I Z A

IZA DP No. 9214

The Competitive Earning Incentive for Sons: Evidence from Migration in China

Wenchao Li
Junjian Yi

July 2015

The Competitive Earning Incentive for Sons: Evidence from Migration in China

Wenchao Li
National University of Singapore

Junjian Yi
*National University of Singapore
and IZA*

Discussion Paper No. 9214
July 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

ABSTRACT

The Competitive Earning Incentive for Sons: Evidence from Migration in China[*]

This paper first finds a clear pattern of child gender difference in family migration in China. Specifically, our estimates show that on average, the first child being a son increases the father's migration probability by 25.2 percent. We hypothesize that the family's competitive earning incentive for sons drives this child gender effect on family migration: parents migrate to earn more money in an attempt to improve their sons' relative standing in response to the ever-rising pressure in China's marriage market. This competitive-earning-incentive hypothesis is then supported by additional empirical evidence. We further find that, facing heavier financial pressure from the marriage market, parents spend less on their sons' education and more on marriage and buying houses and durable goods. This gender difference in resource allocation, together with the absentee-father problem resulting from paternal migration, may unexpectedly adversely affect boys' long-run human capital development in China.

Corresponding author:

Junjian Yi
Department of Economics
Faculty of Arts & Social Sciences
National University of Singapore
AS2 Level 6, 1 Arts Link
Singapore 117570
Singapore
E-mail: junjian@nus.edu.sg

1   Introduction

Migration, normally referring to the movement of workers from the low-productivity sector to the high-productivity sector, improves labor-allocation efficiency. Migration in contemporary China is unprecedentedly large scaled. Specifically, in 2010, China had more than 150 million migrant workers. As such, migration is one of the most crucial forces driving China's fast economic growth (Meng, 2012; Song, Storesletten and Zilibotti, 2011). Figuring out its determinants is therefore of great significance.[1] The high number of explaining factors documented in the existing literature, however, is still not able to fully explain the migration in China.

In this paper, we add a new and important determinant of migration that most of the existing studies overlook but that might be at work in China: parents migrate to earn more money in order to improve their sons' relative standing in response to the ever-rising pressure in the marriage market. The decision to add this factor is motivated by the co-rising of the sex ratio and the number of migrant workers, as illustrated graphically in Figure 1.[2] The sex ratio at birth in China has risen drastically in recent decades. With a biologically normal range of 103–108, it reached 111.9 in 1990, rose to 119.9 in 2000, and remained at that high level until 2010. Sharygin, Ebenstein and Das Gupta (2013) predict that, by 2030, more than 20 percent of men in China aged 30–39 will never have married. The competition in the marriage market has thus become increasingly fierce for men.

We hypothesize that as the marriage market has become considerably competitive for males in China, parents of boys tend to work away from their hometown in an attempt to earn more money to improve their sons' competitiveness. We call this tendency the "competitive-earning-incentive" hypothesis. Some related work argues that as the sex ratio rises, Chinese parents with a son raise their savings in a competitive manner in order to improve their son's relative attractiveness for marriage (Wei and Zhang, 2011). Saving, however, has a smaller and secondary effect on material wealth, whereas earnings

---

1. See, e.g., Meng (2012) and Zhao (1999).
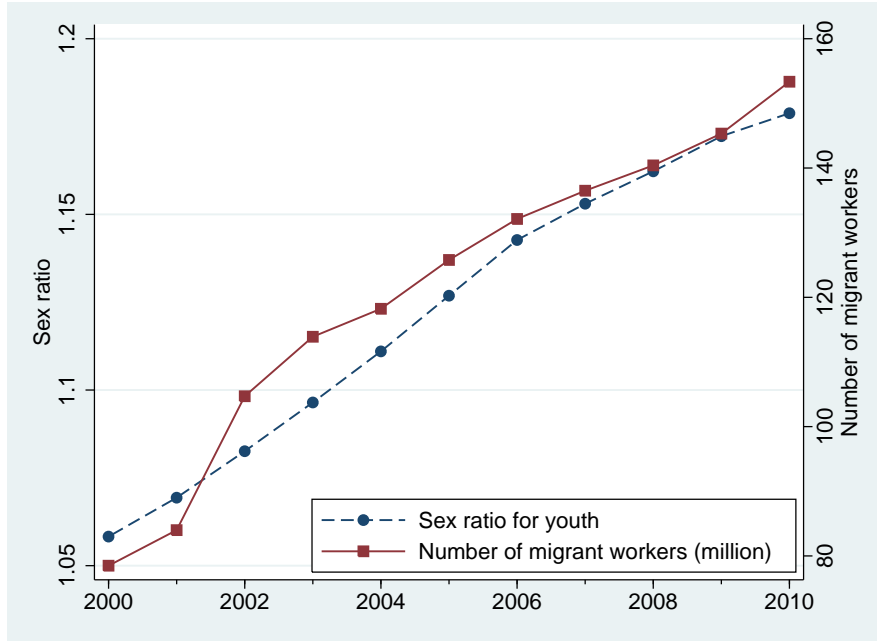2. Sex ratio refers to the number of males to the number of females.

FIGURE 1
Trends of sex ratio and the number of migrant workers in China: 2000–2010

Notes: The sex ratio is calculated for each year for the age cohort from 0 to 15 years old (inclusive). Data for the sex ratios are based on *the 2010 China Census*. For example, the sex ratio for the group of children from 0 to 15 years old in 2009 is inferred from the age cohort from 1 to 16 in *the 2010 China Census*, because these two groups are supposed to be the same. Data for the number of migrant workers (in million) are from *the China Statistical Yearbook 2008–2010*. The correlation coefficient between the sex ratio and the number of migrant workers across years from 2000 to 2010 is 0.98 (with the 95 percent confidence interval: 0.923 to 0.995).

play a larger and primary role in accumulating wealth. Therefore, in the sense that migration considerably improves individual income, the competitive earning incentive resulting from the intensified competition in the marriage market should be a crucial factor in explaining the large-scale migration in China.

The purpose of this paper is to address a causal relationship between family migration and the competitiveness of the marriage market for males, by verifying the competitive-earning-incentive hypothesis. First of all, to provide some preliminary exploration, Figures 1 and 2 respectively illustrate the time series and cross-sectional relationships between the sex ratio for youth and the migration intensity in the nation. Figure 1 graphically illustrates the nationwide time-series correlation between the steadily rising sex ratio for youth and the increasing number of migrant workers. The sex ratio for youth in China and the number of migrant workers are saliently rising together. Apart from
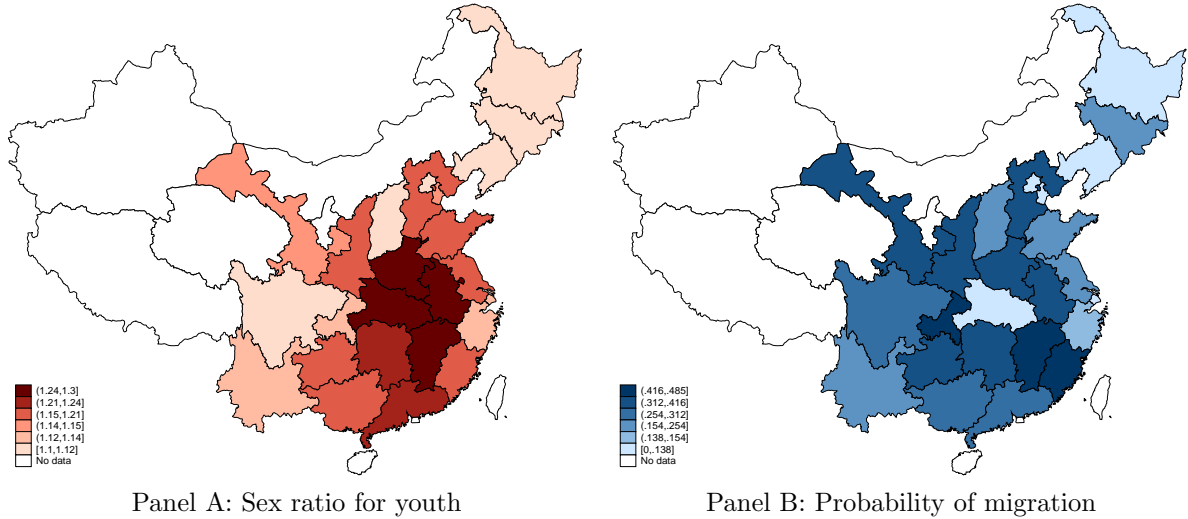
2

Panel A: Sex ratio for youth   Panel B: Probability of migration

FIGURE 2

Cross-sectional sex ratios and probabilities of migration in China in 2010

Notes: In Panel A, the province-specific sex ratios are calculated for the age cohort from 0 to 14 years old (inclusive) using *the 2010 China Census* data. In Panel B, data of the province-specific probabilities of some family member working outside the hometown are from *the 2010 CFPS*. Data are available for 25 provinces. A darker color in an area indicates a higher value of the respective variable of interest in both panels. The correlation coefficient between sex ratios and migration probabilities across provinces in China is 0.59 (with the 95 percent confidence interval: 0.247 to 0.796).

rising together across time, the sex ratio and the number of migrant workers also turn out to be highly correlated cross-sectionally. In Figure 2, we present for 25 provinces in China the province-specific sex ratio for youth and the probability of the family having migrant workers. Although these figures are not rigorous proofs, they are in line with our hypothesis that the competitive earning incentive for sons plays an important role in explaining migration in China.

We then rigorously test the competitive-earning-incentive hypothesis by examining the impact of child gender on household migration using Chinese household data. To address the causal relationship between child gender and family migration, this paper makes use of the first-born child's gender as the key explanatory variable. The exogeneity of this variable is not only documented by the literature (Dahl and Moretti, 2008; Ebenstein, 2010, 2011; Wei and Zhang, 2011), but also supported by the data used in this paper. In particular, the first-child sex ratio is 1.06 in our data set, which is well within the normal range. We show more evidence on the randomness of the first child's gender in China in

our analysis below.

To begin with, we present the estimation results of the effect of having a first-born son on family migration. We find the father is significantly more likely to become a migrant worker if he has a first-born son compared with a daughter. This effect is economically substantial, accounting for a 25.2 percent higher probability of a migrant father. We estimate that in China, roughly 1.7 million first-born daughters younger than 16 years old would have driven their fathers out of their hometowns if they had been sons. These findings are robust to checks of restricting the sample, adding additional control variables, and using alternative measurements of child gender.

How should we interpret these findings? Although the competitive-earning-incentive hypothesis is consistent, it is not the only possible explanation. Another possible interpretation, for instance, is that raising boys in China may cost more money than raising girls, so parents have to try to earn more money to foster their sons. To help sort out the alternative explanations, we go through a series of empirical examinations. First, we find the child-gender effect exhibits clear heterogeneity. In general, having a first-born son has a larger and more significant impact in rural areas and for low-education households. More importantly, when the local sex ratio is higher, the first-born child being a son has a larger impact on family migration. We also find that when the first-born child is a son, migrants in the family tend to send back more remittance. Additionally, the family tends to spend the remittance as well as other available resources on children's marriage and buying houses and other durable goods. The housing conditions are also better in general. Taking all results together, the competitive-earning-incentive hypothesis as an explanation for migration in China is strongly favored.

Our findings are important for several reasons. First, regardless of how one interprets our findings on family migration decisions as well as other decisions, we show child gender matters. The results suggest the effect on parental migration when the first-born child is a son is substantial. As migration stimulates the economic growth, our findings indicate the rising sex ratio is potentially beneficial in terms of economic development. Therefore, the government may announce child-gender-oriented incentives for migration, specially

encouraging girls' parents to work away from their hometowns.

Second, a high sex-ratio imbalance is potentially harmful in terms of children's, in particular boys', educational outcomes in the long run.[3] We find that, facing the heavy finance pressure of their sons' marriage, parents spend more resources on their sons' marriage and on buying houses and other durable goods, while allocating less on their sons' education.

Furthermore, the absentee-father problems resulting from paternal migration may be detrimental to boys' long-run human capital development in China. Growing up without a father has important negative consequences for children (Dahl and Moretti, 2008), and the possible negative emotional consequences are likely to explain adolescents' school performance (Powdthavee and Vernoit, 2013). Zhang et al. (2014) show that one in three children under age 17 in rural China is living without one or both parents who have migrated in search of work in cities. Over 61 million children were left-behind resulting from parents migrating for work in rural China in 2005. The authors further find significant negative impacts of being left-behind on children's cognitive development. Our findings indicate the first-born child being a son drives fathers outside of home. Boys and their siblings thus may be overall more likely to be exposed to these negative effects. Taking into account that fathers play a more important role in modeling the traditional social role for sons than for daughters (Lundberg and Rose, 2002), the absentee-father problems for boys are further magnified.

Last, because the sex-ratio imbalance at birth has been increasing steadily since the mid-1980s, the imbalance for the premarital-age cohort will almost surely continue to increase over the next decade, even if the sex ratio at birth begins to reduce soon. This increase in the sex-ratio imbalance implies the increasing migration intensity that is stimulated by the competitive earning incentive as documented in this paper will rise in significance and notableness in the near future.

The remainder of the paper is organized as follows. In Section 2, we provide background information for our study by describing some basic facts about migration and the

---

3. See, e.g., Dahl and Moretti (2008); Fleisher, Li and Zhao (2010); Oswald and Powdthavee (2010).

marriage market in China. In Section 3, we describe the data and specify the regression equation, and more importantly, we discuss the randomness of the key explanatory variable, the first child's gender, to address a causal inference. In Section 4, we present empirical evidence on the differential impacts of a first-born son versus a first-born daughter on family migration. In Section 5, we discuss all possible interpretations of our findings and present additional empirical evidence attempting to rule out the competing explanations. Section 6 concludes.

2   Background

This section provides the research background. First, we review China's *hukou* system and contemporary migration. Next, we describe the increase in the sex ratio and the implied consequences on the marriage market. We argue that in China, working outside of home significantly improves the family income, and the increased material wealth enhances a young person's, especially a male's, marriage prospects.

2.1   China's *hukou* system and contemporary migration

China has strictly implemented a household registration (*hukou*) system since the early 1950s. Under this system, every person is registered where he or she is born. All administrative activities, such as land distribution, registration of a child in schools, and old-age pension, are based on registration status. Until the early 1990s, the *hukou* system was usually used to distribute food, cooking oil, and coupons. More importantly, moving across localities was very restrictive in both urban and rural areas and across these areas.

Since the 1990s, China government has eased restrictions on the *hukou* system. This easing, along with the substantial increases in foreign and domestic investments, has greatly stimulated China's internal migration (Song, Storesletten and Zilibotti, 2011). The resulting extraordinary surge of migration within China over the past two decades helps reallocate the labor force more efficiently and enhance its economic growth. For individual households, family migration improves family wealth significantly because it

increases family salary income.[4] Ge and Yang (2014), for instance, argue that rural-urban migration is one of the major driving forces behind the wage growth in China. Zhao (1999) documents a positive effect of migration activities on family income.

Temporary migration represents the majority of China's migration, which is of our particular interest. According to China's unique *hukou* system, migrants, in many cases the rural-urban ones, tend to be treated differently. For instance, they have limited access to local unemployment support, health insurance, retirement pensions, or the minimum living allowance scheme available to local *hukou* holders. In addition, migrant children are often denied access to public schools (Meng, 2012). As a result of this institutionalized discrimination attributed to the specific *hukou* system, most migrants do not see their long-run future in cities away from their hometowns. Instead, they leave their children and other family members behind and migrate, hoping to earn as much as possible before returning home (Meng, 2012; Zhao, 1999). The majority of the migration that does occur is indeed largely circular. According to the province-level sample-survey research, the majority of migrants spent less than nine months outside their hometowns (Zhao, 1999).

Because figuring out the determinants of China's large-scaled temporary migration is important, the literature documents many explanatory factors. Research shows various individual (age, gender, marriage status, and years of schooling), household (household size, net income, family wealth, and locality), and community (transportation and communication facilities) characteristics have significant effects (Zhao, 1999). The shortage of farmland and the abundance of household labor also play significant roles in explaining migration, particularly in rural areas. In addition, institutional constraints may contribute as well (Zhao, 1999). The large number of explaining factors documented in the existing literature, however, is still not able to fully explain migration in China.

One crucial factor the literature overlooks is the steadily increasing sex ratio in China, as we can observe the co-rising of the sex ratio and the number of migrant workers (Figure 1). The abnormal sex ratio induces family migration through the channel of the

---

4. We provide empirical evidence on this fact in the next section after introducing the data.

competitive marriage market for males, as we describe below.

2.2 Increasing sex ratio and marriage market

It is documented that the "missing girls" phenomenon can persist with economic development. Many more boys than girls have been born in China despite the country's fast economic growth, resulting in intensified competition in the marriage market for males. In other words, China is experiencing an increasingly severe relative surplus of men in the pre-marital age cohort, foreshadowing a sizable bride shortage. Specifically, the sex ratio for the 0–15 age cohort in China is 1.18 in 2010, implying that by 2025, males will outnumber females at ages 15–30 by about 13 million. The excess males, with no available matched brides, would find it difficult to get married. This issue is a severe social problem in the long run.

A strand of literature documents the possible underlying reasons for China's increasing sex ratio in past decades. Ebenstein (2010) argues that because of China's one-child policy, the chance that parents obtain a son naturally is lowered, and they then turn to sex selection. Almond, Li and Zhang (2013), however, argue that China's land reform increased sex ratios during the early 1980s. Chen, Li and Meng (2013) show empirical evidence that the improved local access to ultrasound technology has also resulted in an increase in the sex ratio at birth. Another strand of literature documents various socioeconomic consequences of the abnormal sex ratio. Several papers, for instance, explore the effects of a sex-ratio imbalance on the marriage and labor markets by gender (Sharygin, Ebenstein and Das Gupta, 2013).

We focus on one of the consequences of the increasing sex ratio, namely, that men have to compete with each other to find a wife. A related key assumption of our story is that a higher level of wealth improves a man's chances in the marriage market. In other words, we assume the material wealth improves competitiveness in the marriage market, especially for males. A strand of literature supports this assumption. For instance, Wei and Zhang (2011) show raising the savings rate, which increases the material wealth, is a channel for a man to improve his standing relative to his competitors in the

marriage market. Households are much less likely to have an unmarried adult son in rural areas if they have a relatively higher-quality house, whereas in urban areas, they are less likely to have an unmarried adult son if they are a homeowner as opposed to a renter (Wei and Zhang, 2011).[5] Studies also show that women exhibit a preference for men who grew up in affluent neighborhoods, and they continue to migrate to wealthier areas (Sharygin, Ebenstein and Das Gupta, 2013). More interestingly, research shows that women put greater weight on a partner's intelligence, which is an indicator for a promising high wealth level, whereas men respond more to physical attractiveness. A couple is less willing to match if her income exceeds his (Bertrand, Pan and Kamenica, 2015). These facts provide convincing evidence that when a man has a higher level of material wealth, he is more competitive in the marriage market, confirming the validity of our assumption.

3   Data and regression specification

This section describes the data and specifies the regression equation. We also emphasize the exogeneity of our key explanatory variable, the first child's gender, to address a clean causal relationship between child gender and family migration. Finally, we discuss migration and family earnings.

3.1   China Family Panel Studies survey

This paper mainly makes use of data from the 2010 baseline survey of *the China Family Panel Studies (CFPS)*, which is a nationally representative, biennially longitudinal survey of Chinese communities, families, and individuals. It is launched in 2010 by the *Institute of Social Science Survey* and designed to collect high-quality individual, family, and community level longitudinal data in contemporary China (Xie, 2012). In the nationwide *CFPS* baseline survey, a total of 14,798 households from 645 communities were

---

5. Generally speaking, house is likely to be the most important piece of household wealth. Those owning their houses are wealthier than those who rent their houses. Thus it confirms the conjecture that a higher level of wealth makes a man more attractive in the marriage market.

successfully interviewed, including 33,600 adults and 8,990 children, in 25 designated provinces.

The CFPS data meet our research needs very well. First, the stratified multi-stage sampling strategy employed ensures that in 2010, *the CFPS* sample represented 95 percent of the total population in China (Xie, 2012). Xu and Xie (2013) carefully compare the CFPS data with *the 2010 China Census* data in terms of some important socioeconomic and demographic variables. They find that characteristics such as distributions of age, sex, rural-urban stratification, educational attainment, and marital status in *the 2010 CFPS* closely resemble those in *the 2010 China Census*. This data-quality assessment confirms that it is reasonable to generalize any empirical finding from *the CFPS 2010* data to all Chinese families. Second, the CFPS provides detailed information on family migration as well as information on family fertility and household structure. Finally, the CFPS data also provide plenty of information on various other family activities, enabling us to examine the child-gender effect on family migration as well as other important family decisions.

In all our regression specifications, the unit of observation is the household. For different purposes, we extract four subsamples throughout the paper. The first one is the whole sample including all households in *the 2010 CFPS* family survey. Then, to analyze the child-gender effect on family migration, we construct the with-children sample by including all households in *the 2010 CFPS* family survey with a first-born child between 0 and 15 years old.[6] Next, to specifically examine the child-gender effect on parents' migration decision, the main sample is extracted by further restricting the with-children sample to households in which both parents are alive, the father is at most 50 years old and the mother is at most 45 years old, and at least one parent participated in *the 2010 CFPS* adult survey.[7] Finally, to closely analyze the economic activities of the migrant families, we restrict the with-children sample to families with at least one migrant worker

---

6. We focus on families with a first-born child between ages 0 and 15 to rule out cases in which the child might start to work. Because we are particularly interested in the first child's gender, we drop those families in which the first births are twins.

7. We impose such a restriction on the ages of fathers and mothers to minimize the probability of their inability to migrate, because of, for instance, health reasons or their retirement out of the labor force.

to obtain the migration sample.

3.2 Regression specification

The regression equation we estimate in this paper is as follows:

$$y_i = \alpha + \beta \times first\_child\_a\_son_i + X_i\gamma + \epsilon_i, \tag{1}$$

where y_i measures the migration status for household i. Six measures of y are examined in our main estimation. They are (1) the number of migrant workers in the family, (2) the migration probability of family members, (3) the migration probability of either parent, (4) the migration probability of the father, (5) the migration probability of the mother, and (6) the migration probability of both parents. These six variables together extensively measure the migration status of the family, exhausting all family migration possibilities. Examining these variables individually is also of interest because understanding how the magnitude of the child gender effect differs across household members might be particularly policy indicative. When the migration probability is considered, y is a dummy variable equal to 1 if the particular household member works outside his hometown. In this case, the regression equation is a linear probability model. The family migration information is based on the survey question in *the 2010 CFPS* asking whether anyone in the family left home last year to work outside. "Outside" means working in a place that is not where the person's household is registered or where the household's permanent address is. In rural areas, it usually refers to working in a different county. In urban areas, it usually refers to working in a different city.[8] In Section 5, which provides additional empirical evidence, the dependent variable y measures various other family economic activities.

The key explanatory variable throughout the paper, $first\_child\_a\_son$, is a dummy equal to 1 if the first-born child is a son. On top of this key explanatory variable, X refers

---

8. Work outside means people do not migrate permanently to work in a different county from where their families reside, as in the case of permanent rural migrant workers. In this paper, the family's temporary migration decision is of particular interest because the majority of the migration in China is circular, not permanent.

TABLE 1
Summary statistics — Family migration and family characteristics

	Mean	Std. Dev.	Min	Max	N
Dependent variables					
Number of migrants	0.433	0.819	0	8	4635
Any member migration (dummy)	0.288	0.453	0	1	4635
Any parent migration (dummy)	0.103	0.304	0	1	3771
Father migration (dummy)	0.091	0.287	0	1	3771
Mother migration (dummy)	0.027	0.162	0	1	3771
Both parents migration (dummy)	0.015	0.121	0	1	3771
Explanatory variables					
First child a son (dummy)	0.516	0.5	0	1	3771
Having at least one son (dummy)	0.662	0.473	0	1	3771
Share of sons	0.551	0.442	0	1	3771
Urban *hukou* (dummy)	0.431	0.495	0	1	3771
Father's age	36.254	5.647	19	50	3771
Mother's age	34.362	5.695	18	45	3771
Father's schooling years	7.94	4.365	0	22	3771
Mother's schooling years	6.748	4.709	0	22	3771

Notes: The first two variables are summarized for the with-children sample, whereas the rest of variables are summarized for the main sample. See notes below Tables 3 and 5 for the data source, sample descriptions, and variable definitions.

to other control variables that are possibly determinants of family migration, including a vector of parents (age, years of schooling, and occupation), household (urban or rural *hukou* and first-child's age), and regional (province fixed effects) characteristics. These characteristics are well documented in the literature as the important explaining factors for migration in China (Zhao, 1999). ϵ represents the measurement error in equation 1.[9] Table 1 reports the summary statistics of the dependent variables and the key explanatory variable for our benchmark regressions.

3.3 Randomness of the first child's gender

We present several pieces of evidence for the randomness of the first child being a son. To infer from equation 1 a causal relationship between child gender and family migration, the exogeneity or the randomness of the first child's gender is crucial.

First, the four most recent Chinese population censuses (1982, 1990, 2000, 2010) show the sex ratio for the fist-born children has been rather stable and falls in the biologically

---

9. All regression models in this paper are weighted by the weights given by *the 2010 CFPS*.

normal range, although the overall sex ratio at birth has been increasing. This fact is well documented in the literature (Dahl and Moretti, 2008; Ebenstein, 2010, 2011; Wei and Zhang, 2011). Thus, whether the first-born child in the Chinese family is a boy or a girl is most likely to be random, because the census data are regarded as nationally representative.

Second, because of the specifics of China's birth control policy, the Chinese families are less likely to practice gender selection on the first birth parity despite the traditional parental preference for sons. Specifically, for households in most rural areas in China, couples are allowed to give birth to another child if their first child is a girl. The parental preference for sons happens to be more severe in these areas, and this policy alleviates the motivation of parents in these areas to abort their first daughter (Scharping, 2013). Therefore, the first child's gender in China is arguably random.

More importantly, our summary statistics obtained from *the CFPS* data strongly suggest the randomness of the first child's gender. The sex ratio of the first-born children is 1.06 as reported in Table 1, which is well within the normal range. More interestingly, the standard deviation of the dummy for the first child being a son is exactly 0.5, again confirming the first-born child gender is most likely random.[10]

Finally, we find the first child's gender is not significantly affected by any control variable used in the empirical analysis throughout the paper. To rigorously test whether the first child's gender is random, we regress it on the full list of control variables. None of these variables have a significant effect on the first child's gender. This result again provides reasoning for our analysis to conduct the reduced-form examinations by directly estimating the impact of having a first-born son.

To sum up, although the overall sex ratio is severely unbalanced in China, we still view the gender of the first-born children as random. As such, using the first child's gender in regressions reveals a causal relationship without incurring the common econometric problems, such as reverse causality or omitted variable bias.[11]

---

10. If Z is a Bernoulli random variable with a mean of 0.5, the standard deviation of Z is equal to 0.5 ($\sqrt{0.5(1-0.5)}$).

11. China's household registration and birth control policies prevent households from migrating for fertility reasons (Li and Zhang, 2009). Thus, the reverse causality is even less likely to be at play.

3.4 Migration and family earnings

Before presenting the estimation results of equation 1, we empirically examine a crucial assumption for our competitive-earing-incentive hypothesis. That is, migrating parents are able to earn more money. Table 2 reports the correlation between family salary income and family migration.[12] The dependent variable is the yearly family salary income (in thousand *yuan*), and the explanatory variable of interest is a dummy equal to 1 if the corresponding household member is a migrant worker. Overall, we can infer from Table 2 that the family income is positively related to family members' migration status. The migration effect on income is considerably large in magnitude and statistically significant. In particular, the father working away from his hometown is associated with an increase in the family salary income by 8.3 thousand *yuan* on average. This amount equals a 30 percent increase to the mean of the family salary income, which is 27.4 thousand *yuan*. Interestingly, the mother's migration status is related to an even larger increase (46 percent or 12.7 thousand *yuan*) in the family salary income. Our findings are in line with Ge and Yang (2014) and Zhao (1999).

4   Child gender and family migration

In this section, we provide empirical evidence on the effect of child gender on family migration. We first estimate the significance and the magnitude of the child-gender effect as well as its differential impacts on different family members' migration status. Subsequently, we present several robustness checks. We provide additional estimates measuring both the direct and indirect effects of child gender on family migration. The indirect effect is through the change in fertility induced by the first child's gender. We also employ other measurements of child gender to check the robustness of the child-gender effect on family migration.

---

12. In this regression model, we are not formally inferring a causal relationship. Instead, we would like to confirm the conception that migration is associated with higher family income.

TABLE 2

Family migration and family salary income — OLS estimation

	Whole sample	With-children sample	Main sample			
	(1)	(2)	(3)	(4)	(5)	(6)
		Dependent variable: family salary income (000 yuan)				
Any member	12.732*** (1.657)	11.471*** (1.012)				
Any parent			8.592*** (1.614)			
Father				8.276*** (1.718)		
Mother					12.706*** (3.050)	
Both parents						17.506*** (4.469)
Observations	14,318	6,056	4,783	4,783	4,783	4,783
Other controls?	YES	YES	YES	YES	YES	YES

Notes: Standard errors are given in parentheses. The whole sample includes all households in *the 2010 CFPS* family survey. See notes below Table 3 for the data source, sample descriptions, and control variables. The dependent variable is the yearly family salary income (in thousand *yuan*). The explanatory variable of interest is a dummy equal to 1 if the corresponding household member(s) worked away from the hometown in the year before the survey.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

4.1 Benchmark results

We begin our empirical analysis by using equation 1 to estimate how child gender affects the family members', especially fathers', migration status.[13] All regressions control for a vector of parents (age, years of schooling, and occupation), household (urban or rural *hukou* and first-child's age), and regional (province fixed effects) characteristics.

Column (1) in Table 3 reports the child-gender effect on the overall number of migrants in the household based on the with-children sample. The explanatory variable of interest, $first\_child\_a\_son$, is a dummy equal to 1 if the first-born child in the family is a son. Its coefficient indicates families in which the first child is a son have 0.04 more migrant workers than families in which the first child is a daughter, on average.

To better understand the magnitude of the estimated marginal effect, throughout the

---

13. Logit models yield virtually identical results compared to OLS models given by equation 1.

TABLE 3

First-child gender and family migration — OLS estimation

	With-children sample		Main sample			
			Dependent variable: migration dummy			
	(1)	(2)	(3)	(4)	(5)	(6)
	Number of migrants	Any member	Any parent	Father	Mother	Both parents
First child a son	0.041*	0.031**	0.022**	0.020**	0.004	0.003
	(0.024)	(0.013)	(0.010)	(0.010)	(0.006)	(0.004)
Observations	4,635	4,635	3,771	3,771	3,771	3,771
R-squared	0.27	0.27	0.21	0.21	0.19	0.21
Other controls?	YES	YES	YES	YES	YES	YES
Daughter baseline	0.41	0.27	0.09	0.08	0.02	0.01
Son percent effect	10.03	11.31	23.65	25.20	18.09	22.71

Notes: Standard errors are given in parentheses. Data are from *the 2010 CFPS*. The with-children sample includes all households in *the 2010 CFPS* family survey with a first-born child between 0 and 15 years old (inclusive), whereas the main sample further restricts the with-children sample to households in which both parents are alive, the father is at most 50 years old and the mother is at most 45 years old, and at least one parent participated in *the 2010 CFPS* adult survey. In column (1), the dependent variable is the number of migrants in the family. In columns (2) to (6), the dependent variable is a dummy equal to 1 if the corresponding household member worked away from the hometown in the year before the survey. The explanatory variable of interest, $first\_child\_a\_son$, is a dummy equal to 1 if the first-born child in the family is a son. Other controls include a vector of parents (age, years of schooling, and occupation), household (urban or rural *hukou* and first-child's age), and regional (province fixed effects) characteristics. The daughter baseline is calculated as the average predicted value of the dependent variable of interest for families with a first-born daughter, using the estimated coefficients on the control variables. The son percent effect is the increase in the value of the dependent variable for a first-born-son family to a first-born-daughter family. It is the odds ratio minus 1.

   ***Significant at the 1 percent level.
  **Significant at the 5 percent level.
 *Significant at the 10 percent level.

paper, we report the "daughter baseline," which is calculated as the average predicted value of the outcome variable of interest for families with a first-born daughter, using the estimated coefficients on control variables.[14] In column (1), where the dependent variable is the number of migrants, the daughter baseline measures the average number of migrants in families with a first-born daughter. In columns (2) to (6), where the dependent variable is a dummy equal to 1 if the corresponding household member works away from hometown, the daughter baseline measures the migration probability of particular members in families with a first-born daughter. More importantly, we also report the "son

---

14. The daughter baseline equals $X\gamma$ in equation 1. These average predicted values turn out to be very close to the raw values in our sample..

percent effect," which is the increase in the value of the outcome variable of interest for families with a first-born son from the daughter baseline.[15] In column (1), the son percent effect indicates the number of migrants increases by 10 percent comparing a family in which the first child is a son to a family in which the first child is a daughter.

In column (2), the son percent effect indicates that for the with-children sample, the migration probability in the family increases by 11.3 percent when the first child is a son. The estimates in columns (3) to (6) are based on the main sample in which parents are in the labor force. In column (3), the son percent effect indicates the migration probability for either the father or mother increases by 23.7 percent when the first child is a son. This result mainly attributes to the migration of fathers, as can be seen from the son percent effect (25.2 percent) in column (4). We view this effect as considerably large. It implies that in 2010, approximately 1.7 million first-born daughters younger than 16 years old in China would have driven their fathers out of their hometowns had they been first-born sons instead.[16] Despite the child-gender effect on mothers' migration being small in magnitude and not statistically significant, the first child being a son still has a positive effect for mothers.

4.2   Robustness checks

4.2.1   Direct and indirect effects of child gender on family migration

We focus on the first child's gender because it has the cleanest causal interpretation, because whether the first child is a boy or a girl can arguably be viewed as random. Estimates based only on the gender of the first child provide the total effect on the relevant outcomes of having a first-born boy versus a first-born girl. As such, our causal estimates may capture several effects, including the direct effect of a first-born son on family migration and the indirect effect due to the subsequent differential fertility choice. The gender of the first child affects parental fertility decisions and then parental labor

---

15. The son percent effect equals $\beta/X\gamma$ in equation 1, which is simply the odds ratio minus 1.

16. China has a total of 417,722,698 households according to *the China 2010 Census*. Among these households, 42.3 percent have at least one child from 0 to 15 years old. Among these families with children, 28.8 percent contain temporary migrant workers and 7.6 percent contain temporarily migrating fathers (calculated by the authors from *the 2010 CFPS*).

supply (Fleisher and Rhodes, 1979).

Although isolating the direct effect of the first-born child's gender is not fully possible with our available data, to aid in the interpretation of our findings, we present additional estimates in Table 4 for family migration. In Panel A, we restrict the sample to families with only one child, and re-estimate equation 1 in the same way as in Table 3. This practice holds constant the family size, although we emphasize that this restricted sample consists of parents who endogenously chose to have only one child. For this restricted sample of one-child families, the probability of a migrating father for a first-born boy versus girl rises to 3 percentage points compared to 2 percentage points in column (4) of Table 3. The estimates for most of the other outcomes are similar to, or larger than, the analogous estimates in Table 3 in magnitude. The estimate for mothers' migration probability remains insignificant.

One way to benchmark our estimates is to compare the first-born-son effect to the family-size effect on family migration. In Panel B of Table 4, we include the number of children as an additional explanatory variable. Several points are worth mentioning. First, the coefficients of the key explanatory variable, $first\_child\_a\_son$, are arguably precisely estimated and do not appreciably change from Table 3 with the addition of the family-size variable. Second, the estimation results suggest that although additional children seem to increase the number of migrants in the family, they do not significantly increase parents' migration probability. Third, apart from the insignificance, the additional child effect is considerably smaller in magnitude than the first-born-son effect for various outcomes. Finally, the indirect fertility effect on family migration could be obtained from the additional child estimates by multiplying the number of reduced births resulting from a first-born son. The fertility effect is even smaller compared to the first-born-son effect, because the number of reduced births is far less than one.[17] To sum up, apart from the endogeneity of the family fertility decision, all these points discussed above suggest it is not necessary to include the number of children in various regressions examining the child-gender effect on family migration.

---

17. The first child being a son decreases the number of children in the family by about 0.228 (calculated using *the 2010 CFPS* data).

TABLE 4

First-child gender and family migration — OLS estimation
Direct and indirect effects

	With-children sample		Main sample			
			Dependent variable: migration dummy			
	(1)	(2)	(3)	(4)	(5)	(6)
	Number of migrants	Any member	Any parent	Father	Mother	Both parents
Panel A: Effects for families with only one child						
First child a son	0.026	0.032*	0.029**	0.030**	0.003	0.004
	(0.029)	(0.017)	(0.013)	(0.012)	(0.007)	(0.005)
Observations	3,058	3,058	2,473	2,473	2,473	2,473
Panel B: Effects controlling for number of children						
First child a son	0.046*	0.031**	0.021**	0.021**	0.005	0.004
	(0.024)	(0.013)	(0.011)	(0.010)	(0.006)	(0.004)
Number of children	0.034*	0.006	0.005	0.010	0.000	0.005
	(0.020)	(0.011)	(0.009)	(0.009)	(0.005)	(0.004)
Observations	4,635	4,635	3,771	3,771	3,771	3,771

Notes: Standard errors are given in parentheses. See notes below Table 3 for the data source, sample descriptions, variable definitions, and control variables. Panel A includes households with only one child. In Panel B, the total number of children is added as an additional explanatory variable.

 ***Significant at the 1 percent level.
 **Significant at the 5 percent level.
 *Significant at the 10 percent level.

4.2.2   Measurements of child gender

The estimation results presented so far show an impact of having a first-born son on the household migration decision-making. Although these results support the assertion that parents of sons are more likely to migrate, we have to consider other measurements of child gender to get a convincing conclusion. In the following, we exploit two different measures of child gender: (1) having at least one son, which is a dummy equal to 1 if the family has one or more sons, and (2) the share of sons, which is the ratio of the number of sons over the number of children. These two variables, however, are less likely to be exogenous, because the gender of children after the first-born child is no longer random, as fertility decisions are likely to be endogenous (Dahl and Moretti, 2008; Ebenstein, 2010).

TABLE 5

Child gender and family migration — OLS and IV estimation
Other measurements

	With-children sample				Main sample							
					Dependent variable: migration dummy							
	Number of migrants		Any member		Any parent		Father		Mother		Both parents	
	OLS (1)	IV (2)	OLS (3)	IV (4)	OLS (5)	IV (6)	OLS (7)	IV (8)	OLS (9)	IV (10)	OLS (11)	IV (12)
Panel A: Having at least one son												
Having at least one son	0.002	0.060*	0.029**	0.045**	0.020*	0.032**	0.023**	0.030**	0.001	0.007	0.004	0.004
	(0.026)	(0.033)	(0.014)	(0.018)	(0.012)	(0.014)	(0.011)	(0.014)	(0.006)	(0.008)	(0.005)	(0.006)
Observations	4,635	4,635	4,635	4,635	3,771	3,771	3,771	3,771	3,771	3,771	3,771	3,771
First stage (Dependent variable: having at least one son)												
First child a son		0.690***		0.690***		0.680***		0.680***		0.680***		0.680***
		(0.009)		(0.009)		(0.010)		(0.010)		(0.010)		(0.010)
Panel B: Share of sons												
Share of sons	0.011	0.053*	0.027*	0.039**	0.031***	0.028**	0.030***	0.026**	0.007	0.006	0.006	0.004
	(0.027)	(0.029)	(0.015)	(0.016)	(0.012)	(0.012)	(0.011)	(0.012)	(0.006)	(0.007)	(0.005)	(0.005)
Observations	4,635	4,635	4,635	4,635	3,771	3,771	3,771	3,771	3,771	3,771	3,771	3,771
First stage (Dependent variable: share of sons)												
First child a son		0.784***		0.784***		0.780***		0.780***		0.780***		0.780***
		(0.006)		(0.006)		(0.007)		(0.007)		(0.007)		(0.007)

Notes: Standard errors are given in parentheses. See notes below Table 3 for the data source, sample descriptions, variable definitions, and control variables. In Panel A, the explanatory variable of interest, having at least one son, is a dummy equal to 1 if the family has one or more sons. In Panel B, the explanatory variable of interest, share of sons, is defined as the ratio of the number of sons over the number of children in the family. In columns (1), (3), (5), (7), (9), and (11), coefficients of the explanatory variables of interest are obtained by the OLS estimation method. In columns (2), (4), (6), (8), (10), and (12), coefficients are obtained by the IV estimation method, the explanatory variables of interest being instrumented by the dummy variable, the first child being a son. The first-stage estimates are reported following the second-stage estimates.

  ***Significant at the 1 percent level.

  **Significant at the 5 percent level.

  *Significant at the 10 percent level.

20

The estimation results reported in Table 5 confirm the child-gender effect on family migration is robust. Although in column (1), the OLS estimates are imprecise compared with the benchmark case, all other OLS estimates turn out to be of similar magnitude to the benchmark estimates. Table 5 also reports the two-stage least square (TSLS) estimates. We use the gender of the first child as an IV for having at least one son and the share of sons. The coefficient in the first stage implies having a first-born son increases the probability of having at least one son by 0.68, which is in line with Li and Wu (2011). Moreover, the first-stage results show the p values of the estimates in all regressions are less than 0.01. According to Staiger and Stock (1997), our TSLS estimates are not subject to the weak instrumental variable problem.

The TSLS estimates in Table 5 further confirm the effect of the first child's gender on parental migration is exclusively through the change in child gender. For example, column (6) in Panel B, on one hand, shows the first child being a son enhances the share of sons by 0.78. On the other hand, the effect of the share of sons on the probability of either parent migrating to work is 0.028. So, the TSLS estimate shows the chain effect of the first child being a son on the parental migration probability is 0.022. This effect is precisely the reduced-form effect of the first child being a son on the parental migration probability reported in column (3) of Table 3.

5 The competitive-earning-incentive hypothesis versus four alternative hypotheses

As the empirical evidence presented in Section 4 reveals a clear pattern of child-gender difference on family migration, we advance our hypothesis that facing an increasing sex ratio for youth, parents' competitive earning incentive for their sons is a significant factor driving them to migrate.

To be more specific, forward-looking parents, and perhaps fathers in particular, would like to earn more money for their sons to improve their competitiveness in the future marriage market. This motivation could persuade some parents near the margin to work

away from their hometowns if they have a first-born son, because doing so improves the family's income. This hypothesis is our preferred explanation for the child-gender effect on family migration. We formally coin it as the competitive-earning-incentive hypothesis. However, several alternative explanations are possible for why the pattern of child-gender difference in family migration may arise. In this section, we present four possible alternative explanations. For each explanation, we derive different implications on various family decisions. We then attempt to rule out the four competing hypotheses by providing additional empirical evidence.

5.1 Alternative interpretations

Several alternative explanations are consistent with the findings of the observed pattern of child gender-difference in family migration as well. The first is the possibility that parents would like to spend more money on their sons, as well-known parental preferences for sons exist in China.[18] This explanation is purely based on parents' tastes: for whatever reason, parents, especially fathers, want to earn more and spend more money for their sons. We call this the "love of sons" hypothesis. Gender-biased parents might decide to migrate to earn more money because they are willing to spend more on their sons.

A closely related explanation is that parents are unbiased regarding child gender but that the monetary cost of raising sons, compared with raising daughters, is relatively higher in China.[19] We call this the "expensive sons" hypothesis.[20] According to this hypothesis, parents care about the well-being of their children equally. Considering whether to work outside to earn money, the parents take into account the asymmetric monetary needs of rearing sons versus daughters. This hypothesis is also consistent with the findings of the child-gender effect on family migration: parents decide to migrate because they realize raising sons costs more money than raising daughters. It is similar in spirit

---

18. The sex ratio itself serves as a valid test for sex preferences (Leung, 1991).

19. Ebenstein (2011) finds that a couple's first son is worth 1.42 years of income more than a first daughter.

20. This expensive-sons hypothesis differs from the preferred competitive-earning-incentive hypothesis, although they both imply sons cost more money. The former hypothesis refers to the rearing cost, whereas the latter refers to the marriage cost.

to the love-of-sons hypothesis, although it states parents are "forced" to work outside the their hometowns instead of doing so voluntarily.[21]

A third possible explanation is that working outside is much more difficult for parents when they have a first-born daughter compared to a first-born son. Parents who realize this fact are more willing to work outside when they have sons. We call this the "ease of migrating" hypothesis. This explanation posits that when the family has a first-born son, it incurs a lower monetary or time cost for parents leaving home to work outside. For instance, in rural areas of China, sons are expected to help their parents in the heavy farm-work, because males are regarded as being more helpful and having higher productivity in farming, making their fathers' migrating to urban areas relatively easier. As such, the first child being a son makes it easier for parents working outside, explaining to some extent the child-gender effect on family migration.

Alternatively, parents, especially fathers, might like to, for some exogenous reasons, live separately when their first-born child is a son. For instance, the family with a first-born son might be more likely to have considerable housing difficulties, driving parents out of their homes. These housing difficulties may be due to the fact that the presence of sons significantly increases the probability of grandparents living in the household. In this story, parents leave the hometown not just for earning reasons. We call this the "love of separation" hypothesis. Because the first child being a son makes it more preferable for parents to live separatively, it would increase the probability of parents working away from home. This hypothesis is also consistent with the child-gender effect on family migration.

To sum up, each of these four hypotheses could potentially explain the evidence in Section 4 that fathers with a first-born son are more likely to work outside the hometown than fathers with a first-born daughter. We have provided the intuition on why this is so. We are not arguing that the alternative explanations presented above and our preferred competitive-earning-incentive hypothesis are necessarily mutually exclusive. More than

---

21. See Davies and Zhang (1995) for a detailed discussion about two types of son preferences. One is directly derived from parental taste, and the other is indirectly derived from the household budget constraint.

one explanation could very well be at play. Furthermore, the explanations that we have discussed above are fairly general. Several variants of these stories could readily fit within our framework.

5.2 Heterogeneity of the child-gender effect and the competitive-earning-incentive hypothesis

This section examines the child-gender effect on family migration across different groups. If the competitive earning incentive is the main reason, parents are more likely to migrate if their sons have more difficulty finding a wife. For example, Sharygin, Ebenstein and Das Gupta (2013) find that households in rural and poor regions face more pressure to find wives for their sons. Furthermore, men in higher sex-ratio regions also face a more competitive marriage market. Our results show the heterogeneity is closely related to the relative difficulty a man faces in finding a wife in different areas or from different groups. This pattern supports our competitive-earning-incentive hypothesis.

5.2.1 Urban versus rural households

In Table 6, Panel A1 (A2) includes households in the urban (rural) area according to the *China Census Bureau* records. Control variables included are comparable to the benchmark regression, except here we exclude the dummy control variable indicating whether the family is registered as an urban household to avoid collinearity problem. Moreover, for rural households in Panel A2, we include in regression columns (7) and (8) the land ownership, which is a dummy equal to 1 if the household possesses any type of farmland, and the family size, which is the number of household members, as additional control variables.[22] The estimation results in Panel A of Table 6 show child gender does not play a significant role in urban areas. In rural areas, it has great importance, probably because of the relatively larger difficulty rural males face in finding a wife. This result is

---

22. The shortage of farmland and the abundance of household labor are among the most important determinants of rural labor migration because they reduce the relative marginal income from labor in farming (Zhao, 1999). The land ownership proxies the accessibility of the farmland for the household, and the family size proxies the availability of the household labor.

in line with Sharygin, Ebenstein and Das Gupta (2013). They find the proportion of men in China that have never been married is especially high among poor men in low-income regions, such as rural areas, that are least able to provide social protection programs. Particularly, the son percent effect for the father's migration is 24.4 percent, which is comparable to the benchmark case.[23]

5.2.2   High- versus low-education households

The proxy variable closest to men's economic status, which is a general indicator of their relative competitiveness in the marriage market in most societies, is educational status (Sharygin, Ebenstein and Das Gupta, 2013). In Table 6, Panel B1 (B2) includes households in which either (neither) parent has a college degree, representing the high- (low-)education subsample. The estimation results reported in Panel B of Table 6 show that whereas the daughter baseline is relatively larger for low-education households than high-education households (and than the benchmark case), the son percent effect differs qualitatively. For the low-education groups, the son percent effect (for the father's migration in column (10), 29.9 percent) is larger than the benchmark case. This finding is consistent with the fact that most single men come from low-education households. For the high-education groups, however, the son percent effect turns out to be surprisingly negative.[24]

5.2.3   Sex ratio and family migration

We finally consider whether the local sex ratio for youth affects the family migration pattern.

---

23. The daughter baseline is obviously larger in rural areas than the benchmark estimates, implying the overall migration probability is much higher in rural areas in China. This observation helps us understand why the son percent effect for rural households is not much larger than the benchmark contrary to our expectation.

24. One highly possible explanation is that family members, especially fathers, in high-education households, facing less financial pressure, attach much more attention to their sons compared with daughters. This gender difference may be due to the parental preferences for sons or the fact that boys are more prone to educational or health problems without parents' accompanies (Dahl and Moretti, 2008). This gender difference may also due to the comparative advantage of a parent in raising a child of the same gender: fathers may be more productive in molding the character of their sons compared with their daughters (Lundberg and Rose, 2002).

TABLE 6

First-child gender and family migration — OLS estimation
Heterogeneity across households

	With-children sample		Main sample				With-children sample		Main sample			
	Dependent variable: migration dummy						Dependent variable: migration dummy					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Number of migrants	Any member	Any parent	Father	Mother	Both parents	Number of migrants	Family member	Any parent	Father	Mother	Both parents
	Panel A1: Urban households						Panel A2: Rural households					
First child a son	0.021	0.026	0.015	0.013	0.004	0.002	0.099***	0.044**	0.025	0.028*	0.003	0.006
	(0.032)	(0.019)	(0.013)	(0.012)	(0.007)	(0.005)	(0.031)	(0.018)	(0.016)	(0.015)	(0.008)	(0.006)
Observations	1,773	1,773	1,520	1,520	1,520	1,520	2,862	2,862	2,251	2,251	2,251	2,251
Daughter baseline	0.20	0.14	0.04	0.03	0.01	0.00	0.53	0.36	0.13	0.12	0.04	0.02
Son percent effect	10.43	18.65	39.67	38.73	45.60	44.36	18.65	12.42	18.89	24.43	8.88	33.16
	Panel B1: High-education households						Panel B2: Low-education households					
First child a son	-0.036	-0.012	-0.032**	-0.032**	0.000	—	0.048*	0.035**	0.028**	0.027**	0.006	0.005
	(0.046)	(0.031)	(0.016)	(0.016)	(0.003)	—	(0.027)	(0.015)	(0.012)	(0.011)	(0.006)	(0.005)
Observations	590	590	525	525	525	—	4,045	4,045	3,246	3,246	3,246	3,246
Daughter baseline	0.14	0.10	0.04	0.04	0.00	—	0.46	0.30	0.10	0.09	0.03	0.01
Son percent effect	-25.43	-11.83	-85.43	-88.07	25.00	—	10.44	11.54	27.23	29.92	20.61	30.73
	Panel C1: High-sex-ratio areas						Panel C2: Low-sex-ratio areas					
First child a son	0.088*	0.048**	0.043**	0.037**	0.025**	0.019**	0.018	0.024	0.008	0.008	-0.004	-0.004
	(0.049)	(0.024)	(0.020)	(0.018)	(0.012)	(0.009)	(0.027)	(0.017)	(0.013)	(0.013)	(0.006)	(0.005)
Observations	1,794	1,794	1,375	1,375	1,375	1,375	2,841	2,841	2,396	2,396	2,396	2,396
Daughter baseline	0.54	0.33	0.09	0.07	0.02	0.01	0.34	0.24	0.09	0.09	0.02	0.01
Son percent effect	16.35	14.60	49.70	51.70	108.72	234.76	5.33	9.90	8.25	9.58	-18.66	-27.21

Notes: Standard errors are given in parentheses. See notes below Table 3 for the data source, sample descriptions, variable definitions, control variables, and meanings of the daughter baseline and the son percent effect. Panel A1 (A2) includes households in the urban (rural) area according to the *China Census Bureau* records. Panel B1 (B2) includes households in which either (neither) parent has a college degree. Panel C1 (C2) includes households in high- (low-)sex-ratio areas. In Panel B1, results are not reported for the regression model (6), because of the small number of observations in the corresponding subsample. ***Significant at the 1 percent level; **Significant at the 5 percent level; *Significant at the 10 percent level.

The competitive-earning-incentive hypothesis implies that when the local sex ratio for youth is high, the child-gender difference in family migration would be magnified.[25] In other words, the first-born child being a son has a significantly larger impact on family migration in high-sex-ratio areas.[26] If other stories are at play, we would be less likely to observe this sex-ratio effect. Panel C in Table 6 reports the heterogeneity across regions with different sex-ratio levels for youth. Columns (1)-(6) include households in China's top nine provinces (Jiangxi, Henan, Anhui, Hubei, Guangdong, Hunan, Fujian, Jiangsu, and Guangxi) with the highest sex ratio for youth (see notes below Figure 2 for details).[27] Columns (7)-(12) include all households in the remaining 16 provinces.[28] We find the daughter baseline does not differ much in these two subsamples, implying the local sex ratio for youth does not affect the migration behavior of parents with daughters. The son percent effect, however, is much higher when the local sex ratio is higher. In particular, in high-sex-ratio areas, the first child being a son increases the father's migration probability by 52 percent (compared to 25 percent in the benchmark estimation), whereas in low-sex-ratio areas, the effect is only 10 percent. More importantly, the coefficients for low-sex-ratio areas are not statistically significant, possibly because in the low-sex-ratio regions, the marriage market competitiveness for males is not intensified. The estimation results on the heterogeneity across regions are consistent with the competitive-earning-incentive hypothesis.

5.3 Additional evidence

To make further progress in differentiating between the competitive-earning-incentive hypothesis and the four alternative hypotheses, we first illustrate how child-gender composition is predicted to affect various other family decisions. Then we present empirical

---

25. Here we assume exogenous sex ratios to parents when they make the migration decision. This assumption is also invoked in Wei and Zhang (2011).

26. For the sex ratio to affect household behavior, parents do not have to know local sex-ratio statistics (Wei and Zhang, 2011). In other words, even without the knowledge of local sex-ratio statistics, parents with a son may make migration decisions that reflect the local sex ratio.

27. Note this subsample includes both well-developed regions, such as Guangdong and Jiangsu, and less-developed regions, such as Guangxi and Henan. The heterogeneity is thus not due to the wealth-level difference.

28. In total, *the 2010 CFPS* data include 25 provinces.

evidence on each of these family decisions.[29] All results in this subsection favor our competitive-earning-incentive hypothesis.

5.3.1 Migration remittance

We first examine the amount and the main usage of the money sent back by family members who worked outside their hometowns. If our preferred competitive-earning-incentive hypothesis is at play, parents are expected to bring back more remittance and spend more on children's marriages and buying houses and other durable goods when they have a first-born son, attempting to increase their son's competitiveness in the marriage market. On the contrary, parents would spend more on the daily living expense or children's schooling when they have a first-born daughter, because they do not have to worry too much about her marriage.[30] However, if the love-of-sons or the expensive-sons hypothesis is at play, we would see parents spending more of the remittance for sons for all usages, possibly including daily expenses and children's schooling. Further, if the love-of-separation or the ease-of-migration hypothesis is at play, we would neither observe migrants bringing back more remittance when the first-born child is a son nor observe the remittance-usage difference.

As reported in column (1) in Table 7, the first child being a son increases the remittance sent back by the migrants by 1.3 thousand *yuan* per year. We then put the first child's gender into the household expenditure equation as in Pollak and Wales (1981). We find that when the first-born child is a son, the family spends the remittance mainly on children's marriage, house construction, and buying expensive durable goods.

---

29. Couples are unlikely to consider the migration arrangements and other family decisions in isolation. In all likelihood, they simultaneously determine the above arrangements. However, to keep things relatively simple, we examine the outcome one by one to differentiate between the preferred competitive-earning-incentive and other alternative hypotheses.

30. Such parents' behavior may be motivated by altruism or guilt. See Li, Rosenzweig and Zhang (2010) for more details.

TABLE 7

First-child gender and migration remittance — OLS estimation

Amount and the main usage

	(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
			Dependent variable: remittance usage dummy						
	Remittance sent back (*000 yuan*)		Children's marriage	Buy expensive durable goods	House construction	Pay debt	Children's schooling	See doctors	Daily living expenses
First child a son	1.331**		0.011*	0.016*	0.044**	0.029*	-0.037*	-0.009	-0.071**
	(0.666)		(0.007)	(0.010)	(0.018)	(0.017)	(0.022)	(0.017)	(0.034)
Observations	1,374		1,154	1,154	1,154	1,436	1,356	1,154	1,154
Other controls?	YES		YES	YES	YES	YES	YES	YES	YES
Daughter baseline	7.51		0.00	0.01	0.05	0.10	0.21	0.06	0.49
Son percent effect	17.74		312.32	184.35	95.31	27.91	-18.13	-14.13	-14.61

Notes: Standard errors are given in parentheses. See notes below Table 3 for the data source, sample descriptions, variable definitions, control variables, and meanings of the daughter baseline and the son percent effect. The estimates are based on the migration sample, which further restricts the with-children sample to families with at least one migrant worker. In column (1), the dependent variable is the amount of remittance sent back by the migrants (in thousand *yuan*). In columns (2) to (8), the dependent variable is a dummy equal to 1 if the remittance sent back is mainly expended on the corresponding usage.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

This finding is consistent with the competitive-earning-incentive hypothesis. Parents building a house and buying other expensive durable goods such as a car could be motivated by a desire to help their sons improve their marriage prospects. In China, most families believe that having a house is essential for their son's prospects in successfully finding a wife (Wei and Zhang, 2011). Also, a groom's family is often responsible for paying his bride's family a onetime transfer that compensates the latter for rearing their daughter (Zhang and Chan, 1999). Finally, the groom's family bears most of the financial cost of holding a wedding ceremony. When the first-born child is a girl, however, the family tends to spend more remittance on children's schooling and daily living expenses.[31] As such, the competing hypotheses mentioned above are less likely to be at play.

5.3.2 Family expenditure pattern

We further examine the effects of child gender on the overall family expenditure. The competitive-earning-incentive hypothesis implies the children's education and the consumption expenditure are lower when the first-born child is a son, whereas the total mortgage amount is higher.

Table 8 shows first-child being a son increases the total mortgage amount in the family by 1.5 thousand *yuan*, whereas it decreases the yearly education and consumption expenditure by 0.5 and 2 thousand *yuan*, respectively. This finding indicate that parents trade off between spending on children's marriage and education outcomes, without a clear preference for either gender. More importantly, the family expenditure pattern on children's education sheds some light on the moderate contribution of the one-child policy in China to its human capital development.[32] Although this policy helps reduce the number of children in a family, it lowers the sex ratio and as a result induces parents to spend more money on children's marriage, indirectly hampering the human capital

---

31. A strand of literature seeks to account for gender differentials in the levels of and returns to schooling observed in developing countries in recent decades (Pitt, Rosenzweig and Hassan, 2012; Rosenzweig and Zhang, 2013). In China in particular, enrollment rates of girls in secondary and tertiary schools have been higher than those of boys since 2005 (Pitt, Rosenzweig and Hassan, 2012).

32. Rosenzweig and Zhang (2009) document that despite the evident significant trade-off between number of children and child quality in China, the contribution of the one-child policy to the development of Chinese human capital was modest at best.

TABLE 8

First-child gender and family expenditure pattern — OLS estimation

	(1)	(2)	(3)
	Total mortgage _(000 yuan)_	Education expenditure _(000 yuan)_	Consumption expenditure _(000 yuan)_
First child a son	1.512*	-0.450**	-2.000***
	(0.900)	(0.203)	(0.663)
Observations	4,553	5,633	5,325
Other controls?	YES	YES	YES
Daughter baseline	4.29	3.13	27.26
Son percent effect	35.28	-14.35	-7.34

Notes: Standard errors are given in parentheses. The estimates are based on the main sample. See notes below Table 3 for the data source, sample descriptions, variable definitions, control variables, and meanings of the daughter baseline and the son percent effect. In column (1), the dependent variable is the family's monthly expenditure (in thousand _yuan_) on education (and entertainment). In column (2), the dependent variable is the household residents' consumption expenditure (in thousand _yuan_) in total. In column (3), the dependent variable is the total amount of mortgage (in thousand _yuan_) for all houses owned by the family.

  ***Significant at the 1 percent level.
  **Significant at the 5 percent level.
  *Significant at the 10 percent level.

development, in particular, for sons.

5.3.3 Household helpers

The next consideration to differentiate between the hypotheses is whether extra helpers are hired because of the absence of family members who work away from their hometowns, and if yes, the associated expense. Although the preferred competitive-earning-incentive hypothesis does not offer a clear prediction, the ease-of-migrating hypothesis certainly predicts the migrant families with a first-born son are less likely to hire extra helpers and consequently incurs a lower associated cost.

  The first two columns in Panel A of Table 9 show that for migrant families, contrary to the ease-of-migrating hypothesis, the possibility of hiring extra workers and associated costs are both higher when the first-born child is a son. In other words, despite the higher migration costs, parents of sons are still more willing to migrate. As such, this evidence rules out the ease-of-migrating hypothesis. If the underlying reason for the higher migration probability for parents with sons facing higher migration costs is due to

TABLE 9

First-child gender, household helpers, and housing condition — OLS estimation

	(1)	(2)	(3)	(4)
	Panel A: Household helpers			
	Migration sample		With-children sample	
	Whether hire extra workers *(dummy)*	Cost of hiring extra workers *(000 yuan)*	Whether hire workers *(dummy)*	Cost of hiring workers *(000 yuan)*
First child a son	0.023**	0.050*	-0.008**	-0.004*
	(0.011)	(0.026)	(0.004)	(0.002)
Observations	1,894	1,839	4,654	4,655
Other controls?	YES	YES	YES	YES
Daughter baseline	0.06	0.06	0.02	0.01
Son percent effect	40.48	88.00	-35.80	-65.09
	Panel B: Housing conditions			
	With-children sample			
	Own current house *(dummy)*	Own a second house *(dummy)*	Building area *(sq. meter)*	How much rent *(000 yuan)*
First child a son	0.022**	0.019*	3.732*	0.063**
	(0.010)	(0.011)	(2.209)	(0.026)
Observations	4,655	4,438	5,464	4,022
Other controls?	YES	YES	YES	YES
Daughter baseline	0.85	0.16	124.31	0.47
Son percent effect	2.63	12.03	3.00	13.28

Notes: Standard errors are given in parentheses. See notes below Table 3 and Table 7 for the data source, sample descriptions, variable definitions, control variables, and meanings of the daughter baseline and the son percent effect. Panel A measures the child gender effect on family's hiring of household helpers. In column (1), the dependent variable is a dummy equal to 1 if the family hires extra workers because of some family member working away from home. In column (2), the dependent variable is the associated cost (in thousand *yuan*) of hiring extra workers (it equals 0 if the family does not hire extra workers). In column (3), the dependent variable is a dummy equal to 1 if the family hires a housekeeper for daily life. In column (4), the dependent variable is the monthly cost (in thousand *yuan*) of hiring housekeepers (it equals 0 if the family does not hire any housekeeper), including both wage paid and other expenses paid. Panel B measures the child-gender effect on families' housing conditions. In column (1), the dependent variable is a dummy equal to 1 if the property right of the current house is solely owned by the family. In column (2), the dependent variable is a dummy equal to 1 if the family owns a second house elsewhere. In column (3), the dependent variable is the building area of the family's current house (in square meters). In column (4), the dependent variable is the amount of monthly rent (in thousand *yuan*) the family could get had the house been rented.

  ***Significant at the 1 percent level.
  **Significant at the 5 percent level.
  *Significant at the 10 percent level.

marriage market pressure, this evidence is in line with the competitive-earning-incentive hypothesis. We notice that the higher probability of hiring workers in the family with a first-born son is not unrelated to family migration. This fact can be revealed from the last two columns in Panel A of Table 9, which indicate that, in general, families with a first-born son are less likely to hire a housekeeper. The fact that families with a first-born son are less likely to hire a housekeeper is itself in line with the competitive-earning-incentive hypothesis: with much more financial pressure on children's marriage, parents of sons try to lower their daily living expenses. Thus, this evidence again supports the competitive-earning-incentive hypothesis.

5.3.4   Housing

We then take a closer look at the family's housing condition. In China, having better housing conditions is considered essential for a man's prospects in successfully finding a wife (Wei and Zhang, 2011). If the competitive-earning-incentive hypothesis is at play, parents tend to prepare their sons for marriage by getting a better and larger house. Specifically, because males are facing an considerably competitive marriage market, their parents tend to migrate in an attempt to earn more money and spend more on improving their housing conditions. On the other hand, the expensive-sons hypothesis may imply worse housing conditions for families with first-born sons, because of the financial pressure or the budget constraint resulting from the higher rearing cost of sons. Also, the love-of-separation hypothesis is expected to imply families with first-born sons have worse housing conditions, possibly because of the higher co-residence probability with grandparents.

Column (1) in Panel B of Table 9 shows the property right of the current house is more likely solely owned by the family when the first-born child is a son. In China, families with sons are much less likely to have an unmarried adult son at home if they are a homeowner as opposed to a renter (Wei and Zhang, 2011). Because the higher probability of owning a house reflects that parents with a son facing more pressure from the marriage market attempt to prepare their son for marriage, this information supports our competitive-

TABLE 10

First-child gender and the family entrepreneurship — OLS estimation

	With-children sample		Main sample			
			Dependent variable: running business dummy			
	(1)	(2)	(3)	(4)	(5)	(6)
	Number of companies	Any member	Any parent	Father	Mother	Both parents
First child a son	0.014	0.015*	0.018**	0.014**	0.002	-0.002
	(0.010)	(0.009)	(0.007)	(0.006)	(0.005)	(0.003)
Observations	4,638	4,638	3,777	3,777	3,777	3,777
Other controls?	YES	YES	YES	YES	YES	YES
Daughter baseline	0.10	0.10	0.03	0.02	0.02	0.01
Son percent effect	14.64	15.99	53.94	57.23	11.85	-21.76

Notes: Standard errors are given in parentheses. See notes below Table 3 for the data source, sample descriptions, variable definitions, control variables, and meanings of the daughter baseline and the son percent effect. In column (1), the dependent variable is the number of enterprises run by the family. In columns (2) to (6), the dependent variable is a dummy equal to 1 if the corresponding household member runs any business.
        ***Significant at the 1 percent level.
        **Significant at the 5 percent level.
        *Significant at the 10 percent level.

earning-incentive hypothesis. More importantly, the first child being a son significantly increases the probability of the family owning a second house by 12 percent. Finally, the housing conditions measured by building area or monthly rent are much better when the first-born child is a son, again supporting the competitive-earning-incentive hypothesis.

5.3.5 Family entrepreneurship

We finally consider the family entrepreneurship choice. The preferred competitive-earning-incentive hypothesis predicts the likelihood of the family running its own business is higher when the first-born child is a son, because it is widely believed that starting one's own business earns more money for the family, although it is risky compared with other choices. If parents of sons turn out to prefer to work outside their hometowns not because of financial reasons (love-of-separation or ease-of-migrating hypothesis), we are less likely to find the above-mentioned pattern of child-gender difference in family entrepreneurship.

Table 10 reports the child-gender effect on family entrepreneurship. When the first-born child in the family is a son, family members, especially fathers, are more likely

to run their own business.[33] This pattern is similar to Table 3. Because migration and entrepreneurship are both regarded as attempts to make more money, this finding supports the competitive-earning-incentive hypothesis.

6   Conclusions

This paper shows child gender affects family migration in China by giving ample empirical evidence. We find that fathers of first-born boys are more likely to work outside their hometowns in an attempt to earn more money to equip their sons in the increasingly more competitive marriage market. A first-born son is 25.2 percent more likely to drive his father to work outside the hometown compared to a first-born daughter. This child-gender effect influences a significant number of households in China. We estimate that in 2010, roughly 1.7 million first-born daughters younger than 16 years old would have driven their fathers away from their hometowns if they had been boys. In this sense, we document a new significant factor that accounts for the unprecedented large-scale migration in China. Because migration significantly contributes to China's economic development (Song, Storesletten and Zilibotti, 2011), our findings call for government policies that are specially made to encourage girls' parents to migrate, possibly by providing them more monetary benefits.

We next provide additional empirical evidence to differentiate between the preferred competitive-earning-incentive hypothesis and alternative hypotheses. We find that having a first-born son has a larger and more significant impact in rural areas and for low-education households. More importantly, when the local sex ratio is higher, the first-born child being a son has a larger impact on family migration. We further find that when the first-born child is a son, migrants in the family tend to send back more remittance. Additionally, the family tends to spend the remittance as well as other available resources on

---

33. Wei and Zhang (2011) find that in response to a rise in the sex ratio, parents with a son may be more prepared to take on relatively dangerous or unpleasant but higher-return activities, and work more days, stimulating economic growth by inducing more entrepreneurship and hard work. Lundberg and Rose (2002) find that men's labor supply and wage rates increase more in response to the birth of sons than to the birth of daughters.

children's marriage and buying houses and other durable goods. The housing conditions are also better in general. Taken individually, each piece of evidence on family economic activities is consistent with the competitive-earning-incentive hypothesis. Taken together, our findings suggest other stories are less likely to be at play.

Regardless of how one interprets our findings of the child-gender effect on family migration, the significant increase in the migration probability for parents whose first-born child is a boy is interesting in its own right. Resulting from the higher probability of parents working outside their hometowns, the first-born boys and their siblings are more likely to live in families in which either or both parents are absent. Zhang et al. (2014) find that over 61 million children are left-behind by parents migrating for work in rural China. The human capital development of these left-behind children has become a major policy issue in China.

The findings of the child-gender difference in parental behavior in this paper seem to be paradoxical. While parents attempt to accumulate more material wealth for their sons in response to the intensified marriage market competitiveness, they seem to put less attention on their sons' human capital development as well as other long-run outcomes. In particular, we show that parents may invest less in the education of their sons compared to daughters. This resource-allocation pattern is well in line with the current strand of literature documenting differential investments in schooling across men and women in developing counties (Pitt, Rosenzweig and Hassan, 2012; Rosenzweig and Zhang, 2013). As such, in spite of the original motive for parental migration being to enhance boys' future marriage prospects, the real consequence may be the opposite: females with more education become even less willing to marry males with less education.

Taking into account the negative effect on boys and the possible spillover effect on siblings of boys regarding to human capital development, we further conclude that social costs of China's increasing sex ratio are highly likely to have been understated. A more detailed investigation of these possibilities is left for future research.

References

Almond, Douglas, Hongbin Li, and Shuang Zhang, "Land reform and sex selection in China," Technical Report, NBER Working Paper No. 19153. 2013.

Bertrand, Marianne, Jessica Pan, and Emir Kamenica, "Gender identity and relative income within households," *Quarterly Journal of Economics (forthcoming)*, 2015.

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

Dahl, Gordon B and Enrico Moretti, "The demand for sons," *Review of Economic Studies*, 2008, *75* (4), 1085–1120.

Davies, James B and Junsen Zhang, "Gender bias, investments in children, and bequests," *International Economic Review*, 1995, pp. 795–818.

Ebenstein, Avraham, "The "missing girls" of China and the unintended consequences of the one child policy," *Journal of Human Resources*, 2010, *45* (1), 87–115.

— , "Estimating a dynamic model of sex selection in China," *Demography*, 2011, *48* (2), 783–811.

Fleisher, Belton and George Rhodes, "Fertility, women's wage rates, and labor supply," *American Economic Review*, 1979, pp. 14–24.

— , **Haizheng Li, and Min Qiang Zhao**, "Human capital, economic growth, and regional inequality in China," *Journal of Development Economics*, 2010, *92* (2), 215–231.

Ge, Suqin and Dennis Tao Yang, "Changes in China's wage structure," *Journal of the European Economic Association*, 2014, *12* (2), 300–336.

Leung, Siu Fai, "A stochastic dynamic analysis of parental sex preferences and fertility," *Quarterly Journal of Economics*, 1991, pp. 1063–1088.

Li, Hongbin and Junsen Zhang, "Testing the external effect of household behavior: The case of the demand for children," *Journal of Human Resources*, 2009, *44* (4), 890–915.

— , **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*, 2010, *118* (1), 1–38.

Li, Lixing and Xiaoyu Wu, "Gender of children, bargaining power, and intrahousehold resource allocation in China," *Journal of Human Resources*, 2011, *46* (2), 295–316.

Lundberg, Shelly and Elaina Rose, "The effects of sons and daughters on men's labor supply and wages," *Review of Economics and Statistics*, 2002, *84* (2), 251–268.

Meng, Xin, "Labor market outcomes and reforms in China," *Journal of Economic Perspectives*, 2012, pp. 75–101.

Oswald, Andrew J and Nattavudh Powdthavee, "Daughters and left-wing voting," *Review of Economics and Statistics*, 2010, *92* (2), 213–227.

Pitt, Mark M, Mark R Rosenzweig, and Nazmul Hassan, "Human capital investment and the gender division of labor in a brawn-based economy," *American economic review*, 2012, *102* (7), 3531.

Pollak, Robert A and Terence J Wales, "Demographic variables in demand analysis," *Econometrica*, 1981, pp. 1533–1551.

Powdthavee, Nattavudh and James Vernoit, "Parental unemployment and children's happiness: A longitudinal study of young people's well-being in unemployed households," *Labour economics*, 2013, *24*, 253–263.

Rosenzweig, Mark R and Junsen Zhang, "Do population control policies induce more human capital investment? twins, birth weight and china's one-child? policy," *Review of Economic Studies*, 2009, *76* (3), 1149–1174.

＿ **and** ＿ , "Economic growth, comparative advantage, and gender differences in schooling outcomes: Evidence from the birthweight differences of Chinese twins," *Journal of Development Economics*, 2013, *104*, 245–260.

Scharping, Thomas, *Birth Control in China 1949-2000: Population Policy and Demographic Development*, Routledge, 2013.

Sharygin, Ethan, Avraham Ebenstein, and Monica Das Gupta, "Implications of China's future bride shortage for the geographical distribution and social protection needs of never-married men," *Population Studies*, 2013, *67* (1), 39–59.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti, "Growing like China," *American Economic Review*, 2011, *101* (1), 196–233.

Staiger, Douglas O and James H Stock, "Instrumental variables regression with weak instruments," *Econometrica*, 1997, *65* (3), 557–586.

Wei, Shang-Jin and Xiaobo Zhang, "The competitive saving motive: Evidence from rising sex ratios and savings rates in China," *Journal of Political Economy*, 2011, *119* (3), 511–564.

Xie, Yu, "The User's Guide of the China Family Panel Studies (2010)," *Beijing: Institute of Social Science Survey, Peking University*, 2012.

Xu, Hongwei and Yu Xie, "The causal effects of rural-to-urban migration on children's wellbeing in China," in "Annual Meeting of the Population Association of America, New Orleans" Citeseer 2013, pp. 10–13.

Zhang, Hongliang, Jere R Behrman, C Simon Fan, Xiangdong Wei, and Junsen Zhang, "Does parental absence reduce cognitive achievements? Evidence from rural China," *Journal of Development Economics*, 2014, *111*, 181–195.

Zhang, Junsen and William Chan, "Dowry and wife's welfare: A theotrical and empirical analysis," *Journal of Political Economy*, 1999, *107* (4), 786–808.

Zhao, Yaohui, "Labor migration and earnings differences: The case of rural China," *Economic Development and Cultural Change*, 1999, *47* (4), 767–782.