A Darwinian Perspective on Entrepreneurship and Economic Growth:
Evidence from the People's Republic of China

Shang-Jin Wei (Columbia University, CEPR, and NBER)

and

Xiaobo Zhang (IFPRI)

Abstract

Chinese growth in recent decades is a miracle not only because its rate of growth is higher than most countries in the world but also because the growth takes place in a political setting that is not always known to be favorable to economic development. We argue that a key pro-growth element in the Chinese story is a rising sex ratio imbalance (relative surplus of men in the pre-marital age cohort). The skewed sex ratio is an unintended byproduct of the spread of ultrasound machines and a coercive population control policy. Since family wealth is an important status variable in the marriage market, the sex ratio imbalance has created a powerful additional incentive for wealth creation. We provide evidence that the imbalance promotes entrepreneurship and hard work. First, new domestic private firms – an important engine of growth – are more likely to emerge from regions with a higher sex ratio imbalance. Second, the likelihood for parents with a son to be entrepreneurs rises with the local sex ratio. Third, households with a son in regions with a more skewed sex ratio demonstrate a greater willingness to accept relatively dangerous or unpleasant jobs and supply more work days. In contrast, the labor supply pattern by households with a daughter is unrelated to the sex ratio. Finally, regional GDP tends to grow faster in provinces with a higher sex ratio. We estimate that the skewed sex ratio contributes about 20% to the overall growth rate in recent years. Since the sex ratio imbalance will become worse in the near future, this growth effect is likely to persist.

1. Introduction

The craving for a partner of the opposite sex is one of the most basic and powerful human desires besides survival. We study in this paper the effects of mating competition on entrepreneurship and economic growth. We dissect the Chinese experience for this purpose.

The recent Chinese growth is called a miracle not only because the annual rates of growth have been higher than most countries in the world for more than 30 years by now but also because the pace of growth has increased in recent years in spite of its substantially higher base relative to its own past. If the high growth rates during 1980-2000 could be understood through the lens of economic convergence or a catch-up story (see, for example, Barro and Sala-i-Martin, 1992 and 2004; and Sachs and Woo, 1997), the acceleration of growth rates in more recent years is perhaps more surprising in the context of cross-country growth experience (although there are theoretical models that could generate this pattern). The growth miracle is all the more puzzling since it takes place in a political setting that is not always known to be favorable to economic development. If one wants a list of factors that are likely to retard growth, one thinks of corruption, over-regulation, insufficient judicial independence, weak protection of property rights, over-supply of government discretion, under-development of the financial system, inefficient state-owned firms, and unpredictability of government policies. China seems to have most items on the list. There must be something else going on in the country that is powerful enough to counteract these anti-growth factors.

In this paper, we argue that a key pro-growth factor in the Chinese story that is not shared by most other economies and has not been discussed in the standard growth theory is a rising sex ratio imbalance (relative surplus of men in the pre-marital age cohort). The skewed sex ratio itself is an unintended byproduct of the spread of ultrasound machines and a coercive population control policy. Since family wealth is an important status variable in the marriage market, the sex ratio imbalance has created a powerful additional incentive for wealth creation. Other things equal, this leads to more entrepreneurial activities and increased supply of work effort.

The idea that mating competition can motivate people to exert stronger effort and undertake greater risk has a Darwinian flavor, and therefore is not entirely new. By examining variations in the sex ratio across immigrant groups in the United States, Angrist (2002) documented that a higher male/female sex ratio within an ethnic group has a large positive effect on the female marriage rate. Interestingly, he found that "higher sex ratios also appear to have

raised male earnings and the incomes of parents with young children." While these results are consistent with what we report in this paper, he did not directly study the effects of sex ratio imbalance on entrepreneurial activities and economic growth. Charles and Luoh (2010) studied how rising incarceration of African American males has created an effective sex ratio imbalance in the marriage market (relative surplus of African American females). A very interesting finding is that African American females appear to respond to this shock by increasing labor supply and schooling. However, the paper did not study entrepreneurship. Concerns for social status have been argued to increase people's propensity to engage in gambling (Robson, 1992) and to be more tolerant of under-diversified activities including entrepreneurship (Roussanov, 2010). A very interesting paper by Roussanov and Savor (2011) documented that unmarried CEOs and mutual fund managers are more willing to take risks than their married counterparts. They interpreted this as evidence that competition in marriage market (through ranking by relative wealth) is what motivates these people to be more tolerant of risks. It is important to note that the underlying marriage market in their sample does not feature a sex ratio imbalance. Therefore, they did not explore variation in the degree of marriage market competition in terms of different sex ratios. Instead, they concentrated on comparing single versus married CEOs and money managers. As the authors acknowledged, marital status of CEOs and fund managers cannot be completely accurately ascertained. For example, on the one hand, divorced CEOs may be classified as married in their sample. On the other hand, married CEOs could be mistakenly classified as unmarried if their marital status is not revealed by any news reports. Still, the spirit of the study is very much in line with the hypothesis in this paper that mating competition could encourage risk-taking activities including entrepreneurship.

The idea that mating competition intensified by a rising sex ratio for the young cohort could be an important part of China's aggregate growth story has not been examined before. This is a potentially important hypothesis. First, within the short span of a decade, China has steadily surpassed the United Kingdom, France, Germany and Japan in moving from the world's 6th largest to the 2nd largest economy in terms of GDP (even without PPP adjustment). Its contribution to incremental world GDP in a given year has already surpassed the United States (IMF, 2009). Therefore, anything that can enhance our understanding of China's growth performance and its future outlook has global implications. Second, the uneven distribution of the sex ratio across its vast territory in combination with the fact that most people marry locally

provides an unusually good opportunity to investigate the economic consequences of mating competition. In addition, a within-country study such as this one has advantages over cross-country studies as the legal system, culture, macroeconomic policies, and other institutions can be more plausibly held constant across regions within a country than across countries. As a very large country, there are many sub-national geographic units in China that allow us to have sufficient statistical power when exploring regional variations.

A sex ratio imbalance in the marriage age cohort – too many men relative to women – is a common demographic feature in many Asian economies, including Korea, India, Vietnam, Singapore, Taiwan, and Hong Kong. In many such economies, parents voluntarily limit the number of children they wish to have. This, together with a strong preference for sons, and the availability of inexpensive technology to screen the gender of a fetus (most commonly by Ultrasound B) to abort the unwanted pregnancy, leads parents to engage in sex selective abortions in favor of sons. But nowhere else has a more skewed sex ratio than China today, where a strict family planning policy has restricted the number of children most families can have to one or two and has greatly reinforced the incentive for sex selective abortions. In 1980, when the strict family planning policy was first introduced in China, its sex ratio at birth was 1.07 boys per girl, which was basically in line with the natural rate observed in most countries. The Chinese sex ratio deteriorated steadily to 1.12 boys per girl in 1990, 1.18 in 2000, and 1.22 in 2007 (Li, 2007; Zhu, Lu, and Hesketh, 2009). As a result, roughly one out of every nine young men today has no realistic hope to get married, mathematically speaking. In some provinces, one out of every six men cannot get married. This situation is projected to deteriorate in the next ten years based on the population census data.

The existing literature has identified several negative consequences of a serious sex ratio imbalance. First, the scale of involuntarily single men is frightening. For example, the number of excess Chinese men under age 20 exceeded 32 million in 2005 (Zhu, Lu, and Hesketh, 2009). This number is greater than the entire male population of Italy or Canada. Second, the imbalance may cause crimes. Using data across Chinese provinces, Edlund, Li, Yi, and Zhang (2007) estimate that every one basis point increase in the sex ratio (e.g., from 1.10 to 1.11 boys per girl) raises violent and property crime rates by 3%, and the rise in the sex ratio imbalance may account for up to one-seventh of the overall rise in crime in China. Den Boer and Hudson (2004) boldly hypothesize that the sex ratio imbalance should generate security concerns for other

countries since the one with a high sex ratio "might actively desire to send its surplus young males to give their lives in a national cause," although they provide no rigorous data analysis to back up their theory. Third, the imbalance may also trigger competitive savings among households – men and households with sons forego current consumption to accumulate wealth in order to improve a young man's standing in the marriage market relative to other men. This increase in the savings rate is inefficient since it does not alter the number of unmarried men in the aggregate. By raising the aggregate savings, the sex ratio imbalance contributes to China's current account surplus, which is a source of international frictions.

In this paper, we study a possibly positive effect of the sex ratio imbalance on economic growth. If the family wealth of a man relative to those of other men is a sorting variable for a man's relative standing in the marriage market, then a rise in the sex ratio can inspire men and parents with a son to find ways to accumulate more wealth. Working harder or longer, and becoming more entrepreneurial are ways to achieve this objective. As a result, the economy may grow faster than it would have otherwise. As far as we know, this effect has never been investigated before. The savings effect of higher sex ratios documented by Wei and Zhang (2009) could also help entrepreneurship: To the extent that a borrowing constraint is a barrier to entrepreneurship and a higher savings rate relaxes the borrowing constraint, higher sex ratios may promote entrepreneurship through this channel. This, however, is not the only or even the main channel for higher sex ratios to have a positive effect on economic growth. In cross-regional growth regressions, we will explicitly control for local savings rate and demonstrate that the effect of higher sex ratios goes beyond the savings channel.

We conduct the empirical analysis using data from censuses of firms, censuses and surveys of population, and household surveys. The empirical results not only support the hypothesis qualitatively, but also are significant quantitatively. Based on the two most recent censuses of firms, we estimate that an increase in the sex ratio by one standard deviation can account for about half the extensive margin of private sector growth (i.e., the birth of new private firms) across regions.

To address concerns about possible biases due to measurement errors, endogeneity, and missing regressors, we employ four different approaches that complement each other. Both an instrumental variable approach and a placebo test suggest that there is a causal effect from a higher sex ratio to more entrepreneurship. Two different sets of household-level data provide

additional information. For example, from a random sample of a national population survey in 2005, we document that parents with a son are more likely to be entrepreneurs in regions with a higher sex ratio. In comparison, the likelihood for parents with a daughter to be entrepreneurs is uncorrelated with the local sex ratio. From a separate survey of rural households in 2002, we estimate that households with a son respond to a rise in the sex ratio by a combination of working more days off farms (including as migrant workers) and becoming more willing to accept unpleasant or relatively dangerous jobs. In contrast (but consistent with our hypothesis), the labor supply pattern of daughter-households is not linked to the local sex ratio.

Finally, to capture the general equilibrium effect, we also directly check whether the growth of per capita GDP across regions is linked to the local sex ratio imbalance and find that the answer is affirmative. We estimate that about 20% of the growth rate of GDP per capita in recent years can be attributed to a rise in the sex ratio. Since the sex ratio imbalance is projected to become worse in the next decade, this effect may become relatively more important over time.

The rest of the paper is organized in the following way. In Section 2, we develop the argument more fully and connect to related literatures. In Section 3, we provide statistical evidence. Finally, in Section 4, we conclude and discuss possible future research.

2. The connection between the hypothesis and the existing literature

The hypothesis that a higher sex ratio can be an important driver for entrepreneurial activities in China is related to four sets of literature: (a) status goods, (b) economics of family, (c) entrepreneurship, and (d) causes and consequences of sex ratio imbalance. Each of them is too vast to be referenced comprehensively here. Instead, we selectively discuss some of them, with a view to highlight the insight most relevant for our empirical investigation.

Several theoretical papers have pointed out a connection between concerns for status (one's relative position in a society), the savings rate, and the economic growth rate (Cole, Mailath and Postlewaite, 1992; Cornero and Jeanne, 1999; and Hopkins and Kornienko, 2009). When wealth defines one's status in the marriage market, a greater concern for status may lead to an increase in the growth rate. In principle, concerns for status could also produce the opposite effect on savings and growth. In particular, if status is enhanced by conspicuous consumption, then a greater concern for status can translate into a reduction in savings (Frank, 1985 and 2005).

It is interesting to note that, while many papers on the topic of status use competition in the marriage market to illustrate the idea, the sex ratio is always assumed to be balanced. In other words, no explicit comparative statistics are derived in terms of a rise in sex ratio imbalance.[1]

Du and Wei (2011b) develop a model that explores the effect of a higher sex ratio on entrepreneurial activities. As it is the only model that explicitly studies such a topic, we review it in some detail. The model features overlapping generations with two genders and a desire to marry. Everyone lives for two periods, and can marry a member of the opposite sex at the beginning of the second period. There are two benefits associated with marriage. First, a husband and wife can pool their income and enjoy a partial public good feature of their joint consumption. That is, the sum of their consumptions can be more than the sum of their incomes. Second, a married person derives additional emotional utility (or "love") from his/her spouse. The amount of emotional utility that a person can bring to her/his future spouse is a random variable in the first period, but its value is revealed and becomes public knowledge when she/he enters the marriage market. Men and women are identical in their utility function and the form of the budget constraint, except that by assumption only men can be entrepreneurs (a simplification that is relaxed in an extension).

At the beginning of the second period, men and women voluntarily participate in the marriage market. The matching between men and women is assumed to follow the Gale-Shapely algorithm, which produces positive assortative matching in equilibrium. More precisely, the best woman (defined by a combination of her wealth and the amount of emotional utility she can bring to her spouse) and the best man (also defined by a combination of wealth and emotional utility) are matched; the second best woman and the second best man are matched; and so on. If there are more men than women, the least attractive men are not married. By Gale and Shapely (1962) and Roth and Sotomayer (1990), this equilibrium is both unique and stable.

In the first period, both men and women solve an optimization problem that takes into account the effect of their choices on the outcome in the marriage market. A representative man makes a sequential choice, first deciding whether to be an entrepreneur and then deciding on a savings rate. Deciding to be an entrepreneur means to pay a fixed fee in order to get a random draw on his productivity, which follows a binominal distribution. If his productivity is high, he becomes a monopolistically competitive producer. If it is low, he becomes a worker. A

---

[1] Edlund (1996) showed that a higher sex ratio imbalance may have a nonlinear impact on women's status and dowry price.

representative woman simply decides on a savings rate (as she is always a worker by a simplifying assumption). Even with a balanced sex ratio, the Nash equilibrium would feature a fraction of men becoming entrepreneurs. This is because entrepreneurship is a higher risk and a higher return activity (relative to being a worker). Similar to a portfolio choice problem, there is generally an interior solution in which a certain fraction of men choose the riskier activity.

Du and Wei (2011b) derive the following key proposition: As the sex ratio rises, as long as it is beyond a (low) threshold, more men choose to become entrepreneurs. Here is the intuition. By the structure of the model, successful entrepreneurs can always succeed in getting married, but failed entrepreneurs do not. When the sex ratio exceeds a threshold, an increase in the sex ratio raises the probability that a male worker will not get married, while it does not alter the expected utility of being an entrepreneur (to a first-order approximation). If the utility from a marriage is sufficiently large, more men would respond to a higher sex ratio by becoming entrepreneurs.

There is an extensive literature in demography that documents the phenomenon of unbalanced sex ratios in Asia (for example, Gu and Roy, 1995; Guilmoto, 2007; and Li, 2007). Several papers have examined the determinants of sex ratio imbalance (including Das Gupta, 2005; Ebenstein, 2009; Edlund, 2009; Li and Zheng, 2009; and Bulte, Heerink and Zhang, 2011). In an influential paper, Oster (2005) proposes that the prevalence of Hepatitis B is a significant cause of the sex ratio imbalance in Asia. But this conclusion is later shown to be incorrect, including by Lin and Luoh (2008) and Oster, Chen, Yu and Lin (2008). In a paper with a clever instrumental variable approach, Qian (2008) shows that an improvement in the economic status of women tends to reduce the sex ratio imbalance. Her instrument for the economic status of women is the world price of tea, whose production is apparently particularly suitable for women laborers. Wei and Zhang (2009) document that a higher sex ratio induces more savings. However, the paper does not examine how labor supply and entrepreneurship respond to a change in the sex ratio, which is the central focus of the current paper.

This discussion has clear implications for the empirical work in this paper. First, it is interesting to find out if entrepreneurial activities are indeed linked to local sex ratios. Second, given the hypothesized mechanisms, it is informative to check whether and how households with a son and those with a daughter respond differently to a rise in the sex ratio. Third, given the possibility that a higher sex ratio could also raise crime rates and have other consequences that

are potentially negative for economic growth, it is important to check the general equilibrium effect – whether economy-wide entrepreneurial activities and work effort increase on net, as reflected in a higher overall growth rate, in response to a rise in the sex ratio.

Because Wei and Zhang (2009) have provided extensive evidence on the effects of a higher sex ratio on savings behavior, we will not look at savings in this paper, and concentrate instead on evidence related to entrepreneurship, work effort, and economic growth. In our view, pursuing entrepreneurial activity (for higher income) and raising the savings rate (out of a given level of income) are two alternative ways to improve a household's future wealth. (We will also control for the effect of savings on economic growth.)

3. Statistical Evidence

We start by providing some basic facts about Chinese growth, which are summarized by two 70% rules. We then use data from the two most recent censuses of manufacturing firms (in 1995 and 2004) to investigate whether local sex ratio imbalance is a predictor of the extent of local entrepreneurial activities. To zoom in on possibly distinct responses by families with a son versus those with a daughter, we turn to household-level evidence. Finally, to capture the general equilibrium effect of a rise in the sex ratio, we conduct panel growth regressions across Chinese provinces over 1980-2005.

Background information: the two 70% rules about Chinese growth

Since our first piece of evidence has to do with regional variations in entrepreneurial activity, we work with the two most recent censuses of firms in 1995 and 2004, respectively, so we can compute the growth in the number of firms by region. During this period, the country's industrial value added (at the current price) grew by 266%.

The growth of the private sector is a major part of the overall growth story. The private sector is not just restricted to firms that were legally registered private firms. In fact, very few firms were registered as private firms in the 1980s and 1990s. According to Huang (2009), many private entrepreneurs at that time found it necessary to set up firms as nominally owned by local governments (in the form of "township-and-village enterprises," or "collectively owned firms"). The goal was presumably to buy "protection" from the local government and to minimize the

risk of state expropriation. Such a practice was widespread and was called "private entrepreneurs wearing a red hat." Most entrepreneurs later engineered or attempted to engineer a change in firm ownership through which they would become a majority shareholder without injecting much additional personal capital. Wu (2007) provides fascinating accounts of many entrepreneurs both when they first "wore a red hat," setting up a nominally collectively owned firm, and when they tried to take off the hat, with uneven success rates. Because of the recognition that most newly established "collectively owned firms" were private firms in disguise, we adopt a broad definition of domestic private firms to include all such firms.

In Table 2, we report a simple exercise that decomposes the contributions to growth by firm ownership type (domestic private firms, majority state-owned firms, and foreign-invested firms). Let X(total, t) be the industrial value added for the country as a whole in year t. Define X(private, t), X(FDI, t), and X(SOE, t) to be the industrial value added in year t by the domestic private sector, foreign invested firms, and state-owned firms, respectively. X(total, 04) = X(private, 04) + X(FDI, 04) + X(SOE, 04). Let s(private, 95), s(FDI, 95), and s(SOE, 95) be the share of the domestic private sector, foreign firms, and state-owned firms, respectively, in natural industrial output in 1995. We can decompose the overall growth rate into a weighted average of the growth rates from the three types of firms:

(1) G(total) = X(total, 04)/X(total, 95) – 1
$$= g(private)*s(private, 95) + g(FDI)*s(FDI, 95) + g(SOE)*s(SOE, 95)$$

From this equation, we can compute the contribution of the domestic private sector to overall growth as: Private sector's share of the contribution = g(private)s(private,95)/g(total) = 6.22*30.7%/2.66 = 71.9%. Similarly, foreign invested firms account for 30.8% of overall growth. The state sector accounts for -2.7% as many state-owned firms were either closed or taken over by private firms. (Note that the decomposition of real or nominal growth rates gives the same result.)

We next decompose the private sector growth into the extensive margin (growth in the number of firms) and the intensive margin (the growth of average output per firm):

(2) Ln[X(private, 04)/X(private, 95)] = Ln[N(private, 04)/N(private, 95)] +
     Ln{[X(private,04)/N(private, 04)]/ [X(private, 95)/N(private, 95)]}

The first term on the right hand side denotes the extensive margin, while the second term denotes the intensive margin growth. In Table 3, we report the result of the decomposition. The contribution of the extensive margin = the first term on RHS/LHS = 0.499/0.728 =68.5%.

To summarize, a little over 70% of Chinese growth is attributable to the rise of the private sector. In addition, almost 70% of private sector growth is attributable to the birth and growth of new private firms[2]. Therefore, the birth and growth of new private firms are a significant part of the Chinese growth story.

Where are domestic private firms most likely to emerge?

We now examine whether there is any connection between the sex ratio and the extensive margin of private sector growth. To reduce noise, we sort all counties into bins based on their sex ratios for the age cohort of 5-19 in 1995. All counties in a bin have an identical sex ratio. In Figure 1, we plot the average growth rate in the number of private firms across all counties in a bin against the initial sex ratio imbalance for that bin. There is a strong positive relationship between the growth of the count of private firms and the sex ratio. That is, regions with a more skewed sex ratio are also places where new private firms are more likely to emerge.

Many factors could affect the birth and growth of new firms. The age structure of the local population, the growth rate of the population, local income and education levels, local industrial structure, and initial scale of the private sector could all matter. We are interested in investigating whether the local sex ratio also plays a role. We do so by looking at variations in the growth rate of the count of private firms and local sex ratios across 1790 counties, conditional on other factors. The specification is as follows:

$$(3) \quad \text{Growth\_in\_firm-count}_{k, 95\text{-}04,} = \beta \, \text{Sex\_ratio}_{k, 95,} + X_k \, \Gamma + e_k$$

The result is reported in Column 1 of Table 4. The coefficient on initial sex ratio for the age cohort 5-19 is 0.019 and statistically significant. To see if the effect of the local sex ratio comes entirely from the local savings rate, we add log local bank savings balance per capita in 1995 as an additional control. The new regressor is positively related to local growth. This is consistent

---

[2] Using a provincial level panel data, Li et al. (2009) show entrepreneurship is a key engine of Chinese economic growth.

with the idea that credit constraints might deter the birth of new firms, and savings helps to partially overcome the constraints. ~~is not statistically significant. In any case,~~ However, the coefficient on the sex ratio is virtually unaffected. This suggests that the effect of higher sex ratios on entrepreneurship goes beyond the savings channel.

To see which sex ratio imbalance in terms of age cohort matters the most, in Columns 3-5, we restrict the sex ratio to the age cohorts 5-9, 10-14, and 15-19, respectively. Interestingly, the first two age cohorts appear to matter more. (While the coefficient on the sex ratio for the cohort 15-19 is not statistically significant here, it becomes significant in another specification in the last column of the same table.) This is indicative that much of entrepreneurial response comes from parents with young children. It is also noteworthy that the coefficient on the sex ratio for each age cohort is smaller than the one for the combined age cohort 5-19. Because most Chinese families have one or two children, relatively few families with children in the three age brackets overlap. As a result, the effect of the sex ratio for the combined age cohort of 5-19 on the growth of private firms is almost the same as the sum of the effects of the sex ratio for the three age cohorts.

The initial sex ratio does not fully capture the sex ratio imbalance for the young cohort over the entire ten-year period. As an alternative measure, we use the average of the sex ratio for the age cohort 5-19 in 1995 and the sex ratio for the age cohort of 4-18 in 2004 (inferred from the 1990 and 2000 censuses, respectively). Due to the way that the sex ratios are reported in the two censuses (at five-year intervals), we are not able to make an exact match in age cohort. In any case, the regression results are reported in the last five columns of Table 4. In all regressions, the coefficients on the sex ratio are positive and statistically significant. In particular, the coefficient on the sex ratio for the age cohort of 15-19 is also positive and significant. Overall, the results indicate that more domestic private firms were established in regions with a higher sex ratio.

Possible problems with the OLS estimation and solutions

The OLS estimation may produce biased estimates. First, there could be errors in measuring the sex ratio for the pre-marital age cohort. For example, with migration in and out of a county (in spite of the policy restrictions), the sex ratio recorded in the population census may not exactly correspond to the sex ratio in the local marriage market. The measurement errors tend

to produce a downward bias. Second, the sex ratio might be endogenous. In particular, the positive association between the local sex ratio and the rate of growth of private firms may reflect a reverse causality. For example, if private entrepreneurs have a stronger urge to have a male heir to take over their business when they retire, then regions that happen to see a lot of private firms may also exhibit a strong son preference and a high sex ratio imbalance. Third, the sex ratio may be endogenous if it is correlated with some missing regressors. For example, in spite of our best efforts to control for determinants of the growth of private firms, there may be other variables that are good predictors of future profitability in a region that are not captured by our list of control variables. If these variables happen to be correlated with the local sex ratio, we may find a positive association between the local sex ratio and the growth of local private firms even when there is no direct economic association between the two.

To address these problems, we adopt four approaches that we hope would complement and reinforce each other. First, we implement a two-stage least squared (2SLS) procedure in which the local sex ratio is instrumented by variables that affect regional variations in the sex ratio but are otherwise unlikely to affect directly the growth of local private firms. Second, we use data from a population survey and check if the likelihood for parents with a son to become entrepreneurs differs from parents with a daughter when the sex ratio rises. Third, we adopt a placebo test on the growth of other profit-seeking firms. If the local sex ratio is simply a proxy for missing regressors that help to forecast local growth potential, then the sex ratio should also forecast the extensive margin growth of foreign-invested firms. Finally, we go to household-level data where we can check possible interactions between local sex ratios and son-families in ways that will also help us to rule out the endogeneity story. In particular, our theory suggests that son families and daughter families may alter their work effort in different ways in response to a common rise in the local sex ratio. We will discuss these approaches in turn.

While none of the approaches individually may address all concerns, we hope the combination of the approaches would enhance the confidence in our interpretations.

Approach 1: Instrumental variable approach

A strategy to address both the measurement error problem and the endogeneity problem is to employ an instrumental variable approach. A key determinant of the sex ratio imbalance is a

strict family planning policy introduced at the beginning of the 1980s[3]. We explore three determinants of local sex ratios that are unlikely to be affected by the growth of local private firms, and for which we can get data. First, while the goals of family planning are national, the enforcement is local. Ebenstein (2009) proposes to use regional variations in the monetary penalties for violating the birth quotas, originally collected by Scharping (2003), as instruments for the local sex ratio. The idea is that, in regions with stiff penalties, parents may engage in more sex-selective abortions, rather than paying a penalty and having more children. The monetary penalty is often on the order of between one to five times the local average annual household income. In addition, Ebenstein (2008) coded a dummy for the existence of extra fines for violations at higher-order births. For example, an additional penalty may kick in on a family for having the 3rd or 4th child in a one-child zone, or the 4th or 5th child in a two-child zone. Such a non-linear financial penalty scheme was introduced by different local governments in different years (if at all), generating variations across regions and over time. These two monetary penalty variables constitute the first two candidates for our instrumental variables. Edlund et al. (2007) conduct some diagnostic checks and conclude that the level of financial penalties is uncorrelated with a region's current economic status. We will perform and report a formal test on whether the proposed instruments and the error term in the second stage regressions are correlated.

The third instrumental variable explores the legal exemptions in the family planning policy. While the policy imposes a strict birth quota on the Han ethnic group (the main ethnic group in the country), the rest of the population (i.e., some 50 ethnic minority groups) do not face or face much less stringent quotas. (The government allowed the exemption, possibly to avoid criticism for using the family planning policy to marginalize the minority groups.) As a result, the share of non-Han Chinese in the total population has risen from 6.7% in 1982 to 8.5% in 2000 (Bulte, Heerink, and Zhang, 2011). Non-Han Chinese are not uniformly distributed across space. In regions with relatively more ethnic minorities, marriages between Han and non-Han peoples are not uncommon, reducing the competitive pressure for men in the marriage

---

[3] China's family planning policy, commonly known as the "one-child policy," has many nuances. Since 1979, the central government has stipulated that Han families in urban areas should normally have only one child (with some exceptions). Ethnic Han families in rural areas can have a second child if the first one is a daughter (this is referred to as the "1.5 children policy" by Ebenstein, 2008). Ethnic minority (i.e., non-Han) groups are generally exempted from birth quotas. Non-Han groups account for a relatively significant share of local populations in Xinjiang, Yunnan, Ganshu, Guizhou, Inner Mongolia, and Tibet.

market (Wei and Zhang, 2009). Therefore, the share of non-Han Chinese in the local population offers another possible instrument.[4]

The first stage regressions are reported in Table 5. The dependent variable in the first two regressions is the initial sex ratio for the age cohort 5-19, whereas that for the last two regressions it is the average sex ratio of the same age cohort over 1995 and 2004. We omit the share of non-Han Chinese in the local population as a possible instrument in regressions 2 and 4.

The financial penalties for violating birth quotas generate a positive coefficient in all four regressions (and significant in three of them). The dummy for the existence of extra penalties for violations at higher-order births also produces a positive coefficient in all four regressions. These results imply that a more severe penalty for violating legal birth quotas tends to induce parents to more aggressively abort girls, resulting in a higher sex ratio imbalance. In other words, when the penalties are light, many couples with daughters may opt to keep the daughter, pay the penalties, and have another child, rather than abort the female fetus. The coefficients on the share of the local population not subject to birth quotas are negative and statistically significant in both regressions in which it appears. This is consistent with the notion that sex selective abortions are less prevalent when birth quotas apply to less people. (In case ethnic minorities have a different entrepreneurial propensity for reasons unrelated to the sex ratio, we will also implement a 2SLS estimation that omits the minority share in the local population from the set of instruments and find that the results are largely the same.)

The F statistics (for the null that all slope parameters are jointly zeros) ranges from 25.0 to 39.7. The Kleibergen-Paap statistics (for weak instruments) for all four regressions are greater than the Stock-Yogo 10% critical value of 19.9.

The second stage regressions are reported in Table 6. In the first four regressions, the sex ratio is measured by its initial value (in 1995). In the last four regressions, the sex ratio is measured by its average value over 1995 and 2004. In odd-numbered regressions, we use all three instrumental variables. In the even-numbered ones, we use only the two financial penalties as instruments. The minority share in the population is excluded from the set of instruments in half of the regressions to alleviate concerns that Hans and non-Hans may have different

---

[4] In principle, variations in the cost of sex screening technology especially the use of an Ultrasound B machine (as documented by Li and Zheng, 2009), and the economic status of women (such as that documented in Qian 2008) could also be candidates for instrumental variables. Unfortunately, we do not have the relevant data. Note, however, for the validity of the instrumental variable regressions, we do not need a complete list of the determinants of the local sex ratio in the first stage.

entrepreneurial propensities for other reasons (although we also perform a formal test to see if the instruments are correlated with the error term in the main regression). In Regressions 1, 2, 5 and 6, we do not include households savings as a control, whereas we do so in the other four regressions.

The Durbin-Wu-Hausman test easily rejects the null that the 2SLS and OLS estimates are the same in all four regressions, implying that the sex ratio variable is likely to be either measured with errors or endogenous. The Hansen's J statistics do not reject the null that the instruments and the error term are uncorrelated in any of the regressions.[5] At least in a pure statistical sense, the instruments are valid. The point estimates in Table 6 are generally much larger than their OLS counterparts in Table 4. This suggests that the downward bias in Table 4 generated either by missing regressors or by measurement errors is substantial.

We can compute the economic significance of the estimates. Using the most conservative estimate in Column 4 of Table 6, an increase in the sex ratio by 3 basis points (e.g., from 1.08 to 1.11), which is equal to the increase in the average sex ratio from 1995 to 2004 (see Table 1A), generates an increase in the natural log number of private firms by 0.39 (=0.13x3). Since the actual increase in log number of firms in this period is log 0.83 (see Row 3 of Table 1A), the rise in the sex ratio can potentially explain 47% (=0.39/0.83) of the actual increase in the number of private firms in China during this period. In other words, the economic impact of the rise in sex ratio in promoting entrepreneurial activities in rural China is potentially very big.

We may take a special note of the role of savings in Columns 3-4 and 7-8. The positive coefficients on local savings indicate that higher local savings are associated with higher rates of firm birth. This is consistent with the notion that credit constraints are a barrier to entrepreneurship and higher household savings may be a way to mitigate the negative effect of credit constraints. At the same time, we note that the coefficient on the sex ratio is virtually unchanged when local savings is added. This suggests that the positive effect of higher sex ratios on entrepreneurship goes beyond the savings channel. (We note, however, that household savings is also likely to be endogenous, and we do not have separate instruments for it.)

We now perform similar regressions for the urban sample, and report the results in Table 7. The results are qualitatively similar to the rural sample. This is consistent with the notion that

---

[5] In household-level regressions to be reported in Tables 11 and 12, we check if the ethnic minorities have a different labor supply pattern from Han Chinese, holding constant local sex ratio imbalance and other determinants of labor supply, and cannot reject the null that there is no difference.

a higher sex ratio imbalance in a city has also induced more people to engage in entrepreneurial activities. It is noteworthy that the IV coefficients on the sex ratio variable for the urban sample in Table 7 (0.11 and 0.12) are smaller than the corresponding coefficients for the rural sample in Table 6 (between 0.13 and 0.14). Using the point estimate in the last column of Table 7 (0.12), an increase in the sex ratio in the urban area by 3 basis points (e.g., from 1.08 to 1.11) would generate an increase in the natural log number of private firms by 0.36 (=12*0.03), which is about 40% of the actual increase in the number of private firms during this period. While this number is smaller than for the rural sample, it is still economically significant. It is also interesting to note that the savings variable is not significant in the urban sample. This could mean that credit constraints are not the most binding constraint for urban entrepreneurship as credits are relatively more available to urban entrepreneurs than to rural entrepreneurs. In any case, the effect of higher sex ratios on entrepreneurship is not affected by controlling for the savings effect.

Approach 2: What types of families are most susceptible to stimulation from a higher sex ratio?

Our story is about how a rise in the sex ratio may encourage more parents with a son, but not necessarily parents with a daughter, to take up entrepreneurial activities. Since the firm census data do not have information on the number and gender composition of the children of the firm owners, we now look at the China Population 1% Survey in 2005 which contains some relevant information.[6] To maximize comparability, we focus on nuclear families with both parents alive and below the age of 40, and with one or two children. We cap the age of the household head at 40 in order to minimize the probability that an adult child with unknown gender has moved away and therefore is not counted in the household survey.

We define entrepreneur $_{kl}$ = 1 if at least one parent in household k in location l is a business owner or self-employed, and zero otherwise. We run the following Probit regression:

$$\text{Prob (entrepreneur }_{kl} = 1) = \lambda \text{ sex\_ratio}_l + \beta \text{ sex\_ratio}_l * \text{dunmmy for son}_k + X_{kl} \Gamma + e_{kl}$$

---

[6] The population census is done once a decade with the latest two rounds in 2000 and 2010. In between the censuses, a stratified random household survey of 1% of the population was conducted in 2005. We were given access to a 20% random sample of the 2005 survey, which covers 325 cities and 343 rural prefectures.

where *sex_ratio$_l$* is the sex ratio for the age cohort of 5-19 in location l, *dummy for son$_k$* takes the value of one if family k's first child is a son and zero otherwise, and X$_{kl}$ is a vector of control variables including household wealth (proxied by the value of the house[7]), household head's age, gender, years of education, and years of education squared, whether there is any household member who has a severe health problem, and the number and the age of the children. λ, β and Γ are parameters to be estimated. e$_{kl}$ is an iid normally distributed random variable.

Because the 2005 population survey does not have information that allows us to compute household savings, we are not able to include it as a regressor. However, we can approximate household wealth by the value of the household dwelling (if it owns the house/apartment). We include the proxy for household wealth as a control variable.

The key parameter of interest is β: If a combination of having a male first child and living in a region with a high sex ratio makes it more likely for parents to become entrepreneurs, we would expect β > 0. The regression results for this specification for urban and rural nuclear families, respectively, are reported in Columns 1 and 4 of Table 8. In both cases, coefficient β is indeed positive and statistically significant.

The coefficient λ describes how families without a male first child respond to the local sex ratio. It is statistically not different from zero in both the rural and urban samples. This means that having a male first child per se does not make parents more likely to be entrepreneurs. Instead, it takes a combination of having a male child and living in a region with a skewed sex ratio for parents to be more inclined to be entrepreneurs.

The above specification requires that the effects of all control variables such as the household head's age and gender on parental inclination to be entrepreneurs are identical regardless of the sex composition of children and the number of children. This requirement may not be realistic. One way to relax this (unnecessary) restriction is to run separate regressions for different types of households (so that the coefficients on the control variables are allowed to take different values for different household types). Running separate regressions this way would reduce possible bias in the estimates of β at the cost of a lower efficiency. Given the relatively

---

[7] This is the house value at the time of purchase. The population survey asks for the value at the time of purchase and the year of construction but not the year of purchase. If we pretend that the year of construction is the same as the year of purchase, and adjust the house value by the national housing price index, we obtain similar results for the rural sample, but the sign on housing wealth turns negative for the urban sample. In both samples, the patterns on the sex ratio remain unchanged.

large sample size, we can afford to sacrifice some efficiency in exchange for an improved chance to obtain unbiased estimates of the key parameters.

We focus on two specific sub-samples: nuclear families with a son versus nuclear families with a daughter. A major advantage of these subsamples is that they can help rule out the possibility that some unobserved family characteristics simultaneously determine the gender of a child and the parental propensity to become entrepreneurs. As Ebenstein (2009) documented, sex selections are mostly done on the second or higher-order births[8]. There are few sex selections on the first-born children, especially in rural areas where the official policy allows for a second child if the first child is a girl. Since most parents prefer having a daughter and a son to having two sons, there is very little reason to select gender on the first child. In other words, the gender of the first child is the choice of nature, not that of the parents. This allows us to focus on the effect of the child's gender on the parental propensity to engage in entrepreneurial activity.

We run separate Probit regressions for nuclear families with a son and nuclear families with a daughter, respectively,

$$\text{Prob (entrepreneur }_{kl} = 1) = \beta \text{ sex\_ratio}_l + X_{kl} \Gamma + e_{kl}$$

where the variables are similar to those explained earlier. The results for the rural sample are reported in Columns 2 and 3 of Table 8. For families with a son, the local sex ratio is a positive and significant predictor for whether the parents are entrepreneurs. In comparison, for families with a daughter, the local sex ratio is not significant. The results for the urban sample are reported in Columns 5 and 6 of Table 8. Again, the local sex ratio is associated with a greater likelihood for parents to be entrepreneurs for families with a son, but not for families with a daughter.

While the gender of the first child may be determined by nature, the number of children in a family is still the choice of parents. To ascertain whether this is important for our conclusion, we augment the above regression with an additional Heckman selection equation that models parental choice to stop at one child. This becomes a system of two Probit regressions. For the selection equation (which families are likely to stop at one child), we add the age of the first

---

[8] The ratio of boys and girls for first born children is close to being natural, whereas the ratio becomes progressively more skewed when one looks at second born children and third born children, respectively (Ebenstein, 2009).

child, a dummy for whether the first child has a disability, and whether the household head is a minority to the list of regressors represented in the main regression. The additional regressors are motivated by features of the national family planning policy. First, if a family decides to have a second child (in regions where a second child is allowed), the family planning policy requires the parents to wait for the first child to be at least five years old. Second, if the first child is disabled, then the family can have a second child (even in areas where normally only one child is allowed). Third, minority families are often exempted from birth quotas.

We report these regressions in Appendix Table A. With the Heckman selection equation, we still find that the local sex ratio has a positive and significant effect on the parental propensity to be entrepreneurs for son families. For daughter families, the sex ratio is insignificantly different from zero. The pattern holds for both the rural and the urban samples.

Approach 3: Placebo tests

We next turn to a placebo test. The basic idea is to examine the birth of new foreign invested firms, and to check if they are related to the local sex ratio imbalance. If the positive association between the local sex ratio and the growth in the number of domestic private firms is purely an artificial outcome of missing regressors that predict relative profitability across regions and happen to be correlated with the local sex ratio, we would expect to also find a similarly positive association between the growth in the number of foreign firms and the local sex ratio. On the other hand, if our theory is right that a higher sex ratio imbalance drives more risk-taking by local Chinese for a given level of growth opportunity, then the local sex ratio won't necessarily affect how foreign-invested firms choose to locate their production in China.

The placebo tests are reported in Table 9. In the first four regressions, the dependent variable is the growth in the number of foreign-invested firms from 1995 to 2004. The right-hand-side regressors are identical to those in Table 4. In none of the cases can we reject the null that the coefficient on the sex ratio variable is zero. In other words, statistically speaking, the location of new foreign-invested firms is uncorrelated with the local sex ratio imbalance. In Columns 5-8 of Table 9, we perform 2SLS regressions with the same set of instruments for the sex ratio as in Table 6. In all cases, the coefficient on the sex ratio is not statistically different from zero. Taken together, the placebo tests make it unlikely that the local sex ratio is a proxy for missing regressors that predict future profitability in a region. This bolsters our confidence in the

interpretation that a higher sex ratio imbalance stimulates more entrepreneurship.

Approach 4: Differential work effort at the household level

Not everyone can be an entrepreneur. However, if the desire to pursue wealth is strong enough, virtually anyone can make more money by working harder or longer. We now turn to a second household data that allows us to examine a household's supply of labor and willingness to accept a relatively dangerous job (in exchange for relatively good pay). The data comes from the Chinese Household Income Project (CHIP) of 2002, which covers 9,200 households in 122 rural counties in 22 provinces.

To make the households as comparable as possible, we construct a sub-sample of households with two living parents and a child.[9] This sub-sample consists of 480 families with a son and 262 families with a daughter in 122 rural counties. Since most unmarried young people live with their parents, the survey does not contain many unmarried young men or women as the household head. Therefore, we are not able to analyze single-person households directly.

We focus on two key aspects of a household's labor supply that are captured in the survey. The first is a household's willingness to accept a relatively dangerous (or unpleasant) job. A dangerous/unpleasant job is defined as one in the mining or construction sector, or one with exposure to extreme heat, extreme cold, or hazardous materials. While the survey does not contain occupation-specific wage information, we may expect that, in equilibrium, the wage rate is higher for a dangerous (or unpleasant) job than other jobs, holding constant skill requirement and other determinants of the wage. In other words, people presumably accept a more dangerous (or a less pleasant) job in exchange for higher pay.

The second variable that we look at is the total number of days in a year that members of a household worked off the farm (mostly as a migrant worker). Off farm work usually pays better, but one has to endure all the difficulties and inconvenience associated with working away from the hometown. Given the policy restrictions on internal migration in China, most migrant workers treat out-of-town jobs as temporary, do not expect to settle in the cities where they work, and likely return to their hometowns eventually.

The summary statistics on these two variables across the rural counties are reported in Table 10. The last panel (first row) indicates that, on average, 27.6% of all three-person

---

[9] We also place an age limit of 40 for household heads to exclude possibly uncounted adult children who have moved away.

households in a county have at least one family member working in a relatively dangerous job. The average fraction is only moderately higher for households with a son (27.7%) than households with a daughter (27.4%). However, the standard deviation across the counties (around 45%) is big. As for the total number of days members of a household worked off the farm, the unconditional average is 35.6 days per household. The son families worked significantly more days off farm (41.4 days) than the daughter families (24.9 days). From the summary statistics, we cannot rule out the possibility that the differences across the two types of households simply reflect a greater ability for a man to work away from home than for a woman. Our theory, however, implies a particular regional variation in the labor supply: son families are more willing to take a relatively dangerous job or work more days if they are located in a region with a more unbalanced sex ratio. Therefore, to test our theory, we have to explore interactions between a household's labor supply and the local sex ratio, while controlling for other determinants of labor supply.

In Table 11, we start with three Probit regressions on household propensity to accept a relatively dangerous job. They are for families with a son, families with a daughter, and a combined sample of families with a child. All the regressions control for family income, children's ages, and characteristics of the head of the household (age, education, and ethnic background). It also controls for health shocks to the family by a dummy denoting "poor health" if the family has a disabled or severely ill member. In Column 1, we focus on households with a son. The local sex ratio has a positive and significant coefficient, implying that a son-family is more willing to take a relatively dangerous job if it lives in a region with a higher sex ratio imbalance. An increase in the sex ratio by 4.3 basis points (which is equal to one standard deviation across the rural counties in the sample as reported in Table 10B) is associated with an increase in the probability for a son-family to accept a dangerous job by 4.1 percentage points (e.g., an increase from 20% to 24.2%). Column 2 of Table 11 looks at households with a daughter. The coefficient on the sex ratio is not statistically significant. In other words, the willingness to accept a dangerous job for a daughter-family is unrelated to the local sex ratio.

In the third column of Table 11, we combine the two sets of households and add a dummy for households with a son and an interaction term between the dummy and the local sex ratio. The local sex ratio is insignificant while the dummy for son-families has a negative coefficient. Most interestingly, the interaction term between the local sex ratio and the dummy

for son families is positive and statistically significant. Our interpretation is that it is not having a son per se that motivates families to be more willing to accept a dangerous job. Rather, it takes a combination of having a son and living in a region with a high sex ratio imbalance to induce families to be more eager to accept a relatively dangerous job.

One may wonder if the intensity of the work effort response to a given rise in the sex ratio depends on the household's initial income. In particular, do poorer households exert stronger effort? To see this, we classify all households into four brackets based on their income and use four dummies ("the poorest income quartile," "the second income quartile," and so on) to denote them. In the last three columns of 11, we interact the local sex ratio with each of the income quartile dummies. As one can see, there is no real difference across income groups. In particular, poor and rich households with a son all increase their willingness to accept a dangerous or less pleasant job equally to a given rise in the sex ratio. Similarly, households with a daughter in all income quartiles are equally insensitive to the sex ratio in their willingness to accept dangerous jobs.

Because this data set records both family income and expenditure, we can investigate the role of savings in the propensity to accept an intrinsically unpleasant job. To reduce the influence of outliers, we define household savings as log(income/expenditure) and add it to the list of control variables. The resulting regressions are reported in Table 11A. While the savings rate appears to be positively correlated with the daughter families' propensity to take an unpleasant job, it does not alter the basic conclusion that a combination of having a son and living in a region with a skewed sex ratio motivates a family to be more willing to take an unpleasant job.

We now turn to another dimension of the rural household labor supply, namely the willingness to tolerate the hardship and inconvenience of working away from home village (mostly as a migrant worker.) Table 12 performs Tobit estimations on the total number of days in the year before the survey that household members worked off farm. In the first column, we look at households with a son. The coefficient on the local sex ratio is positive and statistically significant. An increase in the sex ratio by 4.3 basis points is associated with an increase in the supply of off farm labor by 89.4 days/year (=20.8x4.3). Since the unconditional mean in the sample is 35.6 days per year per household (Column 5, second row, of Table 10A), this represents a huge effect. In the second column of Table 12, we look at households with a

daughter. The coefficient on the sex ratio is not statistically significant. This implies that the supply of off-farm labor by daughter families is uncorrelated with the sex ratio. In the third column of Table 12, we combine the two sets of households, and add a dummy for son families and an interaction term between the dummy and the sex ratio. Similar to Table 11, only the interaction term is positive and statistically significant. In other words, a combination of having a son and living in a region with a high sex ratio motivates these households to be more willing to work away from home.

We again check if the supply of labor in response to a higher sex ratio varies by the income level of households. This is done in the last three columns of Table 12. It turns out that there is no statistical difference across income groups. We also investigate the consequence of adding household savings rate as a control and report the results in Table 12A. The basic conclusion remains the same: a combination of having a son at home and living in a region with a skewed sex ratio makes a household more willing to endure the hardship of having a job away from the home village.

Over all, the patterns in Tables 11, 11A, 12 and 12A are consistent with each other, and consistent with our hypothesis. Of course, accepting a relatively dangerous job and working more days away from home are not mutually exclusive. Taken together, the estimation results suggest that, as the sex ratio imbalance increases, son families respond by increasing both the number of days in off-farm work and the willingness to accept a relatively dangerous job, presumably in pursuit of higher pay. There is no effect of a higher sex ratio on the work effort of daughter families.

General Equilibrium Effect: Sex ratios and per capita GDP growth

So far, we have discussed evidence on how a higher sex ratio stimulates the extensive margin of economic growth in the form of the birth of new private firms, and have also presented some evidence on how it increases the intensive margin of economic growth in the form of a greater supply of work effort and a greater tolerance of hardship and a hazardous work environment. There can be other partial equilibrium effects of a higher sex ratio; some may be negative (e.g., a higher crime rate) and some may be positive (e.g., more creativity). To capture the general equilibrium effect, we now examine the overall relationship between sex ratios and income growth by using panel data on provincial GDP per capita from 1980 to 2005. We

organize the data into five 5-year periods, 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05. Let $y(k, t)$ be the log GDP per capita for province k in period t. We run the following regression:

$$[y(k, t+5)-y(k, t)]/5 = \beta \, sr(k,t) + X(k,t)\Gamma + \text{province fixed effects} + \text{period fixed effects} + e(k,t)$$

where the dependent variable is the average annual growth rate in a 5-year period, $sr(k,t)$ is the sex ratio for the age cohort 5-19 in province k and period t (inferred from the 2000 Population Census), and $X(k,t)$ is a vector of control variables which includes the beginning-of-period log income, $y(k,t)$, the share of working age population in local population, the ratio of local investment to local GDP, the ratio of local foreign trade to local GDP, and birth rate. β is a scalar parameter and Γ is a vector of parameters to be estimated, and $e(k,t)$ is an error term that is assumed to be independent and identically normally distributed. The choice of the control variables is based on the set of robust predictors of growth from the empirical growth literature (Barro and Sala-i-Martin, 1992; Li and Zhang, 2007). One key missing regressor is human capital, of which we do not have a good measure that is both across provinces and over time. We will implement a 2SLS estimation that aims to address this (and other) problems.

Some summary statistics for the panel are reported in Table 13. During 1980-2005, the average annual growth rate of per capita GDP across the provinces was 8.8% with a standard deviation of 2.7%. The average sex ratio in 1980 was 107 boys/100 girls (only slightly higher than the normal ratio), but there were already variations across the provinces with the standard deviation being 3.5 and the maximum ratio being 114 boys/100 girls. As indicated earlier, the sex ratio deteriorates over time.

The panel growth regressions with both province fixed effects and period fixed effects are reported in Table 14. In Columns 1-3, we use the initial sex ratio for the age cohort of 5-19 within a five-year interval. The first column includes the sex ratio, initial income, the ratio of investment to local GDP, population birth rate, and both the province fixed effects and the period fixed effects. The coefficient on the sex ratio is positive and significant: on average, income growth is faster in regions/periods with a higher initial sex ratio. The coefficients on the first two control variables are consistent with the standard growth regressions. In particular, poor regions tend to grow faster; regions with heavier investment also tend to grow faster. The coefficient on the birth rate is negative (consistent with the Malthusian idea) but statistically insignificant.

In Column 2, we expand the list of controls to include share of working age adults in the local population and a measure of trade openness. Neither of the new regressors is significant. The coefficient on the sex ratio remains positive and significant. In Column 3, we add the savings rate as a control. Given the findings in Wei and Zhang (2009), we wish to find out if the sex ratio has a positive effect on economic growth beyond raising the household savings rate. The positive and significant coefficient on the savings variable implies that higher savings rate and higher income growth tend to go together. However, holding constant the local savings rate, the coefficient on the sex ratio is still positive and significant. This implies that a higher savings rate is not the exclusive conduit for a higher sex ratio to affect economic growth.

In Columns 4-6, we use the average sex ratio during a 5-year period (instead of the beginning-of-period value) but otherwise replicate the previous three regressions. In all cases, the coefficients on the sex ratio are positive and statistically significant. This is consistent with the idea that a higher sex ratio is associated with a higher growth rate.

One might worry about endogeneity of the sex ratio, birth rate, and savings rate, and possible correlation between the initial income and the error term in the dynamic panel specification. In addition, measurement errors for the sex ratio variable can also be present. To deal with these concerns, we implement a generalized method of moments (GMM) estimation. In addition to the standard instruments (i.e. lagged regressors) in a GMM setting, we utilize some possibly exogenous instruments as discussed in Table 5.

The results are reported in Table 15. In the first two columns, we report GMM results with no exogenous IVs. In the middle two columns, we add the two penalties for violating birth quotas in the set of instruments. In the last two columns, we further expand the set of instruments to include the share of local population not subject to birth quotas. For robustness, in each case, we use both initial sex ratio and average sex ratio.

All six coefficients on the sex ratio variable are positive and statistically significant. Strikingly, the point estimates are also stable across the three sets of IVs. To understand the economic significance of the estimates, we take the point estimate in Column 2 (0.51) at the face value. An increase in the sex ratio by 4 basis points (which is the level of increase from 2000 to 2005 according to Table 13), holding other variables constant, would raise the growth rate by 2.04 percentage points per annum (= 0.51x 0.04 x 100). This accounts for about 20% (=2.04/10.2) of the actual mean increase in the annual income growth during this period. This means that the

effect of the sex ratio is economically significant. Because the sex ratio for the pre-marital age cohort is projected to be higher over the next decade, and because the "natural" growth rate expected from the convergence force in the Solow model will decline, the relative importance of the sex ratio effect on economic growth is likely to rise in the medium term.

4. Concluding Remarks

Robert M. Solow, the Nobel Prize winner for his pioneering work on the theory of economic growth, once said, "Everything reminds Milton (Friedman) of the money supply. Well, everything reminds me of sex, but I keep it out of the paper."[10] Well, Solow might have missed something economically significant by not linking sex with economic growth. This paper proposes that an unbalanced sex ratio may be one of the significant drivers for economic growth.

A strong sex ratio imbalance in the pre-marital age cohort has emerged in China, Vietnam, Korea, India, Taiwan, Singapore and several other economies since the early part of this century due to a combination of a parental preference for sons and the increasing availability of technology (ultrasound B machines in particular) to screen the sex of a fetus. In the Chinese case, a family planning policy that places a strict limit on the number of children a family can have has exacerbated a desire to engage in sex-selective abortions. As men face a diminishing prospect of finding a wife, parents of a son or the son himself are more eager to do something to improve his standing in the marriage market relative to other men in the same cohort. Since wealth is a significant determinant of one's relative standing, this produces a powerful additional incentive to create wealth. Parents with a son and men, in particular, respond to a rise in the sex ratio by engaging in more entrepreneurial activities, supplying more labor, and becoming more willing to take unpleasant or dangerous jobs, all in pursuit of higher expected pay.

We find strong supportive evidence across regions and households in China. Using data from two censuses of industrial firms in 1995 and 2004, we find that the local sex ratio is a significant predictor of which regions are more likely to have new domestic private firms (beyond other determinants of the birth of new firms). The economic impact is also significant: an increase in the sex ratio by one standard deviation can potentially explain 50% of the difference in the rates of growth of new private firms across regions. Across households, we find

---

[10] "The Concise Encyclopedia of Economics: Robert Merton Solow (1924-)." http://www.econlib.org/library/Enc/bios/Solow.html.

that a combination of having a son and living in a region with a skewed sex ratio raises the likelihood for parents to be business owners or self-employed. We also find that families with a son respond to a higher sex ratio by increasing both the number of days that they work off farm (mostly as migrant workers) and their willingness to take a relatively dangerous job, presumably in exchange for higher pay. Households with a daughter do not respond to a higher sex ratio in the same way. These patterns are consistent with our story.

In the long run, the sex ratio has to be endogenous. For example, if the sex ratio reaches 150 boys per 100 girls, parents would very likely re-evaluate their preference for sons. However, very few regions have so far shown signs of a reversal. This means that if the sex ratio follows a mean-reverting process, the speed of reversion must be slow. In any case, we know with a high degree of confidence that the sex ratio for the pre-marital age cohort in China will be getting worse in the medium term, since the sex ratio at birth today is significantly worse than the ratio for today's 10-year-olds, which in turn is worse than the ratio for today's 20-year-olds. This means that the effect of the sex ratio imbalance on growth will continue to be a force to reckon with in the foreseeable future. This will partially offset the natural force of a declining growth rate that one may expect from the standard convergence story.

Accumulating more wealth is not the only way for men or households with a son to compete in the marriage market. Parents may also invest more in the education of their sons, and push them to work harder in school. There may also be a spillover from a boy's education to a girl's education. Such mechanisms have not been empirically investigated. In addition, as noted earlier, several other economies also have a strong sex ratio imbalance. Some of them are also known to have a high rate of economic growth. We leave a rigorous investigation of these topics to future research.

References:

Angrist, Joshua, 2002, "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *Quarterly Journal of Economics*, 117 (3): 997-1038.

Barro, Robert J. and Xavier Sala-i-Martin, 1992, "Convergence," *Journal of Political Economy*, 100 (2): 223-251.

Barro, Robert J. and Xavier Sala-i-Martin, 2004, *Economic Growth*, MIT Press.

Bhaskar, V, 2011, "Parental Sex Selection and Gender Balance," *American Economic Journal: Microeconomics*, 3(1): 214-44.

Blanchard, Olivier J. and Francesco Giavazzi, 2005, "Rebalancing Growth in China: A Three-Handed Approach." MIT Department of Economics Working Paper No. 05-32.

Bulte, Erwin, Nico Heerink, and Xiaobo Zhang, 2011, "China's One-Child Policy and 'the Mystery of Missing Women': Ethnic Minorities and Male-Biased Sex Ratios," *Oxford Bulletin of Economics and Statistics*, 73 (1): 21-39.

Charles, Kerwin K., and Ming Ching Luoh, 2010, "Male Incarceration, the Marriage Market, and Female Outcomes," *The Review of Economics and Statistics*, 92(3): 614-627.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite, 1992, "Social Norms, Savings Behavior, and Growth," *Journal of Political Economy*, 100(6): 1092-1125.

Corneo, Giacomo and Olivier Jeanne, 1999, "Social Organization in an Endogenous Growth Model," *International Economic Review*, 40(3): 711-25.

Den Boer, Andrea, and Valerie M. Hudson, 2004, "The Security Threat of Asia's Sex Ratios," *SAIA Review* 24(2): 27-43.

Du, Qingyuan, and Shang-Jin Wei, 2010, "A Sexually Unbalanced Model of Current Account Imbalances," NBER working paper 16000.

Du, Qingyuan, and Shang-Jin Wei, 2011a, "Sex Ratios and Exchange Rates," working paper, Columbia University.

Du, Qingyuan, and Shang-Jin Wei, 2011b, "Sex Ratios and Entrepreneurship," working paper, Columbia University.

Gu, Baochang and Krishna Roy, 1995, "Sex Ratio at Birth in China, with Reference to Other Areas in East Asia: What We Know," *Asia-Pacific Population Journal*, 10(3): 17-42.

Das Gupta, Monica, 2005, "Explaining Asia's 'Missing Women': A New Look at the Data," *Population and Development Review* 31(3): 529-535.

Ebenstein, Avraham, 2009, "Estimating a Dynamic Model of Sex Selection in China," Harvard University. Available from pluto.huji.ac.il/~ebenstein/Ebenstein_Sex_Selection_March_09.pdf.

Edlund, Lena, 1996, "Dear Son – Expensive Daughter: Do Scarce Women Pay to Marry?" Essay Two in *The Marriage Market: How Do You Compare*. Dissertation, Stockholm School of Economics.

Edlund, Lena, and Chulhee Lee, 2009. "Son Preference, Sex Selection and Economic Development: Theory and Evidence from South Korea," Discussion Papers 0910-04, Columbia University, Department of Economics.

Edlund, Lena, Hongbin Li, Junjian Yi, and Junsen Zhang, 2007, "Sex Ratios and Crime: Evidence from China's One-Child Policy," IZA DP no. 3214.

Frank, Robert H., 1985, "The Demand for Unobservable and Other Nonpositional Goods," *American Economic Review*, 75 (1): 101-116.

Frank, Robert H., 2004. "Positional Externalities Cause Large and Preventable Welfare Losses," *American Economic Review*, 95 (2): 137-141.

Gale, D., and L. S. Shapley, 1962, "College Admissions and the Stability of Marriage", *American Mathematical Monthly*, 69: 9-14.

Greenspan, Alan, 2009, "The Fed didn't cause the housing bubble," *The Wall Street Journal*, March 11.

Guilmoto, Christophe Z. 2007, "Sex-ratio imbalance in Asia: Trends, consequences and policy responses." Available from: www.unfpa.org/gender/docs/studies/summaries/regional_analysis.pdf.

Hopkins, Ed, and Tatiana Kornienko, 2004, "Running to Keep in the Same Place: Consumer Choice as a Game of Status," *American Economic Review* 94(4): 1085-1107.

Hopkins, Ed, and Tatiana Kornienko, 2009, "Which Inequality? The Inequality of Endowments versus the Inequality of Rewards," University of Edinburgh working paper.

Huang, Yasheng, 2009, *Capitalism with Chinese Characteristics*, Cambridge University Press.

International Monetary Fund, 2009, *World Economic Outlook*, April. Washington DC: International Monetary Fund.

Li, Hongbin and Junsen Zhang, 2007, "Do High Birth Rates Hamper Economic Growth?" *Review of Economics and Statistics* 89: 110-117.

Li, Hongbin, Zheyu Yang, Xianguo Yao, Junsen Zhang, 2009, "Entrepreneurship and Growth: Evidence from China," Discussion Papers 00022, Chinese University of Hong Kong.

Li, Hongbin, and Hui Zheng, 2009, "Ultrasonography and Sex Ratios in China," *Asian Economic Policy Review*, 4(1): 121-137.

Li, Shuzhuo, 2007, "Imbalanced Sex Ratio at Birth and Comprehensive Interventions in China," paper presented at the 4th Asia Pacific Conference on Reproductive and Sex Health and Rights, Hyderabad, India, October 29-31.

Lin, Ming-Jen and Ming-Ching Luoh, 2008, "Can Hepatitis B Mothers Account for the Number of Missing Women? Evidence from Three Million Newborns in Taiwan," *American Economic Review*, 98(5): 2259–73.

Mechoulan, Stephane, 2011, "The External Effects of Black-Male Incarceration on Black Females," unpublished working paper, Dalhousie University.

Mei Fong, 2009, "It's Cold Cash, Not Cold Feet, Motivating Runaway Brides in China," *Wall Street Journal*, June 5.

Modigliani, Franco, 1970, "The Life Cycle Hypothesis of Saving and Intercountry Differences in the Saving Ratio," Introduction, *Growth and Trade, Essays in Honor of Sir Roy Harrod*, W.A. Elits, M.F. Scott, and J.N. Wolfe, eds. Oxford University Press.

Modigliani, Franco, and Shi Larry Cao, 2004, "The Chinese Saving Puzzle and the Life Cycle Hypothesis," *Journal of Economic Literature, 42: 145-170.*

Mu, Ren and Xiaobo Zhang, 2011, "Why Does the Great Chinese Famine Affect the Male and Female Survivors Differently? Mortality Selection versus Son Preference," *Economics and Human Biology*, 9: 92–105.

National Bureau of Statistics, 1988, *China Population Statistical Material Compilation 1949-1985*. Beijing: China Finance and Economic Press.

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

Oster, Emily, Gang Chen, Xinxen Yu, and Wenyao Lin, 2008, "Hepatitis B Does Not Explain Male-Biased Sex Ratios in China," University of Chicago working paper.

Qian, Nancy, 2008, "Missing Women and the Price of Tea in China: The Effect of Sex-specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 123(3): 1251-1285.

Rajan, Raghuram G., 2010, *Fault Lines: How Hidden Fractures Still Threaten the World Economy*, Princeton University Press.

Roth, A. E., and Sotomayor, M. A. O., 1990, *Two-sided Matching: A Study in Game-theoretic Modeling and Analysis*, Cambridge University Press.

Roussanov, Nikolai, 2010, "Diversification and Its Discontents: Idiosyncratic and Entrepreneurial Risk in the Quest for Social Status," *Journal of Finance* 65: 1755-1788

Roussanov, Nikolai, and Pavel Savor, 2011, "Status, Marriage, and Managers' Attitudes to Risk," Unpublished working paper, The Wharton School, University of Pennsylvania.

Sachs, Jeffrey and Wing Thye Woo, 1997, "Understanding China's Economic Performance," NBER working paper 5935.

Scharping, Thomas, 2003, *Birth Control in China 1949-2000,* New York, NY: Routledge Curzon.

Wei, Shang-Jin, and Xiaobo Zhang, 2009, "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," NBER working paper 15093, and forthcoming, *Journal of Political Economy*.

Wu, Xiaobo, 2007, *Thirty Years of China Business,* China CITIC Press, in Chinese.

Zhu, Wei Xing, Li Lu, and Therese Hesketh, 2009, "China's excess males, sex selective abortion, and one child policy: analysis of data from 2005 national intercensus survey," *British Medical Journal,* 2009 (338): 1211.

Table 1A: Summary Statistics for Variables Used in County Level Analysis

	Mean	Median	Standard deviation
Left hand variables			
# of private firms in 1995 per 10,000 people	2.80	2.20	3.68
# of private firms in 2004 per 10,000 people	6.82	4.34	8.76
log(# of private firms in 2004 / # of private firms in 1995)	0.83	0.80	0.84
# of SOEs in 1995 per 10,000 people	0.73	0.55	0.88
# of SOEs in 2004 per 10,000 people	0.21	0.12	0.29
log(# of SOEs in 2004 / # of SOEs in 1995)	-1.42	-1.39	0.82
# of foreign firms in 1995 per 10,000 people	0.16	0.03	0.57
# of foreign firms in 2004 per 10,000 people	0.34	0.04	1.34
log(# of foreign firms in 2004 / # of foreign firms in 1995)	0.52	0.47	0.87
Right hand variables			
Sex ratio for the age cohort 5-19 in 1995 inferred from 1990 census	107.55	106.74	4.61
Sex ratio for the age cohort 4-18 in 2004 inferred from 2004 census	111.24	109.44	7.58
Averaged sex ratio of 1995 and 2004	109.40	108.25	5.59
GDP in 1995 (million yuan)	2061	1174	2748
Average year of schooling based on 2000 census	6.90	7.21	1.26
Share of agricultural output in gross output values in 1995	0.44	0.41	0.24
The ratio of local revenues to total government employees in 1995 (yuan/person)	4355	4126	2420
Population growth from 1990 to 2000	0.06	0.05	0.18
Share of labor force (aged 20-64) in total population in 1995	0.62	0.62	0.04
Share of population aged 65 and above	0.08	0.08	0.02
Residential bank deposit per capita in 1995 (yuan)	1516	1096	1691
Instrumental variables			
Share of minority population in 1990	0.20	0.01	0.32
Average penalties for family planning violations	1.00	1.06	0.18
Average extra penalty for higher order births	0.42	0.33	0.28

Note: Authors' calculation. The sex ratio, population growth, share of labor force, and year of schooling variables are based on China Population Censuses in 1990 or 2000. Because the available age data at the county level are in five-year interval, the cohort aged 0-14 in China Population Censuses in 1990 and 2000 would be 5-19 and 4-18 by 1995 and 2004. Local GDP, gross output values, revenues and total government employees in 1995 are from China Local Public Finance Statistical Yearbook 1995. Residential bank deposits are from Chinese County Database on Social Economic Statistics.

Table 1B: Summary Statistics for Variables Used at the City Level Analysis

	Mean	Median	Standard deviation
Left hand variables			
# of private firms in 1995 per 10,000 people	10.40	6.32	16.24
# of private firms in 2004 per 10,000 people	21.42	9.73	35.98
log(# of private firms in 2004 / # of private firms in 1995)	0.91	0.85	0.69
# of SOEs in 1995 per 10,000 people	2.26	1.76	2.37
# of SOEs in 2004 per 10,000 people	0.52	0.39	0.42
log(# of SOEs in 2004 / # of SOEs in 1995)	-1.11	-1.08	0.64
# of foreign firms in 1995 per 10,000 people	1.31	0.37	2.76
# of foreign firms in 2004 per 10,000 people	2.02	0.37	4.81
log(# of foreign firms in 2004 / # of foreign firms in 1995)	0.52	0.45	0.79
Right hand variables			
Sex ratio for the age cohort 5-19 in 1995 inferred from 1990 census	108.10	107.12	4.45
Sex ratio for the age cohort 4-18 in 2004 inferred from 2004 census	111.42	109.18	6.69
Averaged sex ratio of 1995 and 2004	109.76	108.30	4.99
Industrial GDP in 1995 (million yuan)	80601	37939	118000
Average year of schooling based on 2000 census	6.75	6.73	0.99
Population growth from 1990 to 2000	0.35	0.22	0.62
Share of labor force (aged 20-64) in total population in 1995	0.68	0.70	0.04
Share of population aged 65 and above	0.08	0.08	0.02
Instrumental variables			
Share of minority population in 1990	0.20	0.01	0.32
Average penalties for family planning violations	1.00	1.06	0.18
Average extra penalty for higher order births	0.42	0.33	0.28

Note: Authors' calculations. The sex ratio, population growth, share of labor force, and year of schooling variables are based on China Population Censuses in 1990 or 2000. Because the available age data at the county level are in five-year interval, the cohort aged 0-14 in China Population Censuses in 1990 and 2000 would be 5-19 and 4-18 by 1995 and 2004. Industrial GDP is from China Industrial Census 1995.

Table 2: Contributions to the growth of Chinese industrial output by ownership

		Private	SOEs	Foreign	total
1	Value added in 1995 (billion yuans)	2956	5470	1210	9636
2	Value added in 2004 (billion yuans)	21350	4790	9080	35220
3	growth rate =(2)/(1)-1	6.22	-0.12	6.50	2.66
4	Initial share in total VA (%)	30.7	56.8	12.6	100.0
5	**Share in growth rate (%)**				
	= [(3)*(4)]/Growth of total VA	71.9	-2.7	30.8	100.0

Table 3: The extensive versus intensive margins in the growth of the Chinese private sector

		Number of firms (a)	Average output (million yuans/firm) (b)	Total VA (billion 1995 yuans) (c)
1	1995	807821	3.66	2956
2	2004	2549888	6.20	15815
3	Growth rate = log(row 2)-log(row 1)	0.499	0.229	0.728
4	Share in growth (%)	68.5	31.5	100

Note: The cumulative CPI inflation rate over 1995-2004 is 35%.

Table 4: Sex Ratios and the Growth in the Number of Private Firms at the County Level from 1995 to 2004

	Sex ratio for the age cohort in 1995					Sex ratio for the average age cohort between 1995 and 2004				
	5-19	5-19	5-9	10-14	15-19	5-19	5-19	5-9	10-14	15-19
Initial sex ratio	0.019**	0.015**	0.008**	0.009**	0.004	0.021**	0.017**	0.012**	0.016**	0.011**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)
Log number of firms in 1995	-0.498**	-0.576**	-0.572**	-0.574**	-0.568**	-0.521**	-0.594**	-0.602**	-0.599**	-0.574**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Log GDP in 1995	0.294**	0.530**	0.526**	0.538**	0.539**	0.291**	0.524**	0.518**	0.525**	0.536**
	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Average year of schooling based on 1990 census	-0.066**	-0.170**	-0.171**	-0.172**	-0.176**	-0.052*	-0.157**	-0.162**	-0.157**	-0.166**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Share of agricultural output in gross output values in 1995	-0.893**	-0.902**	-0.915**	-0.902**	-0.898**	-0.897**	-0.903**	-0.914**	-0.905**	-0.893**
	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)	(0.12)	(0.11)	(0.11)	(0.11)	(0.11)
The ratio of local revenues to total government employees (log) in 1995	0.098**	0.002	0.003	0.001	0.001	0.095**	0.001	0.007	0	0
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Population growth from 1990 to 2000	0.420**	0.177*	0.182*	0.172*	0.165*	0.417**	0.178*	0.171*	0.179*	0.170*
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)
Share of labor force (aged 20-64) in total population in 1995	-0.469	-0.699	-0.838*	-0.69	-0.737	0.087	-0.237	0.068	-0.04	-0.668
	(0.50)	(0.45)	(0.46)	(0.46)	(0.47)	(0.51)	(0.46)	(0.47)	(0.48)	(0.46)
Share of population aged 65 and above in total population in 1995	13.175**	11.285**	11.485**	11.417**	11.480**	14.006**	12.032**	12.300**	12.239**	11.587**
	(1.36)	(1.28)	(1.28)	(1.29)	(1.29)	(1.38)	(1.29)	(1.28)	(1.29)	(1.29)
Log local household savings in 1995		0.402**	0.401**	0.405**	0.406**		0.396**	0.388**	0.398**	0.405**
		(0.04)	(0.04)	(0.04)	(0.04)		(0.04)	(0.04)	(0.04)	(0.04)
Adjusted R square	0.32	0.39	0.39	0.38	0.38	0.33	0.39	0.39	0.39	0.38
AIC	3786	3548	3545	3560	3565	3767	3534	3522	3530	3560
N	1790	1764	1764	1764	1764	1790	1764	1764	1764	1764

Notes: The growth in the number of private firms is measured by the increase in the log number of firms from 1995 to 2004. The definition of private firms includes township-and-village enterprises and other "collectively owned" firms. The sex ratio for the age cohort 5-19 in 1995 is inferred from the age cohort 0-14 in the 1990 population census; the sex ratio for the age cohort 4-18 in 2004 is inferred from the age cohort 0-14 in the 2000 population census. Local household savings is measured by log(residential bank deposits/GDP). Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively.

Table 5: Instrumenting for the Local Sex Ratio - First Stage Regressions at the County Level

	R1	R2	R3	R4	R5	R6
	Initial sex ratio			Average sex ratio		
Share of minorities in local population	-2.55**	-2.39**		-3.97**	-3.76**	
	(0.72)	(0.72)		(0.83)	(0.83)	
Penalties for family planning violations	3.59**	3.15**	3.55**	1.44	0.970	1.59
(averaged over the period when the 5-19 cohort was born)	(1.21)	(1.24)	(1.17)	(1.07)	(1.13)	(1.09)
Dummy for extra penalty for higher order births	3.07**	2.89**	3.51**	4.33**	4.15**	5.12**
(averaged over the period when the 5-19 cohort was born)	(0.57)	(0.59)	(0.50)	(0.71)	(0.74)	(0.67)
Initial number of firms (log) in 1995	0.25*	0.26*	0.43**	0.96**	0.92**	1.20**
	(0.13)	(0.13)	(0.13)	(0.17)	(0.18)	(0.17)
Log GDP in 1995	0.41**	0.52**	0.56**	0.66**	0.94**	1.01**
	(0.15)	(0.17)	(0.18)	(0.19)	(0.22)	(0.22)
Average year of schooling inferred from 2000 census	-0.49**	-0.50**	-0.38**	-1.08**	-1.19**	-1.00**
	(0.12)	(0.13)	(0.13)	(0.16)	(0.18)	(0.18)
Share of agricultural output in gross output values in 1995	-0.30	0.05	-0.03	0.14	0.38	0.26
	(0.76)	(0.70)	(0.71)	(0.78)	(0.75)	(0.76)
The ratio of local revenues to total government employees (log)	-0.01	-0.02	-0.03	0.04	-0.08	-0.10
	(0.17)	(0.17)	(0.18)	(0.22)	(0.22)	(0.23)
Growth rate of local population	-0.22	-0.25	-0.42	0.28	0.08	-0.20
(=increase in log population)	(0.43)	(0.45)	(0.43)	(0.53)	(0.57)	(0.56)
Share of labor force (aged 20-64) in local population in 1995	-3.43	-4.27	-3.33	-29.59**	-30.56**	-29.08**
	(5.99)	(6.02)	(5.87)	(4.31)	(4.33)	(4.32)
Share of population aged 65 and above in total population in 1995	5.85	11.62*	15.87**	-38.18**	-36.16**	-29.48**
	(8.35)	(6.83)	(6.59)	(8.77)	(8.51)	(8.35)
Log local household savings in 1995		0.15	0.17		0.47*	0.51**
		(0.19)	(0.19)		(0.24)	(0.25)
Adjusted R square	0.10	0.11	0.10	0.19	0.19	0.18
F Statistic	29.64	26.96	23.6	39.73	36.65	35.1
Kleibergen-Paap Wald F statistic	23.16	20.03	19.14	38.13	34.76	32.17
N	1790	1764	1764	1790	1764	1764

Notes: The dependent variables in R1, R2, and R3 are the sex ratio for age cohort aged 5-19 in 1995, while those in R4, R5, and R4 are the average sex ratio for the age cohort 5-19 in 1995 and 4-18 in 2004. The share of minorities in local population is computed from the 1990 population census at the county level. The two financial penalty variables are only available at the province level and we use their mean values over the same period of the sex ratio for the age cohort 5-19 in 1995. Robust standard errors are in parentheses. The dummy for coastal regions defines Beijing, Liaoning, Tianjin, Hebei, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong and Guangxi as coastal regions. * and ** denote statistically significant at the 10% and 5% levels, respectively. For the Kleibergen-Paap Wald F test for weak instruments, the Stock-Yogo critical values are: 5% maximal IV relative bias 13.91; 10% maximal IV relative bias 9.08; 20% maximal IV relative bias 6.46; 10% maximal IV size 19.93; 15% maximal size 11.59; 20% maximal size 8.75; 25% maximal IV size 7.25.

Table 6: 2SLS Estimation on Sex Ratios and Growth in the Number of Private Firms at the County Level

	R1	R2	R3	R4	R5	R6	R7	R8
	Three IVs	Two IVs	Three IVs	Two IVs	Three IVs	Two IVs	Three IVs	Two IVs
Initial sex ratio for the age cohort 5-19 in 1995	0.20**	0.19**	0.20**	0.19**				
	(0.03)	(0.04)	(0.03)	(0.04)				
Average sex ratio (1995 and 2004)					0.13**	0.13**	0.13**	0.13**
					(0.02)	(0.02)	(0.02)	(0.02)
Log number of firms in 1995	-0.61**	-0.60**	-0.68**	-0.67**	-0.71**	-0.70**	-0.77**	-0.76**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
Log GDP in 1995	0.23**	0.23**	0.43**	0.44**	0.24**	0.24**	0.44**	0.44**
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Average year of schooling of local population	-0.01	-0.02	-0.10**	-0.11**	0.05	0.05	-0.04	-0.04
(inferred from the 2000 census)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Share of agriculture in gross output values in 1995	-0.78**	-0.79**	-0.87**	-0.87**	-0.85**	-0.86**	-0.90**	-0.90**
	(0.17)	(0.17)	(0.16)	(0.16)	(0.14)	(0.14)	(0.13)	(0.13)
The ratio of local government revenue to	0.09*	0.09*	0.00	0.00	0.07*	0.07*	-0.01	-0.01
Log government employees	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Increase in log population	0.55**	0.54**	0.30**	0.29**	0.47**	0.47**	0.25**	0.24**
	(0.13)	(0.13)	(0.14)	(0.13)	(0.11)	(0.10)	(0.11)	(0.11)
Share of labor force (aged 20-64)	0.46	0.41	0.34	0.28	3.56**	3.47**	3.29**	3.17**
in local population in 1995	(1.13)	(1.13)	(1.11)	(1.13)	(0.83)	(0.93)	(0.80)	(0.90)
Share of population aged 65 and above	10.55**	10.69**	7.58**	7.78**	16.98**	16.90**	14.77**	14.68**
in total local population in 1995	(2.02)	(1.95)	(1.87)	(1.82)	(1.66)	(1.71)	(1.62)	(1.67)
Log local household savings in 1995			0.35**	0.36**			0.33**	0.34**
			(0.05)	(0.05)			(0.04)	(0.04)
Adjusted R square	-0.56	-0.46	-0.46	-0.37	-0.14	-0.12	-0.07	-0.04
AIC	5265	5157	5073	4964	4710	4669	4519	4469
Durbin-Wu-Hausman test for endogeneity	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Hansen's J statistic for over identification	0.57	0.55	0.25	0.30	0.41	0.20	0.41	0.20
N	1790	1764	1764	1790	1764	1764	1764	1764

Notes: The share of minority population in 1990 at the county level is excluded as an instrument variable in regressions R2, R4, R6, and R8. Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively.

Table 7: Sex Ratios and Growth in Number of Private Firms in the Urban Sample

	OLS		2SLS	
Sex ratio for average cohort 5-19 in 1995 and 4-18 in 2004	0.03**	0.03**	0.11**	0.12**
	(0.01)	(0.01)	(0.03)	(0.04)
Log number of firms in 1995	-0.51**	-0.51**	-0.46**	-0.45**
	(0.07)	(0.07)	(0.07)	(0.08)
Log industrial GDP in 1995	0.47**	0.48**	0.43**	0.45**
	(0.06)	(0.06)	(0.06)	(0.06)
Average year of schooling based on 1990 census	-0.02	-0.01	-0.02	0.01
	(0.03)	(0.04)	(0.03)	(0.05)
Population growth from 1990 to 2000	0.15*	0.16**	0.08	0.08
	(0.08)	(0.08)	(0.10)	(0.10)
Share of labor force (aged 20-64) in total population in 1995	-0.02	-0.45	2.720	2.37
	(1.39)	(1.48)	(1.91)	(2.06)
Share of population aged 65 and above in total population in 1995	11.49**	9.77**	15.25**	13.83**
	(3.03)	(3.48)	(3.64)	(4.03)
Log local household savings in 1996		-0.11*		-0.11
		(0.07)		(0.08)
Adjusted R square	0.39	0.38	0.26	0.22
AIC	389	361	436	413
Durbin-Wu-Hausman test for endogeneity			0.01	0.01
Hansen's J statistic for over identification			0.59	0.57
N	238	219	238	219

Notes: The growth in the number of private firms is measured by the increase in the log number of firms from 1995 to 2004 at the city level. The sex ratio for the age cohort 5-19 in 1995 is inferred from the age cohort 0-14 in the 1990 population census; the sex ratio for the age cohort 4-18 in 2004 is inferred from the age cohort 0-14 in the 2000 population census. The definition of private firms includes township-and-village enterprises and other "collectively owned" firms. Local household savings refers to log(residential bank deposits/GDP in 1996) (no 1995 data for the urban sample is available). Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively.

Table 8: Probit Estimation on Who Are More Likely to Be Entrepreneurs

	Rural nuclear families			Urban nuclear families		
	One or two children	one son	one daughter	One or two children	one son	one daughter
Local sex ratio for age cohort 5-19	0.26	0.44**	0.08	-0.10	0.47**	0.10
	(0.22)	(0.17)	(0.20)	(0.19)	(0.20)	(0.21)
Sex ratio*dummy for first child being a son	0.21*			0.55**		
	(0.12)			(0.22)		
Log household wealth	0.19**	0.18**	0.16**	0.03**	0.03*	0.02*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Year of education	0.20**	0.24**	0.29**	0.23**	0.24**	0.24**
	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Year of education (squared)	-0.99**	-1.26**	-1.47**	-1.58**	-1.72**	-1.61**
	(0.11)	(0.13)	(0.11)	(0.09)	(0.12)	(0.10)
Household head age	0.00	0.00	-0.01**	0.01**	0.01**	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Dummy for female household head	0.23**	0.24**	0.15**	-0.03	(0.01)	-0.06*
	(0.03)	(0.05)	(0.05)	(0.02)	(0.03)	(0.03)
Dummy for poor health by family members	-0.28**	-0.33**	-0.25**	-0.34**	-0.31**	-0.17
	(0.04)	(0.08)	(0.08)	(0.09)	(0.15)	(0.14)
Dummy for having a child aged 6-10	0.04**	0.07**	0.09**	0.01	-0.03	-0.02
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Dummy for having a child aged 11-15	0.03	0.05	0.08**	-0.04	-0.05	-0.10**
	(0.02)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
Dummy for having a child aged 16-20	0.01	0.05	0.07	0.09**	0.01	0.11
	(0.03)	(0.05)	(0.06)	(0.04)	(0.07)	(0.08)
Number of children	-0.10**			0.19**		
	(0.02)			(0.03)		
First child being a son	-0.23*			-0.58**		
	(0.13)			(0.24)		
Pseudo R squared	0.07	0.07	0.08	0.09	0.09	0.08
Number of geographic units	343	343	343	325	324	325
N	81311	29315	22323	34913	15848	13538

Notes: The dependent variable in the Probit is defined as 1 if either the household head or the spouse is a business owner or self employed. The sample is restricted to those with household heads younger than 40 years old. The sex ratio for the age cohort 5-19 is inferred from the age cohort 0-14 in the 2000 population census at either the city or the prefecture level. Other data are computed by the authors from a 20 percent random sample of the China 1% Population Survey in 2005. Household wealth is proxied by the value of house at the time of purchase (for the urban sample) or construction (for the rural sample). Standard errors are clustered at the city (or prefecture) level. * and ** denote statistically significant at the 10% and 5% levels, respectively.

Table 9: Placebo Tests on the Growth in the Number of Foreign-invested Firms at the County Level

	OLS					2SLS			
	R1	R2	R3	R4		R5	R6	R7	R8
Sex ratio for the age cohort 5-19 in 1995	0.01	0.01				0.02	0.02		
	(0.00)	(0.00)				(0.03)	(0.03)		
Sex ratio for the cohort 5-19			0.01	0.00		0.00			
averaged over 1995 and 2004			(0.01)	(0.01)				(0.02)	(0.02)
Log number of firms in 1995	-0.20**	-0.22**	-0.20**	-0.22**		-0.20**	-0.22**	-0.20**	-0.23**
	(0.03)	(0.03)	(0.03)	(0.03)		(0.03)	(0.03)	(0.03)	(0.03)
Log GDP in 1995	0.14**	0.24**	0.14**	0.24**		0.13**	0.23**	0.18**	0.28**
	(0.05)	(0.05)	(0.05)	(0.05)		(0.05)	(0.06)	(0.06)	(0.06)
Average year of schooling	-0.07	-0.13**	-0.06	-0.13**		-0.06	-0.12**	-0.10*	-0.17**
based on 2000 census	(0.05)	(0.05)	(0.05)	(0.05)		(0.05)	(0.06)	(0.05)	(0.06)
Share of agricultural output	-0.63**	-0.63**	-0.62**	-0.63**		-0.61**	-0.63**	-0.68**	-0.68**
in gross output values in 1995	(0.16)	(0.16)	(0.16)	(0.16)		(0.17)	(0.16)	(0.17)	(0.17)
The ratio of local revenues to	0.30**	0.30**	0.30**	0.30**		0.30**	0.30**	0.31**	0.30**
total government employees	(0.06)	(0.06)	(0.06)	(0.06)		(0.06)	(0.06)	(0.06)	(0.06)
Increase in log population	0.23	0.16	0.23	0.15		0.23	0.16	0.24	0.16
	(0.15)	(0.19)	(0.15)	(0.19)		(0.15)	(0.19)	(0.15)	(0.19)
Share of labor force (aged 20-64)	-0.43	-0.5	-0.24	-0.35		-0.3	-0.36	-1.47	-1.59
in local population in 1995	(0.66)	(0.66)	(0.69)	(0.70)		(0.74)	(0.79)	(1.13)	(1.18)
Share of population aged 65 and above	12.14**	10.92**	12.40**	11.16**		12.17**	10.83**	11.09**	9.96**
in total population in 1995	(1.89)	(1.90)	(1.93)	(1.93)		(1.89)	(1.92)	(2.15)	(2.11)
Log local household savings in 1995		0.17**		0.17**			0.17**		0.18**
		(0.06)		(0.06)			(0.06)		(0.06)
Adjusted R square	0.15	0.16	0.15	0.15		0.14	0.15	0.12	0.13
AIC	2580	2525	2580	2528		2585	2532	2615	2560
Durbin-Wu-Hausman test for endogeneity						0.01	0.08	0.03	0.71
Hansen's J statistic for over identification						0.02	0.11	0.03	0.02
Kleibergen-Paap Wald F statistic						21.32	16.20	21.32	18.22
N	1071	1052	1071	1052		1071	1052	1071	1052

Notes: Foreign invested firms refer to both wholly foreign owned and Sino-foreign joint ventures. Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively. Two financial penalty variables are used as instruments in the 2SLS regressions.

Table 10: Some Summary Statistics for Three-person Households

Variables	Families with a son (480)		Families with a daughter (262)		All 3-person households in the sample (742)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Share of households with at least one member taking a dangerous/unpleasant job	0.277	0.45	0.274	0.45	0.276	0.45
Total number of days that a household worked off farms	41.44	96.26	24.97	69.92	35.62	88.17
Share of households with positive number of off-farm working days	0.19	0.39	0.14	0.34	0.17	0.38
Per capita income (yuan)	3388	2459	3147	2106	3303	2341
Year of education of household head	8.02	2.10	8.19	2.07	8.08	2.09
Either household head or spouse as a minority ethnic group	0.04	0.20	0.06	0.24	0.05	0.22
Having at least a family member with bad health	0.03	0.17	0.04	0.19	0.03	0.18
Head younger than 35	0.53	0.50	0.59	0.49	0.55	0.50
Age of a child 5-9	0.41	0.49	0.47	0.50	0.43	0.50
Age of child 10 or older	0.48	0.50	0.45	0.50	0.46	0.50

Table 10B: Summary Statistics for Sex Ratio in 122 Counties

Mean	Median	Min	Max	Standard deviation
108.89	108.17	100.92	123.13	4.33

Note: The sample consists of households with two living parents and a child. The child is at least 4 years old and the household head is 40 or younger. A dangerous/unpleasant job is defined as one in a mining or construction sector, or with exposure to an extremely high or low temperature or hazardous material at work. The working day count refers to the number of days members of the household working off-farms (as a migrant worker or in a factory). Bad health is a dummy that takes the value of one if at least one family member is disabled or sick over three months in the year before the survey. Since the household survey was conducted in 122 counties, we tabulate the sex ratio for age cohort 0-9 based on the 1990 population census for these counties.

Table 11: Probit Estimation of Household Propensity to Take a Dangerous Job in 2002 (Marginal Effect)

	One son	One daughter	Total	One son	One daughter	Total
Local sex ratio for age cohort 12-21	0.96*	-0.63	-0.70			
	(0.50)	(0.69)	(0.67)			
Having a son			-0.99**			-0.99*
			(0.03)			(0.03)
Sex ratio*son			1.65**			1.59*
			(0.83)			(0.83)
Sex ratio* a dummy for poorest income quartile				0.87*	-0.77	-0.79
				(0.51)	(0.69)	(0.67)
Sex ratio* a dummy for second income quartile				0.90*	-0.73	-0.75
				(0.50)	(0.70)	(0.67)
Sex ratio* a dummy for third income quartile				0.93*	-0.59	-0.68
				(0.50)	(0.70)	(0.67)
Sex ratio* a dummy for richest income quartile				0.94*	-0.6	-0.68
				(0.50)	(0.69)	(0.67)
Log per capita income	0.02	0.00	0.01	-0.01	-0.08	-0.04
	(0.03)	(0.04)	(0.03)	(0.05)	(0.06)	(0.04)
Year of education	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Dummy for household head as an ethnic minority	-0.07	-0.11	-0.09	-0.08	-0.13	-0.1
	(0.09)	(0.10)	(0.07)	(0.09)	(0.09)	(0.07)
Dummy for poor health by at least one family member	-0.07	-0.19	-0.12	-0.06	-0.2	-0.11
	(0.11)	(0.10)	(0.08)	(0.12)	(0.09)	(0.08)
Dummy for household head younger than 35	-0.08*	-0.01	-0.05	-0.09*	0	-0.06
	(0.05)	(0.06)	(0.04)	(0.05)	(0.06)	(0.04)
Dummy for having a child aged 5-9	0.19**	0.07	0.14**	0.18**	0.07	0.13*
	(0.08)	(0.12)	(0.07)	(0.08)	(0.11)	(0.07)
Dummy for having a child aged 10 or older	0.11	0.12	0.11*	0.10	0.13	0.10*
	(0.07)	(0.11)	(0.06)	(0.07)	(0.11)	(0.06)
Pseudo R squared	0.02	0.02	0.02	0.02	0.04	0.02
AIC	572	320	882	577	321	884
N	480	262	742	480	262	742

Notes: The sex ratio for age cohort 12-21 is inferred from that for age cohort 0-14 in the 1990 population census. Other data are derived from the rural sample of CHIP 2002. The income quartiles are defined based on per capita income at the county level using the whole sample. P-values are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively. The sex ratio variable is deflated by 100 so that the reported point estimates and standard errors in rows 1 and 3-6 would not have two extra zeros after the decimal point.

Table 11A: Probit Estimation of Household Propensity to Take a Dangerous Job - Adding Saving as a Control

	One son	One daughter	Total	One son	One daughter	Total
Local sex ratio for age cohort 12-21	1.04**	-0.67	-0.70			
	(0.50)	(0.71)	(0.67)			
Having a son			-0.99**			-0.99*
			(0.03)			(0.03)
Sex ratio*son			1.65**			1.60*
			(0.83)			(0.83)
Sex ratio* a dummy for poorest income quartile				0.93*	-0.78	-0.79
				(0.51)	(0.72)	(0.67)
Sex ratio* a dummy for second income quartile				0.98*	-0.74	-0.75
				(0.51)	(0.73)	(0.67)
Sex ratio* a dummy for third income quartile				1.01**	-0.6	-0.68
				(0.50)	(0.73)	(0.67)
Sex ratio* a dummy for richest income quartile				1.02**	-0.62	-0.68
				(0.50)	(0.72)	(0.67)
Log per capita income	0.05	-0.08	0.01	0.01	-0.15**	-0.04
	(0.04)	(0.06)	(0.03)	(0.05)	(0.07)	(0.04)
Year of education	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Household head as minority ethnic group	-0.07	-0.13	-0.09	-0.08	-0.15	-0.1
	(0.09)	(0.09)	(0.07)	(0.09)	(0.08)	(0.07)
Poor health among at least one family member	-0.07	-0.19	-0.12	-0.06	-0.2	-0.11
	(0.11)	(0.09)	(0.08)	(0.12)	(0.08)	(0.08)
Head younger than 35	-0.08*	-0.01	-0.05	-0.08*	-0.01	-0.06
	(0.05)	(0.06)	(0.04)	(0.05)	(0.06)	(0.04)
Age of a child 5-9	0.21**	0.04	0.14**	0.20**	0.04	0.13**
	(0.09)	(0.12)	(0.07)	(0.09)	(0.12)	(0.07)
Age of child 10 or older	0.12	0.10	0.11*	0.12	0.11	0.10*
	(0.07)	(0.11)	(0.06)	(0.08)	(0.11)	(0.06)
Household saving rate	(0.05)	0.13**	0.00	-0.06	0.12**	-0.01
	(0.04)	(0.06)	(0.03)	(0.04)	(0.06)	(0.03)
Pseudo R squared	0.02	0.03	0.02	0.03	0.05	0.02
AIC	573	318	884	577	319	886
N	480	262	742	480	262	742

Note: The saving variable is defined as log(income/expenditure).

Table 12: Tobit Estimation on the Number of Off-farm Working Days

	One son	One daughter	Total	One son	One daughter	Total
Local sex ratio for age cohort 12-21	20.75**	4.43	2.76			
	(4.99)	(6.93)	(6.89)			
Having a son (coeff x 100)			-19.34**			-18.26**
(se x 100)			(9.15)			(9.16)
Sex ratio*son			18.33**			17.32**
			(8.40)			(8.42)
Sex ratio* a dummy for the poorest income quartile				19.80**	3.23	2.04
				(5.15)	(7.11)	(7.00)
Sex ratio* a dummy for second income quartile				19.49**	3.81	2.00
				(5.08)	(7.05)	(6.98)
Sex ratio* a dummy for third income quartile				20.47**	5.12	3.13
				(5.06)	(7.07)	(6.98)
Sex ratio* a dummy for the richest income quartile				20.47**	5.10	3.17
				(5.02)	(6.96)	(6.94)
Log household income	81.56**	-44.66	33.1	39.53	-122.77**	-25.31
	(35.48)	(40.68)	(28.58)	(52.28)	(59.52)	(42.16)
Year of education	-3.45	13.39	1.25	-3.51	14.74	1.41
	(10.29)	(15.69)	(8.81)	(10.09)	(15.02)	(8.58)
Household head as minority ethnic group	87.7	45.8	68.7	82.16	23.91	53.98
	(104.4)	(111.1)	(78.4)	(101.08)	(108.18)	(76.17)
Poor health among at least one family member	53.53	93.80	-71.29	-49.00	-128.31	-71.65
	(138.4)	(146.47)	(105.60)	(137.69)	(141.20)	(105.09)
Head younger than 35	-52.06	-54.17	-52.69	-60.22	-49.85	-58.05
	(48.47)	(67.54)	(40.02)	(47.93)	(64.32)	(39.38)
Age of a child 5-9	93.48	52.41	47.47	83.06	-54.68	34.04
	(82.48)	(120.67)	(69.35)	(82.18)	(114.66)	(68.51)
Age of child 10 or older	47.87	62.97	13.69	38.210	-56.05	3.23
	(75.67)	(111.23)	(63.27)	(75.96)	(106.66)	(62.82)
Pseduo R squared	0.02	0.00	0.01	0.02	0.01	0.02
AIC	1591	667	2247	1594	668	2246
N	480	262	742	480	262	742

Notes: The coefficients are marginal effects on the latent dependent variables. The sex ratio for age cohort 12-21 in 2002 is inferred from the age cohort 0-14 in the 1990 population census. Other data are derived from the rural sample of CHIP 2002. Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively. All regressions have a constant which is not reported.

Table 12A: Tobit Estimation on the Number of Off-farm Working Days Including Saving as a Control Variable

	One son	One daughter	Total	One son	One daughter	Total
Local sex ratio for age cohort 12-21	19.33**	4.47	2.55			
	(4.94)	(6.89)	(6.95)			
Having a son (coeff x 100)			-18.85**			-17.93*
(se x 100)			(9.22)			(9.20)
Sex ratio*son			17.85**			17.00**
			(8.46)			(8.45)
Sex ratio* a dummy for the poorest income quartile				18.72**	3.28	1.93
				(5.06)	(7.07)	(7.05)
Sex ratio* a dummy for second income quartile				18.31**	3.9	1.82
				(5.00)	(7.00)	(7.02)
Sex ratio* a dummy for third income quartile				19.18**	5.2	2.93
				(4.99)	(7.03)	(7.02)
Sex ratio* a dummy for the richest income quartile				19.22**	5.18	2.99
				(4.95)	(6.92)	(6.99)
Log household income	43.6	-37.26	9.87	10.37	-109.93	-41.5
	(43.75)	(48.62)	(34.29)	(56.48)	(67.16)	(45.42)
Year of education	-3.96	12.81	1.77	-4.01	13.71	1.79
	(10.18)	(15.64)	(8.81)	(10.03)	(15.01)	(8.58)
Household head as minority ethnic group	71.62	48.45	58.76	69.22	28.77	46.93
	(103.89)	(111.57)	(77.99)	(100.71)	(108.55)	(75.69)
Poor health among at least one family member	(47.53)	(93.64)	-68.82	-44.97	-128.4	-69.84
	(136.60)	(146.39)	(105.20)	(135.26)	(141.13)	(104.57)
Head younger than 35	-57.05	-54.77	-54.03	-64.43	-51.01	-59.07
	(48.26)	(67.50)	(39.97)	(47.69)	(64.39)	(39.33)
Age of a child 5-9	68.18	(49.90)	35.51	62.08	-49.46	24.99
	(84.70)	(120.32)	(70.80)	(84.19)	(114.70)	(70.07)
Age of child 10 or older	23.49	(61.35)	2.99	18.380	-52.5	-4.75
	(76.27)	(110.57)	(63.77)	(76.51)	(106.20)	(63.55)
Household saving rate	80.94*	(10.23)	41.71	74.39	(18.55)	32.76
	(45.90)	(61.47)	(41.69)	(45.25)	(57.27)	(39.71)
Pseduo R squared	0.02	0.00	0.01	0.02	0.01	0.02
AIC	1590	669	2248	1593	670	2247
N	480	262	742	480	262	742

Note: See Table 12.

Table 13: Summary Statistics on Variables Used in Provincial Level Regressions, 1980-2005

	Mean	Median	Standard deviation
Annual real growth rate of GDP/per capita (1980-2005)	8.81	8.72	2.71
Annual real growth rate of GDP/per capita (2000-2005)	10.20	10.21	1.48
Per capita GDP (yuan in 1985)	2737	1849	2732
Share of labor force (aged 15-64) in local population	68.02	67.80	4.47
Investment/local GDP (%)	35.62	32.72	11.82
Foreign trade/local GDP (%)	25.22	10.41	39.12
Outstanding loans per capita (yuan)	1631	288	3558
Birth rate (‰)	17.31	17.10	4.20
Sex ratio for the 5-19 age cohort (%)			
Whole sample (1980-2005)	106.79	106.44	3.03
1980	105.12	105.05	2.84
1985	105.44	105.07	2.26
1990	106.68	106.28	1.56
1995	108.48	108.77	1.90
2000	109.58	109.50	4.20
2005	113.57	113.04	4.14

Table 14: Sex Ratios and Income Growth - Provincial Panel Regressions from 1980 to 2005

	R1	R2	R3	R4	R5	R6
Initial sex ratio for age cohort 5-19	0.22**	0.30**	0.33**			
	(0.07)	(0.12)	(0.12)			
Average sex ratio for age cohort 5-19				0.29**	0.40**	0.44**
				(0.08)	(0.10)	(0.10)
Initial log per capita GDP	-8.48**	-9.61**	-11.3**	-8.84**	-10.2**	-11.7**
	(1.48)	(1.76)	(1.9)	(1.28)	(1.6)	(1.6)
Investment/local GDP	0.14**	0.12**	0.11**	0.15**	0.13**	0.12**
(average in a five-year interval)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Share of labor force (aged 15-64) in total population		-0.08	0.02		-0.08	-0.01
(average in a five-year interval)		(0.13)	(0.14)		(0.14)	(0.14)
Foreign trade/local GDP		0.02	0.02		0.02	0.02
(average in a five-year interval)		(0.01)	(0.01)		(0.01)	(0.01)
Household savings rate			0.19**			0.18**
(average in a five-year interval)			(0.05)			(0.05)
Birth rate at the beginning of a 5-year interval	-0.07	-0.05	-0.02			
	(0.08)	(0.08)	(0.08)			
Birth rate (averaged over a 5-year interval)				-0.19*	-0.14	-0.11
				(0.10)	(0.11)	(0.10)
Province fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted R-square	0.64	0.64	0.67	0.67	0.68	0.70
AIC	584	584	571	572	571	558
N	145	145	144	145	145	144

Notes: The panel regressions are performed over five 5-year periods, 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05. The dependent variable is the average annual growth rate of per capita GDP over a 5-year period. Each regression includes both province fixed effects and period fixed effects. The sex ratio for the first four periods is inferred from the 1990 population census, while the sex ratio for the period of 2000-2005 is inferred from the 2000 population census. Robust standard errors are in parentheses. The household savings rate is defined as log(income/consumption), which is taken from Wei and Zhang (2009). * and ** denote statistically significant at the 10% and 5% levels, respectively

Table 15: GMM Estimation with Both Exogenous Instruments and Lagged Regressors – Sex Ratios and Income Growth over 1980 - 2005

	R1	R2	R3	R4	R5	R6
Set of instrumental variables	*Lagged regressors*		*+ Financial penalties for violating birth quotas*		*+ Share of local population exempted from birth quotas*	
Initial sex ratio for age cohort 5-19	0.28**		0.28**		0.28**	
	(0.13)		(0.12)		(0.12)	
Average sex ratio for age cohort 5-19		0.51**		0.51**		0.54**
		(0.15)		(0.15)		(0.15)
Initial log per capita GDP	-12.13**	-13.36**	-12.03**	-13.34**	-12.02**	-13.52**
	(2.22)	(2.10)	(2.17)	(2.09)	(2.13)	(2.06)
Investment/local GDP	0.05	-0.03	0.05	-0.03	0.05	-0.02
(average in a five-year interval)	(0.16)	(0.19)	(0.16)	(0.19)	(0.16)	(0.19)
Share of labor force (aged 15-64) in total population	0.12**	0.13**	0.12**	0.13**	0.12**	0.14**
(average in a five-year interval)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Foreign trade/local GDP	0.01	0.02	0.01	0.02	0.01	0.02
(average in a five-year interval)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)
Household savings rate	0.23**	0.24**	0.23**	0.24**	0.23**	0.24**
(average in a five-year interval)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
Birth rate at the beginning of a 5-year interval	0.04		0.04		0.04	
	(0.12)		(0.11)		(0.11)	
Birth rate (averaged over a 5-year interval)		0.02		0.02		0.00
		(0.17)		(0.17)		(0.16)
AR (1) in first difference (p-value)	0.00	0.00	0.00	0.00	0.00	0.00
AR (2) in first difference (p-value)	0.64	0.47	0.64	0.47	0.64	0.44
Hansen's J statistic for over identification	0.77	0.76	0.85	0.82	0.87	0.82
N	114	114	114	114	114	114

Notes: The data is a panel of five 5-year periods, 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05. The dependent variable is the average annual growth rate of per capita GDP over a 5-year period. The instrumental variables for regressions R1-R2 include the standard lagged regressors in GMM estimation. Two additional instruments, penalty for violating family planning policy (% of local yearly income), and a dummy for extra penalty for higher order births, are included in regressions R3 and R4. The share of minority population is further added in regressions R5 and R6. Both province and time fixed effects are included but not reported here. Robust standard errors are in parentheses. * and ** denote statistically significant at the 10% and 5% levels, respectively.

Appendix Table: Heckman Estimations on the Likelihood That Families Choose to Have Only One Child

First Probit regression: Dependent variable in the main regression = likelihood that parents are entrepreneur (for one-child families);
Second Probit regression: Dependent variable in the selection equation = likelihood that parents stop at the first child

| | Rural nuclear families with | | | | Urban nuclear families | | | |
| | First child a son | | First child a daughter | | First child a son | | First child a daughter | |
	Parents being entrepreneurs	Stop at 1st child	Parents being entrepreneurs	Stop at 1st child	Parents being entrepreneurs	Stop at 1st child	Parents being entrepreneurs	Stop at 1st child
Local sex ratio for age cohort 5-19	0.61**	-1.08	0.34	-1.23*	0.71**	-3.14**	0.50	-3.08**
	(0.21)	(0.83)	(0.27)	(0.70)	(0.30)	(0.39)	(0.35)	(0.51)
Log household wealth	0.16**	0.06**	0.15**	0.04**	0.03**	0.05**	0.02	0.03
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)
Year of education	0.24**	-0.07**	0.29**	-0.11**	0.25**	-0.16**	0.25**	-0.19**
	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.02)	(0.03)
Year of education (squared)	-1.32**	0.91**	-1.57**	1.02**	-1.78**	1.37**	-1.72**	1.60**
	(0.13)	(0.08)	(0.12)	(0.07)	(0.11)	(0.15)	(0.12)	(0.16)
Household head age	0.00	0.00**	-0.01**	0.01**	0.01**	0.01**	0.01**	0.01**
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female household head	0.20**	0.19**	0.12**	0.29**	-0.02	0.18**	-0.08**	0.24**
	(0.05)	(0.05)	(0.06)	(0.05)	(0.04)	(0.06)	(0.04)	(0.04)
Poor health among at least one parent	-0.32**	0.13**	-0.29**	0.11**	-0.31**	0.34**	-0.18	0.03
	(0.07)	(0.05)	(0.09)	(0.04)	(0.15)	(0.15)	(0.14)	(0.11)
Age of a child 6-10	0.23**	-1.09**	0.27**	-1.10**	0.02	-1.07**	0.08	-1.00**
	(0.06)	(0.03)	(0.11)	(0.02)	(0.07)	(0.05)	(0.10)	(0.05)
Age of a child 11-15	0.23**	-1.04**	0.27**	-0.94**	0.01	-0.91**	0.02	-0.76**
	(0.07)	(0.03)	(0.11)	(0.03)	(0.08)	(0.06)	(0.11)	(0.06)
Age of a child 16-20	0.21**	-0.79**	0.22**	-0.55**	0.07	-0.61**	0.22*	-0.36**
	(0.08)	(0.05)	(0.10)	(0.04)	(0.11)	(0.08)	(0.12)	(0.06)
Age of the oldest child		-0.03**		-0.04**		-0.07**		-0.08**
		(0.00)		(0.00)		(0.01)		(0.00)
First child being disabled		-0.71**		-0.38**		-1.07**		-0.19
		(0.15)		(0.17)		(0.26)		(0.34)
Minority household		-0.45**		-0.16**		-0.34**		-0.22**
		(0.05)		(0.04)		(0.09)		(0.07)
Log likelihood	-31857		-29709		-12107		-11966	
N	40244		38255		17152		16026	

Notes: The dependent variable in the first Probit (main) regression is defined as 1 if either the household head or the spouse is a business owner or self employed. The dependent variable in the second Probit (selection) regression takes the value of one if the household has only one child and zero if it has more than one child. The sample is restricted to those with household heads younger than 40 years old. The sex ratio for the age cohort 5-19 is inferred from the age cohort 0-14 in the 2000 population census at either the city or the prefecture level. Other data are computed by the authors from a 20 percent random sample of the China 1% Population Survey in 2005. Household wealth is proxied by the value of the house at the time of purchase (for the urban sample) or construction (for the rural sample). Standard errors are clustered at the city (or prefecture) level. * and ** denote statistically significant at the 10% and 5% levels, respectively.
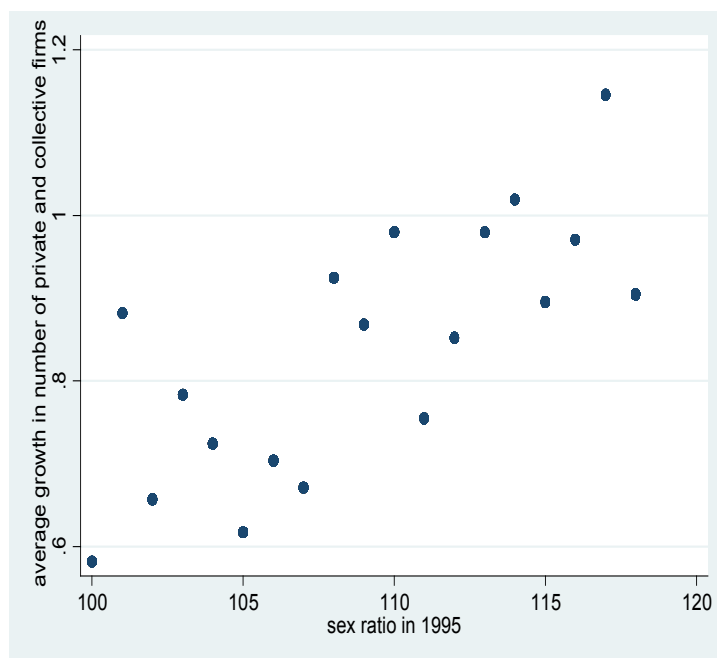
Figure 1: Initial Sex Ratios and Growth Rates of Private Firms during 1995-2004

On the horizontal axis is the sex ratio for the age cohort 5-19 in 1995 inferred from the 1990 Population Census. On the vertical axis is the growth rate in the number of private firms from 1995 to 2004, averaged over all counties that had the same value of sex ratio (up to a basis point).