We then discuss the distribution of fertility in the data. Finally, we describe main variables for our empirical analysis.

2.1. Chinese child twins survey

In this paper, we use the CCTS data. The Urban Survey Unit (USU) of the National Bureau of Statistics conducted the survey in late 2002 and early 2003 in Kunming, China. Kunming, which is the capital city of Yunnan Province, has 14 county-level administrative units and a total population of approximately 5.3 million, with 65% of them living in rural areas. Yunnan is located in the far southwestern corner of China and is a relatively under-developed province. The average per capital GDP was RMB 5800 in Yunnan, in contrast to an average RMB 10,000 in China in 2002.

To the best of our knowledge, CCTS is the first census-type survey on households with twin children. Its sampling and survey design exhibit three important features. First, the survey includes nearly all households with twin children 6–18 years old living in Kunming in 2002. The USU initially identified the households with twin children on the basis of the 2000 population census, according to whether the children have the same birth year and month and whether they have the same relationship with the household head. The addresses of these households are then obtained from the census office, and a visit to the household verifies the presence of twins. Starting from 2300 pairs of potential twins identified in the census, 1694 households with twin children are successfully interviewed. See [Rosenzweig and Zhang \(2009\)](#) and [Yi et al. \(2015\)](#) for a detailed description of the CCTS.

Second, for comparison, the survey contains a probability sample of households with non-twin children in the same age range. Specifically, for every twin household identified, the fourth household on the right-hand side of the same block was chosen to locate a non-twin household. If the fourth household does not have at least one child in the age range of 6–18, interviewers would continue going to the fifth, sixth, and so on. As many as 1693 non-twin households provide information in the survey.

Finally, the survey covers an extensive range of information about family structure and parental labor supply, in addition to a wide range of demographic, social, and economic information at both the individual and household levels. The questionnaire contains household, parents, and children modules, which are answered by household heads, parents, and children, respectively. More importantly, it contains information on children's birth weight. We can thus control for differences in endowments between twins and singletons in estimating the effect of fertility on household structure and parental labor supply.

We restrict our sample to households living in rural areas and with the eldest child less than 16.⁸ Thus, all children are born after the one-child policy in 1979. We have 969 non-twin households and 697 twin households in the rural sample. We also have 624 non-twin households and 542 twin households in the urban sample.

2.2. Distribution of fertility by twinning status

The specifics and enforcements of the one-child policy vary from one place to another in China ([Li et al., 2011](#)). In Kunming, rural households are exempted from the strict one-child policy. The population policy regulated that rural households can have at most two

⁷ [Ebenstein \(2010\)](#); [Ebenstein and Estimating a dynamic model of sex selection in \(2011\)](#) studies the socioeconomic consequences of the one-child policy on sex ratio imbalance in China.

⁸ Age 16 is the legally minimum age for working in China. We minimize the possibility that a child has contributed to household production or to family farm and enterprise by restricting children's age to be less than 16. Our results are robust if we limit the sample to families with the eldest child under the age of 12. We replicate the main results of this paper in [Tables A1 and A2](#) in the appendix.

Table 1
Distribution of fertility.

| | Rural areas | | Urban areas | |
|--------------------|-----------------|-------------------------|-----------------|-------------------------|
| | No twins (1) | First-born twins (2) | No twins (3) | First-born twins (4) |
| One child | 612 (63.16) | | 605 (96.96) | |
| Two children | 355 (36.64) | 650 (93.26) | 19 (3.04) | 531 (97.97) |
| Three children | 2 (0.21) | 47 (6.74) | | 11 (2.03) |
| Number of families | 969 (100) | 697 (100) | 624 (100) | 542 (100) |

Notes: The data source is the CCTS. Percentages are in parentheses for each category (column).

children, although they are encouraged to have only one (Province, 2003). So the population policy in rural Kunming is actually a “two-child” policy. Urban households are under strict enforcement of the one-child policy.

As twins at the second birth parity can hardly be observed in urban China, we only consider twins born at the first birth parity to facilitate a comparative study of rural and urban China.⁹ Table 1 shows the distribution of fertility based on the CCTS data. In this table, we categorize the sample into four groups: non-twin households in rural areas (Column (1)), households with first-born twins in rural areas (Column (2)), non-twin households in urban areas (Column (3)), and households with first-born twins in urban areas (Column (4)). Column (1) shows that in rural areas approximately 37% of non-twin households have more than one child, but less than 1% of the group have more than two children. For households with first-born twins in rural areas, more than 93% have two children only (Column (2)). In urban areas, nearly 97% of non-twin households have one child only (Column (3)), and almost 98% of households with first-born twins have two children only (Column (4)).

In sum, the “two-child” policy and one-child policy have been strictly enforced in rural and urban Kunming, respectively. Therefore, twinning accounts for a substantial part of the variation in fertility in the CCTS data.¹⁰

2.3. Variable description

Table 2 summarizes the descriptive statistics of the main variables. The CCTS covers detailed information on parental labor supply. The survey includes not only parental work status, but also work time. The CCTS also provides information on parental work type, migration status, and earnings. Columns (1)–(6) summarizes the characteristics of non-twin and twin households in rural areas, and columns (7)–(12) summarizes the characteristics of non-twin and twin households in urban areas. Given that extended families are more prevalent in rural China than in urban China, our discussions focus on rural households as described by columns (1)–(6).

Columns (1)–(6) of Table 2 summarizes the characteristics of non-twin and twin households in rural areas. The variable “Worked” is a dummy variable indicating whether a parent worked on his or her own farm or took a non-agricultural job in 2002. Columns (1) and (3) show that around 79% of mothers and 85% of fathers in rural households worked. The probability of working was higher for fathers than mothers. Regarding working time, we employ two variables. One is measured by the number of days per month, and the other is by the number of hours per week. We find that both mothers and fathers worked roughly 20 days per month, but mothers worked fewer hours per week: 37 versus 41 for fathers. Column (5) shows the difference in means between twin and non-twin households, and column (6) shows the *p*-value of the *t*-test on the difference in means. We find little differences in work status and work time for mothers between these two types of families. Regarding paternal work status, the probability of working is three percentage points lower in twin families than that in non-twin families. However, this difference is barely statistically significant ($p = .14$). Finally, the difference in paternal working hours between twin and non-twin families is very small.

Fertility may affect not only parental work status and work time but also work type. In the survey, respondents also provide information about the work type. Specifically, we consider whether parents run private businesses.¹¹ We call parents who run private businesses entrepreneurs. We find that 25% and 26% of mothers and fathers, respectively, were entrepreneurs in rural areas (columns (1) and (3)). Columns (5) and (6) show no significant differences between parents in non-twin and twin families.

We also check the effects of fertility on parental migration status and earnings due to the richness of the data. The variable “migrant” equals one if a parent worked and left the home village for at least six months in 2002. Otherwise, it equals zero. We find that in rural areas 3% of mothers and 5% of fathers were migrants. Regarding earnings, Columns (1) and (3) shows that fathers earned substantially more than mothers in rural Kunming. Mothers earned RMB 390 per month. By contrast, fathers earned RMB 620. Columns (5) and (6) show differences in these two variables between twin and non-twin families. Specifically, both parents from twin

⁹ Twinning at the first birth parity is also considered a better natural experiment than twinning at higher birth parities. See, e.g., Rosenzweig and Wolpin (1980).

¹⁰ Almost all families have three children or fewer in the CCTS. We thus exercise caution in generalizing our results based on China to countries with very high fertility.

¹¹ Private business does not include working on families' own land.

Table 2
Summary statistics.

| | Rural areas | | | | | | Urban areas | | | | | |
|---|-------------|-------|-----------|-------|----------------------|---------|-------------|-------|-----------|-------|----------------------|---------|
| | Non-twin | | Twins | | t-test | | Non-twin | | Twins | | t-test | |
| | (n = 969) | | (n = 697) | | | | (n = 624) | | (n = 542) | | | |
| | Mean | S.D. | Mean | S.D. | (3) ⁻ (1) | p-value | Mean | S.D. | Mean | S.D. | (3) ⁻ (1) | p-value |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | |
| Maternal labor supply | | | | | | | | | | | | |
| Worked (dummy) | 0.79 | 0.41 | 0.79 | 0.41 | 0.00 | [.97] | 0.61 | 0.49 | 0.62 | 0.49 | 0.01 | [.61] |
| Days worked per month | 20.00 | 11.31 | 20.24 | 11.41 | 0.24 | [.67] | 14.25 | 12.08 | 14.95 | 12.21 | 0.70 | [.32] |
| Hours worked per week | 36.50 | 22.95 | 37.24 | 23.45 | 0.74 | [.52] | 26.67 | 23.66 | 28.30 | 24.48 | 1.62 | [.25] |
| Entrepreneur (dummy) | 0.25 | 0.43 | 0.26 | 0.44 | 0.01 | [.60] | 0.06 | 0.25 | 0.12 | 0.32 | .05*** | [.00] |
| Migrant | 0.02 | 0.15 | 0.05 | 0.21 | .02*** | [.01] | 0.02 | 0.14 | 0.04 | 0.20 | .02** | [.02] |
| Earnings per month (RMB 1,000) | 0.38 | 0.41 | 0.40 | 0.45 | 0.02 | [.33] | 0.60 | 0.55 | 0.63 | 0.73 | 0.02 | [.58] |
| Paternal labor supply | | | | | | | | | | | | |
| Worked (dummy) | 0.86 | 0.34 | 0.84 | 0.37 | -0.03 | [.14] | 0.75 | 0.43 | 0.75 | 0.44 | -0.01 | [.71] |
| Days worked per month | 22.21 | 9.81 | 21.89 | 10.50 | -0.32 | [.52] | 17.97 | 11.14 | 18.22 | 11.42 | 0.25 | [.71] |
| Hours worked per week | 41.54 | 20.78 | 41.65 | 22.64 | 0.11 | [.92] | 34.86 | 23.66 | 35.74 | 24.66 | 0.88 | [.54] |
| Entrepreneur (dummy) | 0.27 | 0.45 | 0.25 | 0.43 | -0.03 | [.21] | 0.11 | 0.31 | 0.12 | 0.33 | 0.02 | [.39] |
| Migrant | 0.03 | 0.17 | 0.07 | 0.26 | .04*** | [.00] | 0.05 | 0.22 | 0.06 | 0.24 | 0.01 | [.65] |
| Earnings per month (RMB 1,000) | 0.64 | 0.82 | 0.59 | 0.52 | -0.05 | [.13] | 0.87 | 0.70 | 0.93 | 1.01 | 0.06 | [.22] |
| Household structure | | | | | | | | | | | | |
| Co-residence of grandparents (#) | 0.26 | 0.58 | 0.31 | 0.62 | 0.05* | [.09] | 0.26 | 0.56 | 0.26 | 0.61 | 0.01 | [.87] |
| Co-residence of grandparents (dummy) | 0.19 | 0.39 | 0.23 | 0.42 | .04* | [.07] | 0.19 | 0.40 | 0.18 | 0.39 | -0.01 | [.62] |
| Co-residence of other adults (#) | 0.02 | 0.16 | 0.02 | 0.16 | 0.00 | [.99] | 0.02 | 0.20 | 0.01 | 0.11 | -0.01 | [.22] |
| Co-residence of other adults (dummy) | 0.01 | 0.12 | 0.02 | 0.12 | 0.00 | [.83] | 0.01 | 0.11 | 0.01 | 0.07 | -0.01 | [.20] |
| Family characteristics | | | | | | | | | | | | |
| Number of children | 1.37 | 0.49 | 2.07 | 0.25 | .70*** | [.00] | 1.03 | 0.17 | 2.02 | 0.14 | .99*** | [.00] |
| Maternal age | 34.75 | 4.13 | 36.57 | 4.94 | 1.81*** | [.00] | 36.36 | 4.27 | 36.79 | 4.78 | 0.43 | [.10] |
| Maternal age at the first birth | 23.48 | 3.04 | 24.85 | 4.09 | 1.37*** | [.00] | 25.00 | 3.11 | 25.54 | 3.68 | .54*** | [.01] |
| Maternal schooling years | 8.28 | 2.90 | 7.90 | 2.98 | -.37** | [.01] | 10.82 | 3.09 | 10.29 | 3.31 | -.53*** | [.00] |
| Maternal ethnicity (Han = 1) | 0.83 | 0.38 | 0.84 | 0.37 | 0.01 | [.78] | 0.91 | 0.29 | 0.93 | 0.25 | 0.02 | [.12] |
| Average birth weight of children (k.g.) | 3.13 | 0.44 | 2.45 | 0.42 | -.68*** | [.00] | 3.10 | 0.49 | 2.43 | 0.46 | -.67*** | [.00] |

Notes: The data source is the CCTS.

families were more likely to be migrants, but fathers from twin families earned less ($p = .13$).

We use four variables to measure family structure. The first one is the number of both paternal and maternal grandparents living in the family in 2002. The average number of coresiding grandparents is 0.28 in rural areas. The second one is a dummy variable indicating whether a household had at least one coresiding grandparent. Table 2 shows that 21% of rural households live with at least one grandparent. The third and fourth variable is about the coresidence of other adult relatives such as aunts and uncles.¹² Columns (5) and (6) show that the probabilities of coresiding with grandparents are significantly higher for twin than for non-twin families in rural areas. For example, the probability of living with grandparents is approximately four percentage points higher in twin families than non-twin families in rural areas. We do not observe statistical significant difference in the probability of coresiding with other adult relatives between non-twin and twin households in rural areas.

We use four other variables to measure parental characteristics: maternal age, age at the first birth, schooling years, and ethnicity. Table 2 shows that mothers on average were 36 years old in 2002, had their first children at age 24, and have 8 years of schooling; 82% are Han Chinese. Comparing Column (3) with (1), we find that mothers with twin children were older in 2002, had their first children at older ages, and have less schooling years. We carefully control for these pre-determined characteristics in the regression analysis.

Finally, we use birth weight as a proxy variable of child prenatal endowments. Twins usually have inferior prenatal endowments such as birth weight than singletons. As our data show, twins are on average more than 0.6 kg lighter than singletons at birth. If we do not control for birth weight, the endowment effects may contaminate estimates using twinning as a natural experiment to identify the causal effect of fertility on household outcomes (Rosenzweig and Zhang, 2009). We address this concern using the CCTS data because we can directly control for birth weight in all regressions.

3. The effect of fertility on parental labor supply in rural China

To test the effects of fertility on parental labor supply, we estimate the following equation:

¹² The CCTS data contain no families that formally hired domestic helpers.

Table 3
OLS and IV estimates of fertility and parental labor supply

| | OLS (rural, non-twin) | | IV (rural) | | IV (urban) | |
|--------------------------------|-----------------------|---------------------|--------------------|--------------------|--------------------|--------------------|
| | Mother (1) | Father (2) | Mother (3) | Father (4) | Mother (5) | Father (6) |
| Dependent variable | | | | | | |
| Worked (dummy) | -.062** (0.030) | -.053** (0.026) | 0.038 (0.035) | 0.023 (0.031) | 0.050 (0.033) | -0.020 (0.032) |
| Days worked per month | -1.392* (0.843) | -1.343* (0.742) | 1.065 (0.984) | 0.845 (0.883) | 1.631* (0.857) | 0.395 (0.839) |
| Hours worked per week | -2.013 (1.721) | -1.869 (1.581) | 2.792 (2.004) | 2.302 (1.881) | 3.151* (1.693) | 0.355 (1.781) |
| Entrepreneur (dummy) | -.137*** (0.030) | -.146*** (0.031) | 0.060 (0.038) | 0.010 (0.037) | .064*** (0.022) | 0.036 (0.024) |
| Migrant (dummy) | -.034*** (0.010) | -.022** (0.011) | .036** (0.015) | .057*** (0.019) | 0.008 (0.013) | -0.001 (0.016) |
| Earnings per month (RMB 1,000) | -.069*** (0.024) | -0.066 (0.051) | .083** (0.034) | -0.000 (0.051) | .098** (0.046) | .175*** (0.056) |
| First-stage | | | | | | |
| Twinning | | | .717*** (0.024) | | .996*** (0.012) | |
| Observations | 969 | | 1,666 | | 1,166 | |

Notes: The data source is the CCTS. Each cell reports the estimate of α_1 in Eq. (1) from a separate regression. We control for children's average birth weight and maternal age, age squared, age at the first birth, schooling years, and ethnicity in all regressions. Columns (1) and (2) include the non-twin households in rural areas; Columns (3) and (4) include both non-twin households and households with twin children at the first birth parity in rural areas; and Columns (5) and (6) include both non-twin households and households with twin children at the first birth parity in urban areas. Heteroskedasticity-robust standard errors are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

$$y_i = \alpha_0 + \alpha_1 n_i + X_i \alpha_2 + \alpha_3 bw_i + \epsilon_i. \quad (1)$$

The dependent variable, y_i , is a measure of parental labor supply in household i . The measures of parental labor supply are discussed in the section above. The key independent variable, n_i , is fertility or the number of children. The variable X_i refers to a vector of maternal characteristics, including age, age squared, age at the first birth, schooling years, and ethnicity. bw_i is children's average birth weight in the family. ϵ_i is an error term.

3.1. OLS estimates using the non-twin sample in rural areas

We estimate the effects of fertility on parental labor supply in rural areas. We first conduct ordinary least squares (OLS) estimations of Eq. (1) using the sample of non-twin households only. We have a total of 969 non-twin households in the sample (Columns (1) and (2) in Table 2). We are interested in the estimate of α_1 , which measures the effect of fertility on labor supply.

The first two columns of Table 3 report the estimates of Eq. (1). Given the space limitations, we only report the estimated coefficients on α_1 in Eq. (1) for different dependent variables (labor supply measures). In all regressions, we control for children's average birth weight and maternal age, age squared, age at the first birth, schooling years, and ethnicity.

Our OLS estimates show that fertility is negatively correlated with parental labor supply, which is consistent with the prior literature (Angrist and Evans, 1998; Agüero and Marks, 2011). In terms of work status, with an additional child, the probability of working decreases by six percentage points for mothers and five percentage points for fathers. The estimates are statistically significant at the 5% level. Regarding working time, both parents worked one day less per month with an additional child. The estimates are statistically significant at the 10% level. Parental working hours per week also decreases with fertility, but the estimates are statistically insignificant.

Fertility is also negatively correlated with entrepreneurship, migration, and earnings. Both parents with more children are less likely to be entrepreneurs. The effects are statistically significant at the 1% level. With an additional child, the probability of running their own business decreases by about 14 percentage points for both parents. Regarding migration status, with the increase in fertility, parents, particularly mothers, are less likely to be migrants. The probability of being migrants for mothers decreases by three percentage points with an additional child. Finally, mothers earn substantially less with the increase in fertility. With an additional child, mothers' monthly earnings decrease by RMB 69, from the average amount of RMB 370. However, the correlation between fertility and fathers' earnings is not statistically significant.

3.2. IV estimates using the rural sample

The OLS estimates based on the non-twin sample do not have a causal interpretation because of the endogeneity of fertility. Fertility is endogenous either because it is a household decision determined by unobserved household characteristics, or because labor supply and earnings could affect fertility (Browning, 1992).

Table 4
OLS Estimates of the determinants of fertility.

| Dependent variable: | Number of children | | | |
|---------------------------------|--------------------|---------------------|--------------------|---------------------|
| | Rural sample | | Urban sample | |
| | (1) | (2) | (3) | (4) |
| Twinning | .697*** (0.018) | .704*** (0.018) | .990*** (0.009) | .988*** (0.009) |
| Maternal age | | 0.009 (0.019) | | 0.015 (0.010) |
| Maternal age squared/100 | | 0.009 (0.026) | | -0.018 (0.013) |
| Maternal age at the first birth | | -.032*** (0.004) | | -0.002 (0.002) |
| Maternal schooling years | | -.024*** (0.003) | | -.008*** (0.002) |
| Maternal ethnicity(Han = 1) | | -.153*** (0.028) | | -.027 (0.021) |
| Durbin–Wu–Hausman tests | | | | |
| χ^2 -statistics | 0.80 | | 1.23 | |
| <i>p</i> -values | 0.37 | | 0.27 | |
| Observations | 1,666 | 1,666 | 1,166 | 1,166 |
| R-squared | 0.42 | 0.48 | 0.91 | 0.91 |

Notes: The data source is the CCTS.

Following the literature (Rosenzweig and Wolpin, 1980), we use twin birth to identify the effect of fertility on parental labor supply. Specifically, we conduct the instrumental variable (IV) estimations by using twinning at the first birth as an IV. For this purpose, we use a sample that includes non-twin households and households that have twins at the first birth parity in rural areas. This sample contains a total of 1666 households (Columns (1) and (2) in Table 1). According to Imbens and Angrist (1994), this IV estimate can be interpreted as a local average treatment effect (LATE). It identifies the treatment effect of an increase in fertility induced by twinning at the first birth parity on parental labor supply. The IV estimate of α_1 identifies the causal effect of an increase in fertility from one to two induced by twinning on labor supply.¹³

We first examine whether twinning at the first birth is a valid IV by estimating the effect of twinning on fertility. The results are reported in Table 4. The point estimates suggest that twin families on average have 0.7 more children. The heteroscedasticity-robust *t*-statistics are almost 40 in both columns, suggesting the IV is highly correlated with fertility (Stock and Yogo, 2002). The R-squared is 0.42 in Column (1), suggesting twinning alone can explain 42% of the variation in fertility. To test the exogeneity of twinning at the first birth parity, we carry out a Durbin–Wu–Hausman test. The Durbin–Wu–Hausman test statistic suggests no statistically significant difference between the estimates of α_1 in Columns (1) and (2).

Oster (2016) suggests that the stability of coefficients, as examined by the Durbin-Wu-Hausman statistic, is not sufficient condition for the exogeneity of the treatment variable. We also need to consider the increase in R^2 after adding controls. Following the procedure of Oster (2016),¹⁴ we assume equal selection in the observables and unobservables. We also assume that controlling for the unobservables increases the R^2 by the same amount as controlling for the observables. Allowing for selection on the unobservables under these two assumptions, the coefficient of twinning on the number of child is 0.711. The bootstrapped ($n = 1,000$) standard error is 0.018. The estimate using the procedure of Oster (2016) suggests that selection on the unobservables does not drive the estimate of the effect of twinning on the number of children.

The IV estimates reported in Columns (3) and (4) in Table 3 suggest that fertility does not reduce parental labor supply in rural areas, which is in stark contrast to the OLS estimates (Columns (1) and (2)). Most of the estimates of α_1 are statistically insignificant. In three cases, the variable fertility is statistically significant but the signs are positive. Fertility induced by twinning at the first birth parity increases the probability of migration for both parents and increases mothers' earnings.¹⁵

Using twinning as an identification for fertility raises concerns. First, prenatal endowments, such as child birth weight, are generally inferior for twins than for singletons, and thus twinning may have an effect on family behavior through endowments besides fertility (Rosenzweig and Zhang, 2009). Following Rosenzweig and Zhang (2009), we address this concern directly by controlling for child birth weight in estimating Eq. (1).

¹³ Twins are different from singletons primarily because twins exhibit inferior prenatal endowments than singletons. As discussed in the Data section, we follow Rosenzweig and Zhang (2009) to control for children's birthweight in all regressions. Another concern is that twins have closer birth spacing compared with singletons. The economy of scale implies that the cost of raising a pair of twins might be lower than the cost of raising two singletons. However, this economy of scale applies similarly to parental and grandparental childcare. To the extent that twinning does not affect the marginal rate of substitution between parental and grandparental childcare, twinning does not introduce further bias in our estimates because of the economy of scale in raising twins.

¹⁴ The procedure of Oster (2016) is implemented through the command "psacalc" in Stata. We thank an anonymous referee for suggesting this procedure.

¹⁵ The unexpected increase in fertility raises household financial burdens. As the childcare is partly compensated by grandparents, parents may migrate from rural to urban areas and earn more.

Second, the incidence of twinning is higher among mothers undergoing fertility treatments, which may invalidate it serving as an IV. However, this concern does not apply to our sample. Kunming is a less developed area in China. Fertility drugs were unavailable in the 1980s and 1990s.¹⁶ In addition, allowing for selection on the unobservables barely changes the estimate of the effect of twinning on the number of children, suggesting that the incidence of twinning is uncorrelated with unobserved maternal characteristics.

Third, Rosenzweig and Wolpin (1980, 2000) suggest twinning does not affect parental life-cycle labor supply, especially at the first birth, because the subsequent fertility is lower for families with twinning, and complete fertility is unaffected. However, twinning does affect complete fertility in China, because most households have fewer children under the one-child policy than their optimal fertility level without the constraint. Thus, twinning serves as a good natural experiment in increasing fertility in China.

Summing up the results in Columns (1)–(4) of Table 3, the exogenous increase in fertility induced by twinning does not reduce parental labor market outcomes in our sample. This result is consistent with the recent literature on fertility and parental labor supply in rural areas in developing countries (Agüero and Marks, 2008, 2011; Cáceres-Delpiano, 2012). He and Zhu (2016) also find that fertility does not reduce female labor force participation using the 2000 census. The 2000 census is comparable to CCTS which is conducted in 2002. So, we replicated the findings of He and Zhu (2016) using an alternative data source. However, the existing literature does not provide a good answer to why the OLS estimates of the effect of fertility on labor supply variables are negative whereas the IV estimates are not. We try to provide an answer in the next section.

4. Effects of fertility on household structure in rural china

We propose that the negative effect of fertility on parental labor supply is mediated by the childcare provided by grandparents. When the number of children unexpectedly increases due to twinning, the family structure may be adjusted. As grandparents' time cost is lower than that of parents, grandparents move into the household, take care of the children, and do the household chores. In this scenario, parental labor supply does not necessarily decrease with the increase in fertility. To test this hypothesis, this section examines the effect of fertility on the coresidence of grandparents and other adult relatives.

We estimate Eq. (1) using the coresidence of grandparents and other adult relatives as dependent variables. Panel A of Table 5 reports IV estimates of α_1 in Eq. (1) using the rural sample. The instrumental variable is twinning at the first birth parity. We find that an increase in the number of children increases the coresidence of grandparents (Columns (1) and (2)). Households with one more child have on average 0.128 more coresiding grandparents, and are seven percentage points more likely to have coresiding grandparents. The effects are statistically significant at the 5% level. Having one more child also increases the coresidence of other adult relatives, but the effects are not statistically significant (Columns (3) and (4)).

The results in Table 5 can explain the positive effects of fertility on the probability of parents being migrants and entrepreneurs. If grandparental time is less valuable than parental time in the labor market, fertility reduces grandparents' labor supply (Rupert and Zanella, 2014) and substitutes parental childcare by grandparental childcare. When grandparents move into the household with the increase in fertility, parental time spent on childcare and household chores probably decreases. Therefore, parental time in the labor market might be more and flexible. Summarizing the results reported in Table 5, the exogenous increase in fertility induced by twinning enhances the coresidence of grandparents. This result is consistent with our prediction. With the increase in fertility, grandparents move into the household to take care of their grandchildren.

One limitation with the CCTS is that it does not contain information on grandparental childcare and household chore time. Using the Chinese nutrition and health survey in 1989, a simple tabulation shows that hours that grandparents spent on childcare and household chores per week doubled when the number of children younger than six increased from one to two (Li et al., 2013). In addition, by using the Chinese population policy as a natural experiment, Li et al. (2013) show grandparents' time spent on childcare and household chores increased by 23 h/week when fertility increased from one to two and more. By contrast, parental time on childcare and household chores increased by only six hours. These results support this paper's proposed mediation role of the childcare provided by grandparents.

5. A comparative study of urban China

Rural and urban areas of China are in different stages of economic development. While mutual help within extended families is prevalent in rural China, parents in urban China may not seek help from members of the extended families. Parents in urban China have access to formal childcare instead.

The CCTS contains an urban sample, allowing us to conduct a comparative analysis of rural and urban China. See Columns (7)–(12) of Table 2 for summary statistics of the urban sample. A notable feature of the urban sample is that almost all non-twin households have only one child, and almost all twin households have only two children. Thus, the non-twin sample in urban areas contains little variation in fertility. OLS estimates of Eq. (1) using the non-twin sample alone is not informative. As in the rural sample, we use twinning at the first birth parity as the instrumental variable for fertility. Columns (3) and (4) of Table 4 show that twinning shifts fertility up by almost one child. As in the rural sample, we implemented the procedure of Oster (2016) by assuming equal selection in observables and unobservables. The coefficient of twinning on the number of children is 0.986 with a bootstrapped

¹⁶ During a follow-up survey of the CCTS project in 2013, one of the authors interviewed many mothers and staff members in the local bureau of statistics in the field work. No one had ever heard about the fertility drugs that can be used to increase the incidence of twinning.

Table 5
IV Estimates of fertility and household structure.

| Dependent Variable | Co-residence of grandparents | | Co-residence of other adults | |
|--------------------------------|------------------------------|-------------------|------------------------------|-------------------|
| | Number (1) | Dummy (2) | Number (3) | Dummy (4) |
| Panel A. Rural sample | | | | |
| Number of children | .128** (0.050) | .070** (0.035) | 0.012 (0.015) | 0.012 (0.011) |
| First-stage Twinning | | | .717*** (0.024) | |
| Observations | | | 1,666 | |
| Panel B. Urban sample | | | | |
| Number of children | 0.008 (0.045) | -0.007 (0.031) | -0.009 (0.009) | -0.010 (0.007) |
| First-stage Twinning | | | 0.996*** (0.012) | |
| Observations | | | 1,166 | |

Notes: The data source is the CCTS. We control for children's average birth weight and maternal age, age squared, age at the first birth, schooling years, and ethnicity in all regressions. Columns (1) and (2) include both non-twin households and households with twin children at the first birth parity in rural areas; and Columns (3) and (4) include both non-twin households and households with twin children at the first birth parity in urban areas. Heteroskedasticity-robust standard errors are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

($n = 1,000$) standard error of 0.016.

Columns (5) and (6) of Table 3 show that fertility has positive effects on parental labor supply in some dimensions. Although fertility does not increase parental labor force participation (He and Zhu, 2016), fertility affects the intensive margin of parental labor supply. Having one more child increase makes the mother work two more days per month and three more hours per week. The effects are statistically significant at the 10% level. An increase in fertility also makes mothers six percentage points more likely to be entrepreneurs. The effect is statistically significant at the 1% level. Having one more child also increases both parents' earnings per month. Mothers earn ten percent more, and fathers earn 18 percent more. The effects are statistically significant at the 5% level. Overall, Columns (5) and (6) of Table 3 indicate that in urban areas both parents increase their labor supply if they have one more child.

Does fertility increases the coresidence of grandparents in urban China? Panel B of Table 5 shows IV estimates of the effect of fertility on the coresidence of grandparents and other adult relatives using the urban sample. We do not detect statistically significant effect of fertility on the coresidence of grandparents and other adult relatives. As parents in urban China have access to formal childcare, they do not seek help from members of the extended family. To pay for formal childcare, parents with more children need to earn more. Thus, it is not surprising that the positive response of labor supply is even stronger in urban China than that in rural China.

The comparison of the results using rural and urban samples of CCTS illustrates the substitution between informal and formal childcare in the process of economic development. In the less developed rural areas of China, grandparents supply informal childcare within the extended family. With the help of grandparents, parents maintain their labor supply when they have an additional child. In the more developed urban areas of China, parents do not seek informal childcare from the children's grandparents. To cover the expense of formal childcare, parents increase their labor supply and earn more to cover the increasing family expenditures. The accessibility of formal childcare substitutes for the informal childcare provided by grandparents.

6. Discussion and conclusion

This study examines the effects of fertility on household structure and parental labor supply in rural China where extended families are prevalent. To address the endogeneity problem, we use the CCTS data. The CCTS is a unique survey on households with twin children and a comparison group of non-twin households as well. Based on the rural sample of this data set, although a negative correlation exists between fertility and parental labor supply, we find no evidence about the negative effects of fertility on parental labor supply by using twinning as a natural experiment. By contrast, we find that the twinning-induced increase in fertility significantly enhances the coresidence of grandparents in rural China. Therefore, the childcare provided by grandparents mitigates the negative effects of fertility on parental labor supply in rural China.

We also conduct a comparative analysis using data from urban China. We find that although fertility does not reduce parental labor supply, fertility does not induce coresidence of grandparents in urban China. The contrast between rural and urban China suggests that formal childcare can substitute for informal childcare provided by grandparents throughout different stages of economic development.

The findings in this paper have important policy implications. The one-child policy is currently subject to heated debate in both academic research and government policy making. We show that reducing fertility cannot increase labor supply, and thus, it does not help raise the immediate income. This finding indicates that China's birth control policy may not contribute to the growth of GDP in the short or medium run. The lower fertility may have led to better child quality, and this improved human capital will promote the economic growth of China in the long run. However, Rosenzweig and Zhang (2009) conclude that the contribution of the one-child policy in China to the development of its human capital is modest at best.

The results of our paper also shed light on the understanding of the aggregate socioeconomic changes in developing countries, particularly the comovements of fertility, household structure, labor supply, and economic development. The prevalence of extended families decreases with economic development (Chu and Yu, 2009). With the decrease in extended families, fertility would have a bigger effect on labor supply if formal childcare is inaccessible to parents. Thus, the decline of the extended family can be a factor weighing against economic development (or women's labor force participation) because it allows prime-age adults not to have to withdraw from the labor market to have children. Furthermore, the decline of extended families could be one reason for the decline of fertility with economic development. That is, having a child now means more parents leave the labor force because grandparents are no longer around to take care of the child while parents work. Formal childcare can compensate for the negative effects of the decrease in childcare service from grandparents on parental labor supply. However, the causality can also run from the opposite direction as shown in the paper. The decrease in fertility causes the decline in extended family, because of the less demand for childcare from grandparents. In this sense, our paper is only a starting point to explore the interplays between fertility, household structure, and labor supply during economic development. Future research in this direction is highly desirable.

Appendix A

Table A1

OLS and IV estimates of fertility and parental labor supply, using the sample of families with the eldest child younger than twelve years old.

| | OLS (rural, non-twin) | | IV (rural) | | IV (urban) | |
|--------------------------------|-----------------------|---------------------|-------------------|--------------------|--------------------|---------------------|
| | Mother (1) | Father (2) | Mother (3) | Father (4) | Mother (5) | Father (6) |
| Dependent variable | | | | | | |
| Worked (dummy) | -0.045 (0.039) | -0.028 (0.032) | 0.026 (0.045) | 0.038 (0.037) | 0.064 (0.043) | 0.028 (0.039) |
| Days worked per month | -1.258 (1.096) | -1.067 (0.919) | 0.999 (1.255) | 1.126 (1.086) | 2.287** (1.144) | 1.747 (1.074) |
| Hours worked per week | -1.687 (2.212) | -0.929 (1.910) | 3.608 (2.571) | 3.912* (2.320) | 4.762** (2.300) | 3.456 (2.322) |
| Entrepreneur (dummy) | -.172*** (0.040) | -.170*** (0.041) | 0.070 (0.049) | 0.055 (0.049) | .089*** (0.031) | .074** (0.034) |
| Migrant (dummy) | -.034*** (0.012) | -.021 (0.014) | .034** (0.017) | .037* (0.021) | 0.008 (0.018) | -0.010 (0.020) |
| Earnings per month (RMB 1,000) | -.079** (0.033) | -.051 (0.057) | 0.052 (0.046) | 0.042 (0.056) | 0.066 (0.067) | .238*** (0.081) |
| First-stage | | | | | | |
| Twinning | | | | .729*** (0.030) | | 1.013*** (0.016) |
| Observations | | 624 | | 1,033 | | 733 |

Notes: The data source is the CCTS. Each cell reports the estimate of α_1 in Eq. (1) from a separate regression. We control for children's average birth weight and maternal age, age squared, age at the first birth, schooling years, and ethnicity in all regressions. Columns (1) and (2) include the non-twin households in rural areas; Columns (3) and (4) include both non-twin households and households with twin children at the first birth parity in rural areas; and Columns (5) and (6) include both non-twin households and households with twin children at the first birth parity in urban areas. Heteroskedasticity-robust standard errors are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A2

IV estimates of fertility and household structure, using the sample of families with the eldest child younger than twelve years old.

| Dependent Variable | Co-residence of grandparents | | Co-residence of other adults | |
|------------------------------|------------------------------|-------------------|------------------------------|-------------------|
| | Number (1) | Dummy (2) | Number (3) | Dummy (4) |
| Panel A. Rural sample | | | | |
| Number of children | .208*** (0.069) | .119** (0.048) | 0.013 (0.025) | 0.015 (0.017) |
| First-stage | | | | |
| Twinning | | | .729*** (0.030) | |
| Observations | | | 1,033 | |
| Panel B. Urban sample | | | | |
| Number of children | -0.039 (0.061) | -0.041 (0.040) | -0.006 (0.012) | -0.007 (0.008) |
| First-stage | | | | |
| Twinning | | | 1.013*** (0.016) | |
| Observations | | | 733 | |

Notes: The data source is the CCTS. We control for children's average birth weight and maternal age, age squared, age at the first birth, schooling years, and ethnicity in all regressions. Columns (1) and (2) include both non-twin households and households with twin children at the first birth parity in rural areas; and Columns (3) and (4) include both non-twin households and households with twin children at the first birth parity in urban areas. Heteroskedasticity-robust standard errors are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

References

- Agüero, J.M., Marks, M.S., 2008. Motherhood and female labor force participation: evidence from infertility shocks. *Am. Econ. Rev.* 500–504.
- Agüero, J.M., Marks, M.S., 2011. Motherhood and female labor supply in the developing world: evidence from infertility shocks. *J. Hum. Resour.* 46 (4), 800–826.
- Angrist, J.D., Evans, W.N., 1998. Children and their parents' labor supply: evidence from exogenous variation in family size. *Am. Econ. Rev.* 88 (3), 450–477.
- Baker, M., Gruber, J., Milligan, K., 2008. Universal childcare, maternal labor supply, and family well-being. *Rev. Econ. Stud.* 116 (4), 709–745.
- Borjas, G., 2000. *Labor Economics*. McGraw-Hill.
- Bronars, S.G., 1994. The economic consequences of unwed motherhood: using twin births as a natural experiment. *Am. Econ. Rev.* 1141–1156.
- Browning, M., 1992. Children and household economic behavior. *J. Econ. Lit.* 1434–1475.
- Cáceres-Delpiano, J., 2012. Can we still learn something from the relationship between fertility and mothers employment? Evidence from developing countries. *Demography* 49 (1), 151–174.
- Cain, G.G., 1966. *Married Women in the Labor Force: An Economic Analysis*. University of Chicago.
- Chu, C.Y., Kim, S., Tsay, W.J., 2014. Coresidence with husbands parents, labor supply, and the duration to first birth. *Demography* 51, 185–204.
- Chu, C.Y., Xie, Yu, Yu, R.R., 2011. Coresidence with elderly parents: a comparative study of southeast china and taiwan. *J. Marriage Family* 73 (1), 120–135.
- Chu, C.Y.C., Yu, R.R., 2009. *Understanding Chinese Families: A Comparative Study of Taiwan and Southeast China*. Oxford University Press.
- Cramer, James, 1980. Fertility and female employment: problems of causal direction. *Am. Sociol. Rev.* 167–190.
- Cruces, G., Galiani, S., 2007. Fertility and female labor supply in latin America: new causal evidence. *Labour Econ.* 14 (3), 565–573.
- Ebenstein, A., 2009. When is the local average treatment close to the average? Evidence from fertility and labor supply. *J. Hum. Resour.* 44 (4), 955–975.
- Ebenstein, A., 2010. The “missing girls” of china and the unintended consequences of the one child policy. *J. Hum. Res.* 45 (1), 87–115.
- Ebenstein, A., 2011. Estimating a dynamic model of sex selection in China. *Demography* 48 (2), 783–811.
- Fleisher, B.M., Rhodes, G.F., 1979. Fertility, women's wage rates, and labor supply. *Am. Econ. Rev.* 69 (1), 14–24.
- Gong, X., Van Soest, A., 2002. Family structure and female labor supply in mexico city. *J. Hum. Resour.* 163–191.
- He, X., Zhu, R., 2016. *Fertility and Female Labour Force Participation: Causal Evidence From Urban China*. The Manchester School, pp. 664–674.
- Imbens, G.W., Angrist, J.D., 1994. Identification and estimation of local average treatment effects. *Econometrica* 467–475.
- Killingsworth, M.R., Heckman, J.J., 1986. Female labor supply: a survey. *Handbook of Labor Economics*. vol. 1. pp. 103–204.
- Lefebvre, P., Merrigan, P., 2008. Child-care policy and the labor supply of mothers with young children: a natural experiment from canada. *J. Labor Econ.* 26 (3), 519–548.
- Li, H., Yi, J., Zhang, J., 2011. Estimating the effect of the one-child policy on the sex ratio imbalance in china: identification based on the difference-in-differences. *Demography* 48 (4), 1535–1557.
- Li, H., Yi, J., Zhang, J., 2013. Fertility and parental labor supply in a developing country: Identification based on a unique population policy. Working paper.
- Li, H., Zhang, J., Zhu, Yi, 2008. The quantity-quality trade-off of children in a developing country: identification using chinese twins. *Demography* 223–243.
- Maurer-Fazio, M., Connelly, R., Chen, L., Tang, L., 2011. Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000. *J. Hum. Resour.* 46 (2), 261–294.
- Nakamura, A., Nakamura, M., 1992. The econometrics of female labor supply and children. *Econ. Rev.* 11 (1), 1–71.
- Oster, E., 2016. Unobservable selection and coefficient stability: theory and evidence. *J. Bus. Econ. Stat.* (forthcoming).
- Posadas, J., Vidal-Fernández, M., 2013. Grandparents' childcare and female labor force participation. *IZA J. Labor Policy* 2–14.
- Family Planning Commission of Yunnan Province, 2003. *Regulations on Population and Family Planning in Yunnan Province*. China Population Press.
- Rosenzweig, M.R., Wolpin, K.I., 1980. Life-cycle labor supply and fertility: causal inferences from household models. *J. Polit. Econ.* 88 (2), 328–348.
- Rosenzweig, M.R., Wolpin, K.I., 1985. Specific experience, household structure, and intergenerational transfers: farm family land and labor arrangements in developing countries. *Q. J. Econ.* 961–987.
- Rosenzweig, M.R., Wolpin, K.I., 2000. “Natural” natural experiments. *J. Econ. Lit.* 38 (4), 827–874.
- Rosenzweig, M.R., Zhang, J., 2009. Do population control policies induce more human capital investment? Twins, birth weight and China's one-child policy. *Rev. Econ. Stud.* 76 (3), 1149–1174.
- Rupert, P., Zanella, G., 2014. *Grandchildren and Their Grandparents' Labor Supply*. Working Paper.
- Schultz, T.P., 2010. Population policies, fertility, women's human capital, and child quality. *Handbook of Development Economics*. vol. 4. pp. 3249–3303.
- Stock, J.H., Yogo, M., 2002. Testing for Weak Instruments in Linear IV Regression. National Bureau of Economic Research.

- Glass, J., 1985. Household structure and labor force participation of black, hispanic, and white mothers. *Demography* 22 (3), 381–394.
- Willis, R.J., 1987. What have we learned from the economics of the family? *Am. Econ. Rev.* 77 (2), 68–81.
- Levine, R.E., 1992. The effect of household structure on womens economic activity and fertility: evidence from recent mothers in urban mexico. *Econ. Dev. Cult. Change* 41 (1), 89–102.
- Yi, J., Heckman, J.J., Zhang, J., Conti, G., 2015. Early health shocks, intrahousehold resource allocation, and child outcomes. *Econ. J.* F347–F371.
- Zeng, Z., Xie, Y., 2014. The effects of grandparents on childrens schooling: evidence from rural china. *Demography* 51 (2), 599–617.