An Endogenous Sex Selection Model and a Test Using the Chinese Census

Wei Yang

October, 2013

Abstract

The parental sex selection behaviour of altruistic parents without gender specific preference in a monogamous marriage market is studied. Using an OLG model with intra-household resource allocation determined by marriage market bargaining, we find that the equilibrium with a balanced male/female ratio is unstable. Depending on the effect of the marriage market sex ratio on intra-household allocation, there exist stable equilibria with unbalanced sex ratios. Specifically, if the marriage market responsiveness elasticity is small enough and males earn more than females, an equilibrium with a male/female birth ratio in excess of one is stable. From that equilibrium, a reduction in the value of a marriage will tend to increase the male/female birth ratio. This foreshadows the result from the extended model in Yang (2013a), where a reduction in the value of a marriage is attributed to China’s one-child policy, which is empirically associated with an increase in the male/female birth ratio. Another empirical implication of the model is the requirement for equilibrium that the marriage market male/female sex ratio has a negative effect on the number of male births. We find empirical support for this proposition in the 2000 Chinese census.

Keywords: parental sex selection, missing women, intra-household allocation

JEL: J11 J12 J13

I. Introduction

It is largely agreed that males outnumber females at birth with a natural ratio of 103 to 106 males per 100 females. Because of higher death rates for males, we typically observe roughly equal proportions of men and women in most populations.³ However, it is observed that, in recent decades, the sex ratio increased significantly in some developing countries, notably India and

---

¹ Department of Economics, McMaster University, Kenneth Taylor Hall, Room 437, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4M4. E-mail: yangw7@mcmaster.ca, Tel:1-905-9662896
² Thanks to Luke Chan, Byron Spencer, and Mike Veall for comments on this paper. Financial support was received from the Council of Library and Information Resources (CLIR). The funders had no role in the research and the author has no conflicts of interest to declare.
³ Coale (1991) finds that the ratio slightly favours men (1.022) when population grows at a 2% annual rate, and women (0.997) in stationary populations.
China, a phenomenon referred to by Sen (1992) as the "missing women". In India, the sex ratio for ages 0-6 was 107.8 in 2001, with some states reaching as high as 120-125.\textsuperscript{4} Recent waves of population censuses in China reveal more astonishing observations. The sex ratio at birth has increased from 108.5 in 1982 to 113.8 in 1990 to 119.9 in 2000. This demographic imbalance has many far-reaching impacts on many aspects of economic life. Studies have found the sex ratios imposes effects on both the marriage market and the labour market (Angrist (2002); Becker (1991); Chiappori et al. (2002), and Rao (1993)), on household savings (Wei and Zhang (2011)), and on crime (Edlund et al. (2007)).

The skewed sex ratio has drawn extensive research attention because the deviation from the biologically normal population sex ratio usually indicates the practice of either prenatal or postnatal sex selection (or both). The former includes various mechanisms, e.g. coital frequency and sex-selective abortions\textsuperscript{5}. The latter is associated with sex-differential mortality rates across ages, which, during childhood, is linked to parental investments in the health of the child, and, in extreme cases, infanticide. Chu (2001) finds that prenatal sex selection, through sex determination by ultrasound examination and selective abortion, is the primary cause for the rise of the sex ratio in China. Bhaskar (2011) provides anecdotal evidence about historical infanticide practices in India. This behaviour of parents is usually thought to be motivated by a prevailing son preference rooted in cultural factors in Asia.\textsuperscript{6} However, Anderson and Ray (2010) find that the number of missing women in sub-Saharan Africa is comparable to those for India and China due to the relatively high female mortality rate during adolescence and early adulthood, even though the sex ratio at birth is biologically normal. Moreover, they find evidence that a comparable proportion of women was missing at the start of the 20th century in the United States as they are in India and China today. It seems that "missing women" and son preference may not be associated with a specific type of culture, and may change over time. In this paper, we explore a model in which sex-selection may be exhibited by parents with no inherent gender preference and there is a market-driven explanation for high male-female sex ratios.

It is often assumed in the literature on human sex ratios that parents prefer married sons to married or single daughters, and married daughters to single sons, e.g. Bhaskar (2011) and Edlund (1999). This is arguably consistent with the Trivers and Willard (1973) evidence for the hypothesis that natural selection favours parental ability to adjust the probability of a male or a female offspring according to the individual parents' ability to invest in them. They argue that a male in good condition would be expected to have more offspring than a female in good condition but a male in poor condition would be expected to have fewer offspring than a female in poor condition. Hence parents in good condition would want male offspring given they are likely to be in good condition like their parents. This suggests that if parents can affect the sex of

\textsuperscript{5} See Edlund (1999) for a summary.  
\textsuperscript{6} Higher than normal sex ratios, e.g. 108-109, are also observed in other Asian areas such as South Korea and Taiwan.
their children, those parents who do not have a set target for children of one specific sex but rather who adjust their target according to circumstances would gain a margin in natural selection. Hence, we assume parents have neutral preferences with respect to the specific gender of their children but rather want to have a child of the sex associated with the higher reproductive value.

Our approach is to assume that parents are altruistic in that they value the life-time utility of each child. The utility of an unmarried child, male or female, is determined by his/her consumption and leisure given the wage rate. The utility of a married child, male or female, is a function of household income, the intra-household resource distribution, and the value of his/her offspring (a household public good). Since the value of the child includes the value of the child’s offspring if he/she gets married, and hence iteratively includes the weighted value of all descendents, the reproduction concern of parents is implicitly taken account. We do not consider in our model the case that parents value transfers from their children, e.g. financial supports. While it is common for parents to receive support from their children especially in old age, the net transfer from the parents to the offspring is still likely to be positive. In evolution, as per Bergstrom (1995), successful strategies (defined either by a genetic type or by a social norm) are replicated more frequently than unsuccessful strategies. Since transfers from parents to offspring generally increase the survival of the latter, we expect pure altruistic behaviour towards offspring.

A male and a female must marry to have offspring. Monogamy is assumed. Subject to the household budget constraint, a husband and a wife choose not only their consumption and leisure, but also the time they spend on child-raising. Males and females are allowed to have different wage rates. (Qian (2008) finds that the relative wages affects female’s survival rate.) Based on studies on collective household behaviour, e.g. Chiappori (1992), Browning et al. (1994), Li and Wu (2011), and Zhang and Chan (1999) and following Siow (1998), we assume that husbands and wives can enforce marriage agreements on intra-household resource allocation. Finally, the relative resource allocation of a husband and a wife within a household is determined by the relative bargaining power of males and females in the marriage market when the marriage is formed, which, as per Rao (1993), is associated with the tightness of marriage market, i.e. the sex ratio of potential males and females.

Under these assumptions, parents will prefer a married child to an unmarried one (where the child can choose to remain single if there is no utility gain to marriage), and they will want a child of the sex associated with higher expected life-time utility. Those of the more plentiful sex are more likely not to find a spouse and have a reduced share of resources within the household if they successfully get married. In equilibrium, the sex ratio is such that the expected life-time utility at birth of a male is exactly the same as that of a female. In this paper, we find that multiple equilibria exist. The balanced sex ratio (100 males per 100 females) equilibrium with equal intra-household allocation is unstable. Depending on how the sex ratio affects the intra-household allocation, there could be two stable equilibria, one with the sex ratio smaller than 100, another with the sex ratio larger than 100. Specifically, if the marriage market responsiveness
elasticity, defined as the percentage change in the husbands' share of the total marriage benefit in response to a one per cent change in the sex ratio, is small enough, there exists a stable solution with more males than females at equilibrium.

This paper provides a new economic perspective to understanding parents' sex selection behaviour under the market setting, where households make forward-looking family plans exploiting their knowledge about the market structure, market signals, and modern medical technologies. The model suggests that the imbalanced sex ratio can well be caused by altruistic parents with no inherent preference for males seeking maximal welfare for their children, where the parental desire for sons is endogenously determined. Using this framework to study the increase in sex ratios in developing countries, we find that the marriage squeeze caused by population growth could be an important cause. In China, the one-child policy played a crucial role in the increase in sex ratios through two mechanisms: 1) it depresses the value of a marriage by imposing a birth quota; 2) it increases men's relative bargaining power in the marriage market.

In the empirical part, of this paper, one aspect of the model is tested with the 2000 Chinese census data. This model requires (as we expect would many models with a stable sex-ratio equilibrium), that parents tend to favour a sex when it become more scarce in the marriage market. We find that marriage market tightness, e.g. the sex ratio of males aged between 20 and 23 to females aged between 17 and 20 in a year, has a statistically significant effect on the determination of the gender of children born in that year. Our estimates suggest that a one percentage point increase in the marriage market sex ratio decreases the probability of having a boy by 0.02 to 0.05 percentage points, which is equivalent to decreasing the sex ratio of the birth cohort by 1.02 to 1.15 percentage points.

The paper is organized as follows: Section 2 presents the theoretical model and its implications. In Section 3, we describe our data and outline our estimation methods. The empirical results are presented in Section 4. Finally, Section 5 offers conclusions and discussions.

II. The Basic Model

A. Household utility maximization

In this model, each individual lives for two periods, i.e. childhood and adulthood. An individual lives with parents his/her during childhood and forms his/her own family in adulthood. An adult can either stay single or get married. Only married adults can reproduce. At any period $\tau$, the sex ratio, $\theta_\tau$, is defined as the number of male adults to that of female ones.

The utility in childhood is determined by the child's market goods consumption, $x_{ct}$, and the time spent with father (male), $t_{m\tau}$, and mother (female), $t_{f\tau}$, where subscript $m$ and $f$ stands for male and female, and defined as

$$u_{ct} = q(x_{ct}, l(t_{m\tau}, t_{f\tau}))$$
where \( q(.) \) is increasing in \( x_{ct} \) and \( l(.) \) is increasing in \( t_m \) and \( t_f \) \((l_m > 0 \text{ and } l_f > 0)\), with the second derivatives, \( l_{mm} < 0 \text{ and } l_{ff} < 0 \). As the time of father's and mother's can be either substitutes or complements, either \( l_{mf} \leq 0 \text{ or } \geq 0 \).

The utility of an adult depends on his/her marital status, i.e. single or married. If one remains single, it is determined solely by his/her consumption of market goods

\[
u_{at} = x_{at}
\]

Marriage is assumed to be monogamous. If married, an adult acquires utility from his/her private consumption and from his/her offspring, a lone child\(^7\) who is a public good for the parents. Parents care about the lifetime welfare of their child. Utility can be defined as

\[
u_{at} = x_{at} + \beta g(q(x_{ct}, l(t_{mt}, t_{ft})), V_{ct+1})
\]

where \( \beta \) \((0 < \beta < 1)\) is the weight on the welfare of the offspring and \( V_{ct+1} \) is the adulthood welfare of a child in period \( \tau + 1 \). Therefore, the lifetime utility of an individual, \( i \), who is born at period \( \tau \) and married at period \( \tau + 1 \) is

\[
U_{it} = u_{ct} + u_{at+1} = q(x_{ct}, l(t_{mt}, t_{ft})) + x_{at+1} + \beta g(q(x_{ct+1}, l(t_{mt+1}, t_{ft+1})), V_{ct+2})
\]

We assume \( g(.) \) is increasing in \( q \) and \( V \) with second derivatives \( g_{qq} < 0 \) and \( g_{VV} < 0 \). For simplicity, in the benchmark model, we assume \( g_{qV} = 0 \). This precludes the cases where parents treat the transfers to their children during childhood as a substitute or complement for their children's welfare during adulthood. However allowing \( g_{qV} \) to be nonzero does not qualitatively affect our results regarding sex selection as long as it is the same for males and females.

The wage rate of males, \( W_m \), and females, \( W_f \), are assumed to be fixed and exogenously given. Each individual has one unit of time endowment in adulthood, which can be used either to work for income or to take care of his/her children. Since leisure is excluded, the budget constraint for adults, e.g. male or female, who remain single is

\[
x_{mt} = W_m
\]

\[
x_{ft} = W_f
\]

---

\(^7\) This assumption does not have a significant effect on the results of the model because all children in a household are assumed to be born simultaneously and we focus our analysis on the equilibria where sons and daughters have the same value to parents. We generalize the model to allow for more than one child in order to study China's one-child policy (Yang, 2013a) and parents' sex selection behaviour of subsequent children conditional on the sex of the first child (Yang, 2013b).
For a married couple, the husband and wife make joint decisions on the amount of time spent with the child and hence time in the job market, and their own and hence their child’s consumption subject to the budget constraint of a household:

\[ x_{mt} + x_{ft} + x_{ct} = (1 - t_{mt})W_m + (1 - t_{ft})W_f \]

I assume that husbands and wives can enforce marriage agreements contingent on their child’s adulthood welfare, i.e. \( V_{ct+1} \). Given that within-household choices are efficient, we focus on the males’ choices subject to the constraint that females acquire at least their reservation utility. So a male with wage \( W_m \) who marries a female with wage \( W_f \) will choose the time spent with their child by both the husband and wife, \( t_{mt}, t_{ft} \), and the consumption of market goods by each family member, \( x_{mt}, x_{ft}, x_{ct} \), to maximize his adulthood utility

\[
u^*_m(W_m, W_f, Z, V_{ct+1}) = \max_{(t_{mt}, t_{ft})x_{mt}x_{ft}x_{ct}} x_{mt} + \beta g(q(x_{ct}, l(t_{mt}, t_{ft})), V_{ct+1})
\]

subject to constraints

\[(c1) \quad x_{ft} + \beta g(q(x_{ct}, l(t_{mt}, t_{ft})), V_{ct+1}) \geq Z\]

\[(c2) \quad x_{mt} + x_{ft} + x_{ct} \leq (1 - t_{mt})W_m + (1 - t_{ft})W_f\]

where \( Z \) is the reservation utility of females. \((c1)\) implies that a male has to provide for a female at least \( Z \) for her to be willing to marry him, otherwise he has to remain unmarried. \((c2)\) defines the income budget constraint of a household, which shows that a couple gives up total consumption of \( t_{mt}W_m + t_{ft}W_f + x_{ct} \) for their child. In order to capture the reproduction motive, we assume the marginal utility of the transfer to the child, when the transfer is zero, is much larger than the marginal utility of parents’ consumption (which is one in the current setting). Both \((c1)\) and \((c2)\) should hold with equality at the optimum; otherwise the male would increase his consumption. Therefore, \((c1)\) and \((c2)\) imply

\[(1) \quad x_{mt} = (1 - t_{mt})W_m + (1 - t_{ft})W_f - x_{ct} - Z + \beta g(q(x_{ct}, l(t_{mt}, t_{ft})), V_{ct+1})\]

Substituting equation \((1)\) into the objective function transforms the optimal problem into an unconstrained one

\[
u^*_m(W_m, W_f, Z, V_{ct+1}) = \max_{(t_{mt}, t_{ft})W_m + (1 - t_{ft})W_f - x_{ct} + 2\beta g(q(x_{ct}, l(t_{mt}, t_{ft})), V_{ct+1}) - Z\]

where a male, given the wage of males and females, chooses the transfer schemes of the household to the child, including the time the husband and the wife spent with the child and the
child's consumption, contingent on the reservation utility of the wife determined by the marriage market, $Z$, and the adulthood welfare of the child, $V_{ct+1}$. However, by assuming a transferable utility function (see Bergstrom (1997) for a detailed discussion), the reservation utility of a wife, $Z$, is additively separable in the husband's objective function so that the optimal choices of $t_{mt}$, $t_{ft}$, and $x_{ct}$ are independent of $Z$. The first-order conditions with respect to $t_{mt}$, $t_{ft}$, and $x_{ct}$ are

\begin{align*}
2βg_qqlm &= W_m \\
2βg_qqlf &= W_f \\
2βg_qqx &= 1
\end{align*}

Equation (2) to (4) define the optimal choices of the time that parents spend with their child, $t_{mt}^*$ and $t_{ft}^*$, as well as the child's consumption, $x_{ct}^*$. Recall we assume $g_{qV} = 0$ so that these choices are also independent of $V_{ct+1}$, and are only functions of $W_m$ and $W_f$, i.e. $t_{mt}^* = t_{mt}^*(W_m,W_f), t_{ft}^* = t_{ft}^*(W_m,W_f), x_{ct}^* = x_{ct}^*(W_m,W_f)$, and hence $q^* = q^*(W_m,W_f)$. $t_{mt}^*$ and $t_{ft}^*$ define the gender role of a husband and a wife in a household, e.g. the main income earner and care provider for children. Equations (2) and (3) show that it is generally determined by the comparative advantage of providing care and the relative opportunity cost of males and females. $l_m/l_f = W_m/W_f$ should hold for an interior solution. If the wage is higher for males than females or females are more effective care givers, the model predicts the gender differences, i.e. $t_{mt}^* < t_{ft}^*$, that are usually observed, i.e. on average, males work more in the labour market and females more in households.

Substituting the first-order conditions back to the budget constraints, we have the equations for the consumptions of the husband and the wife

\begin{align*}
(5) \quad x_{ft}^*(W_m,W_f,Z,V_{ct+1}) &= Z - βg(q^*(W_m,W_f),V_{ct+1}) \\
(6) \quad x_{mt}^*(W_m,W_f,Z,V_{ct+1}) &= (1 - t_{mt}^*)W_m + (1 - t_{ft}^*)W_f - x_{ct}^*(W_m,W_f) - x_{ft}^*(W_m,W_f,Z,V_{ct+1})
\end{align*}

Besides the wages, consumptions are also affected by the contribution to the parents’ utility of their child’s adulthood utility and the reservation value of females. Equation (5) and (6) show that, holding $W_m,W_f,V_{ct+1}$ fixed, the division of consumption between a husband and a wife is determined by the reservation utility of females, which is independently determined by the equilibrium of the marriage market.

Denoting the utility of a husband and a wife during their adulthood as $V_{mt}$ and $V_{ft}$, by definition, they are

\begin{align*}
V_{mt}(V_{ct+1}) &= x_{mt}^* + βg(q^*,V_{ct+1})
\end{align*}
\[ V_{f\tau}(V_{c\tau+1}) = x_{f\tau}^* + \beta g(q^*, V_{c\tau+1}) \]

and hence

\[ (7) \quad V_{mr}(V_{c\tau+1}) + V_{f\tau}(V_{c\tau+1}) = (1 - t_{mr}^*)W_m + (1 - t_{f\tau}^*)W_f - x_{ct}^* + 2\beta g(q^*, V_{c\tau+1}) \]

which is the household utility possibility frontier conditional on \( V_{c\tau+1} \). At the time when a male makes a marriage proposal, the adulthood welfare of the future child, \( V_{c\tau+1} \) is unknown, e.g. a son might be different from a daughter in general. Taking the expectation with respect to \( V_{c\tau+1} \), we have

\[ V_{mr} + V_{f\tau} = (1 - t_{mr}^*)W_m + (1 - t_{f\tau}^*)W_f - x_{ct}^* + E_{V_{c\tau+1}}(2\beta g(q^*, V_{c\tau+1})) \]

Let \( p_m \) denote the probability of having a boy, \( V_{ct+1}^m \) the adulthood welfare of their child if it is a boy, and \( V_{ct+1}^f \) if a girl, then, at the time of marriage, the expected household utility possibility frontier is

\[ V_{mr} + V_{f\tau} = (1 - t_{mr}^*)W_m + (1 - t_{f\tau}^*)W_f - x_{ct}^* + 2p_m\beta g(q^*, V_{ct+1}^m) + 2(1 - p_m)\beta g(q^*, V_{ct+1}^f) \]

Define the household income net of the transfer to the child \( I^* = (1 - t_{mr}^*)W_m + (1 - t_{f\tau}^*)W_f - x_{ct}^* \) and let \( p_m = 1/2 \), then

\[ (8) \quad V_{mr} + V_{f\tau} = I^* + \beta g(q^*, V_{ct+1}^m) + \beta g(q^*, V_{ct+1}^f) \]

Based on equation (8), we define the steady state where the value of an adult male or of an adult female, i.e. the utility during adulthood, does not change over time and over generations, i.e.

\[ V_m = V_{mr} = V_{ct+1}^m \text{ and } V_f = V_{f\tau} = V_{ct+1}^f \text{ for all } \tau, \text{ then} \]

\[ (9) \quad V_m + V_f = I^* + \beta g(q^*, V_m) + \beta g(q^*, V_f) \]

Equation (9) is the expected long run value of a household as a basic reproduction unit for altruistic parents who maximize their utility when parental sex selection is infeasible. Nonetheless, households prefer the sex associated with higher value. Hence we are interested in the equilibria where males and females have the same value, i.e. \( V_m = V_f = V^* \). In a stationary equilibrium with no incentive for sex selection, equation (9) implies that the value function of an individual satisfies

\[ ^8 \text{ When the value gap between males and females is high enough, for example, when the value of females is lower than that of males, household and social resources are likely to be redirected biased toward males, e.g. food, shelters, and healthcare resources, which eventually leads to the observation of "missing women", and hence increases the relative scarcity of females. Lin et al. (2012) find that access to sex-selective abortion increases the share of male children born and decreases the relative neo-natal mortality, which is interpreted as allowing parents to abort unwanted girls and increasing the average care that girls receive when they are born.} \]
which leads to Proposition 1 on the existence of stationary solution without sex selection. Figure 1 provides a graphical illustration of Proposition 1.

**Proposition 1.** Despite the wage gap ($W_m$ and $W_f$) and potential difference in child care ($l_m$ and $l_f$) between males and females, there exists a solution with no incentive for sex selection if

$$
\lim_{V \to \infty} \frac{\partial \beta g(q^*, V)}{\partial V} < 1.
$$

**Proof.** Let $T$ denote operator defined in equation (10), $V^* = T(V^*) = I^*/2 + \beta g(q^*, V^*)$. (1)

Monotonicity: by assumption $\frac{\partial g(q^*, V)}{\partial V} > 0$, so $T(V_1) > T(V_2)$ if $V_1 > V_2$. (2) Discounting: for any real number, $a$, $T(V + a) = \frac{I^*}{2} + \beta g(q^*, V + a) = \frac{I^*}{2} + \beta g(q^*, V(1 + \frac{a}{V}))$, when $V$ approaches infinity. $T(V + a) = \lim_{V \to \infty} \left[ \frac{I^*}{2} + \beta g(q^*, V) + \frac{\partial \beta g(q^*, V)}{\partial V} * V * \frac{a}{V} \right] = T(V) + \lim_{V \to \infty} \frac{\partial \beta g(q^*, V)}{\partial V} * a$. If

$$
\lim_{V \to \infty} \frac{\partial \beta g(q^*, V)}{\partial V} < 1,
$$

then operator $T$ satisfies $T(V + a) \leq T(V) + \delta a$ where $0 < \delta < 1$. By (1) and (2), operator $T$ is a contraction mapping (Blackwell sufficient condition), and hence, by the Banach fixed point theorem, there exist solutions for $V^* = T(V^*)$. ■

**B. Marriage market clearance**

Proposition 1 shows that rational individuals, who care strongly for the wellbeing of their offspring could end up at stationary equilibria where males and females are of the same value. Since a male and a female have to form a family to have offspring, the intra-household resource allocation and the dynamic of the marriage market also play a crucial role in the determination of the relative status of males and females, in spite of gender roles defined by the gender comparative advantages.

The marriage market is assumed to be monogamous as it is the commonest marriage setting. Since a marriage only lasts for one period, there are no divorces or remarriages. At any period $\tau$, each male adult tries to match with one female adult, so that, with sex ratio $\theta_\tau$, i.e. the number of male adults over that of female ones, the marriage market clearance implies that the possibility of a male to find a spouse is $\min(1, 1/\theta_\tau)$, i.e. 1 if $\theta_\tau \leq 1$; $1/\theta_\tau$ if $\theta_\tau > 1$, and that of a female is $\min(1, 1/\theta_\tau)$.

The marriage market condition, the sex ratio, affects not only the possibility of finding a spouse but also, as per Weiss (1997), the division of gains from a marriage, which could be in the form of either a dowry or a bride-price in traditional societies or the division of labour within families in modern societies. In our model, the payoff to a male is $W_m$ (works one unit of time and consumes $W_m$) if he remains single, and $x_{m\tau}^* + \beta g(q^*, V_{ct+1})$ if married. For a female, the payoff is $W_f$ if single and $x_{f\tau}^* + \beta g(q^*, V_{ct+1})$ if married. Therefore, the net benefit of a marriage is the difference in the payoff between a married couple and two singles (a male and a
female), \(V_m + V_f - W_m - W_f\). We assume that a husband and a wife reach an agreement on each party's share of the net benefit when they get married, with the husband taking \(s\) (\(0 \leq s \leq 1\)) portion of the net benefit with the remainder to the wife.

The marriage market is assumed to clear through bargaining processes so \(s\) is determined by the relative bargaining power of males and females in the market. Following Siow (2008), who finds that marriage market tightness (defined as the sex ratio of unmarried men to unmarried women) is a sufficient statistic for marriage market conditions, we assume that the relative bargaining power of men and women, and hence the resource allocation between husbands and wives is a function of the sex ratio. Hence the share of husbands, \(s\), is a differentiable and non-decreasing function of \(\theta\) so \(s'(\theta) \leq 0\). In a marriage, a husband takes \(s(\theta)\)(\(V_m + V_f - W_m - W_f\)) and a wife takes \([1 - s(\theta)]\)(\(V_m + V_f - W_m - W_f\)) of the total benefit, so that the welfare difference between a married individual and a single one, i.e. \(V_{mt} - W_m\) for men and \(V_{ft} - W_f\) for women, is the share of the total benefit one takes, given the market tightness \(\theta\).

\[
V_{mt} - W_m = s(\theta)(V_{mt} + V_{ft} - W_m - W_f)
\]
\[
V_{ft} - W_f = [1 - s(\theta)](V_{mt} + V_{ft} - W_m - W_f)
\]

Recall equation (7) the household utility possibility frontier

\[
V_{mt}(V_{ct+1}) + V_{ft}(V_{ct+1}) = (1 - t_{mt}^*)W_m + (1 - t_{ft}^*)W_f - x_{ct}^* + 2\beta g(q^*,V_{ct+1})
\]

Combining both equations, we have

\[
(11) \quad V_{mt} = s(\theta)(I^* + 2\beta g(q^*,V_{ct+1}) - W_m - W_f) + W_m
\]
\[
(12) \quad V_{ft} = [1 - s(\theta)](I^* + 2\beta g(q^*,V_{ct+1}) - W_m - W_f) + W_f
\]

Equation (11) shows that the value of a married man depends on the value of his child, who can be either a son, \(V_{ct+1}^m\), or a daughter, \(V_{ct+1}^f\). Recall the possibility of having a son is 0.5 and taking the expectation

\[
(13) \quad V_{mt} = s(\theta)((I^* - W_m - W_f) + \beta[g(q^*,V_{ct+1}^m) + g(q^*,V_{ct+1}^f)] + W_m
\]
\[
(14) \quad V_{ft} = [1 - s(\theta)]((I^* - W_m - W_f) + \beta[g(q^*,V_{ct+1}^m) + g(q^*,V_{ct+1}^f)] + W_f
\]

Moreover, the value of a son or a daughter depends on whether he or she can get married in adulthood. Conditional on \(\theta_{t+1}\), the probability of a son getting married is \(\min(1, 1/\theta_{t+1})\). If married, his value, \(V_{ct+1}^m\), is \(V_{mt+1}\), if single, \(V_{ct+1}^m = W_m\). For a daughter, \(V_{ct+1}^f = V_{ft+1}\) with probability \(\min(1, \theta_{t+1})\), and \(V_{ct+1}^f = W_f\) probability \(1 - \min(1, \theta_{t+1})\).
When there are fewer adult men than women, i.e. $\theta_{t+1} \leq 1$, all men can find spouses and get married so that $V_{ct+1}^{m} = V_{mt+1}$, but only $\theta_{t+1}$ proportion of women get married and $1 - \theta_{t+1}$ proportion remain single with $V_{ct+1}^{f} = V_{ft+1}$ and $V_{ct+1}^{f} = W_{f}$ respectively. Therefore, when $0 \leq \theta_{t+1} \leq 1$, equations (13) and (14) transform to

$$
V_{mt} = s(\theta_{t})[\{(l^{*} - W_{m} - W_{f}) + \beta g(q^{*}, V_{mt+1}) \}
+ (1 - \theta_{t+1})g(q^{*}, W_{f})] + W_{m}
$$

$$
V_{ft} = [1 - s(\theta_{t})][\{(l^{*} - W_{m} - W_{f}) + \beta g(q^{*}, V_{mt+1}) \}
+ (1 - \theta_{t+1})g(q^{*}, W_{f})] + W_{f}
$$

When $1 \leq \theta_{t+1} < \infty$, all women get married but $1 - 1/\theta_{t+1}$ proportion of men remain single so that equations (13) and (14) transform to

$$
V_{mt} = s(\theta_{t})[\{(l^{*} - W_{m} - W_{f}) + \beta [(1/\theta_{t+1})g(q^{*}, V_{mt+1}) + (1 - 1/\theta_{t+1})g(q^{*}, W_{m})] \}
+ \beta g(q^{*}, V_{ft+1})] + W_{m}
$$

$$
V_{ft} = [1 - s(\theta_{t})][\{(l^{*} - W_{m} - W_{f}) + \beta [(1/\theta_{t+1})g(q^{*}, V_{mt+1}) + (1 - 1/\theta_{t+1})g(q^{*}, W_{m})] \}
+ \beta g(q^{*}, V_{ft+1})] + W_{f}
$$

As shown in equations (15) to (18), the adulthood welfare of a man, or a woman, as a part of a household is affected by both the total value of the household and the allocation with the spouse. While the latter is determined by the marriage market tightness at time of marriage, the former relies on the expectation of the welfare of their offspring, especially the likelihood of finding spouses and building their own families. The sex ratio in the future marriage market, $\theta_{t+1}$, plays a crucial role in the expectation. In the case when a shock increases the value of men, $V_{mt+1}$, parents would prefer sons to daughters for higher household values. This increases the sex ratio of their offspring, which will, in turn, drive down the expected value of a son by decreasing the probability of finding a spouse and having offspring.

C. Equilibrium sex ratio with marriage market clearance

Our model defines a system where parents would like to have a child of the sex associated with higher value deriving either from better marriage prospects or receiving a higher share of the benefit of a marriage that does occur. When it is feasible, e.g. technically possible and at low enough cost, this preference will translate into sex selection behaviour. This process keeps biasing the sex ratio until the value of males and females reaches a new balance through the marriage market. As it is observed that, in most countries, parental sex selection is rare and the sex ratio remains stable even after prenatal screening and abortion became accessible and of relatively low cost, we focus on the stationary equilibrium with no incentive for sex selection.
For an equilibrium to have no incentive for sex selection, parents will not benefit from switching their children’s sex. Under the assumption that the cost of sex selection is zero, it implies that the expected value of a son and a daughter are the same, i.e. in the case \(0 \leq \theta_{t+1} \leq 1\),
\[
g(q^*, V_{mt+1}) = \theta_{t+1} g(q^*, V_{ft+1}) + (1 - \theta_{t+1})g(q^*, W_f),
\]
and in the case \(1 \leq \theta_{t+1} < \infty\),
\[
(1/\theta_{t+1}) g(q^*, V_{mt+1}) + (1 - 1/\theta_{t+1}) g(q^*, W_m) = g(q^*, V_{ft+1}).
\]
With this condition, equations (15) to (18) give
\[
\begin{align*}
(19) \quad \frac{V_{mt} - W_m}{s(\theta_t)} &= \begin{cases} 
(I^* - W_m - W_f) + 2\beta g(q^*, V_{mt+1}) & \forall 0 \leq \theta_{t+1} < 1 \\
(I^* - W_m - W_f) + 2\beta[(1/\theta_{t+1}) g(q^*, V_{mt+1}) + (1 - 1/\theta_{t+1}) g(q^*, W_m)] & \forall 1 \leq \theta_{t+1} < \infty
\end{cases}
\end{align*}
\]
\[
\begin{align*}
(20) \quad \frac{V_{ft} - W_f}{1 - s(\theta_t)} &= \begin{cases} 
(I^* - W_m - W_f) + 2\beta[\theta_{t+1} g(q^*, V_{ft+1}) + (1 - \theta_{t+1})g(q^*, W_f)] & \forall 0 \leq \theta_{t+1} < 1 \\
(I^* - W_m - W_f) + 2\beta g(q^*, V_{ft+1}) & \forall 1 \leq \theta_{t+1} < \infty
\end{cases}
\end{align*}
\]
The LHS of equations (19) and (20) show the effect of marriage market tightness (\(\theta_t\)) on the allocation of the benefit of a marriage between a husband and a wife. For a given net benefit, an increase in \(s(\theta_t)\) increases the value to the husband, \(V_{mt}\), and decreases that to the wife, \(V_{ft}\). The RHS is the expected value of a marriage under the assumption that a son and a daughter bring about the same value, which shows the effect of future marriage market tightness (\(\theta_{t+1}\)).

When \(\theta_{t+1} > 1\) (\(\theta_{t+1} < 1\)), the risk of having a son (daughter) who cannot get married drives down the value of a marriage so the net benefit of a marriage is highest at the point \(\theta_{t+1} = 1\) given the value of a son and a daughter. Equation (19) and (20) define a dynamic system for \(V_m, V_f\), and \(\theta\) over time. Since \(\theta_{t+1}\) affects only the majority sex, it can be simplified and defined by
\[
\begin{align*}
(21) \quad \frac{V_{ft} - W_f}{1 - s(\theta_t)} &= (I^* - W_m - W_f) + 2\beta[\theta_{t+1} g(q^*, V_{ft+1}) + (1 - \theta_{t+1})g(q^*, W_f)] & \forall 0 < \theta_{t+1} < 1 \\
(22) \quad \frac{V_{mt} - W_m}{s(\theta_t)} &= (I^* - W_m - W_f) + 2\beta[(1/\theta_{t+1}) g(q^*, V_{mt+1}) + (1 - 1/\theta_{t+1}) g(q^*, W_m)] & \forall 1 \leq \theta_{t+1} < \infty
\end{align*}
\]
For equation (21), the LHS gives the relationship between the value to the wife, \(V_{ft}\) and her benefit share, \(1 - s(\theta_t)\), given the total marriage benefit. Holding the net marriage benefit fixed, it implies
\[
\frac{dV_{ft}}{d\theta_t} = -\frac{F_{\theta_t}}{F_{V_{ft}}} = -(V_{ft} - W_f) \frac{s'(\theta_t)}{1 - s(\theta_t)}
\]
which is non-negative \( \frac{dV_{fr}}{d\theta} \geq 0 \) since \( s'(\theta) \leq 0 \). Define the marriage market responsiveness elasticity for females as the percentage change in the share of the net marriage benefit of wives in response to one per cent change in the sex ratio, or \( \rho_f = \frac{d(1-s)/d\theta}{\theta} = \frac{-\theta s'}{1-s} \), we have

\[
(23) \quad \frac{dV_{fr}}{d\theta} = \frac{1}{\theta} (V_{fr} - W_{fr}) \rho_{fr}
\]

The RHS of equation (21) gives the expected net benefit of a marriage, which shows how the future sex ratio, \( \theta_{r+1} \), affects the net benefit of a marriage benefit as it changes the probability of the female offspring to get married. Holding the net marriage benefit fixed, it implies

\[
(24) \quad \frac{dV_{fr+1}}{d\theta_{r+1}} = -\frac{F_{\theta_{r+1}}}{F_{V_{fr+1}}} = -\frac{1}{\theta_{r+1}} \frac{g(q^*, V_{fr+1}) - g(q^*, W_{fr})}{g_{V_{fr+1}}}
\]

where \( g_{V_{fr+1}} = \frac{\partial g(q^*, V_{fr+1})}{\partial v_{fr+1}} \). Since \( g_{V_{fr+1}} > 0 \), \( \frac{dV_{fr+1}}{d\theta_{r+1}} < 0 \).

For equation (22) where \( 1 \leq \theta_{r+1} < \infty \), the LHS gives the relationship between the value to the husband, \( V_{mr} \), and his benefit share, \( s(\theta_r) \), given the total marriage benefit.

\[
\frac{dV_{mr}}{d\theta} = (V_{mr} - W_m) \frac{s'(\theta_r)}{s(\theta_r)}
\]

which is non-positive \( \frac{dV_{mr}}{d\theta} \leq 0 \) since \( s'(\theta) \leq 0 \). Similarly, define the marriage market responsiveness elasticity for males as the percentage change in the share of the net marriage benefit to husbands in response to one per cent change in the sex ratio, \( \rho_m = \frac{d(s)/d\theta}{s} = \frac{-\theta s'}{s(\theta)} \), we have

\[
(25) \quad \frac{dV_{mr}}{d\theta} = -\frac{1}{\theta} (V_{mr} - W_m) \rho_{mr}
\]

The RHS of equation (22) implies, holding total marriage benefit fixed,

\[
(26) \quad \frac{dV_{mr+1}}{d\theta_{r+1}} = \frac{1}{\theta_{r+1}} \frac{g(q^*, V_{mr+1}) - g(q^*, W_m)}{g_{V_{mr+1}}}
\]

where \( g_{V_{mr+1}} = \frac{\partial g(q^*, V_{mr+1})}{\partial v_{mr+1}} \). Since \( g_{V_{mr+1}} > 0 \), \( \frac{dV_{mr+1}}{d\theta_{r+1}} > 0 \).

At a stationary equilibrium where the value of an adult male or of an adult female does not change over time, i.e. \( V_m = V_{mr} \) and \( V_f = V_{fr} \) for all \( \tau \), and the degree of marriage market tightness is constant, i.e. \( \theta_r = \theta \) for all \( r \), equations (21) and (22) imply
With equation (27), we can discuss the characteristics of the solutions with no parental sex selection at the stationary equilibrium.

**Proposition 2.** With a marriage market where the market tightness, i.e. the sex ratio of males to females, determines husbands' and wives' shares of the net benefit in marriages, there exists a solution at \( \theta^* = 1 \) if \( \lim_{V \to \infty} 2s(\theta^*) \frac{\partial \beta g(q^*, \nu)}{\partial \nu} < 1 \), where \( s(\theta^*) \) is a husband's share of the net benefit in a marriage.

Proof. At \( \theta^* = 1 \), equation (27) gives \( V_m = [s^* \left( I^* - W_m - W_f \right) + W_m] + 2s^* \beta g(q^*, V_m) \), where \( s^* = s(1) \). Similar to the proof of Proposition 1, there exists a solution if \( \lim_{V \to \infty} 2s(\theta^*) \frac{\partial \beta g(q^*, \nu)}{\partial \nu} < 1 \).

At this symmetric equilibrium, not only does the number of males equal the number of females in the population but the life-time welfare of a man and a women are the same. Recall \( s = \frac{V_m - W_m}{V_m - W_m + V_f - W_f} \), \( V^* = V_m = V_f \) implies

\[
\frac{V_m - W_m}{2V_m - W_m - W_f} \Rightarrow \begin{cases} 
\frac{s^*}{2} & \text{if } W_m > W_f \\
\frac{1}{2} & \text{if } W_m = W_f \\
\frac{s^*}{2} & \text{if } W_m < W_f
\end{cases}
\]

Even with gender differences in labour market opportunities, the marriage market bargaining and intra-household allocation redistribute the consumption and equalize the value to men and women. Figure 2 illustrate the solution at \( \theta^* = 1 \). Point \( O \) depicts the equilibrium where \( \theta^* = 1 \) and \( V^* = V_f = V_m = \frac{s^*(W_m+W_f)-W_m}{2s^*-1} = W_m + \frac{s^*}{1-2s^*}(W_m + W_f) \).\(^9\) Curves \( OA_f \) and \( OA_m \) are isocurves of the value of a marriage derived from the RHS of equation (27) in the range of \( 0 < \theta < 1 \) and \( 1 \leq \theta < \infty \) respectively. The closer \( \theta \) is to one, the more the offspring are likely to get married, and hence the higher the value of a marriage is; therefore the isoquant bends downward with a minimum at \( \theta = 1 \). Curves \( OB_f \) and \( OB_m \) represent the value of a marriage to a wife (with a sex ratio less than one) and to a husband (with a sex ratio greater than one) respectively, as given by the LHS of equation (27). For a given value of a marriage, the share of

\(^9\) Note \( V^* = \frac{s^*(W_m+W_f)-W_m}{2s^*-1} = W_m + \frac{s^*}{1-2s^*}(W_m + W_f) \geq \max(W_m, W_f) \).
a wife (a husband) increases (decreases) as \( \theta \) increases, so curve \( OB_f \) is increasing and \( OB_m \) is decreasing with \( \theta \). This equilibrium is also the solution that maximizes the utilitarian social welfare since every individual is able to get married and enjoy \( V^* \) during adulthood.

However, this equilibrium is unstable since shocks are likely to drive it either side of \( \theta = 1 \). Firstly, if \( \theta > 1 \) for example, on curve \( OB_m \), a male gets the value implied by his share, \( s(\theta) \), given the sex ratio \( \theta \). So all points above \( OB_m \) depict cases where males gains extra value, and hence parents would prefer sons to daughters in this area. On the contrary, all points below \( OB_m \) are cases where females gains extra value and parents prefer daughters. Similarly, when \( \theta < 1 \), in the area above \( OB_f \), parents prefer daughters, while they prefer boys in the area below \( OB_f \).

This highlights the importance of marriage market clearance. For example, in a case where a shock causes \( \theta > \theta^* = 1 \), parents do not necessarily prefer daughters to sons in response. The value of a son decreases but that of a daughter also decreases because it is affected by the value of her male offspring. The relative value of a daughter to a son is determined by the intra-household resource allocation, i.e. above or under curve \( B_f OB_m \).

Secondly, all points on \( OA_m \) are the value of males and sex ratio combinations, i.e. \( (V_m, \theta) \), that satisfy the expected value of a marriage. All points above it have a marriage value higher than the expectation, which implies that \( V_m \) increases. Similarly, \( V_m \) decreases in the area below \( OA_m \). When \( \theta < 1, V_f \) increases above \( OA_f \) and decreases below it. In figure 2, curve \( B_f OB_m, A_f OA_m \), and line \( \theta^* = 1 \) divide six areas, where we can see that any shock is likely to lead to divergence from the original \( \theta^* = 1 \).

**Proposition 3.** With a marriage market where the sex ratio of males to females determines husbands’ and wives’ shares of the net benefit in marriages, there exists a solution at either 

\[
0 < \theta^* < 1 \quad \text{or} \quad 1 < \theta^* < \infty, \quad \text{if} \quad \lim_{v \to \infty} 2\theta^* s(\theta^*) \frac{\partial \beta g(q^*,v)}{\partial v} < 1 \quad \text{for} \quad 0 < \theta^* < 1 \quad \text{and} \\
\lim_{v \to \infty} 2 \frac{1}{\theta^*} s(\theta^*) \frac{\partial \beta g(q^*,v)}{\partial v} < 1 \quad \text{for} \quad 1 < \theta^* < \infty.
\]

Proof. Equation (27) implies

\[
\begin{align*}
V_f &= \left[ s^* \left( l^* - W_m - W_f \right) + W_f + (1 - \theta) g(q^*,W_f) \right] + 2\theta^* s^* \beta g(q^*,V_f) \quad \forall 0 < \theta < 1 \\
V_m &= \left[ s^* \left( l^* - W_m - W_f \right) + W_m + (1 - 1/\theta) g(q^*,W_m) \right] + 2 \frac{1}{\theta^*} s^* \beta g(q^*,V_m) \quad \forall 1 \leq \theta < \infty
\end{align*}
\]

where \( s^* = s(\theta^*) \). Similarly to the proof of Proposition 1, there exist a solution if 

\[
\lim_{v \to \infty} 2\theta^* s(\theta^*) \frac{\partial \beta g(q^*,v)}{\partial v} < 1 \quad \text{for} \quad 0 < \theta^* < 1 \quad \text{and} \quad \lim_{v \to \infty} 2 \frac{1}{\theta^*} s(\theta^*) \frac{\partial \beta g(q^*,v)}{\partial v} < 1 \quad \text{for} \quad 1 < \theta^* < \infty.
\]

This solution is illustrated in Figure 3. In the case when the intra-household resource allocation favours husbands, i.e. \( s_2(\theta) > s(\theta) \), the value of married men increases for all \( \theta \) so that curve \( OB_m \) moves up to curve \( O_2B_m \). The intersection of \( O_2B_m \) and \( OA_m, O_2 \), provides the equilibrium
at $1 < \theta^*$ because $(1/\theta)g(q^*V_m) + (1 - 1/\theta)g(q^*W_m) = g(q^*V_f)$. Similarly, if the intra-household resource allocation favours wives, i.e. $s_1(\theta) < s(\theta)$, the curve $OB_f$ moves up to curve $O_1B_f$. The intersection of $O_1B_f$ and $OA_f$, $O_1$, provides the equilibrium at $0 < \theta^* < 1$ with $V_f > V_m$.

The slope of curve $O_2B_m$ at the equilibrium is $\frac{dV_m}{d\theta}|_{O_2B_m} = -\frac{1}{\theta^*}(V_m - W_m)\rho_m$ as shown in equation (25) and that of $O_2A_m$ is $\frac{dV_m}{d\theta}|_{O_2A_m} = -\frac{1}{\theta^*}\frac{g(q^*V_m) - g(q^*W_m)}{g_{V_m}}$ as in equation (26). When the size of the slope of curve $O_2B_m$ is smaller than that of curve $O_2A_m$, i.e. $\left|\frac{dV_m}{d\theta}|_{O_2B_m}\right| < \left|\frac{dV_m}{d\theta}|_{O_2A_m}\right|$, the two curves define a stable oscillation. Rearranging the condition, it is equivalent to $\rho_m < \frac{g(q^*V_m) - g(q^*W_m)}{g_{V_m}(V_m - W_m)}$. So the stability of the equilibrium is determined by the responsiveness of the marriage market and the functional form of the utility function. When $g(q^*, \cdot)$ is close to linear with respect to $V$ in the range between $W_m$ and $V_m$, i.e. individuals are close to risk neutral, the RHS of the condition is close to one. It increases when the curvature increases, i.e. when there is greater risk aversion, and goes to infinity in extreme cases. Similarly, it can be shown that the equilibrium with $\theta^* < 1$ ($O_1$ in figure 3) is stable when $\rho_f > \frac{g(q^*V_f) - g(q^*W_f)}{g_{V_f}(V_f - W_f)}$. So for a given degree of marriage market responsiveness, a population with higher risk aversion is more likely to settle at an equilibrium with more males than females ($\theta^* > 1$). Or for a given level of risk aversion, increasing the marriage market responsiveness can change the stability of these solutions, and eventually change the equilibrium sex ratios.

Since we assume the natural probability of having a boy (or a girl) is 0.5, equilibrium sex ratios that deviate from $\theta^* = 1$ imply the existence of sex selection. Parents prefer the sex associated with the higher value until the probability of not being married balances out the sex premium. While an equilibrium with sex selection could be stable, it could be inefficient as well. For example, in an equilibrium with more men than women, welfare can be improved by "switching" a man into a woman. While the expected value of the switched man remained the same, by "releasing" a wife it increases the welfare of an unmarried man in the population.

III. Empirical issues

Our model provides an explanation for the sharp increase in the sex ratio in developing countries, e.g. India and China. While it is often thought to be caused by the son preference from cultural factors, our model shows that it can equally be driven by altruistic parents with no inherent preference for a specific gender. In the Indian case, marriage squeeze along with population growth puts women in a disadvantageous position on the marriage market, and hence decreases wives' share of resources within a household (Rao (1993) and Neelakantan and Tertilt (2008)). This drives down the value of a daughter (pushes up the value of a married son), and also the
expected value of a marriage. Graphically, it implies that curve $O A_m$ (in figure 3) moves downward and generates a new equilibrium to the right of $\theta^*_2$.

In China, the sex ratio increased dramatically from about the biological normal level in 1980 to about 1.20 in 2000. Research, e.g. Ebenstein (2010) and Li et al. (2011), has found that the one-child policy can explain the majority of the increase. Our model provides a framework to analyze the mechanisms through which the policy affected the sex ratio. Firstly, it decreases the value of a marriage since it imposes a constraint on the number of children in a household. In figure 3, it means curve $O A_m$ moves downward. Secondly, the reduced value of the marriage increases the value of the outside option (remaining unmarried) of men relative to women (because men are higher paid) and hence increases men's bargaining power in the marriage market. Graphically, it implies that curve $O_2 B_m$ moves upward. The two effects combined drive the equilibrium further right. Yang (2013a) further investigates the two mechanisms in a more general model.

While this empirical evidence seems consistent with our model, there are a number of roadblocks to a complete test involving the observability of key variables and the availability of data. In this paper we only begin the process by testing one condition required for equilibrium in our model (or any model where the sex-selection effect is contained by the marriage market). This is the hypothesis that the ratio of number of males to females in the marriage market has a negative effect on the probability of a male birth. We test this hypothesis using the Chinese census.

A. Data

The 1% sample of the 2000 Chinese population census is employed, which contains around 12 million observations across provinces and municipalities in the mainland China. The large sample size is essential for the study. Firstly, the identification relies on a good measure of the marriage market tightness, i.e. the sex ratio of males to females, which requires large enough sample sizes in each cell. Secondly, the large sample size increases statistical power so that we can accurately measure small effects. In this study, we focus on the gender determination of individuals who were born between 1971 to 1990, i.e. observations aged 10 to 30 in 2000.\(^{10}\) Firstly, this is the period when the sex ratio at birth started to deviate upward from the normal rate. Secondly, the registered residence (Hukou) system in China was still strictly regulated and intra-province mobility was low. In order to calculate the sex ratio in the marriage market, we use the number of males and females around age 20 when these individuals were born.

The measure of the marriage market tightness, i.e. the marriage market sex ratio ($MMSR$), is the key independent variable in this study. We use various methods to calculate this variable. In China, the legal marriage age is 20 for females and 22 for males and more than 40% of females

\(^{10}\) Individuals aged between 0 to 9 in 2000 census are excluded since Chan et al. (2013) compare them to those aged between 10 to 19 in 2010 census and find that the former is underreported, with greater shortage for girls than for boys.
choose to get married before they reach age 21. Husbands are on average 2.15 years older than wives. Table 1 provides more details about the marriage statistics in China. In order to capture the number of unmarried males and females in the marriage market, for the benchmark measure, we assume that the marriage market is centered at age 20 and calculate the sex ratio as the number of males of age 20, 21, 22, and 23 over that of females of age 17, 18, 19, and 20. We call this the $t \pm 3$ measure. Men older than 20 are included because women may marry older men, and women of younger cohorts are included because they are potential competitors if men are willing to wait. Sex ratios of $t \pm 5$ and $t \pm 9$ are also used, which, taking $t \pm 9$ for example, is the ratio of males aged 20 to 29 to females aged 11 to 20.\(^{11}\) Moreover, the sex ratio is also calculated under alternative assumptions that the marriage market is centered at age 18 and 22 respectively. The sex ratio is measured for each year, province, and rural/urban cell.\(^{12}\) The rural and urban sectors in one province are considered as two separate marriage markets because the registered residence type (rural/urban) was strictly regulated by the registered residence system. Rural individuals are excluded from a bundle of public goods available to urban individuals, e.g. education for children, healthcare services, even if they live in urban areas. These barriers make urban individuals reluctant to consider potential rural spouses and likely divide the market.

Other covariates include minority, onechild and rural. Minority is an indicator of the ethnic group, defined as 1 if an individual belongs to any of the minority ethnic groups, 0 if he/she belongs to the majority Han group. Onechild equals 1 if an individual was subject to the one-child policy at birth (meaning one belonged to Han and was born after 1978) and is 0 otherwise. Rural represents the registered residence type and equals 1 if rural, 0 if urban. Table 2 provides summary statistics of these variables in our sample.

**B. Identification and empirical models**

This study focuses on the effect of the marriage market sex ratio on the gender determination of children. We mainly employ the variation in the marriage market sex ratio over time caused by changes in the population growth, even though it varies cross-sectionally as well. As per Neelakanta and Tertilt (2008), when women seek to marry older men, positive population growth makes males of a birth cohort relatively scarce in a marriage market as they match to a younger and larger cohort of females. Figure 4 shows the age (birth year) profile of the population, which indicates the fluctuation in the population growth, and the average marriage market sex ratios for each birth cohort. It can be seen that the marriage market sex ratio was driven up by the population decline in the late 1950s to 1.35 and then decreased due to the sharp population growth in early 1960s. These changes in the population growth were caused by the famine in

\(^{11}\) A similar measure is used in Rao (1993) in the Indian context.

\(^{12}\) The sex ratios are also calculated for year-province cells and year-city cells. While the year-city cell approach has a large enough sample size to derive accurate sex ratio measures, there are two issues. First, migration for marriage purpose is much more common among cities in a province than inter-province. Second, in the 2000 census, we only observe birth province rather than birth city so that the city-level sex ratio cannot be matched to individuals who have moved after birth. In any case, regressions using these alternative sex ratios provide similar results.
China from 1959 to 1961 and the concentration of birth after the economy recovered. Starting from 1963, national birth control policies were introduced and population and family planning offices were established in local authorities. This policy largely accounts for the population declines from 1964 to 1967 and in the 1970s with an interruption from 1968 to 1971 due to the start of the Cultural Revolution. Accordingly, the sex ratio increased and then decreased in the 1960s. In figure 4, we can also see that the $t \pm 9$ sex ratio measure is much smoother than the $t \pm 3$ measure.

Figure 5 presents the sex ratio at birth from 1950 to 1999 for Han and the minority ethnic groups separately. It seems that sex ratios at birth of Han and the minority groups follows similar trends before 1990, but the increase in the 1980s for Han is much larger while the minority group sex ratio fluctuates much more throughout. In this study, we focus on the birth sex ratio from 1970 to 1989, and investigate how it is affected by the marriage market sex ratio, i.e. the MMSR of those who were around age 20 during this period. Assuming the variation in the marriage market sex ratio is exogenous to the gender determination, the regression is specified as

$$y_{igt} = \alpha + \beta_0 MMSR_{gt} + \beta_1 MMSR_{gt} \times \text{minority}_{igt} + \gamma X_{igt} + \epsilon_{igt}$$

where $y_{igt}$ equals one if individual $i$ in province-rural/urban cell $g$ of birth cohort $t$ is male and 0 if female; $MMSR$ is the marriage market sex ratio; $X_{igt}$ is a vector of covariates including onechild, rural, a set of residence province indicators, a set of birth cohort indicators, and a set of ethnic group indicators. Because of the large sample size, we estimate the model using the linear probability model (LPM). Following Bertrand et al. (2004) and Moulton (1990), standard errors are clustered at the province level.

$\beta_0$ and $\beta_1$ are the coefficients of interest, which shows the effect of MMSR on the probability of being a boy at the individual level. While $\beta_0$ gives the effect for Han, $\beta_0 + \beta_1$ is the effect for minority ethnic groups. The effect is estimated separately for Han and minority groups for two reasons. First, the one-child policy, which has been found to influence significantly the sex ratio at birth, is for Han only. Second, as shown in figure 5, the sex ratio variation is much larger for minority groups. The coefficient for onechild gives the difference in the probability of being a boy associated with the one-child policy. Since the birth cohort and ethnic group fixed effects are both controlled for, it is comparable to the difference-in-differences estimates in Li et al. (2011).

Another issue is the technology used by parents to determine the sex of a fetus in China, which is generally believed to be the ultrasound scan. This technology was not publicly available until the 1980s. As Chu (2001) notes, "China manufactured its first ultrasound B-machine in 1979. Since 1982, both Chinese-manufactured and imported ultrasound B-machines have been introduced on a large scale. In 1987, over 13,000 such machines were being used in hospitals, on average about six for each county." In order to further explore this issue, we separate individuals who were born in 1970s from those in 1980s and estimate the effect separately.
IV. Results

Table 2 presents the estimates of equation (28) using different measures of the marriage market sex ratio with the benchmark measure ($t \pm 3$ centered at age 20) in the central panel. Column (4) shows that the marriage market sex ratio has a statistically significant negative effect for both Han and the ethnic minority group. The point estimate for Han is -0.021, which means that one percentage point increase in the marriage market sex ratio decreases the probability of being a boy by 0.021 percentage points. Taking an example with the marriage market sex ratio initially at 1.05, it implies that males are born with a probability slightly higher than 0.5 (i.e. 0.5122). In the case when the marriage market sex ratio increases one percentage point from 1.05 to 1.06, it decreases the probability of being a boy to 0.5120 and leads to a sex ratio of 1.049 for the next generation. For the minority group, the point estimate adds up to -0.04, which implies an adjustment in the sex ratio to 1.048 for the next generation. The larger magnitude of the effect for the minority group is likely a reflection of the larger variation shown in Figure 5.

Table 2 also presents the estimates using alternative marriage market sex ratios. While the benchmark $t \pm 3$ measure assumes a concentrated marriage market, the $t \pm 9$ measure assumes an individual can match a spouse in a much larger age span. Table 2 shows that alternative sex ratios generate very similar estimates, which implies that the regression is robust to the age gap assumption. Another assumption on the marriage market is about the age at which it is centered. While the estimates of the sex ratio effect for Han seem robust to this assumption (with increasing magnitudes as the centering age is increased from age 18 to 22), those for the minority group become statistically insignificant.

The estimates in Table 2 show that the sex ratio at birth is negatively correlated with the sex ratio in the marriage market. However, this association could be due to the difference between the rural and urban sectors, because as shown in Table 1 the rural sector is associated with lower marriage market sex ratios but a higher proportion of male births than the urban sector. Columns (1) and (2) of Table 3 investigate this and report the estimates of equation (28) using the rural and urban subsamples separately. The estimates based on the rural subsample provide similar results to the benchmark model in Table 2, which indicates the estimates are not entirely driven by the rural/urban differences. However, the estimates using the urban subsample are not statistically significant.

Table 3 also provides estimates using the 1980s and 1970s subsamples. As discussed in the last section, we expect to observe larger effects if the prenatal screening is mainly through the ultrasound technology. Comparing columns (3) and (6) in Table 3, the estimates support the hypothesis since strong effects are found in the 1980s only.

V. Conclusion and discussion

In this paper, we study the parental sex selection behaviour of altruistic parents without gender specific preference in a monogamous marriage market. An OLG model with the intra-household
resources allocation determined by marriage market bargaining is developed. The model suggests that multiple equilibria exist. While the equilibrium with a balanced sex ratio is unstable, there could be two stable equilibria, one with the sex ratio smaller than 100, another with the sex ratio larger than 100, depending on the responsiveness of intra-household allocation to marriage market tightness. Specifically, if the marriage market responsiveness elasticity, defined as the percentage change in the husbands' share of the total marriage benefit in response to a one percent change in the sex ratio, is small enough, there exists a stable solution with more males than females at equilibrium.

The model can be employed to study the increase in sex ratios in India and China. We find that while the increase in India is likely to be driven by shocks in the marriage market, the one-child policy played a crucial role in the case in China. The policy decreases the value of a marriage and hence causes a decrease in the transfers from a husband to a wife within a household. This results in an increase in the relative value of males to females and leads to a sex ratio more biased to male. The one-child policy is explored more formally in the more general model in Yang (2013a).

In the empirical part, the model is tested with the 2000 Chinese census data. We find evidence that supports a stabilizer effect of the marriage market on sex selection. Marriage market tightness, defined as the marriage market sex ratio of males to females, has a negative and statistically significant effect on the probability of male birth. This evidence is consistent with our model. Our model provides a market-driven explanation for the sharp increase in the sex ratio in developing countries, e.g. India and China, which is often thought to be caused by a son preference based solely on cultural factors. While these alternative theories have very different implications for many policy questions, further empirical research is required to distinguish them.
References


Table 1: Descriptive statistics for selected variables in the 2000 census

<table>
<thead>
<tr>
<th></th>
<th>All (1)</th>
<th>Rural (2)</th>
<th>Urban (3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural</td>
<td>0.7633</td>
<td>0.9494</td>
<td>1.0589</td>
</tr>
<tr>
<td></td>
<td>(0.4251)</td>
<td>(0.2031)</td>
<td>(0.2013)</td>
</tr>
<tr>
<td>MMSR ($t \pm 3$)</td>
<td>1.004</td>
<td>0.9494</td>
<td>1.0589</td>
</tr>
<tr>
<td></td>
<td>(0.2095)</td>
<td>(0.2031)</td>
<td>(0.2013)</td>
</tr>
<tr>
<td>MMSR ($t \pm 9$)</td>
<td>0.8577</td>
<td>0.7987</td>
<td>0.9167</td>
</tr>
<tr>
<td></td>
<td>(0.1778)</td>
<td>(0.1374)</td>
<td>(0.1935)</td>
</tr>
<tr>
<td>Age of first marriage (female) ($^{(i)}$)</td>
<td>21.76</td>
<td>21.35</td>
<td>22.97</td>
</tr>
<tr>
<td></td>
<td>(3.231)</td>
<td>(3.103)</td>
<td>(3.298)</td>
</tr>
<tr>
<td>Age of first marriage (male) ($^{(i)}$)</td>
<td>23.94</td>
<td>23.49</td>
<td>25.07</td>
</tr>
<tr>
<td></td>
<td>(3.915)</td>
<td>(3.894)</td>
<td>(3.733)</td>
</tr>
<tr>
<td>Husband-wife age gap ($^{(ii)}$)</td>
<td>2.145</td>
<td>2.040</td>
<td>2.402</td>
</tr>
<tr>
<td></td>
<td>(3.584)</td>
<td>(3.597)</td>
<td>(3.539)</td>
</tr>
<tr>
<td>Male</td>
<td>0.5114</td>
<td>0.5127</td>
<td>0.5077</td>
</tr>
<tr>
<td></td>
<td>(0.4999)</td>
<td>(0.4998)</td>
<td>(0.4999)</td>
</tr>
<tr>
<td>Minority</td>
<td>0.0985</td>
<td>0.1064</td>
<td>0.0731</td>
</tr>
<tr>
<td></td>
<td>(0.2980)</td>
<td>(0.3084)</td>
<td>(0.2603)</td>
</tr>
<tr>
<td>Onechild</td>
<td>0.5462</td>
<td>0.5588</td>
<td>0.5047</td>
</tr>
<tr>
<td></td>
<td>(0.4979)</td>
<td>(0.4965)</td>
<td>(0.5000)</td>
</tr>
<tr>
<td>Number of observations</td>
<td>4,035,146</td>
<td>3,080,054</td>
<td>955,092</td>
</tr>
</tbody>
</table>

Notes: Standard deviations are reported in parentheses. Statistics are derived using individuals in the 2000 census who were born between 1970 and 1989 except for those rows designated by ($^{(i)}$), where only ever-married individuals are used and the row designated by ($^{(ii)}$), where the calculation is based on the set of marriages of economic family heads.
Table 2: Effects of the sex ratio with various definitions on the probability of having a boy

<table>
<thead>
<tr>
<th></th>
<th>Centered at age 18</th>
<th></th>
<th>Centered at age 20</th>
<th></th>
<th>Centered at age 22</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t ± 3</td>
<td>t ± 5</td>
<td>t ± 9</td>
<td>t ± 3</td>
<td>t ± 5</td>
<td>t ± 9</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td><strong>MMSR</strong></td>
<td>-0.015***</td>
<td>-0.011**</td>
<td>-0.017*</td>
<td>-0.021***</td>
<td>-0.019***</td>
<td>-0.020*</td>
</tr>
<tr>
<td></td>
<td>(0.0057)</td>
<td>(0.0050)</td>
<td>(0.0098)</td>
<td>(0.0065)</td>
<td>(0.0066)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td><strong>MMSR<em>minority</em></strong></td>
<td>-0.027***</td>
<td>-0.032***</td>
<td>-0.029***</td>
<td>-0.019***</td>
<td>-0.025***</td>
<td>-0.030***</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0079)</td>
<td>(0.0084)</td>
<td>(0.0064)</td>
<td>(0.0083)</td>
<td>(0.0094)</td>
</tr>
<tr>
<td><strong>Onechild</strong></td>
<td>0.0058*</td>
<td>0.0052</td>
<td>0.0009</td>
<td>0.0052</td>
<td>0.0057*</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td>(0.0032)</td>
<td>(0.0031)</td>
<td>(0.0031)</td>
<td>(0.0031)</td>
<td>(0.0031)</td>
<td>(0.0032)</td>
</tr>
<tr>
<td><strong>Rural</strong></td>
<td>0.0013</td>
<td>0.0017</td>
<td>0.0012</td>
<td>0.0004</td>
<td>0.0007</td>
<td>0.0006</td>
</tr>
<tr>
<td></td>
<td>(0.0042)</td>
<td>(0.0043)</td>
<td>(0.0048)</td>
<td>(0.0043)</td>
<td>(0.0044)</td>
<td>(0.0050)</td>
</tr>
<tr>
<td><strong>R^2</strong></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of Prov.</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>No. of Obs.</td>
<td>4,035,146</td>
<td>4,035,146</td>
<td>4,035,146</td>
<td>4,035,146</td>
<td>4,035,146</td>
<td>4,035,146</td>
</tr>
</tbody>
</table>

Notes: Standard errors clustered at the province level are reported in parentheses; the associated level of statistical significance is based on a t-test with df=G-2. *p < 0.1, **p < 0.05, ***p < 0.01. The other covariates are sets of province indicators, birth year indicators, and indicators for ethnic groups.
Table 3: Effects of the marriage market sex ratio on the probability of having a boy with subsamples

<table>
<thead>
<tr>
<th></th>
<th>Rural</th>
<th>Urban</th>
<th>1980s</th>
<th></th>
<th>1970s</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>All</td>
<td>Rural</td>
<td>Urban</td>
<td>All</td>
</tr>
<tr>
<td>MMSR</td>
<td>-0.021**</td>
<td>0.002</td>
<td>-0.024</td>
<td>0.007</td>
<td>0.005</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.0096)</td>
<td>(0.0072)</td>
<td>(0.0156)</td>
<td>(0.0365)</td>
<td>(0.0177)</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>MMSR*minority</td>
<td>-0.020**</td>
<td>-0.009</td>
<td>-0.054***</td>
<td>-0.055***</td>
<td>-0.040***</td>
<td>0.009</td>
</tr>
<tr>
<td></td>
<td>(0.0080)</td>
<td>(0.0106)</td>
<td>(0.0111)</td>
<td>(0.0139)</td>
<td>(0.0139)</td>
<td>(0.0074)</td>
</tr>
<tr>
<td>Onechild</td>
<td>0.002</td>
<td>0.017***</td>
<td>0.003</td>
<td>-0.001</td>
<td>-0.012**</td>
<td>0.019***</td>
</tr>
<tr>
<td></td>
<td>(0.0036)</td>
<td>(0.0044)</td>
<td>(0.0044)</td>
<td>(0.0034)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>0.003</td>
<td></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>R²</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
</tr>
<tr>
<td>No. of Prov.</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
<td>31</td>
</tr>
<tr>
<td>No. of Obs.</td>
<td>3,080,054</td>
<td>955,092</td>
<td>2,159,567</td>
<td>1,719,398</td>
<td>440,169</td>
<td>1,875,579</td>
</tr>
</tbody>
</table>

Notes: The MMSR is measured using $t \pm 3$ centered at age 20. Standard errors clustered at the province level are reported in parentheses; the associated level of statistical significance is based on a t-test with df=G-2. *p < 0.1, **p < 0.05, ***p < 0.01. The other covariates are sets of province indicators, birth year indicators, and indicators for ethnic groups. The coefficient for Onechild is omitted for the regressions for the 1980s because the one-child policy was introduced in 1979.
Figure 1

\[ V = V^* \]
\[ V = l'/2 + \beta g(q^*, V^*) \]
\[ V = l'/2 \]

Figure 2

\[ (l^* - W_m - W_f) + 2\beta(\theta g(q^*, V_f) + (1 - \theta)g(q^*, W_f)) \]

Area 1: \( V_m \uparrow, \theta \uparrow \)

Area 2: \( V_m \downarrow, \theta \uparrow \)

Area 3: \( V_m \downarrow, \theta \downarrow \)

Area 4: \( V_f \downarrow, \theta \uparrow \)

Area 5: \( V_f \downarrow, \theta \downarrow \)

Area 6: \( V_f \uparrow, \theta \downarrow \)

\[ 0 \]

\[ V_f - W_f \]

\[ W_m + \frac{s^*}{1 - 2s^*}(W_m + W_f) \]

\[ W_m \]

\[ V_f \]

\[ B_f \]

\[ B_m \]

\[ V_m - W_m \]

\[ \frac{V_m - W_m}{s(\theta)} \]

\[ \theta^* = 1 \]
Figure 3

Population growth and the average marriage market sex ratios from 1950 to 1980

Note: The axis on the left is for population and the axis on the right is for MMSRs.
Figure 5

Sex ratio at birth from 1950 to 1999
Han vs. Minority ethnical groups