Multidimensional Premarital Investments with Imperfect Commitment*

V. Bhaskar† Wenchao Li‡ Junjian Yi§

June 5, 2020

Abstract

This paper studies how imperfect commitment within marriage affects parental investments in their children. Using nationally representative Chinese household survey data, we examine the effect of a more male-biased local sex ratio upon investments in boys, relative to investments in girls. We find that boys' parents migrate more to improve earnings, and increase investment in housing while reducing educational investment. These patterns are consistent with imperfect commitment combined with severe marriage market competition: At the time of marriage, a boy cannot commit a larger share of his labor earnings to his spouse, while housing, as a public good, is shared equally. Our theoretical model of multidimensional premarital investments with imperfect commitment has predictions which are consistent with these empirical findings.

Key words: Premarital investments; Imperfect commitment; Sex ratio imbalance; Public goods; Human capital investments; Human capital development

JEL Codes: J12; J13; J16; J18; J24; D10; O15; J61

*We thank Sangwoo Choi for capable research assistance.

†Department of Economics, University of Texas at Austin; E-mail: v.bhaskar@austin.utexas.edu.

‡Corresponding author: Wenchao Li. School of Business and Management, Shanghai International Studies University; E-mail: wenchao@shisu.edu.cn;

§Department of Economics, National University of Singapore; E-mail: junjian@nus.edu.sg.

1 Introduction

What determine premarital investments in children when marriage market considerations are important? Following Becker (1973), many economists have examined this question. Much of the literature examines the question while assuming transferable utility (see Cole et al., 2001; Iyigun and Walsh, 2007). This effectively assumes full commitment at the time of marriage. However, recent work has begun to examine the implications of imperfect commitment (see Anderson and Bidner, 2015; Galichon et al., 2019). Indeed, some papers assume non-transferable utility, which is an extreme instance of imperfect commitment (see Peters and Siow, 2002; Bhaskar and Hopkins, 2016).[1]

The imperfect commitment assumption is particularly compelling in societies like China. Before marriage, prospective brides are in an enviable position due to excess of males on the marriage market. After marriage, divorce is prohibitively costly and the traditional power of husbands reasserts itself. Imperfect commitment arises from the divergence in the relative bargaining powers of men and women, between the ex-ante stage (before marriage) and the ex-post stage (after marriage).

This paper studies how imperfect commitment within marriage affects premarital investments in children undertaken by their parents, both empirically and theoretically. The empirical analyses are in the setting of China's marriage markets. As Figure 1 shows, the fraction of births in China that are male has been increasing, foreshadowing a sizable bride shortage. In an influential study, Wei and Zhang (2011) document a competitive saving motive: Parents increase investments in their sons in order to improve their marriage market prospects, as the excess of males increases. Many subsequent studies show that the sex ratio imbalance—and the resulting marriage market competition—induces men to increase premarital investments relative to women. However, these papers do not examine whether all components of investments move in the same direction. Our main empirical finding is

[1]The textbook of Browning et al. (2014) treats family economics in a full commitment paradigm. Other examples of marriage matching and investments models include Siow (1998); Choo and Siow (2006); Choo (2015); Low (2017); Chiappori et al. (2018a); Zhang (2019, 2020).

that this is not the case.

We distinguish between two forms of capital in premarital investments: bequeathed physical capital (such as housing) and human capital.[2] While it has been well documented that both forms of capital enhance men's marriage market prospects, we propose that, given the lack of commitment, housing is more effective than human capital.

Specifically, marital partners are unable to commit, at the time of marriage, to share future household resources in a pre-agreed fashion. If a man invests in human capital, which will increase his future labor earnings, the sharing of this between spouses is determined by ex post—rather than ex ante—bargaining. If a man invests in housing, which is a public good and thus non-excludable, spouses jointly consume it without bargaining. Hence, as men traditionally have a high bargaining power after marriage, housing signals a credible commitment and thus is more favorable to women. That is, a man's attractiveness as a marital partner depends not only on the total, but also on the composition of investments. This creates an incentive for parents with sons to shift their investments towards housing and away from human capital when marriage markets are unfavorable to males.

To investigate the relationship between the sex ratio imbalance and investments undertaken by parents, we use data from a nationally representative Chinese household survey—the 2010 China Family Panel Studies (CFPS) survey. We examine the effect of county-level sex ratios (defined as the ratio of males to females) on investment decisions for families with a first-born son, using families with a first-born daughter in the same county as a comparison group.[3] We control for county fixed effects to deal with unobservable heterogeneity. The identification partly relies on a well-recognized demographic regularity that the gender of the first child is plausibly random, and gender selections in China typically occur at second

[2]In countries like China, housing improves a man's marriage market prospects and therefore, has been traditionally considered as an investment in preparation for marriage. Premarital investment in housing is generally made by a man's parents, as housing is the single most important piece of physical capital in Chinese families which represents a large financial burden for the young. Housing bought by parents at the time when their son is young can be regarded as one for his marriage because of the bequeathable nature of housing capital and the prevalence of intergenerational coresidence. See Section 2 for more detailed discussions.

[3]The administrative divisions of China consist of four practical levels—province, prefecture, county, and community.

and higher order births (see Figure 1).

Our results show that parents of boys are more likely to increase labor supply and in particular, to become migrant workers as the sex ratio increases. In China, migration substantially raises family income, thereby permitting larger investments in children.[4] More importantly, the composition of investments is affected by the sex ratio imbalance, with the share invested in housing increasing relative to the share in children' education in families with sons. Specifically, a 0.1 increase in the local sex ratio is associated with a roughly 24.1 percent increase in the probability of having a migrant father, a 4.1 percent increase in house construction area, and a 11.0 percent decline in annual education expenditure per child, for first-son families relative to first-daughter families.[5]

We show that our empirical findings are not driven by differential family structures due to son-preferring fertility stopping rules.[6] We then provide a comprehensive set of robustness checks to address the concerns about the potential endogeneity of sex ratios. In particular, we use recent development in high-dimensional methods to select important sex-ratio confounders and add them into regressions to isolate the independent role of sex ratio imbalance (Belloni et al., 2014). In addition, while mainly focusing on the ordinary least squares (OLS) estimates, we show that estimations using the instrumental variable (IV) method yield similar results when we explore quasi-experimental variation in sex ratios at the county level induced by the implementation of China's family planning policy.

Motivated by these empirical findings, we develop a model of premarital investments that incorporates imperfect commitment and distinguishes between investments in a public good (bequeathed physical capital) and a private good (human capital). We assume that the two partners cannot commit to a future surplus sharing rule at the time of marriage. The public good is shared without bargaining within marriage, while shares in the private good are determined by ex-post bargaining, with bargaining weights that do not depend upon

[4]In China, household registration (hukou) system makes it prohibitively difficult for migrants to settle down in the destination, who instead, usually leave home temporarily to increase earnings (Zhao, 1999).

[5](0.1 is the overall increase in the sex ratio in China between 2002 to 2010. It also corresponds to one standard deviation of county-level sex ratios.

[6]Families are more likely to stop childbearing if they have a son.

marriage market conditions. We show the existence of a unique, stable equilibrium under quite general conditions. The model setting enables us to perform a more general welfare analysis on how equilibrium investments differ from utilitarian efficient investments, as well as rich comparative statics analysis. In particular, we find that bargaining power imbalances distort investment decisions relative to those that maximize utilitarian efficiency—the gender with greater bargaining power, men, overinvests, while the weaker one, women, underinvests. Similar distortions arise if there is greater competition for women by men, and the two distortions reinforce one another. When men are in excess supply but have greater ex post bargaining power, the model reproduces the empirical patterns we have found. In particular, as the sex ratio becomes more unbalanced, men invest more in the public good, while women invest less. When the cost function exhibits sufficient supermodularity, this may cause investments in the private good to move in the opposite direction.

We have already discussed the most relevant literature on premarital investments, and so we will not repeat this discussion here. Most of this literature focuses on one-dimensional investments. An exception is Dizdar (2018), who examines how multidimensional investments give rise to the possibility of coordination failures and inefficiencies even in a transferable utility setting. Nonetheless, there always exists an efficient equilibrium, when utility is transferable. This contrasts with our findings obtained when utility is imperfectly transferable.

This paper is also related to the literature on limited commitment within the household and the implications for household decisions. The pioneering work of Lundberg and Pollak (1993) introduced "separate-spheres" bargaining. The implications of the inability to commit have been examined in several contexts: intertemporal consumption (Mazzocco, 2007), private expenditures and time use (Lise and Yamada, 2018), underinvestment in childcare and child education (Gobbi, 2018), and fertility (Doepke and Kindermann, 2019). Chiappori and Mazzocco (2017) discuss popular household models with limited commitment.

More closely, Nunn (2005) builds a model in which bride price serves as a credible commitment by men. Anderson and Bidner (2015) develop a theory in which marital payment—with property rights that can be clearly defined and easily transferred at the time of marriage—acts as a more effective commitment device and therefore, is valued more than education in the marriage market. Our paper establishes, both empirically and theoretically, that families with sons invest more in housing to signal credible commitment at the expense of lower educational investment, which is related to marriage competition due to sex ratio imbalance.

Finally, this paper contributes to the literature on sex ratio imbalance in China and other societies.[7] Our empirical analysis builds upon Wei and Zhang (2011), who show that parents facing high sex ratios competitively save more to improve marriage market prospects of their sons. We show that, for the same purpose, parents with sons shift the composition of investments towards housing and away from education. Such distorted investment patterns may hurt the human capital development of young-generation men.[8] The detrimental effect is even larger if we take into account parental migration induced by a greater earning incentive.[9] Our data indeed show a negative relationship between sex ratios and human capital outcomes of boys, including cognitive and non-cognitive skills and health. Since early-stage human capital outcomes have cumulative impacts on late-stage achievements and lifetime productivity, this may be especially costly (Heckman, 2006; Cunha and Heckman, 2007; Heckman, 2007; Heckman et al., 2013). Hence, the social costs of sex ratio imbalance may have been largely understated.

The next section provides background information. Section 3 describes the data and empirical strategy. Section 4 reports empirical results and shows the robustness. Section 5 discusses the results, which motivates the theoretical analysis in Section 6. The paper concludes with a brief summary and discussions about the implications for human capital development of the next generation and for the study of marriage.

[7]This literature includes, among others, Chiappori et al. (2002); Ebenstein (2010); Bhaskar (2011); Ebenstein (2011); Wei and Zhang (2011); Bhaskar and Hopkins (2016); Lin et al. (2016).

[8]This may be one of the reasons for the fact that along with rising sex ratios, the level of high school enrollment rate of men relative to women is decreasing in China (see Appendix Figure A1).

[9]Because of the strict household registration system, children of migrant workers are typically left behind in their hometown. Based on the population census, there are more than 60 million left-behind children in China (Zhang et al., 2014).

2 Background: sex ratio imbalance, marriage market, and premarital investments in China

In this section, we first describe sex ratio imbalance in China. We then present evidence for ex ante marriage market competition and ex post imperfect commitment. We also discuss the extent to which housing can be considered as investment in preparation for marriage.

Sex ratio imbalance The sex ratio at birth in China has increased drastically over time, from 1.12 boys per girl in 1990 to 1.2 in 2000. It has stabilized at that high level since then. In the current population under age 15, there are 13 percent more boys than girls. The imbalance primarily is due to sex-selective abortions, which in turn can be attributed to traditional preference for sons (Ebenstein, 2010) and in part to China's family planning policy (Li et al., 2011). Specifically, parents undertake gender selections to satisfy their dual interests in having a son and complying with the birth quota stipulated by the policy.[10]

Ex ante: marriage market competition High sex ratios at birth decades ago have led to an oversupply of marriage-age men in the market, which results in intensified competition for prospective brides and increased marriage expenditures facing parents with sons. The CFPS survey data show that household expenditure on marriage ceremonies increases by about 24 percent as the local sex ratio rises by 0.1 (see Appendix Figure A2 for graphical evidence). In particular, grooms' families are spending more on marriage over time, whereas brides' are less affected (see Appendix Figure A3 and Brown et al., 2011).

Ex post: imperfect commitment Imperfect commitment within marriage is partly reflected by frictions in the marriage market—or more specifically, the difficulty in divorce. In China, the share of the population divorced and the divorce rate are sufficiently low for different genders, age cohorts, and education levels (see Appendix Table A1). As marriage

[10]Li and Pantano (2013) structurally estimate the demographic consequences of gender-selection technology.

is costly to reverse, marriage market conditions have little effect on the bargaining position of husband and wife within marriage. This indicates that once married, divorce is unlikely an outside option if within-marriage negotiation breaks down, consistent with the setting of separate spheres bargaining model which incorporates imperfect commitment (Lundberg and Pollak, 1993).

Given that female bargaining power within the household is traditionally low and the enforcement of female alimony rights is generally weak in developing countries, a woman who is in a favorable position at the time of marriage will lose this advantage after being married. Such an asymmetry between ex ante and ex post bargaining power in marriage gives rise to imperfect commitment.

Premarital housing investments In China, housing as a form of bequeathable physical capital has been traditionally considered as an investment in preparation for marriage. Wrenn et al. (2019) show empirical evidence for housing investment in China as a provision for marriage. In general, a marriage-age man or his family is required to cover the cost, at least substantially, of providing a home for the newlyweds (Huang, 2010; Shepard, 2016; Pierson, 2010). Family housing wealth enhances a man's marriage market prospects: Typically, a man with more housing wealth is a more desirable partner in the marriage market.[11] In rural areas, a household is much less likely to have an unmarried adult son if they have a higher-quality house; in urban areas, a household is less likely so if they are a homeowner as opposed to a renter (Wei and Zhang, 2011).

Housing capital bought by the parents (and potentially used by the parents) at the time when the future groom is young, can be regarded as one for his marriage as well: First, housing purchased by parents—which has a dominant role in household wealth composition—is a form of investments in children to the extent of the bequeathable nature of housing capital (Xie and Jin, 2015). Second, a marriage-age man often has not yet accumulated enough

[11] In some personal interviews, most respondents shared that they would not like to get married if they still had to rent (Xinhua, 2011). According to recent surveys, nearly 70 percent of women indicated that housing consideration was a priority in choosing a husband (Huang, 2010; Beijing, 2013).

wealth to afford a house on his own and needs his parents' assistance. Housing purchase is usually the most important component of household expenditure for children's marriage in China (Pierson, 2010). Third, intergenerational family co-residence is common in China, especially in rural areas. While in some cases the groom's family buys a new house for the couple, more than 70 percent of young couples live with the groom's parents, at least within the first few years after marriage. Therefore, this paper focuses on family housing investments as physical capital investment for children's marriage.

It is also worth noting that while both family housing wealth and the groom's education improve marriage prospects for men, housing turns out to be more important in China. Analyses based on the population census data show that better housing condition is significantly correlated with a larger probability of finding a martial partner for men. High education also does so, yet the correlation becomes insignificant when taking into account the effect of housing (see Appendix Table A2).

3 Data and regression model

To investigate how high sex ratios affect premarital investments and their composition, we use data from the CFPS survey. This section introduces the data, describes our main outcome variables, and presents the regression model.

3.1 The China Family Panel Studies survey

The CFPS survey is widely considered nationally representative due to its large sample size and advanced sampling design. The survey contains datasets with high-quality information at the individual (both adult and child), household, and community levels. It consists of a total of 14,798 households, and includes 33,600 adults and 8,990 children who were successfully interviewed. The CFPS survey covers 645 communities in 25 designated provinces (out of 34 province-level units), representing 95 percent of the entire population in contemporary

China (Xie, 2012).[12] In addition, the survey implements a scientifically stratified multi-stage sampling design.

One strength of the CFPS survey lies in its comprehensive information. The family-level dataset contains details of family activities and household characteristics such as migration, expenditures, investments, income and wealth, as well as fertility. For each family surveyed, detailed information is available on demographic and labor-market characteristics of family members, such as age, gender, schooling years, occupation, and working location, as provided in the individual-level dataset. The community-level dataset offers regional demographic and socioeconomic information. The datasets are linked across different levels by a set of identification numbers, and in particular, a household identification number allows us to group individuals by living unit. Parent-child relationship can also be precisely identified. Outcome variables of interest are thus readily linked with potential covariates, enabling systematic empirical analyses.

Sample construction Our empirical analyses are based on a cross-sectional sample of households drawn from the 2010 nationwide CFPS baseline survey—which provides the most comprehensive set of information on household activities compared with other waves—and exploit cross-county variations in the sex ratio.[13]

Specifically, we extract a sample of households in the 2010 CFPS family dataset in which the first-born child was between the ages of zero and 15 years, both parents were alive and at most 50 years old, and at least one parent participated in the adult survey. We focus on families in which the eldest child was under the age of 15 to minimize the possibility that the children have started work or participated actively in household decision-making. We impose a constraint on the ages of parents to minimize the probability of their retirement or their ineffectiveness in making investment decisions due to, for example, health reasons. By placing age limits on both parents and children, we maximize comparability across families.

[12] Hong Kong, Macao, Taiwan, Xinjiang, Tibet, Qinghai, Inner Mongolia, Ningxia, and Hainan are not included.

[13] This does not render our analyses less strong, partly because variation in sex ratios across survey waves is not that large within a county.

The main sample contains 4,314 observations.

3.2 Main outcome variables

Our empirical analyses mainly use three sets of outcome variables: parental labor supply (as a proxy for stronger earning/investment incentive), housing investments, and child educational investments.

Parental labor supply We construct five measures of parental labor supply. The first three are binary variables that equals one a) the father, b) the mother, and c) either parent works away from the home town. The information comes from a question in the CFPS family survey that asks whether any member in the family works in a place that is not where the household is registered or where its permanent address is. According to China's household registration system that is used to differentiate permits of where people are allowed to live and work—the hukou system—it is prohibitively difficult for migrants to assimilate with the local population.[14] Instead, migrants usually leave home temporarily to increase their earnings. This kind of circular migration is considered a crucial form of labor supply in China (Zhao, 1999), which is typically associated with a large increase in gross family income (see Appendix Table A3); migration remittances sent back by migrant workers can be used for family investments. As Panel A, Table 1 shows, approximately 9.8 percent of fathers and 2.5 percent of mothers in our sample work outside their hometown; 11 percent of households have at least one migrant parent.

The other two measures are yearly working hours of the father and mother. This information is available from the CFPS adult dataset. More working time generally is associated with more labor income, and therefore, can ease the household budget constraint and increase investments in children. In our sample, the average father works 2,466 hours per year, and the average mother works 2,416 hours per year.

[14] The hukou system results in institutionalized discrimination against migrants: They have limited access to various benefit schemes that are available to local residents, and their children are often denied access to public schools (Zhao, 1999).

Housing investments We construct three variables from the CFPS family dataset for housing investments: construction area, an ownership dummy, and mortgage debts. The ownership dummy indicates whether the family owns the property right of any house, and equals one if the property deed and other relevant contracts of one or more houses belong solely to this family; self-constructed houses in rural regions are also counted. Construction area refers to the area for residential use, and is specified for home owners. A larger construction area or a higher mortgage signals houses of higher quality, and typically is indicative of greater housing investments. Panel B, Table 1 shows that 83.1 percent of the families in our sample own a house. Mean construction area is 126 square meters among homeowners. An average household has a mortgage of 5,392 yuan.

Child educational investments We construct two variables of child educational investments, focusing on the first-born child in the family. The first, education expenditure, is yearly total expenses on the child's education, including tuition fees, book and stationery costs, after-class tutoring expenses, accommodation fees, and commuting fees, yet excluding living expenses. The second variable, an education funding dummy, equals one if the family has put aside a specialized fund for the child's education. Both variables are defined for children who are at least two years old, and the information comes from the CFPS child dataset. As in panel C, Table 1, the average yearly education expenditure is 1,507 yuan per child, and 29.7 percent of families have an education fund for the child.

3.3 Regression model

We estimate the following regression model:

$$k_{ic} = \beta_0 + \beta_1 FirstSon_{ic} + \beta_3 FirstSon_{ic} * SexRatio_c + X_{ic}\Gamma + \lambda_c + \epsilon_{ic}, \quad (1)$$

where k_{ic} represents outcome variables for household i in county c, $FirstSon_{ic}$ is a binary indicator that equals one if the first-born child in the family is a boy, and $SexRatio_c$ refers to the county-level sex ratio. A vector of additional control variables, X_{ic}, includes various parental and household characteristics: both parents' age, schooling years, hukou status, political status, plus age of the first child, region of residence, and ethnicity (column 1 of Table 2 reports the summary statistics).[15] Regressions also control for county fixed effects, λ_c, to account for unobservable cross-county heterogeneity. The error term is denoted by ϵ_{ic}.

We assume that parents infer the local sex ratio from the premarital-age cohort. In our main regressions, we use sex ratios for those between the ages of ten and 24 years, which are obtained from the 2010 population census. Later we check whether empirical findings are insensitive to using sex ratios for different age brackets (Section 4.4), as in prior work (Wei and Zhang, 2011). Sex ratios are at the county level, as each county can be treated as a local marriage market. China's hukou system presents a formidable obstacle to marriage migration (Davin, 2005; Wei and Zhang, 2011). The census shows that more than 90 percent of rural residents and 62 percent of urban residents live in their county of birth and, 89 percent of couples are from the same county. Of migrant couples in cities, 82 percent are from the same city, suggesting that migrants often get married before leaving their hometown.

We focus on the interaction-term coefficient β_3, which measures the effect of high sex ratios on outcomes for first-son families, using first-daughter families as a control group. For example, a positive estimate of β_3 when the outcome variable is parental migration, as we may expect, suggests that when sex ratios are high, parents of boys have a desire to earn more and invest more in children than parents of girls. The sign and magnitude of the estimate of β_3 when the outcome variables measure housing investments and child educational investments tell how sex ratio imbalance affects premarital investments in different forms of capital in first-son families relative to first-daughter families.

[15]In robustness checks, we include different sets of controls. For example, we add the number of children, average household income, etc.

3.3.1 Identifying assumptions

Obtaining unbiased OLS estimates of β_3 requires that in equation (1), the error term is not substantially correlated with the interaction between the first-son dummy and sex ratio.

The identification partly relies on a well-recognized demographic regularity: The first child in a family being a boy or a girl is plausibly random. Data from China population censuses (1982, 1990, 2000, and 2010) reveal that high sex ratios in China are driven by imbalances in second- and higher-order births, while the sex ratio for first births is rather stable and falls in the biologically normal range; see Figure 1. Parents are least likely to practice gender selection on the first birth, despite their son preferences. Before 2015, a second child was officially permitted if the first one was a girl for households in most rural areas, where son preferences appear stronger. This "1.5 children" policy was applicable to residents who accounted for more than 60 percent of the Chinese population, and markedly alleviated their motivation to abort the first daughter.[16]

Statistical evidence from our sample also validates the randomness of first-child gender. An average family in our sample has 1.5 children, consistent with the above-mentioned policy. Nearly half of the families have a first son and the other half have a first daughter. Specifically, the mean of the first-son dummy is 0.507, which implies a sex ratio of 1.03, well within the normal range; the standard deviation is 0.5, which suggests that first-child gender is like a random Bernoulli trial in which having a boy or a girl has an equal probability;[17] see Table 2. In addition, we regress the first-son dummy on the full set of control variables used in our analyses, and find no significant correlation with these variables.

The strongest evidence in favor of the randomness of first-child gender is that first-son and first-daughter families have similar predetermined parental and household characteristics in our data, as the balance test in Table 2 shows. For example, 12.1 percent of first-son families and 12.8 percent of first-daughter families belong to minority ethnic groups. The difference

[16]The policy was replaced by a nationwide two-child policy in 2015, which further alleviates the motivation to abort the first daughter. Ebenstein (2011, 2010) shows that most Chinese families prefer one boy and one girl to two boys.

[17]A Bernoulli random variable with a mean of 0.5 has a standard deviation of 0.5.

is -0.007, which is statistically indifferent from zero at the ten percent level or below.

In the next section, we will present a broad range of robustness analyses to test whether our empirical results are driven by identification issues, which mainly come from the potential endogeneity of local sex ratios.

4 Empirical results

This section presents empirical results on how sex ratio imbalance affects parental labor supply, housing investments, and child educational investments for families with a first-born son relative to those with a first-born daughter. The results are shown to be robust to potential issues related to son-preferring fertility stopping rules and potential endogeneity of local sex ratios.

4.1 Sex ratio imbalance and parental labor supply

We begin our empirical analyses by estimating equation (1) based on our sample, using measures of parental labor supply constructed in Section 3.2 as outcome variables. Results are reported in the first panel of Table 3. In these and the following estimations, we mainly use the OLS method;[18] estimations are weighted by the CFPS survey sampling weights; standard errors are clustered at the county level and given in parentheses.

In column (1) where the outcome variable indicates father's migration, the estimate is 0.235 (standard error 0.094) for the coefficient on the interaction between the first-son dummy and sex ratio, β_3, as reported at the top of the column.[19] The positive sign and statistical significance (at the five percent level) of the estimate suggest that a high sex ratio is much more likely to induce the migration of fathers of a first-born son relative to fathers of a first-born daughter. Specifically, the estimate implies that with an increase in

[18]Our OLS estimates are similar to the marginal effects from the Probit estimations when the outcome is a binary variable.

[19]The estimates of coefficients on other control variables, which are unreported for brevity, have the expected sign and magnitude.

the sex ratio from 1.08 in 2002 to 1.18 in 2010, the probability of having a migrant father, on average, significantly increases by 2.4 percentage points for a first-born boy relative to a first-born girl, all else held equal. This difference is economically significant, representing a 24.1 percent difference relative to the baseline father-migration probability of about 0.1 in our sample, as reported in the middle of the column.

We obtain qualitatively similar findings for mother's migration and the migration of at least one parent, according to columns (2) and (3) of panel A, Table 3. The effect of sex ratio imbalance for first-son families relative to first-daughter families is substantial in both percentage point and percentage terms. For example, a 0.1 increase in the sex ratio would increase the probability of a first boy's mother, relative to a first girl's mother, working outside the hometown by about one percentage point or 38.6 percent. In addition, working time increases with a rise in the sex ratio for parents of first sons relative to parents of first daughters, as shown in columns (4) and (5).[20] These results support that compared with first-daughter families, high sex ratios boost parental labor supply among first-son families.

As we discussed earlier, migration significantly increases family income, and so does increasing working time. Results in this section indicate that parents attempt to increase investments in their children when marriage market conditions are disadvantageous, as indicated by having a son together with facing high sex ratios.

4.2 Sex ratio imbalance and premarital investments

We next examine how sex ratio imbalance affects premarital investments in physical capital and human capital. We estimate equation (1) based on our sample, using measures of housing investments and child educational investments constructed in Section 3.2 as outcome variables. Results are reported in panel B, Table 3.

[20]Sample sizes here are smaller than in previous columns, because some respondents did not respond to the survey questions about working hours.

Housing investments We focus on housing investments when considering premarital investments in bequeathed physical capital, for reasons given in Section 2. Housing investment results are reported in the first three columns of panel B, Table 3. In column (1) where the outcome variable is (log) house construction area, the estimate for the interaction-term coefficient is 0.413 (standard error 0.205), positive and statistical significant at the five percent level. This indicates that parents with first-born sons prepare a larger house relative to those with first-born daughters as the sex ratio becomes more biased towards males. Based on the estimate, as the sex ratio rises from 1.08 in 2002 to 1.18 in 2010, house construction area for home possessors with a first son would increase by about 4.1 percent relative to home possessors with a first daughter.

In column (2) that concerns housing ownership and column (3) that concerns housing mortgage loan, the estimates of the interaction-term coefficient are positive and significant in terms of both statistical sense and economic magnitude. All these results imply that, in the presence of a high sex ratio, parents with first-born boys become considerably more aggressive in investments in housing relative to parents with first-born girls.

Child educational investments Results on child educational investments, reported in the remaining two columns of panel B, Table 3, are in contrast with the patterns for housing investments. In column (4) where the outcome variable is annual education expenditure for the first-born child in the family (in thousand yuan), the estimated interaction-term coefficient is -1.663 (standard error 0.800), negative and statistically different from zero at the five percent level. Accordingly, with a 0.1 increase in the sex ratio, annual education expenditure per child is 166 yuan less if the child is a boy. The economic magnitude is sizeable compared with a mean expenditure of 1,507 yuan per child, representing a 11 percent reduction. Result in column (5) shows that a high sex ratio reduces the probability that parents with a first-born boy have put aside a specialized fund for his education relative to parents with a first-born girl.

Taken together, empirical results in panel B, Table 3 show that the effects of sex ratio imbalance on the composition of premarital investments are in opposite directions: A combination of having a son and experiencing a scarcity of prospective brides induces more physical capital investments but less human capital investments. This may be because parents with sons are motivated to invest their available financial resources in the capital form that can more effectively enhance their sons' marriage market prospects.

4.3 Addressing potential issues related to son-preferring fertility stopping rules

While we have shown in the previous section that the gender of the first-born child can be viewed as random, such an argument for identification is potentially not quite complete: Due to son preferences, subsequent fertility decisions may be different depending on the gender of the first child. Families with a first-born daughter typically are more likely to have a second or third child in order to get a boy, while families with a first-born son are more likely to stop childbearing and therefore, have a smaller family (Ebenstein, 2011). This section shows that these concerns do not confound our main results.

Family size One might worry that our findings on parental labor supply, housing investments, and educational investments per child may reflect the effect of family size. Specifically, it is possible that fewer resources are available for a first-born girl—who may have a sibling—compared with a first-born boy—who may have no sibling, as implied by the theory of child quantity-quality trade-off. This is what our results on housing investments show; results on educational investments on the first child, however, show the opposite. That is, our empirical results are not fully in line with this interpretation.

To further address the concern with the family-size effect, we control for the total number of children in a family in robustness checks. Table 4, panel A reports the results. The outcome variables in columns (1), (2), and (3) are the paternal-migration dummy, log house construction area, and education expenditure for the first child, respectively (using other measures of parental labor supply and premarital investments yields similar patterns). Results after controlling for family size are not significantly different from the baseline (the first row of the table). We perform a generalized Hausman test to formally show that the estimated coefficient on the interaction term is statistically indifferent from the baseline estimates.

We then control for the number of children plus its interaction with the first-son dummy, to take into account the possibility that family size may have differential effect depending on the gender of the first child. We still get similar results.

An alternative strategy to deal with the potentially confounding family-size effect is to restrict the sample to families with only one child (about 65 percent of the full sample). Results are reported in Table 4, panel B and show a similar pattern to the baseline in terms of the sign and magnitude of the estimated interaction-term coefficient. The pattern remains similar when we further restrict the sample to families that are less likely to have a second child. Specifically, we restrict the sample to one-child families in which the only child is above the age of four (about 50 percent of the full sample). We note that for these two groups of families, the effect on housing investments is less pronounced than the baseline, while the effect on child educational investments is more pronounced.

Child-gender composition Another issue raised by using the gender of the first child as a regressor is that it may not be an appropriate proxy for marriage market conditions the family faces. As we have discussed, it is common in China that a first daughter is followed by a second son. For these families, despite having a first daughter, they have to worry about the marriage prospects of their second son—similar to first-son families. To isolate the effect of marriage market conditions, we use variables that better represent actual child-gender composition in a family. Specifically, we replace the first-son dummy in equation (1) with a dummy for having any son and the proportion of sons. This yields qualitatively similar results as before.[21] We then control for the number of children (and its interaction with

[21] When we use a dummy for having any son, the estimates essentially compare the effect for families with at least one son with families with no son. When we use the proportion of sons, the estimates essentially compare the effect for all-son families with all-daughter families.

the child-gender measure) in these empirical exercises, and again obtain similar results; see Table 4, panel C.

While these patterns confirm that how families make labor supply and premarital investments decisions in response to high sex ratios depends on child gender, the two child-gender measures we use may be endogenous. To partly address this issue, we use the first-son dummy as an instrument for the two variables and repeat all empirical analyses above (first-stage results are given in Appendix Table A4). We observe that this does not change the pattern of results.

In summary, robustness analyses in this section show that concerns related to son-preferring fertility stopping rules are not likely to be the main driver of our findings.

4.4 Addressing potential issues related to local sex ratios

We have followed the practice of focusing on OLS estimates for the effect of sex ratio imbalance in the literature. But strictly speaking, county-level sex ratios may not be exogenous. Below we provide evidence that helps alleviate this concern and isolate the marriage market effect of sex ratios. We also provide evidence that our results are not sensitive to using sex ratios for different age cohorts.

Unobservable cross-county heterogeneity Counties with higher sex ratios perhaps have unobserved characteristics—for example, culture—that may affect household investment decisions. This concern is partly addressed in our research design, as we control for county fixed effects and focus on the coefficient on the interaction between local sex ratio and first-child gender. That is, we compare first-son families with first-daughter families within a county, which reduces the confounding effects of unobserved county-level characteristics as long as they affect investments of the two types of families in a county in a similar manner.

To check the extent to which unobserved cross-county heterogeneity is an issue in our estimations, we exclude county dummies which are previously controlled for (to saturate the model, we include county-level sex ratios). Results are reported in Table 5, panel A, showing a similar pattern to the baseline (the first row of the table).

Sex-ratio confounders If certain factors are correlated with the sex ratio and affect premarital investments of first-son and first-daughter families in a different manner, our estimates may be biased. Zhang (2017) proposes six factors that may be responsible for high sex ratios in China: traditional son preferences, wealth, gender difference in earnings, old-age support (sons serve as a better source of insurance in old age than daughters), the family planning policy, and gender-selection technology.

We have two arguments to address the concern about traditional son preferences. First, while people in counties with higher sex ratios have, on average, stronger son preferences than those in counties with balanced sex ratios, our findings indicate that the former invest less in sons' education relative to the latter. Second, traditional son preference may lead to differential family structures between high- and balanced-sex-ratio regions. We have discussed in detail in Section 4.3 that issues raised by this son-preferring fertility stopping rule are not a primary concern in our results.

To address the concern about the remaining confounding factors, we control for average household financial wealth—defined as the sum of liquid and illiquid assets, household income, and gender wage differential at the community level (the sub-level of county), as well as the number of children and a variable indicating social insurance at the household level.[22] For each of these factors, we allow it to affect first-son and first-daughter families differently, by controlling for its interaction with the first-son dummy.

In addition to the above-mentioned factors, we consider grandparental coresidence. Child gender may affect the likelihood of living with grandparents differently across high- and balanced-sex-ratio counties. This in turn may affect household decisions. We control for this factor as well as its interaction with the first-son dummy.

Table 5, panel B reports the robustness results using the three representative outcome

[22]We also control for gender wage ratio, which does not make much difference from using gender wage difference. As fertility is the direct target of the family planning policy, and thus can be regarded as a proxy for the implementation of the policy, we control for the number of children. Gender-selection technology progress can be proxied by local average household wealth or income.

variables. It shows that controlling for all above potentially confounding factors—either individually or collectively—leads to virtually no difference in our results. In each robustness regression, the estimated coefficient on the interaction term β_3 is not significantly different from the baseline estimate.

Sex-ratio confounders selected by high-dimensional method The above robustness analyses concern potential sex-ratio confounders discussed in the literature. The current development in methods with high-dimensional data enables us to consider a much more comprehensive set of sex-ratio confounders, and select the most important ones with the help of machine learning (Belloni et al., 2014).

In this robustness analysis using high-dimensional method, the initial set of factors that are considered as being potentially correlated with the sex ratio include local adult residents' age, schooling years, hukou status, political status, marital status, the number of siblings, ethnicity, annual income, social insurance, scores for a word test and a math test, depression score, and coresidence status with parents. For each of these factors, we compute the average, the average for men, the average for women, and the gender difference at the county level. The final set consists of 363 variables made up of the levels and quadratics in each of the initial variables, and interactions of all the preceding variables with each other, as in the example discussed in Belloni et al. (2014). We regress county-level sex ratios on these variables and use high-dimensional methods to select potentially important ones—those are strongly predictive of local sex ratios. Then we control for their interactions with the first-son dummy in estimating equation (1). In this way, we take into account factors whose effects are most likely to be confounded with the effect of sex ratio imbalance. Results show that after partialling out the confounding effects, our main empirical patterns remain; see the last part of Table 5, panel B.

Together, results reported in Table 5, panel B show that our findings are robust to the inclusion of various potentially confounding factors and therefore, mainly driven by marriage market considerations that result from sex ratio imbalance. This implies that potential omitted variable bias associated with sex ratio is not likely to be a primary concern with our estimations.

IV estimations To further alleviate the concern about the possible endogeneity of sex ratios, we use the IV estimation method to estimate equation (1) as a robustness check. Cross-county variation in sex ratios may be accounted for by variation in financial penalties for violating the family planning policy and quota of births stipulated by the policy (depending on whether the household belongs to ethnic minorities, since ethnic minorities are generally exempted from the policy). Therefore, the interactions of these variables (as well as their interactions with a dummy for ethnic minority) with the first-son dummy are used as instruments for the interaction of the sex ratio and the first-son dummy in equation (1).

As reported in Table 5, panel C, IV regression results reveal qualitatively similar patterns to OLS estimates in the benchmark. Results regressing sex ratios on policy-violation penalty, birth quota, and their interactions with a dummy for ethnic minority—in lieu of first-stage results—suggest that heavier financial penalties levied for unauthorized births and more births allowed by the policy are associated with higher sex ratios (see Appendix Table A5).

We, however, exercise caution in interpreting the IV results and still focus on the OLS results, to avoid problems associated with common candidates for instruments of sex ratios (such as the ones used here). China's family planning policy is passed down the administrative chain of command until it is interpreted and adapted to suit local needs (Short and Zhai, 1998). Therefore, financial penalties and birth quotas are likely to be endogenously stipulated by local governments based on local conditions, and may be correlated with household decisions independent of the local sex ratio (Ebenstein, 2011; Wei and Zhang, 2011).

Sex ratios for alternative age cohorts We also check whether our findings are robust to using sex ratios for different age cohorts. Instead of the premarital-age cohort between the ages of ten and 24 years, we recalculate sex ratios for local population in the age brackets of 10–14, 15–19, and 20–24. We find that using sex ratios for each age bracket gives a qualitatively similar pattern of estimates for the interaction-term coefficient: The sign is preserved and the magnitude varies only moderately (panel D, Table 5). This is perhaps due to the persistence of the local level of sex ratio imbalance over time. If we take any potential measurement error in sex ratios—which tends to produce attenuated coefficient estimates—into account, our results may be interpreted as lower bounds of the true effect of sex ratio imbalance.

4.5 Sex ratio imbalance and child human capital development

We have shown that high sex ratios lead to increased parental migration, increased housing investments, and reduced educational investments for families with a first-born son relative to families with a first-born daughter. This would have important implications for human capital development of the next generation. The development of boys relative to girls may be undermined because of underinvestment in education and increased parental migration driven by sex ratio imbalance.

Table 6 presents estimation results on the effect of sex ratio imbalance on the formation of human capital based on equation (1). Dependent variables in columns (1) and (2) are children's cognitive outcomes, defined for first-born children who are at least ten years old. These outcomes are measured by latest class rankings in mathematics and Chinese examinations, and set to one minus the rank over the total number of students in the class so that a larger value implies a better result. (For instance, ranking first in a class of 50 students gives a value of 0.98, while ranking last gives a value of zero.) The negative estimates for the interaction-term coefficient in both columns are large and statistically significant, which implies that a high sex ratio adversely impacts academic achievement of boys relative to girls.

In columns (3) and (4), we turn to children's non-cognitive outcomes—interpersonal communication skills including openness and cooperation, again defined for first-born children who are at least ten years old. After interviewing each child, the CFPS interviewers were asked to evaluate communication skills of the child. Children's behavior was ranked from one to seven, where a larger number means a better performance. We recode interviewers' ratings as binary variables, a value of one including evaluations of five, six, and seven. Estimates show that an increase in the sex ratio from 1.08 in 2002 to 1.18 in 2010 significantly reduces the fraction of boys exhibiting openness by 5.0 percentage points relative to girls. The analogous statistic for cooperation is 5.7 percentage points.

Columns (5) and (6) report results on health outcomes, measured by the child's body weight and height. We use z-scores that are transformed based on international child growth standards—UK reference growth charts. The means of the z-scores are negative, as can be seen from the table. The empirical pattern revealed is similar to the pattern for cognitive and non-cognitive skills. The effect of a 0.1 increase in the sex ratio is weight of a boy being 0.09 standard deviation further below the average of comparable international children, and height being 0.02 standard deviation further below, relative to a girl.

Results in Table 6 suggest that sex ratio imbalance hurts human capital development of boys relative to that of girls. We propose two underlying reasons: (i) distortion in family investments, or specifically, underinvestment in education, and (ii) parental migration. Parental migration results in a shortage of parenting inputs and mental costs related to family separation; therefore, it has large adverse impacts on children's learning outcomes—especially when taking into account the possibility that parental inputs may be complements to children' own efforts—and various other outcomes (Zhang et al., 2014; Lyle, 2006; McKenzie and Rapoport, 2011).

To further show that migration is a potential mechanism through which sex ratio imbalance hurts boys, Table 7 compares human capital outcomes, time allocation, and psychological well-being between children with and without an absent father or mother. The patterns indicate that migration, especially that of fathers, is accompanied by less satisfactory human capital outcomes, less time devoted to at-home studying and physical exercise, as well as lower psychological well-being for children left behind. Moreover, the absentee-father problem may be magnified for boys, partly because fathers are more important in modelling social roles for sons than for daughters.

5 Discussion of results

Our results have shown that when sex ratios are high, parents of boys, relative to those of girls, tend to increase labor supply, and shift investments toward housing and away from children's education. In the following, we provide plausible interpretations of these results and also discuss competing hypotheses.

In line with prior literature, we propose that parents facing steep marriage odds due to sex ratio imbalance increase labor supply—and in particular, work away from home—in a competitive manner in order to increase resources available for investments in their children. Further, we propose that the effect of sex ratio imbalance on the composition of investments is because of imperfect commitment in marriage. As men and women are unable to commit, at the time of marriage, to share future household resources in a pre-agreed fashion, the future labor income will be subject to ex post bargaining, where bargaining power depends on who earns the income. Therefore, a man who brings with him more housing at the time of marriage—which will not be subject to ex post bargaining—is a more desirable marriage partner than one with higher labor earnings but a smaller house. This explains why parents who want to ensure the marriage of their son direct investments towards more housing than education.

One competing hypothesis centers around the possibility that the difference in investments between first-son and first-daughter families is affected by factors other than sex ratio imbalance—such as household structures. In high-sex-ratio regions where son preferences are stronger, a first-born girl is more likely to have sibling(s) relative to a first-born boy, while in low-sex-ratio regions, the first child may be the only child regardless of the gender. Possibly, parents with more children have to devote less time to the labor market, and in particular, are less likely to migrate; they are also poorer and have less residual wealth to invest in real estate, as more children dilute household resources. Although in this respect, the hypothesis is consistent with part of our empirical findings, we have provided various robustness analyses in Sections 4.3 and 4.4 to verify that our results are not mainly driven by issues raised by son-preferring fertility stopping rules. Also, results on educational investments show that when sex ratios are high, parents tend to invest less on a first-born boy relative to a first-born girl, which contradicts the sibling size effect that in first-son families there are more per child resources available.

Another challenge to our interpretation centers around interpreting housing as premarital investments. Household investments in housing may reflect the desire to get higher returns. It is possible that some unobserved county-specific shock leads more boys to be born, and is also linked to higher returns to real estate investment. But any county-specific factors would impact housing investments of local families within the area in similar ways, and the effect would not depend on child gender. Therefore, comparing investments between first-son and first-daughter families differences out the effect of such shocks. It is also possible that wealthier parents—that is, parents with a first son in high-sex-ratio regions, possibly because of higher probability of migration or fewer children—just happen to keep their wealth in the form of housing. But we have controlled for parental education levels in our main regressions, and further added household financial wealth and household income in robustness regressions, to account for any wealth effect.

Therefore, it is primarily due to marriage market considerations that parents with sons attach greater importance to housing investments when the sex ratio gets higher. We have provided supportive evidence in Section 2 that parents see their property as physical capital investments in preparation for their sons' marriage, at the time when their sons are young. A closer look at the stated purpose of migration remittances based on our data also indicates the marriage market effects on parental investment decisions even if children are still young: When faced with a higher sex ratio, son families relative to daughter families are more willing to spend the migration remittances on the son's marriage (building or buying a house, in particular); see Appendix Table A6.

Heterogenous effects of sex ratio imbalance across families also support that our empirical findings are mainly driven by marriage market considerations. We split the main sample into two subsamples based on first-born children's proximity to marriage age. For the subsample that contains households with a first-born child above the age of 11, the effect of sex ratio imbalance on housing investment for son families relative to daughter families appears much more prominent—the magnitude more than double the benchmark (see Appendix Table A7). This finding is in line with the interpretation that housing investment patterns mainly reflect parental considerations for children's marriage market prospects.

Next we consider status seeking as a potential explanation. Perhaps, the level of status competition among son families is higher than among daughter families in counties with high sex ratios. A strong desire to conform to norms in the same social strata induces families with a son to engage in earning activities and housing investments more aggressively. This interpretation is in line with Brown et al.'s (2011) finding that grooms' families spend more on weddings as local competition for status intensifies. However, it is incompatible with our finding that parents facing higher sex ratios invest relatively less in sons' education.

To sum up, while some other stories seem to rationalize part of our findings on parental labor supply and investments, a natural and highly plausible interpretation is that bequeathable physical capital of men is more attractive to potential brides in the marriage market than human capital, as spouses are unable to commit to an agreement regarding the future division of the latter. This interpretation motivates us to build a theoretical model with imperfect commitment within marriage that incorporates two different types of premarital investments, which we turn to next.

6 A model of multidimensional premarital investments with imperfect commitment

In this section we build a model of multidimensional premarital investments with imperfect commitment within marriage. We highlight the difference between physical capital (housing) and human capital. The key characteristic of housing that we focus on is that it is a public good within marriage—both partners share the house equally, even if their valuations of it may differ. Human capital, on the other hand, gives rise to earnings that are individual specific. The share of a spouse in these earnings depends upon relative bargaining powers after marriage, and not on marriage market conditions.[23] Individuals realize that their attractiveness on the marriage market depends on the composition of their ex ante investment, given the lack of commitment. We show the existence of a unique, stable equilibrium in such a setting. The model setting enables us to perform a more general welfare analysis on how equilibrium investments differ from utilitarian efficient investments. It also provides us with rich comparative statics. In particular, when men are in excess supply, and have large ex post bargaining power, the model predictions match the empirical patterns we have found.

6.1 The basic model with a balanced sex ratio

We assume a continuum of men and a continuum of women. Let us consider first a situation with equal measures on both sides of the market. At the ex ante stage, the parents of a boy have to choose a vector of investments for their son, (x_B, y_B). x_B is investment in a private good, such as the son's human capital. y_B is investment in good which is public within marriage, such as the purchase of a house. The costs of investment, in terms of foregone consumption utility for the parents, are given by a function $c : \mathbb{R}_+^2 \to \mathbb{R}_+$.

We normalize the return on investments on each good to one, so that one unit of investment in the good yields, one average, one unit of the good – this is without loss of generality since the cost functions capture any non-linearity.[24] Similarly, the parents of a girl choose a vector of investments (x_G, y_G). For simplicity, we assume that the cost functions are the

[23]Conceptually, our analysis can be applied to any situation with multidimensional investments, where the sharing rules for the returns differ across investment categories. The distinction between public and private goods is a leading example, but by no means the only one. For example, individuals may have greater bargaining power over labor income than over income derived from financial assets. Our model can be extended to such a context. Our model can also be adapted to account for marital transfers such as dowry and bride price, the sharing rules of which are determined ex ante.

[24]Consequently, convexity of the cost function captures both diminishing returns to investment, and decreasing marginal utility of parental consumption.

same for the two sexes, an assumption that is easy to relax.

Finally, we assume that returns to the private good are stochastic: each boy is subject to a zero-mean shock ε, which has distribution function F. Consequently, the realized return on investment in a boy equals $x_B + \varepsilon$. Similarly, each girl is subject to a zero-mean shock η, which is distributed according to G, so that the realized return from equals $x_G + \eta$.

Bhaskar and Hopkins (2016) show that when investments are one-dimensional such the shocks ensure existence and uniqueness of a (quasi-symmetric) equilibrium in investment levels, under appropriate distributional assumptions. Our goal here is to extend the analysis to multidimensional investments, where the public good component gives rise to complementarities between the investments on the two sides.

We assume that two parties who agree to marry cannot commit to a sharing rule of the returns to the parental investments in any private good, including the stochastic component. Instead, the shares in the private good are determined by ex post bargaining. We model this by assuming that a man has a share λ_B in the returns, with his partner securing the remaining share, $1-\lambda_B$. Similarly, a woman has share λ_G of the returns to in her investment in the private good good, the remainder going to her spouse. These simple sharing rules can be derived from cooperative or non-cooperative models of bargaining, as in Nash (1950) or Rubinstein (1982). As Shaked and Sutton (1984) have shown, the shares will not depend upon the outside options of the two parties, as implied by marriage market conditions, unless these outside options are binding.[25] The public goods consumption by the couple equals the sum $y := y_B + y_G$, with a man's payoff being $v_B(y)$ and a woman's payoff being $v_G(y)$.

We assume that the cost function, the public good valuation functions, and shocks satisfy the following assumption.

Assumption 1:

1. $c(.)$ is continuously twice differentiable and strictly convex.

2. $c_x(0, y) = 0$ and $\lim_{x\to\infty} c_x(x, y) = \infty$ for any y. $c_y'(x, 0) = 0$ and $\lim_{y\to\infty} c_y(x, y) = \infty$

[25]Since divorce is rare in China, one may assume that the outside options are unlikely to be binding, and therefore irrelevant in the calculation of the ex ante returns from investment.

for any x.

3. $v_B(.)$ and $v_G(.)$ are increasing, twice differentiable and strictly concave.

4. $f(\varepsilon)$ and $g(\eta)$ are twice continuously differentiable.

Utilitarian efficient investments Consider first a social planner, who chooses the levels of investments, but who cannot dictate the sharing rule. Consider first the case where the social planner maximizes the ex-ante expected utility of the parent, before she observes the sex of her child. Thus, since the child is equally likely to be a boy or a girl, the planner will give their respective utilities equal weight. In this case, it is straightforward to see that the investment profile (x_B^{**}, y_B^{**}) must satisfy the first order conditions:

$$c_x(x_B^{**}, y_B^{**}) = 1, \tag{2}$$

$$c_y(x_B^{**}, y_B^{**}) = v_B'(y_B^{**} + y_G^{**}) + v_G'(y_B^{**} + y_G^{**}). \tag{3}$$

We focus on quasi-symmetric equilibria, where all men invest the same amount (x_B^*, y_B^*), and all women invest (x_G^*, y_G^*). We say that there is *underinvestment* (resp. *overinvestment*) in the private good by one of the sexes, say men, if the marginal cost evaluated at the profile, $c_x(x_B^*, y_B^*) < 1$ (resp. > 1). Similarly, there is underinvestment in the public good by men, if the marginal cost $c_y(x_B^*, y_B^*)$ is less than the sum of the marginal benefits to the spouses. Observe that this notion of over or underinvestment in a good differs from a simple comparison of the equilibrium investment level with the utilitarian level.[26]

Let us now consider equilibrium behavior in the marriage market. Matching takes place after investments and payoff shocks are realized. Suppose that a boy with investment profile (x_G, y_G) and shock value ε matches with girl with profile of investments (x_G, y_G) and shock

[26]This difference arises for the private good since the cost function is not separable—if it is so, and $c_{xy} = 0$, then the two notions coincide. For the public good, strategic interaction between the two sides of the market gives an additional reason for non-coincidence of the two notions.

realization η. Then overall payoff of the boy from this match equals

$$\lambda_B(x_B + \varepsilon) + (1 - \lambda_G)(x_G + \eta) + v_B(y_B + y_G). \tag{4}$$

The overall payoff of the girl from this match equals

$$\lambda_G(x_G + \eta) + (1 - \lambda_B)(x_B + \varepsilon) + v_G(y_B + y_G). \tag{5}$$

Our focus is on a quasi-symmetric equilibrium where all men invest (x_B, y_B) and all women invest (x_G, y_G). In such an equilibrium, men with higher levels of ε are uniformly more attractive to any woman, since $\lambda_B < 1$. Similarly, women with higher values of η are uniformly more attractive to men. Thus, any stable matching must be assortative in the shocks, and since it must also be measure preserving, the matching function $\phi(\varepsilon)$ satisfies $F(\varepsilon) = G(\phi(\varepsilon))$.

Let us examine the incentives for investment. We will focus on the first order conditions for optimal investment. It is routine to verify that the second order conditions are also satisfied. However, we will simply assume that these conditions are sufficient, and that large deviations are unprofitable; Bhaskar and Hopkins (2016) have shown that this is indeed so for one-dimensional investments, under appropriate assumptions on costs and the distributions of shocks.[27]

Let us first consider marginal incentives for investment in a private good, such as human capital. In this case, if a man invests $x_B + \Delta$, the marital return on this investment arises from the fact that he is now more attractive to every woman. In particular, since a woman gets the same fraction $(1 - \lambda_B)$ of this Δ increment, as she does from a larger shock, then for any ε, he is as attractive as a man with a higher shock value ε', which satisfies $\varepsilon' - \varepsilon = \Delta$. Since the shock value of his marriage partner will equal $\phi(\varepsilon')$, and he gets a fraction $(1 - \lambda_G)$

[27]We do not anticipate any difficulty in extending this argument to the two-dimensional case, but this would divert from the focus of the present paper.

of this value, his marginal marriage market benefit from investment equals

$$(1-\lambda_G)\int \phi'(\varepsilon)f(\varepsilon)d\varepsilon,$$

where

$$\int \phi'(\varepsilon)f(\varepsilon)d\varepsilon = \int \frac{f(\varepsilon)}{g(\phi(\varepsilon))}f(\varepsilon)d\varepsilon =: \theta_B. \tag{6}$$

Consequently, the first order condition for optimal investment in the private good by a boy, i.e. the best response $\hat{x}_B$ is given by

$$c_x(\hat{x}_B, y_B) = \lambda_B + (1-\lambda_G)\theta_B. \tag{7}$$

Observe that the return to investment in the private good is independent of the common investment level chosen by girls. Similarly, for women, the first order condition for investment in the private good is

$$c_x(\hat{x}_G, y_G) = \lambda_G + (1-\lambda_B)\theta_G, \tag{8}$$

where

$$\theta_G := \int \frac{g(\eta)}{f(\phi^{-1}(\eta))}g(\eta)d\eta. \tag{9}$$

Consider next the public good. The key point is that free rider problem does not bedevil such investments, since they improve an individual's attractiveness on the marriage market. The marginal benefit from increasing $\hat{y}_B$ to $\hat{y}_B + \Delta$ is that for any ε, the boy is as attractive as type $\hat{\varepsilon}$ that satisfies

$$v_G(\hat{y}_B + \Delta, y_G) + (1-\lambda_B)\varepsilon = v_G(\hat{y}_B, y_G) + (1-\lambda_B)\hat{\varepsilon},$$

Thus, you will be matched with type $\phi(\hat{\varepsilon})$ rather than $\phi(\varepsilon)$, and the benefit of this is $(1-\lambda_G)[\phi(\hat{\varepsilon}) - \phi(\varepsilon)]$. Thus the marginal return on the marriage market at any realization of ε is given by

$$\frac{1-\lambda_G}{1-\lambda_B}\phi'(\varepsilon)v'_G(\hat{y}_B+y_G).$$

Averaging over all realizations of ε, we see that the marriage market return from investment in the public good equals

$$v'_G(\hat{y}_B+y_B)\frac{(1-\lambda_G)}{(1-\lambda_B)}\int\frac{f(\varepsilon)}{g(\phi(\varepsilon))}f(\varepsilon)d\varepsilon.$$

In addition, an individual also benefits from his own consumption of the public good, at rate $v'_B(\hat{y}_B+y_G)$. Thus the first order condition for optimal investment by men in the public good is given by

$$c_y(\hat{x}_B,y_B)=v'_B(\hat{y}_B+y_G)+\frac{1-\lambda_G}{1-\lambda_B}\theta_B v'_G(\hat{y}_B+y_G). \tag{10}$$

In the case of the public good, we see that the best response for men, $\hat{y}_B$ does directly depend upon y_G, the investment level chosen by girls, since the payoff from the public good is strictly concave. In consequence, since the marginal costs of investment in the private good depend upon public good investment (unless the cost function is separable), $\hat{x}_B$ and $\hat{y}_B$ are both functions of y_G. However, they do not depend upon x_G, the level of girls' investments in the private good.

Since the argument is identical for women, the first order condition for optimal investment in the public good by women is

$$c_y(\hat{x}_G,\hat{y}_G)=v'_G(y_B+\hat{y}_G)+\frac{1-\lambda_B}{1-\lambda_G}\theta_G v'_B(y_B+\hat{y}_G). \tag{11}$$

Equations (7), (8), (10), and (11) characterize the best response investments in the private good and the public good by both the sexes.

Equilibrium We will now show that quasi-symmetric equilibrium exists and is unique. Furthermore, such an equilibrium is necessarily and stable, under the usual best response dynamics, which in turn implies that the comparative statics predictions will be intuitive.

Consider the best response on the boys' side to a profile of investments on the girls' side, (x_G, y_G). Since the first order conditions for the boys' investments are unaffected by x_G, we may write this as a pair of functions $\hat{x}_B(y_G)$ and $\hat{y}_B(y_G)$. Thus, $\hat{y}_B(y_G)$ is the best response investment in the public good, when both types of investments are chosen optimally by the boy. Similarly, we may define best responses on the girls' side, $\hat{y}_G(y_B)$. Let ζ denote the composition of the functions $\hat{y}_G$ and $\hat{y}_B$ so that $\zeta(u) = \hat{y}_G(\hat{y}_B(u))$. The fixed points of ζ correspond to quasi-symmetric equilibria. More precisely, y_G is a fixed point of ζ if and only if $y_G = \hat{y}_G(\hat{y}_B)(y_G))$, so that the pair $(\hat{y}_B(y_G), y_G)$ are mutual best responses, with the private good investments being given by $\hat{x}_B(\hat{y}(y_B))$ and $x_G = \hat{x}_G(y_B)$.

Observe that ζ is continuous and differentiable on the positive reals. Since $c_y(x, 0) = 0$, by assumption 1, $\zeta(0) > 0$. Also, since $v'_B(y)$ and $v'_G(y)$ are decreasing, while $c_y(x, y) \to \infty$, $\zeta(y) < y$ for y sufficiently large. Thus there exists a fixed point of ζ, which we denote by y^*_B. Let us denote this quasi-symmetric equilibrium by $((x^*_B, y^*_B), (x^*_G, y^*_G))$.

To show uniqueness and stability, consider the slope of the best responses that compose ζ. The derivative of the boys' best response is

$$\frac{d\hat{y}_B}{dy_G} = \frac{\Omega^B c^B_{xx}(.)}{\Delta^B - \Omega^B c^B_{xx}(.)},$$

where $\Omega^B := v''_B(.) + \frac{1-\lambda_G}{1-\lambda_B}\theta_B v G''(.) < 0$, and Δ^B is the determinant of the Hessian of the cost function, $c(.)$, evaluated at (x^*_B, y^*_B). Since the cost function is strictly convex, $\Delta^B > 0$ and $c^B_{xx} > 0$, so that $\frac{d\hat{y}_B}{dy_G} \in (-1, 0)$.

Similarly,

$$\frac{d\hat{y}_G}{dy_B} = \frac{\Omega^G c^G_{xx}(.)}{\Delta^G - \Omega^G c^G_{xx}(.)},$$

where $\Omega^G := v''_G(.) + \frac{1-\lambda_B}{1-\lambda_G}\theta_G v B''(.) < 0$, and Δ^G is the determinant of the Hessian of the cost function, evaluated at (x^*_G, y^*_G). Thus $\frac{d\hat{y}_G}{dy_B} \in (-1, 0)$. Consequently, at any fixed point, ζ has a slope that is positive and strictly less than 1. Thus, there can be at most one fixed point.

We summarize our results in the following proposition.

Proposition 1 *Under assumption 1, there exists a unique quasi-symmetric equilibrium, which is stable.*

Remark 1 *Uniqueness of equilibrium can be established under weaker conditions than in assumption 1. We do not require* $c_x(x, 0) = 0$ *or* $c_y(0, y) = 0$*. All that is required is that* $\hat{y}_B(0) > 0$ *and* $\hat{y}_G(0) > 0$*, so that equilibria are interior and satisfy the first order (equality) conditions. If this assumption is not satisfied, one may have pair of equilibria, where only one of the two sexes invest in the public good.*

Welfare implications We now proceed the qualitative properties of equilibrium investments and their relation to the utilitarian efficient investments. Our first proposition is an efficiency result when the sexes are symmetric.

Proposition 2 *Suppose that* $F = G$ *and* $\lambda_B = \lambda_G$*, so that both sexes have the same distribution of shocks and the two bargaining powers, over labor income, are equal. Then investments in public goods and private goods are utilitarian efficient.*

The proof of this result follows from a verification that the first order conditions for equilibrium investments coincide with the conditions for utilitarian efficiency, under the two assumptions of the proposition. Consider the first order condition for investments in the public good by men, equation (10), and focus on the coefficient on $v_G'(y)$, the marginal utility of the spouse from the public good. When the shocks have the same distribution, a man of type ϵ is matched to a woman of the same type. Consequently, the ratio of the two densities in expression for θ_B, equation (6), equals one for every ϵ, and thus $\theta_B = 1$. Further, since $1 - \lambda_B = 1 - \lambda_G$, the ratio of the two terms equals one. Thus the weight on female marginal utility equals one, and the first order condition reduces to Samuelson's efficiency condition for public goods. A similar argument applies to the condition for male investments in the private good, and for women's investments in both goods.

There are two notable features of this result. First, there is no trace of the free rider problem that normally bedevils investments in public goods. Even if demands for the public good differ greatly between the sexes, both sides will provide it. This follows from the fact that investments take place prior to matching. Consequently, if one side of the market, say women, value the public good more, an individual man has an incentive to provide it, in order to improve his attractiveness on the marriage market.

Second, the efficiency result may appear unexpected. Even if there is no equal sharing of income, with each partner having a larger bargaining power over the returns from the investments in human capital, this does not result in investment inefficiency, as long as both sexes are in the same position. Indeed, the proposition can be generalized as follows: boys' investments in both public and private goods are efficient as long as $\frac{1-\lambda_G}{1-\lambda_B}\theta_B = 1$; girls' investments are efficient as long as $\frac{1-\lambda_B}{1-\lambda_G}\theta_G = 1$. Of course, if the shock distributions are the same, $\theta_B = \theta_G = 1$, and if bargaining power over own labor incomes are equal, then $\frac{1-\lambda_G}{1-\lambda_B} = 1$.

The next proposition examines the role of differences in bargaining power between the sexes.

Proposition 3 *If men have more bargaining power over own income than women so that $\lambda_B > \lambda_G$, and if $F = G$, then men overinvest in the private good, and also overinvest in the public good, while women underinvest in both types of goods, relative to the utilitarian investment levels.*

This proof also follows from an examination of the first order conditions for equilibrium. When $F = G$, $\theta_B = \theta_G = 1$. Men have a greater incentive to invest in the private good, since they get a larger share of the return, and also a relatively larger share of the return of their partner. Thus the marginal benefit exceeds one. For women, the marginal benefit is less than one, and so they underinvest in the private good. More subtle is the incentive of men to overinvest in the public good. Men invest in the public good since it is an effective way to compete on the marriage market, more effective than investing the private good, where they cannot guarantee their partner a larger share. Effectively, investment in the public good acts as a commitment device for men.

Suppose now that marriage market competition favors women, while men have greater bargaining power within marriage. One would expect that the incentive of men to overinvest (and of women to underinvest) is magnified. To model a balanced marriage market that favors women, let us assume that the shocks for women are more dispersed than the shocks for men. To formalize this, we assume that distribution G is larger in the dispersive order than a distribution F, or $G \geq_d F$, i.e.:

$$g(G^{-1}(z)) \leq f(F^{-1}(z)) \text{ for all } z \in (0,1), \tag{12}$$

with the inequality being strict on a set of z values with positive measure; see Shaked and Shanthikumar (2007, pp148-9).[28]

If $G \geq_d F$, so that $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} \geq 1$ for all values of ε, and is strictly less on a set values of ε of positive measure. Thus $\theta_B > 1$, and $\theta_G < 1$. Observe that the incentive for men to invest in both goods is given by the *product* of θ_B and $1 - \lambda_G$. Thus increased marriage market competition interacts with bargaining power asymmetry to have a multiplier effect. We summarize this discussion in the following proposition.

Proposition 4 *Suppose that $G \geq_d F$, so that the shocks are more dispersed for women than for men, so that marriage market competition favors women. Suppose also that $\lambda_B > \lambda_G$, so that men's bargaining power is greater than women's. Both these factors cause men to overinvest in both goods and women to underinvest in both goods. The two forces, marriage market competition and bargaining power asymmetry, interact to have a larger multiplier effect on investments.*

Comparative statics We are now in a position to examine the effects of a change in bargaining power of one of the sexes upon equilibrium investments on both sides. Since we

[28]For example, if F and G are both uniform distributions, where the support of G is a longer interval than that of F, then $G \geq_d F$. A second example is two normal distributions—the one with the higher variance is larger in the dispersive order.

have shown that the equilibrium is stable, we would expect that the comparative statics would be intuitive. Nonetheless, it is entirely possible that an increase in male bargaining power *reduces* the level of investment in the private good, while increasing investment in the public good. This arises due to fact that the cost function $c(x, y)$ is not separable in its two arguments. Consequently, if $c_{xy}(.)$ is large, and the marriage market competition is very important for males, their desire to be more attractive on the marriage market increases public good investment. This raises the marginal cost of private good investments, and reduces its level, even thought the marginal benefit from such investments also increases. Similarly, women will invest less in the public good, and this may prompt greater private good investments on their part, despite a fall in the return. Observe that the reduction in women's public good investments will induce even greater public good investments by the men, due to the fact that the public good evaluation functions are strictly concave. Similarly, the increase in male public good investments will induce a further reaction in women's investments in the public good. This is summarized in the following proposition.

Proposition 5 *Suppose that marriage market competition favors women, so that $\theta_B > 1$ and $\theta_G < 1$. An increase in men's bargaining power λ_B increases public goods investments by men, an reduces that by women. If c_{xy} is large enough, equilibrium investments by men in the private good fall. Furthermore, if marriage market competition for men increases while that for women men falls, this also increases men's investments in the public good, and may reduce their investments in the private good.*

The proof of this proposition is presented in Appendix B. Apart from the analytical proof, we also present an illustrative numerical example, which illustrates the equilibrium values of investments by both sexes as λ_B changes. The intuition for this illustrates the role of imperfect commitment. As λ_B increases, with λ_G remaining fixed, men get a larger share of the total private good in marriage. But this makes men who invest more in the public good more attractive to women. Consequently, men invest more in the public good. Since c_{xy} is large, this increases the cost of investing in the private good, and men invest less in it, even though the return to such investments has increased.

The example also illustrates that as θ_B increases and θ_G falls, investments in the public good by men increases, while those in the private good may fall. That is, increased marriage market competition induces men to invest more in the public good, as this is a credible way to commit on the marriage market when their bargaining power over the private good is large.

6.2 Unbalanced sex ratios

The distortionary effects on investments arise when bargaining powers are asymmetric, and favor one sex. This is most likely to be men, given their superior legal and customary position in many societies, and in this case, there will be overinvestment in boys and underinvestment in girls, with boys investing in those goods where they can commit to share the rewards more equally. On the other hand, distortions can also arise due to difference in ex ante competitive position, since the sex that faces more competition has a greater incentive to invest, in order to improve its competitive position. We now see that how the two distortions can reinforce each other, in traditional societies where men have greater bargaining power within marriage, and are also in greater number on the marriage market. This is particularly relevant in countries such as India and China.

Our modeling strategy incorporates the following innovation, which allows the sex ratio to affect investment incentives in a continuous fashion.[29] Suppose that the ratio of women to men is $r < 1$. We assume that the overall marriage market is composed of many local marriage markets, where the sex ratio varies. In some of these marriage markets, there is an excess of men, while in the others, the marriage market is balanced. A reduction in r, the aggregate ratio of women to men, increases the likelihood that an individual woman resides in a market where there is an excess of men. A simple way of modeling this is as follows. Fix $\hat{r} < 1$, and let this be sufficiently small so that the aggregate sex ratio, r, lies in the

[29] In previous work such as Bhaskar and Hopkins (2016), the sex ratio has discontinuous effects on investment incentives at $r = 1$.

interval $(\hat{r}, 1]$. The sex ratio in a local market takes one of two values, $\hat{r}$ and 1, where the probability of the first value is $\rho(r)$. Since the aggregate sex ratio equals r, $\rho(r)$ must satisfy the equation:

$$\rho(r)\hat{r} + (1 - \rho(r)) = r,$$

which implies that

$$\rho(r) = \frac{1-r}{1-\hat{r}}.$$

It is straightforward to verify that $\rho(1) = 0$, which is consistent with our analysis of the case of a balanced sex ratio. Further,

$$\rho'(r) = -\frac{1}{1-\hat{r}} < 0.$$

Utilitarian efficient investments Before analyzing equilibrium investments, let us consider the conditions for utilitarian efficiency. For private goods, the return on investment equals 1, and this return is either shared if the individual marries, or accrues entirely to the individual if he remains single. Since the utilitarian planner puts equal weight on both partners, the first order condition for utilitarian efficiency remains $c_x(.) = 1$, for both men and women.

However, utilitarian investments in the public good for a man do depend upon the sex ratio, since the likelihood of marriage determines whether the public good is shared, or consumed singly. Observe that a man is married with probability r. In the event that he is married, the benefit of the public good accrues also to his partner, while if he is not married, it does not. Consequently, utilitarian efficiency requires:

$$c_y(x_B^{**}, y_B^{**}) = r[v_B'(y_B^{**} + y_G^{**}) + v_G'(y_B^{**} + y_G^{**})] + (1-r)v_B'(y_B^{**}). \tag{13}$$

The effects of the sex ratio r upon efficient investments by men is, in general, ambiguous. If men have a greater matching probability, due to an increase in r, then the planner would like them to invest more, since the investments benefit their partner. However, they are also more likely to benefit directly from their partner's investment in the public good, and are less likely to remain single, and this is a force towards reducing men's investments in the public good.

For a woman, her probability of marriage always equals one, and hence the efficiency condition is independent of r:

$$c_y(x_G^{**}, y_G^{**}) = v_B'(y_B^{*,*} + y_G^{**}) + v_G'(y_B^{**} + y_G^{**}). \tag{14}$$

Equilibrium investments Let us now turn to equilibrium investments. The matching function in the local marriage market now takes two different forms, depending upon whether there is an excess of men or not. In a local market where the sex ratio is balanced, the matching function is ϕ, as we have already analyzed. So consider a local market with an excess of men, so that the local sex ratio is $\hat{r}$. Let $\hat{\varepsilon}$ denote the lowest quality boy that is matched, and let the matching function in this case be denoted by ϕ_+. Since the matching must be measure preserving, it now satisfies

$$1 - F(\varepsilon) = \hat{r}[1 - G(\phi_+(\varepsilon)]. \tag{15}$$

The derivatives of $\phi_+'(.)$ is given by

$$\phi_+'(\varepsilon) = \frac{f(\varepsilon)}{\hat{r}g(\phi_+(\varepsilon))}.$$

When boys are in excess supply, so that $\hat{r} < 1$, this magnifies the impact of an increase in the boy's shock value (and his investments) upon his match quality. Intuitively, since there is smaller measure of girls than boys, the qualities of the girls are more dispersed relative to the boys. Thus, the marriage market return to own quality is greater for boys.

Let $\xi_+(\eta) = \phi_+^{-1})'(\eta)$, i.e. ξ_+ is the inverse of the matching function in a market with an excess of boys, and specifies which quality of boy is matched to type η of girl. By the same logic,

$$\xi_+'(\eta) = \frac{\hat{r} g(\eta)}{f(\xi_+(\eta))}.$$

Let us now define θ_{B+}, as follows:

$$\theta_{B+} := \int_{\hat{\varepsilon}} \phi_+'(\varepsilon) f(\varepsilon) d\varepsilon = \frac{1}{\hat{r}} \int_{\hat{\varepsilon}} \frac{f(\epsilon)}{g(\phi_+(\epsilon))} f(\epsilon) d\epsilon.$$

Similarly, we define θ_{G+}, as follows:

$$\theta_{G+} := \int \xi_+'(\eta) d\eta = \hat{r} \int \frac{[g(\eta)]^2}{f(\xi_+(\eta))} d\eta$$

Note that the expressions θ_B and θ_G, that apply to a balanced local marriage market, are as defined previously.

Let $\hat{f} := f(\hat{\varepsilon})$, and let $\bar{U}$ denote the utility gain of the boy from being matched to the lowest quality girl, as compared to being unmatched. The first order condition for a boy's optimal investment in the private goods given by

$$\begin{aligned} c_x(x_B^*, y_B^*) = \rho(r) \Big[(1 - \hat{r}) + \hat{r}[\lambda_B + (1 - \lambda_G)\theta_{B+}] + \hat{f}\bar{U} \Big] \\ + (1 - \rho(r)) \left[\lambda_B + (1 - \lambda_G)\theta_B \right]. \end{aligned} \tag{16}$$

The first line on the right-hand side considers the payoff in a local marriage market where there is an excess of boys. The investment return in such a market consists of three terms. With probability $1 - \hat{r}$ the boy is single, and enjoys the entire return on his investment. With probability $\hat{r}$, he is married, and must share the return with his spouse. However, in this case, an increment to investment also increases his rank in the marriage market, and therefore, there is a marriage market return on his investment. Observe that this marriage market investment return is magnified, since it divided by $\hat{r}$. Intuitively, since the (relative) measure of girls in the local market is only $\hat{r}$, the effective dispersion amongst girls is larger than amongst boys, increasing the marriage market returns to investment for boys. Finally, the third term, reflects the fact that by increasing investment, the boy increases his chances of being married, by overtaking the lowest ranked boy, of quality $\hat{\varepsilon}$. In other words, the desire not to left unmatched magnifies investment incentives.

The second line reflects the payoff in a balanced local marriage market. In this case, the boy gets a fraction λ_B of his own return, plus the marriage market return, which is lower since the effective dispersion on girls' qualities is lower in a local market where there is an excess of girls.

The first order condition for the public good for men is given by

$$c_y(x_B^*, y_B^*) = \begin{array}{c} \rho(r)\left[(1-\hat{r})v_B'(y_B^*) + \hat{r}v_B'(y_B^*+y_G^*) + \hat{r}\frac{1-\lambda_G}{1-\lambda_B}v_G'(y_B^*+y_G^*)\theta_{B+} + \frac{v_G'(y_B^*+y_G^*)}{1-\lambda_B}\hat{f}\bar{U}\right] \\ +(1-\rho(r))\left[v_B'(y_B^*+y_G^*) + \frac{1-\lambda_G}{1-\lambda_B}v_G'(y_B^*+y_G^*)\theta_B\right]. \end{array} \tag{17}$$

For women, the first order condition for investment in the private good is simpler:

$$c_x(x_G^*, y_G^*) = \lambda_G + (1-\lambda_B)\left[\rho(r)\theta_{G+} + (1-\rho(r))\theta_G\right]. \tag{18}$$

Since a woman is always matched, she gets a fraction λ_G of her own return, and with probability $\rho(r)$ she gets the marital return in the market where women are short supply, and with the remaining probability, the marital return in a balanced marriage market.

The first order condition for women's investment in the public good is

$$c_y(x_G^*, y_G^*) = v_G'(y_B^*+y_G^*) + \frac{1-\lambda_B}{1-\lambda_G}v_B'(y_B^*+y_G^*)[\rho(r)\theta_{G+} + (1-\rho(r))\theta_G)]. \tag{19}$$

An interior equilibrium, where both sexes invest positive amounts in the public good, is given by the profile of investments that satisfy the first order conditions (16), (17), (18) and (19). Existence, uniqueness and stability of equilibrium is confirmed by the following proposition.

Proposition 6 *For any $r \leq \hat{r}$,there exists a unique quasi-symmetric equilibrium, which is stable, provided that assumption 1 is satisfied.*

We omit the proof since it follows the same line of argument as that of proposition 1.

Welfare implications As in the analysis of balanced marriage markets, we will say that a sex overinvests in a good, private or public, if the marginal cost of investment exceeds the first-best utilitarian criterion.[30] We therefore compare the first order conditions at an equilibrium, (16), (17), (18) and (19), with the conditions for utilitarian efficiency. Recall that the efficiency condition for private good is that the marginal cost equals 1, while for the public good, these are given by equations (13) and (14). In order to focus on the effects of the sex ratio per se (rather than on differences in bargaining power or the distribution of shocks), let us assume that $\lambda_B = \lambda_G$, and that $F = G$.

Proposition 7 *Suppose that the sexes are otherwise symmetric, except the imbalance in the sex ratio, and that the common distribution of shocks has a density that is weakly increasing. An imbalance in the sex ratio, $r < 1$, leads men to overinvest in both public and private goods, and women to underinvest in both goods.*

The proof of this proposition follows from a comparison of the first-order conditions for equilibrium with those for utilitarian efficiency. Since bargaining powers are equal, $\lambda_B = \lambda_G$, when the shocks are identically distributed for the two sexes, $\theta_B = \theta_G = 1$. Furthermore, when the density of shocks is weakly increasing, it may be verified that $\theta_B^+ \geq 1$, while $\theta_G^+ \leq \hat{r}$. Finally, note that by definition, $\rho(r)\hat{r} + (1 - \rho(r)) = r$.

From the condition for men's investment in the private good, equation (16), we see that $c_x(x_B^*, y_B^*) > 1$, for two reasons. First, since the mass of women is smaller than that of men, women's qualities are effectively more dispersed, intensifying competition by men. Second, men have an additional incentive to compete in order to avoid being left single—this is captured by the term $\hat{f}\bar{U}$. Conversely, the first order condition for women's investment in the private good, equation (18), shows that women underinvest, since the qualities of men are effectively less dispersed.

Turning to the public good, we see from equation (17) that men overinvest in the public good in order to avoid being left single. From equation (19), we see that women invest less

[30] As before, overinvestment in a good is consistent the *level* of investment being lower than the first best level, due the fact that $c_{xy} > 0$.

in the public good, for two reasons. First, they underweight the effect on their utility of their partner by a fraction r, even though efficiency dictates that they give this weight one. Second, there is a crowding out effect: Since men invest more in public goods, women have less incentives to invest.

Comparative statics of the sex ratio Our focus is on the comparative statics effects on the equilibrium investment with respect to r, the sex ratio. This can be derived by differentiating the system of equations given by the first order conditions (16), (17), (18) and (19), with respect to r. Appendix C presents the comparative statics results. These are relatively complex, and are difficult to interpret. Consequently, we have resorted to numerical methods, investigating the pattern of behavior for a large range of parameter values. We are particularly interested in the case where $\lambda_B > \lambda_G$, so that men have greater bargaining power over their own labor market income than women do, although we also explore the case where bargaining powers are equal. We find the following results:

- Boys' investments in the public good are always decreasing in r; girls' investments in the public good are always increasing in r. That is, if a sex faces increased marriage market competition, it invests more in the public good.[31]

- The effects of changes in r upon private good investments are ambiguous, for both sexes. While the direct effect of increased marriage market competition increases incentives for both types of investment. However, the direct effect may be larger for the public good, and the increased marginal costs from investing in the public good may reduce private good investments. This is the case when bargaining powers are asymmetric (i.e. $\lambda_B >> \lambda_G$) and when the cost function exhibits supermodularity, so that c_{xy} is large.

Appendix C presents one illustrative example when bargaining powers are asymmetric. It shows that the effects on private good investments by boys are ambiguous—these investments are decreasing in r for r values near 1, but are increasing in r for smaller values of r. In other

[31] While these are numerical results, this is true for all the parameter values we studied.

words, for some range of r values, the pressure to increase public good investments arising from marriage market competition crowds out private good investments, thereby inducing a reduction in such investments. Similarly, girls' investments in the private good are decreasing in r for r values close to 1, but are increasing in r when the sex ratio is more unbalanced.

To summarize, our analysis of investments in unbalanced marriage markets shows that boys have a direct incentive to invest more, both in the private good and in the public good, so that they overinvest relative to the utilitarian benchmark. Conversely, girls underinvest in both goods relative to the efficient benchmark. These effects are more pronounced for public goods, where investments of the two sides are strategic substitutes. Consequently, boys' public good investments are increasing in the sex ratio, and girls' public good investments are decreasing. Since increased public good investments raise the marginal costs of private good investments, boys may invest less in the private good in response to an unfavorable sex ratio, while girls may invest more.

7 Conclusion

This paper studies how imperfect commitment in marriage affects premarital investments in children made by their parents. Using nationally representative Chinese data, we find that high sex ratios lead to increased parental migration (due to a greater motivation to raise earnings), increased housing investments, and reduced educational investments for families with sons relative to families with daughters. Our empirical work controls for unobserved county-level heterogeneity and compares the effect of sex ratio imbalance for first-son families with the effect for first-daughter families, at the county level. We also provide a variety of robustness checks to address the concerns about potential endogeneity of sex ratios and other concerns.

We propose that bequeathable physical capital such as housing—which will be shared equally between spouses in its consumption after marriage and therefore, indicates a credible commitment—is more attractive to potential marital partners than human capital—the sharing of which will be subject to bargaining after marriage. We provide supportive evidence for this interpretation and also discuss some competing hypotheses.

We then develop a model where imperfect commitment combines with sex ratio imbalance to affect the magnitude and composition of investments in children. The model differentiates investments in public good (housing) and private good, and assumes that after-marriage bargaining powers do not depend on marriage market conditions. With sex ratio imbalance, the model predicts that men—the oversupplied side—would increase investments to compete for marital partners. When men have greater control over private good after marriage, the model predicts that they would increase public good investments in order to improve their attractiveness in the marriage market. This may crowd out private good investment, as we find in our empirical work.

This paper has important implications for human capital development of the next generation. Underinvestment in education as well as increased parental migration driven by marriage market considerations may undermine the development of boys. The associated social loss is likely to be large, especially given the drastic increase in returns to education in China (Heckman, 2003, 2005). The paper also has important implications for the study of marriage. While classic work focuses on one-dimensional assortative marriage matching on income, wage, or education, recent studies begin to realize the importance of matching along multiple dimensions (Galichon and Salanié, 2010; Chiappori et al., 2012; Dupuy and Galichon, 2014; Chiappori et al., 2018b). Our results highlight the different roles in marriage matches of different types of capital when commitment is imperfect. This area deserves more attention in future research.

References

Anderson, Siwan and Chris Bidner, "Property Rights over Marital Transfers," *Quarterly Journal of Economics*, 2015, *130* (3), 1421–1484.

Becker, Gary S., "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, pp. 813–846.

Beijing, "Married to the Mortgage," *Economist*, 2013.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Bhaskar, V., "Sex Selection and Gender Balance," *American Economic Journal: Microeconomics*, 2011, *3* (1), 214–244.

— **and Ed. Hopkins**, "Marriage as a Rat Race: Noisy Premarital Investments with Assortative Matching," *Journal of Political Economy*, 2016, *124* (4), 992–1045.

Brown, Philip H., Erwin Bulte, and Xiaobo Zhang, "Positional Spending and Status Seeking in Rural China," *Journal of Development Economics*, 2011, *96* (1), 139–149.

Browning, Martin, Pierre-André Chiappori, and Yoram Weiss, *Economics of the Family*, Cambridge University Press, 2014.

Chiappori, Pierre-Andre and Maurizio Mazzocco, "Static and intertemporal household decisions," *Journal of Economic Literature*, 2017, *55* (3), 985–1045.

Chiappori, Pierre-André, Bernard Fortin, and Guy Lacroix, "Marriage Market, Divorce Legislation, and Household Labor Supply," *Journal of Political Economy*, 2002, *110* (1), 37–72.

— **, Monica Costa Dias, and Costas Meghir**, "The Marriage Market, Labor Supply, and Education Choice," *Journal of Political Economy*, 2018, *126* (S1), S26–S72.

— **, Sonia Oreffice, and Climent Quintana-Domeque**, "Fatter attraction: Anthropometric and socioeconomic matching on the marriage market," *Journal of Political Economy*, 2012, *120* (4), 659–695.

— **,** — **, and** — , "Bidimensional matching with heterogeneous preferences: education and smoking in the marriage market," *Journal of the European Economic Association*, 2018, *16* (1), 161–198.

Choo, Eugene, "Dynamic marriage matching: An empirical framework," *Econometrica*, 2015, *83* (4), 1373–1423.

— **and Aloysius Siow**, "Who Marries Whom and Why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

Cole, Harold L., George J. Mailath, and Andrew Postlewaite, "Efficient Non-Contractible Investments in Large Economies," *Journal of Economic Theory*, December 2001, *101* (2), 333–373.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *American Economic Review*, 2007, *97* (2), 31–47.

Davin, D., "Marriage Migration in China: The Enlargement of Marriage Markets in the Era of Market Reforms.," *Indian Journal of Gender Studies*, 2005, *12* (2-3), 173–188.

Dizdar, Deniz, "Two-sided investment and matching with multidimensional cost types and attributes," *American Economic Journal: Microeconomics*, 2018, *10* (3), 86–123.

Doepke, Matthias and Fabian Kindermann, "Bargaining over babies: Theory, evidence, and policy implications," *American Economic Review*, 2019, *109* (9), 3264–3306.

Dupuy, Arnaud and Alfred Galichon, "Personality traits and the marriage market," *Journal of Political Economy*, 2014, *122* (6), 1271–1319.

Ebenstein, Avraham, "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy," *Journal of Human Resources*, 2010, *45* (1), 87–115.

— , "Estimating a Dynamic Model of Sex Selection in China," *Demography*, 2011, *48* (2), 783–811.

Galichon, Alfred and Bernard Salanié, "Matching with trade-offs: Revealed preferences over competing characteristics," *SSRN Working Paper*, 2010.

— **, Scott Duke Kominers, and Simon Weber**, "Costly concessions: An empirical framework for matching with imperfectly transferable utility," *Journal of Political Economy*, 2019, *127* (6), 2875–2925.

Gobbi, Paula E, "Childcare and commitment within households," *Journal of Economic Theory*, 2018, *176*, 503–551.

Heckman, James, "China's investment in human capital," *Economic Development and Cultural Change*, 2003, *51* (4), 795–804.

— , "China's human capital investment," *China Economic Review*, 2005, *16* (1), 50–70.

— , "Skill formation and the economics of investing in disadvantaged children," *Science*, 2006, *312* (5782), 1900–1902.

— , "The Economics, Technology, and Neuroscience of Human Capability Formation," *Proceedings of the National Academy of Sciences*, 2007, *104* (33), 13250–13255.

— **, Rodrigo Pinto, and Peter Savelyev**, "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 2013, *103* (6), 2052–2086.

Huang, Shuo, "Housing and Marriage," *Chinadaily*, 2010.

Iyigun, Murat and P. Randall Walsh, "Building the Family Nest: Premarital Investments, Marriage Markets, and Spousal Allocations," *Review of Economic Studies*, 2007, *74* (2), 507–535.

Li, Hongbin, Junjian Yi, and Junsen Zhang, "Estimating the Effect of the One-Child Policy on the Sex Ratio Imbalance in China: Identification Based on the Difference-in-Differences," *Demography*, 2011, *48* (4), 1535–1557.

Li, Qi and Juan Pantano, "The Demographic Consequences of Sex-Selection Technology," *Unpublished manuscript*, 2013.

Lin, Wanchuan, Kathryn McEvilly, and Juan Pantano, "Obesity and sex ratios in the US," *Review of Economics of the Household*, 2016, *14* (2), 269–292.

Lise, Jeremy and Ken Yamada, "Household sharing and commitment: Evidence from panel data on individual expenditures and time use," *The Review of Economic Studies*, 2018, *86* (5), 2184–2219.

Low, Corinne, "A "reproductive capital" model of marriage market matching," *Unpublished manuscript*, 2017.

Lundberg, Shelly and Robert A. Pollak, "Separate Spheres Bargaining and the Marriage Market," *Journal of Political Economy*, 1993, pp. 988–1010.

Lyle, David S., "Using Military Deployments and Job Assignments to Estimate the Effect of Parental Absences and Household Relocations on Children's Academic Achievement," *Journal of Labor Economics*, 2006, *24* (2), 319–350.

Mazzocco, Maurizio, "Household Intertemporal Behaviour: A Collective Characterization and A Test of Commitment," *Review of Economic Studies*, 2007, *74* (3), 857–895.

McKenzie, David and Hillel Rapoport, "Can Migration Reduce Educational Attainment? Evidence from Mexico," *Journal of Population Economics*, 2011, *24* (4), 1331–1358.

Nash, John F., "The bargaining problem," *Econometrica*, 1950, *18*, 155–162.

Nunn, Nathan, "A Model Explaining Simultaneous Payments of Dowry and Brideprice," *Unpublished manuscript*, 2005.

Peters, Michael and Aloysius Siow, "Competing Premarital Investments," *Journal of Political Economy*, 2002, *110* (3), 592–608.

Pierson, David, "China's Housing Boom Spells Trouble for Boyfriends," *Los Angeles Times*, 2010.

Rubinstein, Ariel, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 1982, *50* (1), 97–109.

Shaked, Avner and John Sutton, "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model," *Econometrica*, November 1984, *52* (6), 1351–1364.

Shaked, Moshe and J George Shanthikumar, *Stochastic orders*, Springer Science & Business Media, 2007.

Shepard, Wade, "Demand For Housing In China's Heated Property Market Is Real," *Forbes*, 2016.

Short, Susan E. and Fengying Zhai, "Looking Locally at China's One-Child Policy," *Studies in Family Planning*, 1998, pp. 373–387.

Siow, Aloysius, "Differential Fecundity, Markets, and Gender Roles," *Journal of Political Economy*, 1998, *106* (2), 334–354.

Wei, Shang-Jin and Xiaobo Zhang, "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy*, 2011, *119* (3), 511–564.

Wrenn, Douglas H, Junjian Yi, and Bo Zhang, "House prices and marriage entry in China," *Regional Science and Urban Economics*, 2019, *74*, 118–130.

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

— **and Yongai Jin**, "Household Wealth in China," *Chinese Sociological Review*, 2015, *47* (3), 203–229.

Xinhua, "Housing Pressure Leads to Delayed Marriage," *Chinadaily*, 2011.

Zhang, Hanzhe, "An Investment-and-Marriage Model with Differential Fecundity," *Unpublished manuscript*, 2019.

— , "Pre-matching gambles," *Games and Economic Behavior*, 2020, *121*, 76–89.

Zhang, Hongliang, Jere R. Behrman, C. Simon Fan, Xiangdong Wei, and Junsen Zhang, "Does Parental Absence Reduce Cognitive Achievements? Evidence from Rural China," *Journal of Development Economics*, 2014, *111*, 181–195.

Zhang, Junsen, "The evolution of China's one-child policy and its effects on family outcomes," *Journal of Economic Perspectives*, 2017, *31* (1), 141–60.

Zhao, Yaohui, "Labor Migration and Earnings Differences: The Case of Rural China," *Economic Development and Cultural Change*, 1999, *47* (4), 767–782.

Table 1 Summary statistics of main outcome variables

	Mean	Std. dev.	Min	Max	Observations
A: Parental labor supply					
Paternal migration	0.098	0.297	0	1	4,314
Maternal migration	0.025	0.158	0	1	4,314
At least one parent migration	0.111	0.314	0	1	4,314
Paternal working hours, yearly	2,466	947	400	5,400	1,534
Maternal working hours, yearly	2,416	902	240	5,400	978
B: Housing investments					
Housing construction area, thousand sq.m	0.126	0.086	0.008	1	4,169
Housing ownership	0.831	0.375	0	1	4,314
Housing mortgage, thousand	5.392	32.04	0	750	4,314
C: Child educational investments					
Education expenditure, thousand	1.507	2.629	0	40	3,978
Having an education funding	0.297	0.457	0	1	3,978

Notes: Data are from the 2010 CFPS survey. The main sample includes all households in the 2010 CFPS family dataset in which the first-born child was between the ages of zero and 15, both parents were alive and at most 50 years old, and at least one parent participated in the 2010 CFPS adult survey. In panel C, child educational investments are measured for first-born children who are at least two years old. Descriptive statistics are weighted by the CFPS survey sampling weights.

Table 2 Balance test: First-son versus first-daughter families

	Mean (Std. dev.)				
	All	First-son families	First-daughter families	Difference	SE
	(1)	(2)	(3)	(4)	(5)
First son	0.507				
	(0.500)				
Sex ratio (m/f)	1.077	1.076	1.077	-0.001	0.003
	(0.101)	(0.100)	(0.101)		
Ethnicity (minority=1)	0.124	0.121	0.128	-0.007	0.010
	(0.330)	(0.326)	(0.334)		
Region of residence (urban=1)	0.438	0.452	0.424	0.028	0.015
	(0.496)	(0.498)	(0.494)		
First-child age	8.746	8.623	8.874	-0.251	0.138
	(4.543)	(4.531)	(4.552)		
Father's age	36.14	36.03	36.27	-0.240	0.187
	(6.149)	(6.137)	(6.162)		
Father's schooling years	7.818	7.890	7.745	0.145	0.131
	(4.308)	(4.266)	(4.350)		
Father's political status (party=1)	0.091	0.090	0.092	-0.002	0.009
	(0.287)	(0.286)	(0.289)		
Mother's age	34.30	34.21	34.40	-0.190	0.190
	(6.251)	(6.264)	(6.239)		
Mother's schooling years	6.549	6.591	6.506	0.085	0.143
	(4.693)	(4.652)	(4.735)		
Mother's political status (party=1)	0.026	0.030	0.023	0.007	0.005
	(0.160)	(0.171)	(0.149)		
Observations	4,314	2,186	2,128		

Notes: Data are from the 2010 CFPS survey. In the first three columns, standard deviations are given in parentheses. In the last column are standard errors for the difference between first-son and first-daughter families, none of which are statistically significant at the five percent level.

Table 3 Baseline results: Parental labor supply and premarital investments

A: Parental labor supply					
	Migration			Working hours, log	
Dependent variable	Father	Mother	At least one parent	Father	Mother
	(1)	(2)	(3)	(4)	(5)
First son * Sex ratio (β_3)	0.235**	0.098*	0.264***	0.569***	0.473
	(0.094)	(0.059)	(0.093)	(0.169)	(0.408)
Observations	4,314	4,314	4,314	1,534	978
R-squared	0.109	0.064	0.113	0.164	0.256
Dependent variable mean	0.098	0.025	0.111	7.726	7.701
Percentage difference sex ratio+0.1	24.1	38.6	23.8	5.7	4.7
Model	OLS	OLS	OLS	OLS	OLS
Other controls	YES	YES	YES	YES	YES
County fixed effects	YES	YES	YES	YES	YES

B: Premarital investments					
	Housing investments			Child educational investments	
Dependent variable	Construction area, log sq.m	Ownership	Mortgage, thousand	Education expenditure, thousand	Having an education fund
	(1)	(2)	(3)	(4)	(5)
First son * Sex ratio (β_3)	0.413**	0.233**	15.40**	-1.663**	-0.337**
	(0.205)	(0.117)	(7.141)	(0.800)	(0.161)
Observations	4,169	4,314	4,314	3,978	3,978
R-squared	0.278	0.177	0.145	0.323	0.135
Dependent variable mean	4.650	0.831	5.392	1.507	0.297
Percentage difference sex ratio+0.1	4.1	2.8	28.6	-11.0	-11.3
Model	OLS	OLS	OLS	OLS	OLS
Other controls	YES	YES	YES	YES	YES
County fixed effects	YES	YES	YES	YES	YES

Notes: Data are from the 2010 CFPS survey. In columns (1)–(3) of panel A and columns (2)–(5) of panel B, the difference in the effect of sex imbalance between first-son and first-daughter families is reported in both percentage points (β_3) and percentages (β_3/dependent variable mean); in the remaining columns, the difference is reported in percentages. In columns (4) and (5) of panel B, child educational investments are measured for the first-born child in the family who is at least two years old. Estimations are weighted by the CFPS survey sampling weights. Standard errors given in parentheses are clustered at the county level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 4 Addressing issues related to son-preferring fertility stopping rules

Dependent variable		Paternal migration	House construction area, log sq.m	Education expenditure, thousand
		(1)	(2)	(3)
		Interaction-term coefficient (β_3)		
Benchmark		0.235**	0.413**	-1.663**
A: Family-size effect				
Adding # children		0.240**	0.409**	-1.689**
		[0.218]	[0.478]	[0.285]
Adding # children & Interaction with first son		0.245**	0.410*	-1.689**
		[0.215]	[0.745]	[0.467]
B: Families with one child				
One-child families	No age limit	0.234**	0.336	-1.776**
	Child $\geq$ 4	0.223**	0.217	-2.411**
C: Child-gender compositions				
Having any son	OLS	0.223***	0.310	-1.168*
	OLS, adding # children	0.221***	0.313	-1.151
	OLS, adding # children & interaction	0.220***	0.313	-1.154
	IV	0.355**	0.528**	-2.505**
	IV, adding # children	0.360**	0.522**	-2.644**
	IV, adding # children & interaction	0.356**	0.505**	-2.608**
Share of sons	OLS	0.300***	0.398*	-1.095
	OLS, adding # children	0.302***	0.394*	-1.112
	OLS, adding # children & interaction	0.301***	0.394*	-1.114
	IV	0.305**	0.495**	-2.173**
	IV, adding # children	0.312**	0.493**	-2.231**
	IV, adding # children & interaction	0.308**	0.474**	-2.243**

Notes: Data are from the 2010 CFPS survey. In column (3), education expenditure is measured for first-born children who are at least two years old. The difference in the effect of sex imbalance between first-son and first-daughter families is reported in percentage points (β_3). In panel A, p-values of Hausman's general specification test for the equality of β_3 are given in square brackets. In panel C, the instrument for having any son and the share of sons in IV regressions is the first-son dummy; see Appendix Table A4 for first-stage results. Estimations are weighted by the CFPS survey sampling weights. Standard errors are clustered at the county level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 5 Addressing issues related to county-level sex ratios

Dependent variable		Paternal migration (1)	House construction area, log sq.m (2)	Education expenditure, thousand (3)
		Interaction-term coefficient (β_3)		
Benchmark		0.235**	0.413**	-1.663**
A: Unobservable cross-county heterogeneity				
No county fixed effects		0.233**	0.245	-1.857***
		[0.914]	[0.017]	[0.428]
B: Sex-ratio confounders				
Adding average household financial wealth		0.236**	0.397**	-1.665**
		[0.688]	[0.479]	[0.939]
Adding average household financial wealth		0.236**	0.396*	-1.675**
& Interaction with first son		[0.738]	[0.413]	[0.885]
Adding average household income		0.237**	0.402*	-1.662**
		[0.592]	[0.363]	[0.911]
Adding average household income		0.239***	0.405**	-1.632**
& Interaction with first son		[0.663]	[0.593]	[0.748]
Adding gender wage differential, m-f		0.251***	0.356*	-1.756**
		[0.142]	[0.029]	[0.441]
Adding gender wage differential, m-f		0.252***	0.356*	-1.766**
& Interaction with first son		[0.176]	[0.025]	[0.453]
Adding social insurance		0.236**	0.432**	-1.694**
		[0.911]	[0.418]	[0.560]
Adding social insurance		0.242***	0.429**	-1.679**
& Interaction with first son		[0.494]	[0.550]	[0.858]
Adding grandparental coresidence		0.232**	0.394*	-1.661**
		[0.567]	[0.526]	[0.824]
Adding grandparental coresidence		0.237**	0.393*	-1.664**
& Interaction with first son		[0.857]	[0.532]	[0.966]
Adding all variables above		0.249***	0.347*	-1.794**
		[0.298]	[0.271]	[0.321]
Adding all variables above		0.260***	0.339*	-1.802**
& Interactions with first son		[0.245]	[0.156]	[0.331]
Adding variables selected by high-dimensional method		0.251***	0.519**	-1.734**
& Interactions with first son		[0.786]	[0.359]	[0.844]
C: IV results		0.374*	1.283*	-3.291*
		(0.224)	(0.776)	(1.993)
D: Sex ratios for alternative age cohorts				
Cohort of sex ratio	10–14	0.209**	0.289*	-0.852
	15–19	0.249***	0.284*	-1.166
	20–24	0.106	0.299*	-1.376**

Notes: Data are from the 2010 CFPS survey. In column (3), education expenditure is measured for first-born children who are at least two years old. The difference in the effect of sex imbalance between first-son and first-daughter families is reported in percentage points (β_3). In panels A and B, p-values of Hausman's general specification test for the equality of β_3 are given in square brackets. For IV regressions in panel C, see Appendix Table A5 for results regressing sex ratios on excluded instruments in lieu of first-stage results. Estimations are weighted by the CFPS survey sampling weights. Standard errors are clustered at the county level, and given in parentheses in panel C.

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

Table 6 Sex imbalance and child human capital development

	Cognitive		Non-cognitive		Health	
Dependent variable	Math ranking (1)	Chinese ranking (2)	Openness (3)	Cooperation (4)	Weight, z-score (5)	Height, z-score (6)
First son * Sex ratio (β_3)	-0.734***	-0.567**	-0.498**	-0.572***	-0.907**	-0.179
	(0.237)	(0.246)	(0.250)	(0.200)	(0.412)	(0.605)
Observations	1,154	1,154	2,125	2,125	4,137	3,870
R-squared	0.618	0.641	0.405	0.457	0.265	0.261
Dependent variable mean	0.692	0.702	0.859	0.729	-0.505	-0.639
Percentage difference sex ratio+0.1	-10.6	-8.1	-5.8	-7.9	-18.0	-2.8
Model	OLS	OLS	OLS	OLS	OLS	OLS
Other controls	YES	YES	YES	YES	YES	YES
County fixed effects	YES	YES	YES	YES	YES	YES

Notes: Data are from the 2010 CFPS survey. Human capital outcomes are measured for the first-born child in a family. In columns (1)–(4), the sample excludes families in which the first child is below ten years old. The difference in the effect of sex imbalance between first-son and first-daughter families is reported in both percentage points (β_3) and percentages (β_3/dependent variable mean). Estimations are weighted by the CFPS survey sampling weights. Standard errors given in parentheses are clustered at the county level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 7 The impacts of parental migration on children

	Father			Mother		
	At-home mean (1)	Migration mean (2)	Difference (3)	At-home mean (4)	Migration mean (5)	Difference (6)
A: Human capital outcomes						
School math exam ranking	0.683	0.646	0.037*	0.679	0.686	-0.007
School Chinese exam ranking	0.698	0.673	0.025	0.695	0.688	0.007
Openness	0.862	0.881	-0.019	0.863	0.883	-0.020
Cooperation	0.727	0.678	0.049*	0.723	0.650	0.073
Weight, kg	29.03	27.89	1.140*	28.97	26.43	2.540**
Height, m	1.286	1.259	0.027**	1.284	1.255	0.029
B: Time allocation, hours						
Homework and revision	2.006	1.718	0.288***	1.981	1.803	0.178
After-school tuition	0.399	0.129	0.270***	0.371	0.347	0.024
Extracurricular reading	0.720	0.604	0.116**	0.713	0.521	0.192**
Physical exercise	0.336	0.274	0.062*	0.332	0.252	0.080
Observations						2,245
C: Psychological well-being						
Happiness	0.465	0.369	0.096***	0.459	0.290	0.169***
Optimism about future	0.409	0.398	0.011	0.410	0.323	0.087*
Relationship with others	0.341	0.280	0.061**	0.337	0.242	0.095*
Popularity	0.285	0.233	0.052**	0.281	0.226	0.055
Observations						2,259

Notes: Data are from the 2010 CFPS survey. For more information on variables in panel A, see notes to Table 6. In panels B and C, the sample excludes families in which the first child is below ten years old. Differences between non-migrant and migrant families are reported in columns (3) and (6); H_0 is difference=0 and H_1 is difference>0.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

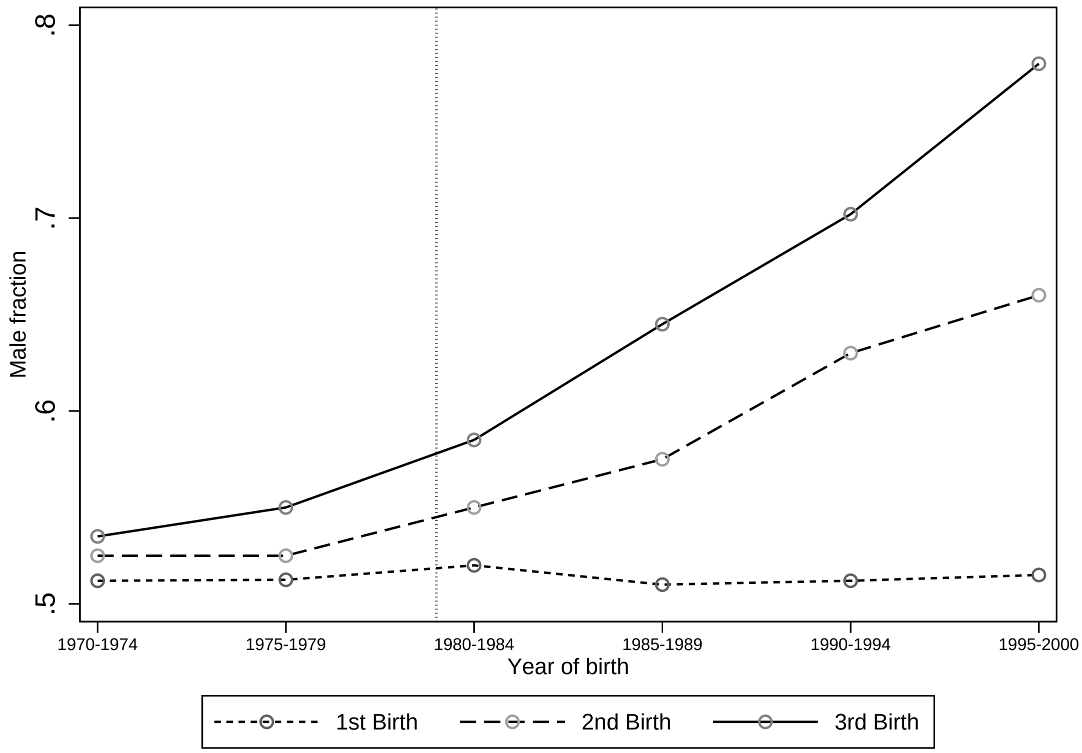


Figure 1 Male fraction of births by birth order in China

Notes: Data are from Ebenstein (2010). The figure shows a steep rise in the sex ratio over the past decades, and the imbalance comes from gender selection among second- and higher-order births, rather than among first-order births.

Online Appendices

A Additional tables and figures

Table A1 Marital status by age, gender, and education level

Age cohort	Secondary school		High school		College and above	
	Male	Female	Male	Female	Male	Female
A: Share of population divorced						
22–31	0.011	0.009	0.007	0.009	0.003	0.004
32–41	0.024	0.018	0.027	0.038	0.018	0.034
42–51	0.024	0.019	0.029	0.047	0.022	0.052
52–61	0.018	0.019	0.019	0.033	0.017	0.042
B: Share of population ever married						
22–31	0.636	0.780	0.505	0.628	0.363	0.453
32–41	0.944	0.984	0.943	0.968	0.945	0.955
42–51	0.979	0.996	0.985	0.992	0.989	0.987
52–61	0.985	0.997	0.992	0.995	0.995	0.990
C: Divorce rate						
22–31	0.018	0.011	0.013	0.014	0.008	0.010
32–41	0.026	0.018	0.029	0.039	0.019	0.036
42–51	0.024	0.019	0.030	0.047	0.022	0.053
52–61	0.018	0.019	0.020	0.033	0.017	0.042

Notes: Data are from the 2010 China population census.

Table A2 Housing, education, and marital status of men

Dependent variable	Marital status of men (married=1)				
	(1)	(2)	(3)	(4)	(5)
High-quality housing	0.019***				0.013***
(costs $\geq$ 50k=1)	(0.004)				(0.004)
High-quality housing		0.045***			0.044***
(private bathroom=1)		(0.004)			(0.004)
High education			0.002		
(high school and above=1)			(0.004)		
High education				0.010**	0.005
(college and above=1)				(0.005)	(0.005)
Age	0.461***	0.460***	0.461***	0.460***	0.460***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Age square	-0.008***	-0.008***	-0.008***	-0.008***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hukou (urban=1)	0.018***	0.015***	0.024***	0.020***	0.008**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Observations	94,457	94,457	94,457	94,457	94,457
R-squared	0.216	0.217	0.216	0.216	0.217
Dependent variable mean	0.440	0.440	0.440	0.440	0.440
Model	OLS	OLS	OLS	OLS	OLS

Notes: Data are from the 2000 China population census. The sample includes men who were: (i) between the ages of 20 and 40; (ii) from families who bought or built a new house between 1997 and 2000; and (iii) unmarried before the house was bought or built. Since the houses were newly got in families with an unwedded son of marriage age, they were most likely for the son's marriage purpose. The outcome variable is a dummy that equals one if the son got married during that period (1997–2000), and zero if he remained single. Housing condition is measured by: (i) a dummy variable that equals one if the new house cost no less than 50,000 yuan (about 6.3 times per capita GDP in 2000); and (ii) a dummy variable that equals one if the new house had a private bathroom (as opposed to a shared bathroom). Education level is measured by: (i) a dummy variable indicating whether the man had at least a high school diploma; and (ii) a dummy variable indicating whether the man had at least a college degree. Standard errors are given in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table A3 Parental migration and gross family income

Dependent variable	Gross family income, thousand			
	(1)	(2)	(3)	(4)
Paternal migration	6.935***			
	(2.447)			
Maternal migration		8.891***		
		(3.093)		
At least one parent migration			7.065***	
			(2.248)	
Both parents migration				11.672***
				(3.702)
Observations	4,314	4,314	4,314	4,314
R-squared	0.191	0.190	0.191	0.189
Dependent variable mean	32.1	32.1	32.1	32.1
Percentage increase (migration=1)	21.6	27.7	22.0	36.4
Model	OLS	OLS	OLS	OLS
Other controls	YES	YES	YES	YES
County fixed effects	YES	YES	YES	YES

Notes: Data are from the 2010 CFPS survey. The migration effect on gross family income is reported in both percentage points and percentages. Estimations are weighted by the CFPS survey sampling weights. Standard errors given in parentheses are clustered at the county level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table A4 First-stage results—Child-gender measures are instrumented

Second-stage dependent variable	Paternal migration	House construction area, log sq.m	Education expenditure, thousand
	(1)	(2)	(3)
A: Endogenous variable is having any son			
First son	1.206***	1.213***	1.224***
	(0.233)	(0.229)	(0.252)
R-squared	0.630	0.638	0.611
B: Endogenous variable is share of sons			
First son	1.113***	1.099***	1.123***
	(0.165)	(0.156)	(0.177)
R-squared	0.821	0.825	0.809
Observations	4,314	4,169	3,978

Notes: Data are from the 2010 CFPS survey. The instrument for having any son and the share of sons is the first-son dummy; see panel C, Table 4 for second-stage results. Estimations are weighted by the CFPS survey sampling weights. Standard errors are clustered at the county level.

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

Table A5 Regressing sex ratios on policy-violation penalty and quota of births

Dependent variable	Sex ratio		
"Second-stage" dependent variable	Paternal migration estimation	House construction area estimation	Education expenditure estimation
	(1)	(2)	(3)
Policy-violation penalty	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)
Quota of births	0.034***	0.031***	0.037***
	(0.005)	(0.006)	(0.006)
Policy-violation penalty * Minority	-0.004***	-0.004***	-0.004***
	(0.000)	(0.000)	(0.000)
Quota of births * Minority	-0.025**	-0.019*	-0.027**
	(0.011)	(0.011)	(0.011)
Observations	4,314	4,169	3,978
R-squared	0.663	0.653	0.663
Other controls	YES	YES	YES

Notes: Data are from the 2010 CFPS survey. Other controls include controls in the respective IV estimations in Table 5, plus province dummies (here both the dependent variable and key explanatory variables are at the county level). Estimations are weighted by the CFPS survey sampling weights. Standard errors given in parentheses are clustered at the county level.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table A6 Stated purpose of migration remittances

Dependent variable	Migration purpose	
	For children's marriage (1)	For children's education (2)
First son * Sex ratio (β_3)	0.179**	0.096
	(0.079)	(0.262)
Observations	1,071	1,071
R-squared	0.213	0.272
Model	OLS	OLS
Other controls	YES	YES
County fixed effects	YES	YES

Notes: Data are from the 2010 CFPS survey. The difference in the effect of sex imbalance between first-son and first-daughter families is reported in percentage points (β_3). Estimations are weighted by the CFPS survey sampling weights. Standard errors are given in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table A7 Heterogenous effects of sex imbalance across families

Dependent variable	Paternal migration	House construction area, log sq.m	Education expenditure, thousand
	(1)	(2)	(3)
Benchmark: First son * Sex ratio (β_3)	0.235**	0.413**	-1.663**
A: Families with a first child above 11			
First son * Sex ratio (β_3)	0.254**	0.846**	-0.265
	(0.119)	(0.392)	(1.073)
Observations	1,811	1,745	1,811
R-squared	0.162	0.265	0.369
Dependent variable mean	0.092	4.656	1.526
B: Families with a first child below 11			
First son * Sex ratio (β_3)	0.284**	0.115	-2.651*
	(0.110)	(0.221)	(1.391)
Observations	2,503	2,424	2,167
R-squared	0.151	0.361	0.357
Dependent variable mean	0.102	4.646	1.492

Notes: Data are from the 2010 CFPS survey. In column (3), education expenditure is measured for first-born children who are at least two years old. Panel A includes a sample of families with a first child above the age of 11, and panel B includes a sample of families with a first child below the age of 11. The difference in the effect of sex imbalance between first-son and first-daughter families is reported in percentage points (β_3). Estimations are weighted by the CFPS survey sampling weights. Standard errors given in parentheses are clustered at the county level.

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

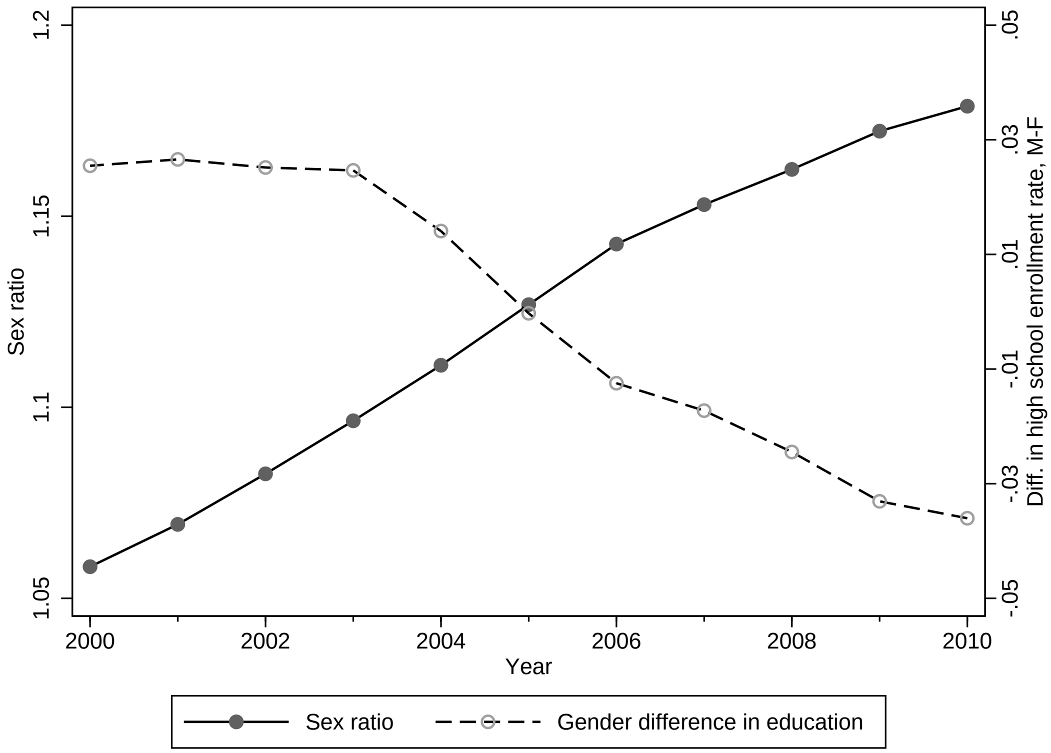


Figure A1 Trends in sex ratio and gender difference in education in China

Notes: Data on sex ratios and high school enrollment rates are projected from the 2010 China population census. For example, the sex ratio for the cohort between the ages of zero and 15 in 2006 is calculated using the cohort between the ages of four and 19 in 2010, since these two cohorts are supposed to be the same. The correlation coefficient between sex ratio and gender difference in high school enrollment rate during 2000–2010 is -0.972 (with the 95 percent confidence interval: -0.993 to -0.893).

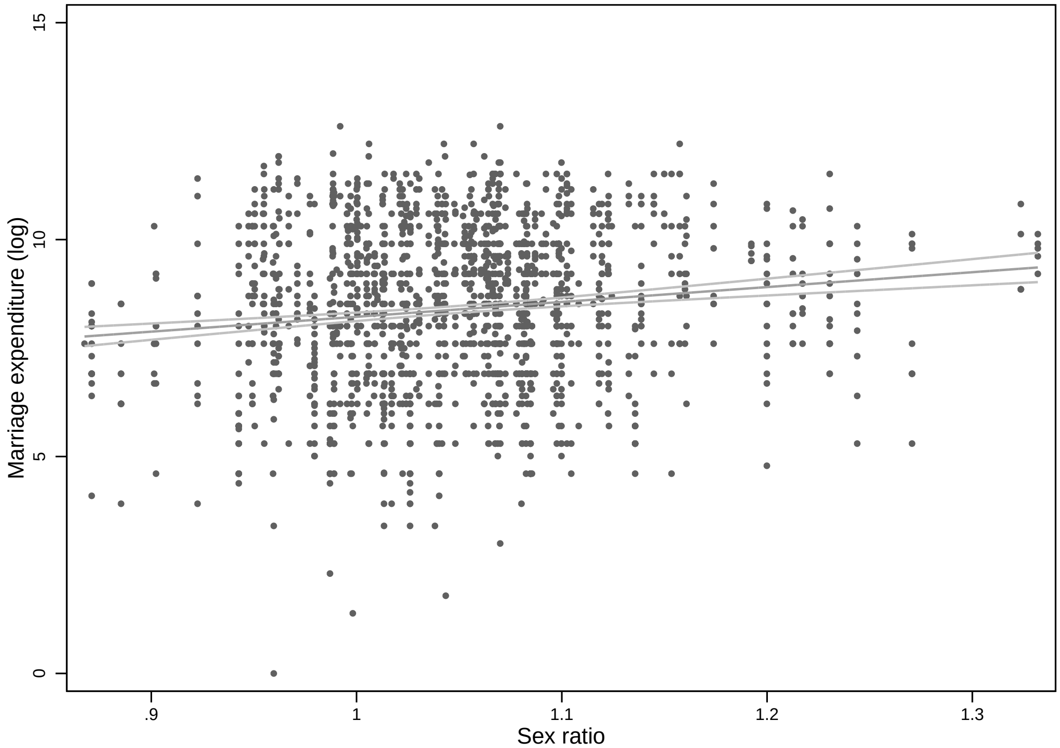
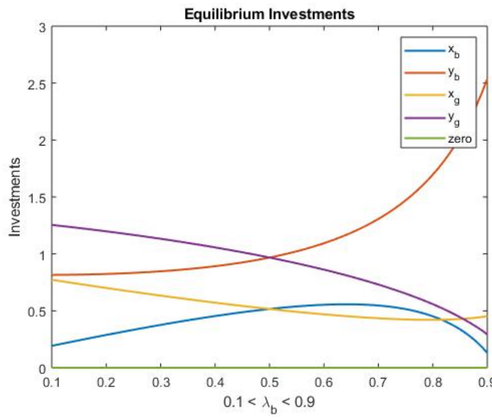


Figure A2 County-level sex ratios and marriage expenditures in China

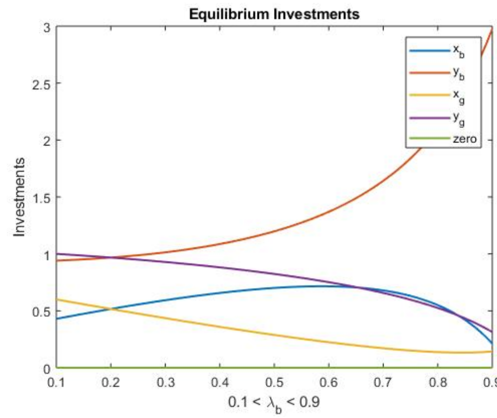
Notes: Data on county-level sex ratios are from the 2010 China population census. Data on marriage expenditures are from the 2010 CFPS survey.

Figure A3 Trends in sex ratio and marriage expenditure in China

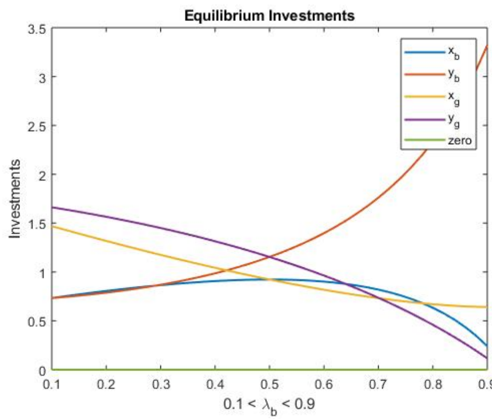
Notes: Data on sex ratios are projected from the 2010 China population census. For example, the sex ratio for the cohort between the ages of zero and 15 in 2006 is calculated using the cohort between the ages of four and 19 in 2010, since these two cohorts are supposed to be the same. Data on marriage expenditures are from Brown et al. (2011).



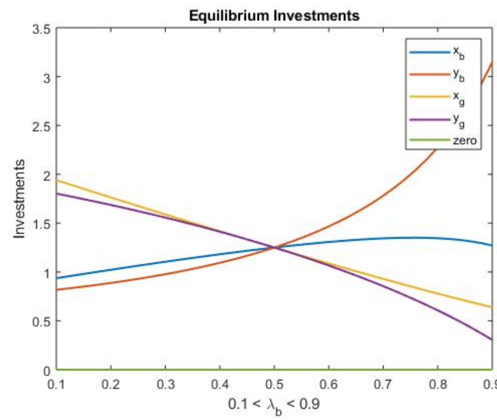
(a) $\gamma = 0.5, \theta_B = 1, \theta_G = 1$



(b) $\gamma = 0.2, \theta_B = 1, \theta_G = 1$



(c) $\gamma = 0.5, \theta_B = 2, \theta_G = 0.5$



(d) $\gamma = 0.2, \theta_B = 2, \theta_G = 0.5$

Figure A4 Equilibrium investments as a function of λ_B

Notes: This figure is generated in Online Appendix B. Each graph depicts the level of equilibrium investments as a function of $\lambda_B \in [0.1, 0.9]$.

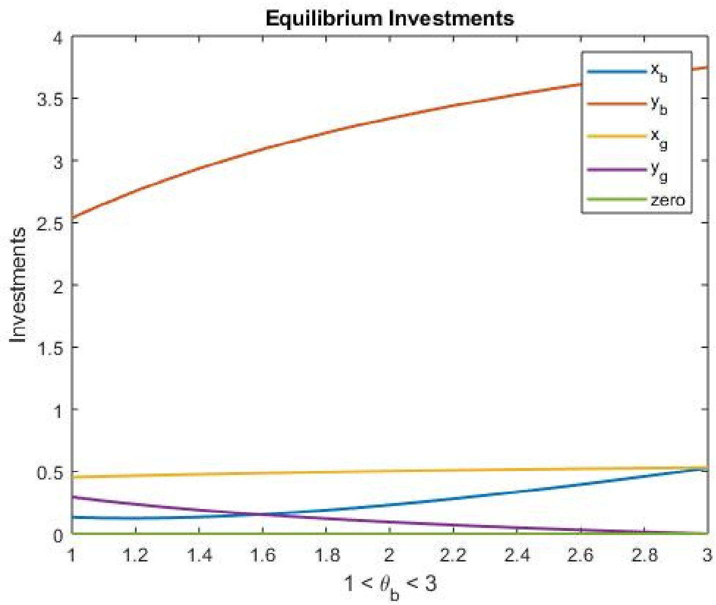


Figure A5 Equilibrium investments as a function of θ_B

Notes: This figure is generated in Online Appendix B. It depicts the level of equilibrium investments while fixing $\theta_B\theta_G$ to 1 with $\lambda_B = 0.9, \lambda_G = 0.5$ and $\gamma = 0.5$.

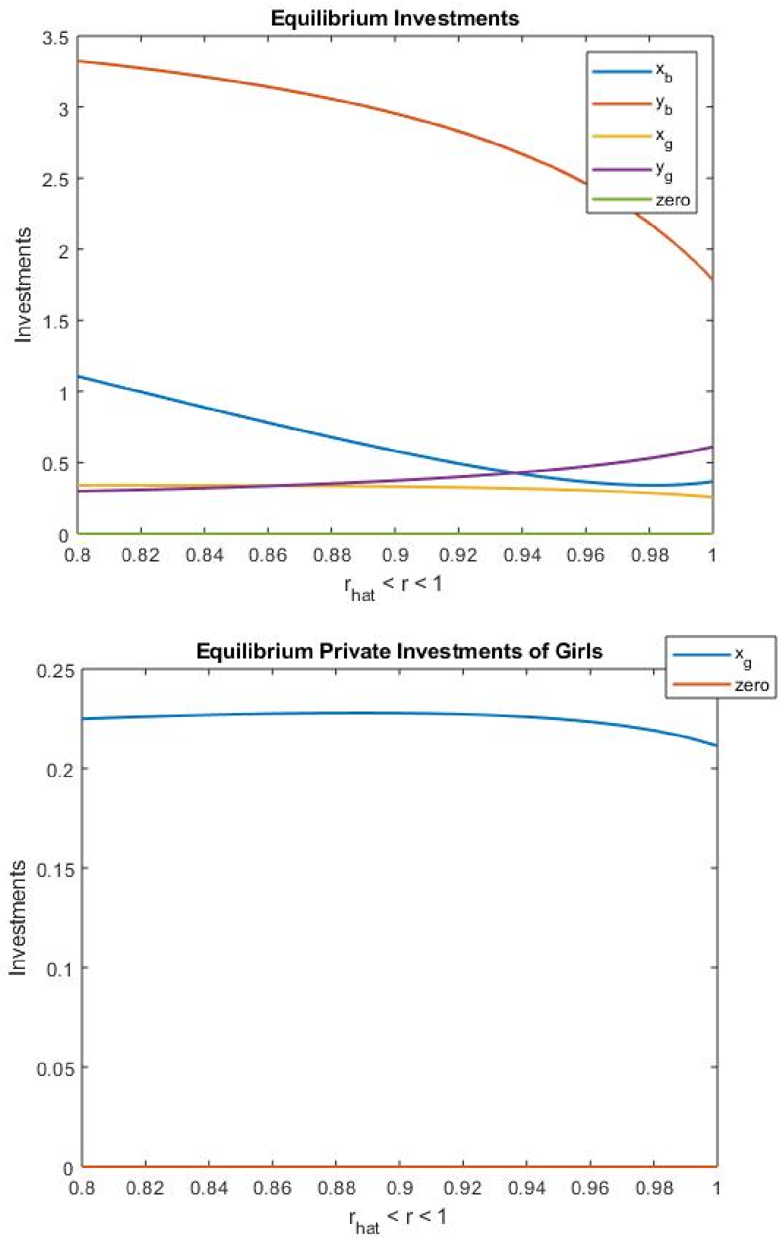


Figure A6 Equilibrium investments with unbalanced sex ratios

Notes: This figure is generated in Online Appendix C. It depicts the level of equilibrium investments when sex ratios are unbalanced.

B Comparative statics in balanced marriage markets

We totally differentiate the first order conditions for optimal investment, to obtain the following system:

$$c^B_{xx}(.)\frac{dx_B}{d\lambda_B} + c^B_{xy}(.)\frac{dy_B}{d\lambda_B} = 1 \quad \text{(A1)}$$

$$c^B_{xy}(.)\frac{dx_B}{d\lambda_B} + c^B_{yy}(.)\frac{dy_B}{d\lambda_B} = v^{''}_B(\hat{y})(\frac{dy_B}{d\lambda_B} + \frac{dy_G}{d\lambda_B}) + \frac{1-\lambda_G}{1-\lambda_B}\theta_B v^{''}_G(\frac{dy_B}{d\lambda_B} + \frac{dy_G}{d\lambda_B}) + \frac{1-\lambda_G}{(1-\lambda_B)^2}\theta_B v^{'}_G \quad \text{(A2)}$$

$$c^G_{xx}(.)\frac{dx_G}{d\lambda_B} + c^G_{xy}(.)\frac{dy_G}{d\lambda_B} = -\theta_G \quad \text{(A3)}$$

$$c^G_{xy}(.)\frac{dx_G}{d\lambda_B} + c^G_{yy}(.)\frac{dy_G}{d\lambda_B} = v^{''}_G(\hat{y})(\frac{dy_B}{d\lambda_B} + \frac{dy_G}{d\lambda_B}) + \frac{1-\lambda_B}{1-\lambda_G}\theta_G v^{''}_B(\frac{dy_B}{d\lambda_B} + \frac{dy_G}{d\lambda_B}) - \frac{\theta_G}{(1-\lambda_G)^2}v^{'}_B \quad \text{(A4)}$$

The equations (A1)–(A4) can be put into a matrix form as follows: $\mathbf{n}$ x $\mathbf{q} = \mathbf{a}$, where $\mathbf{n}$ is a 4×4 matrix and $\mathbf{q} = (\frac{dx_B}{d\lambda_B}, \frac{dy_B}{d\lambda_B}, \frac{dx_G}{d\lambda_B}, \frac{dy_G}{d\lambda_B})$.

We first calculate inv($\mathbf{n}$) and factor out the determinant of $\mathbf{n}$ in order to make the final expression simple. Denote by det the determinant of $\mathbf{n}$. We will present the final answer in the form

$$\mathbf{q} = \frac{1}{det}\hat{\mathbf{q}}$$

, where $\hat{\mathbf{q}} = \mathbf{n}^{-1}a \times det$. det can be simplified in the following way. Define $H_i = (c^i_{xy})^2 - c^i_{xx}c^i_{yy}$, where $i = G, B$. Note that H_i is the determinant of Hessian of the c_i function. In the simplification, we factor out H_i.

$$det = A_0 + B_0\frac{1}{1-\lambda_G} + C_0\frac{1}{1-\lambda_B}, \quad \text{(A5)}$$

where

$$A_0 = (H_B + c^B_{xx}v''_B)(H_G + c^G_{xx}v''_G) - c^B_{xx}v''_B c^G_{xx}v''_G$$

$$B_0 = H_B c^G_{xx}v''_B\theta_G(1-\lambda_B)$$

$$C_0 = H_G c^B_{xx}v''_G\theta_B(1-\lambda_G)$$

Next, we calculate $\hat{\mathbf{q}} \equiv (\hat{q}_1, \hat{q}_2, \hat{q}_3, \hat{q}_4) = \frac{1}{det}(\frac{dx_B}{d\lambda_B}, \frac{dy_B}{d\lambda_B}, \frac{dx_G}{d\lambda_B}, \frac{dy_G}{d\lambda_B})$

$$\hat{q}_i = A_i + B_i\frac{1}{1-\lambda_G} + C_i\frac{1}{1-\lambda_B} + D_i\frac{1}{(1-\lambda_B)^2}, \tag{A6}$$

where $i = 1,2,3,4$ and

$$A_1 = (v''_B - c^B_{yy})H_G - [c^G_{xx}c^B_{yy}v''_G + c^G_{xy}c^B_{xy}v''_B\theta_G]$$

$$B_1 = c^G_{xx}c^B_{xy}v'_B v''_B\theta_G - (1-\lambda_B)c^G_{xx}c^B_{yy}v''_B\theta_G$$

$$C_1 = (H_G - c^G_{xy}c^B_{xy}\theta_G)v''_G\theta_B(1-\lambda_G) + c^G_{xx}\theta_G\theta_B(c^B_{xx}v''_G + v'_G v''_B)$$

$$D_1 = (H_G + c^G_{xx}v''_G)c^B_{xy}v'_G\theta_B(1-\lambda_G)$$

$$A_2 = c^B_{xy}H_G + c^G_{xx}c^B_{xy}v''_G + c^B_{xx}c^G_{xy}v''_B\theta_G$$

$$B_2 = c^G_{xx}c^B_{xy}v''_B\theta_G(1-\lambda_B) - c^B_{xx}c^G_{xx}v'_B v''_B\theta_G$$

$$C_2 = c^B_{xx}c^G_{xx}\theta_B\theta_G(v''_B v'_G - v''_G v'_B) + c^B_{xx}c^G_{xy}\theta_B\theta_G v''_G(1-\lambda_G)$$

$$D_2 = -v'_G\theta_B(c^B_{xx}H_G + c^B_{xx}c^G_{xx}v''_G)$$

$$A_3 = (c^G_{yy} - v''_G)H_B\theta_G + [c^G_{xy}c^B_{xy}v''_G + c^G_{yy}c^B_{xx}v''_B\theta_G]$$

$$B_3 = v''_B\theta_G(1-\lambda_B)(c^G_{xy}c^B_{xy} - H_B\theta_G) - c^G_{xy}v'_B\theta_G(H_B + c^B_{xx}v''_B)$$

$$C_3 = c^B_{xx}c^G_{yy}\theta_B\theta_G v''_G(1-\lambda_G) - c^B_{xx}c^G_{xy}\theta_B\theta_G(v''_B v'_G + v''_G v'_B)$$

$$D_3 = -c^B_{xx}c^G_{xy}v''_G v'_B\theta_B(1-\lambda_G)$$

$$A_4 = -c^G_{xy} H_B \theta_G - [c^G_{xx} c^B_{xy} v''_G + c^G_{xy} c^B_{xx} v''_B \theta_G]$$

$$B_4 = (H_B + c^B_{xx} v''_B) c^G_{xx} v'_B \theta_G - c^G_{xx} c^B_{xy} v''_B \theta_B (1 - \lambda_B)$$

$$C_4 = c^B_{xx} c^G_{xx} \theta_B \theta_G (v''_B v'_G + v''_G v'_B) - c^B_{xx} c^G_{xy} \theta_B \theta_G v''_G (1 - \lambda_G)$$

$$D_4 = c^B_{xx} c^G_{xx} v''_G v'_G \theta_B (1 - \lambda_G)$$

B.1 Numerical results

We now present a numerical example, which shows the parameter range where $\frac{dx_B}{d\lambda_B} < 0$. We assume the following functional forms and parameter values.

- $c(x, y) = \frac{\gamma}{2}(x + y)^2 + \frac{1-\gamma}{2}(x^2 + y^2)$
- $v_G(y = v_B(y) = ay - by^2$
- $\lambda_G = 0.5$
- a = 1, b = 0.1
- $\gamma = \{0.2, 0.5\}$

Each graph in Appendix Figure A4 represents the level of equilibrium investments as a function of $\lambda_B \in [0.1, 0.9]$. Note that we have interior solutions for all $\lambda_B \in [0.1, 0.9]$. Graphs on the left column are set for $\gamma = 0.5$ and those on the right for $\gamma = 0.2$. Each row is for different values of θ_B, θ_G. It is easy to see that x_B decreases in λ_B after some point of λ_B. The table below summarizes those critical values of λ_B after which $\frac{dx_B}{d\lambda_B} < 0$.

	$\gamma = 0.5$	$\gamma = 0.2$
$(\theta_B = 1, \theta_G = 1)$	0.642	0.81
$(\theta_B = 2, \theta_G = 0.5)$	0.535	0.773

For example, when $\theta_B = 1, \theta_G = 1, \gamma = 0.5$, $\frac{dx_B}{d\lambda_B} < 0$ for $\lambda_B \in (0.642, 0.9]$. Note that the result is for $\lambda_B \in [0.1, 0.9]$. The range is chosen to guarantee interior solutions.

Appendix Figure A5 shows the equilibrium investments as a function of θ_B while fixing $\theta_B \theta_G$ to 1 with $\lambda_B = 0.9, \lambda_G = 0.5$ and $\gamma = 0.5$.

Note that we have interior solutions for all $\theta_B \in [1, 3]$. We can see x_B is decreasing in θ_B for some value of θ_B. Solving analytically for $\frac{dx_B}{d\theta_B}$ gives us that $\frac{dx_B}{d\theta_B} < 0$ for $\theta_B \in [1, 1.96049]$ and increasing otherwise.

C Comparative statics in unbalanced marriage markets

We totally differentiate the first-order conditions for equilibrium investments, equations (16), (17), (18) and (19), with respect to r, to obtain the following system of four equations with four unknowns. The second derivative of each function is evaluated at the equilibrium values. To simplify the notation, we denote the utility derived from public good investment of a boy when he remains single by $m_B = v_B(y_B^*)$.

$$\begin{aligned}
c_{xx}^B \frac{dx_B}{dr} + c_{xy}^B \frac{dy_B}{dr} &= \rho^{'}(r)[(1-\hat{r})(1-\lambda_B) + (1-\lambda_G)(\hat{r}\theta_{B+} - \theta_B) + \hat{f}\bar{U}], \\
c_{xy}^B \frac{dx_B}{dr} + c_{yy}^B \frac{dy_B}{dr} &= (v_B^{''} + \frac{1-\lambda_G}{1-\lambda_B} v_G^{''}\theta_B)(\frac{dy_B}{dr} + \frac{dy_G}{dr}) \\
&\quad + \rho^{'}(r)[(1-\hat{r})[m_B^{'} - v_B^{'}] + \frac{1-\lambda_G}{1-\lambda_B} v_G^{'}(\hat{r}\theta_{B+} - \theta_B) + \frac{v_G^{'}}{1-\lambda_B}\hat{f}\bar{U}] \\
&\quad + \rho(r)[(1-\hat{r})[m_B^{''}\frac{dy_B}{dr} - v_B^{''}(\frac{dy_B}{dr} + \frac{dy_G}{dr})] \\
&\quad + (\frac{1-\lambda_G}{1-\lambda_B} v_G^{''}(\hat{r}\theta_{B+} - \theta_B) + \frac{v_G^{''}}{1-\lambda_B}\hat{f}\bar{U})(\frac{dy_B}{dr} + \frac{dy_G}{dr})], \\
c_{xx}^G \frac{dx_G}{dr} + c_{xy}^G \frac{dy_G}{dr} &= (1-\lambda_B)\rho^{'}(r)(\theta_{G+} - \theta_G), \\
c_{xy}^G \frac{dx_G}{dr} + c_{yy}^G \frac{dy_G}{dr} &= v_G^{''}(\frac{dy_B}{dr} + \frac{dy_G}{dr}) \\
&\quad + \frac{1-\lambda_B}{1-\lambda_G}[v_B^{''}(\frac{dy_B}{dr} + \frac{dy_G}{dr})[\rho(r)\theta_{G+} + (1-\rho(r)\theta_G)] + v_B^{'}\rho^{'}(r)(\theta_{G+} - \theta_G)].
\end{aligned} \tag{A7}$$

The above equations can be written in a matrix form as follows:

$$\mathbf{Nq} = \rho'(r)\mathbf{a}, \tag{A8}$$

where $\mathbf{N}$ is a 4×4 matrix and $\mathbf{q} = (\frac{dx_B}{dr}, \frac{dy_B}{dr}, \frac{dx_G}{dr}, \frac{dy_G}{dr})$. We write $\mathbf{q}$ in the form

$$\mathbf{q} = \rho'(r)\mathbf{N}^{-1}\mathbf{a} \equiv \frac{\rho'(r)}{det(N)}\hat{\mathbf{q}}\mathbf{a}, \tag{A9}$$

where $\hat{\mathbf{q}} := det(N)\mathbf{N^{-1}}$. So

$$det(\mathbf{N}) = A_0 + B_0\frac{1-\lambda_G}{1-\lambda_B} + C_0\frac{1-\lambda_B}{1-\lambda_G}, \tag{A10}$$

where

$$\begin{aligned} A_0 &= H_B H_G + c^B_{xx} H_G[\rho(1-\hat{r})(m^{''}_B - v^{''}_B) + v^{''}_B] \\ &+[c^G_{xx} H_B + \rho c^B_{xx} c^G_{xx}(1-\hat{r}) m^{''}_B] v^{''}_G, \\ B_0 &= c^B_{xx} H_G v^{''}_G[(1-\rho)\theta_B - \rho(\hat{f}\bar{U} + \hat{r}\theta_{B+})], \\ C_0 &= c^G_{xx}[H_B + c^B_{xx}\rho m^{''}_B(1-\hat{r})] v^{''}_B[(1-\rho)\theta_G + \rho\theta_{G+}]. \end{aligned} \tag{A11}$$

The expression for $\hat{q}$ takes the following form:

$$\begin{aligned} \hat{q}_1 = &\frac{1}{(1-\lambda_B)(1-\lambda_G)}[c^G_{xx} c^B_{xy} v^{'}_B(\theta_G - \theta_{G+})(1-\lambda_B)\hat{q}_{11} - [c^G_{xx}\hat{q}_{12} + (c^G_{xy})^2(1-\lambda_G)\hat{q}_{13}]\hat{q}_{14} \\ &+ c^B_{xy}\hat{q}_{15}\hat{q}_{16} c^B_{xy} c^G_{xy}(\theta_G - \theta_{G+})(1-\lambda_B)(1-\lambda_G)\hat{q}_{17}], \end{aligned} \tag{A12}$$

where

$$
\begin{aligned}
\hat{q}_{11} &= \hat{f}\bar{U}\rho v_G^{''} + (1-\lambda_B)v_{BB}^{''}((1-\rho)+\rho\hat{r}) + (1-\lambda_G)v_G^{''}((1-\rho)\theta_B + \rho\hat{r}\theta_{B+}).\\
\hat{q}_{12} &= (1-\lambda_B)^2(c_{yy}^B - \rho m_B^{''}(1-\hat{r}))v_B^{''}((1-\rho)\theta_G + \rho\theta_{G+}).\\
&\quad +(1-\lambda_G)[\hat{f}\bar{U}\rho c_{yy}^G v_G^{''} + (1-\lambda_B).\\
&\qquad [c_{yy}^G\rho(1-\hat{r})(m_B^{''} - v_B^{''}) - \rho m_B^{''}(1-\hat{r})v_G^{''} + c_{yy}^G(v_B^{''} - c_{yy}^B) + c_{yy}^B v_G^{''}]]\\
&\quad +(1-\lambda_G)^2 c_{yy}^G v_G^{''}((1-\rho)\theta_B + \rho\hat{r}\theta_{B+})\\
\hat{q}_{13} &= -\hat{f}\bar{U}v_G^{''} + (1-\lambda_B)[c_{yy}^B - \rho(1-\hat{r})(m_B^{''} - v_B^{''})]\\
&\quad -(1-\lambda_G)v''_G[(1-\rho)\theta_B + \rho\hat{r}\theta_{B+}].\\
\hat{q}_{14} &= \hat{f}\bar{U} + (1-\lambda_B)(1-\hat{r}) + (1-\lambda_G)(\hat{r}\theta_{B+} - \theta_B).\\
\hat{q}_{15} &= (1-\lambda_B)c_{xx}^G v_B^{''}[(1-\rho)\theta_G + \rho\theta_{G+}] + (1-\lambda_G)[H_G + c_{xx}^B v_G^{''}].\\
\hat{q}_{16} &= \hat{f}\bar{U}v_G^{'} + (1-\lambda_B)(1-\hat{r})(m_B^{'} - v_B^{'}) + (1-\lambda_G)v_G^{'}(\hat{r}\theta_{B+} - \theta_B).\\
\hat{q}_{17} &= \hat{f}\bar{U}v_G^{''} + (1-\lambda_B)v_B^{''}[(1-\rho)+\rho\hat{r}] + (1-\lambda_G)v_G^{''}[(1-\rho)\theta_B + \rho\hat{r}\theta_{B+}].
\end{aligned}
\tag{A13}
$$

$$
\hat{q}_2 = c_{xx}^B c_{xy}^G(\theta_G - \theta_{G+})\hat{q}_{21}[1 - \frac{v_G^{'}}{1-\lambda_G}] + \hat{q}_{22}[\frac{c_{xx}^B}{(1-\lambda_B)(1-\lambda_G)}\hat{q}_{23} - \frac{c_{xy}^B}{1-\lambda_G}\hat{q}_{24}], \tag{A14}
$$

where

$$
\begin{aligned}
\hat{q}_{21} &= \hat{f}\bar{U}v_G^{''} + (1-\lambda_G)v_G^{''}[(1-\rho)\theta_B + \rho\hat{r}\theta_{B+}] + (1-\lambda_B)v_B^{''}[(1-\rho)+\rho\hat{r}.\\
\hat{q}_{22} &= (1-\lambda_G)[c_{xx}^G v_G^{''} - H_G] - (1-\lambda_B)c_{xx}^G v_B^{''}[(1-\rho)\theta_G + \rho\theta_{G+}].\\
\hat{q}_{23} &= \hat{f}\bar{U}v_G^{'} + (1-\lambda_G)v_G^{''}(\hat{r}\theta_{B+} - \theta_B) + (1-\lambda_B)(1-\hat{r})(m_B^{'} - v_B^{'}).\\
\hat{q}_{24} &= \hat{f}\bar{U} + (1-\lambda_G)(\hat{r}\theta_{B+} - \theta_B) + (1-\lambda_B)(1-\hat{r}).
\end{aligned}
\tag{A15}
$$

$$
\begin{aligned}
\hat{q}_3 &= \frac{1}{(1-\lambda_B)(1-\lambda_G)}[(\theta_G - \theta_{G+})(1-\lambda_b)\hat{q}_{31} + c_{xy}^G v_B^{'}(\theta_G - \theta_{G+})\hat{q}_{32}\\
&\quad -c_{xy}^G\hat{q}_{33}[c_{xy}^B(1-\lambda_B)\hat{q}_{34} + c_{xx}^B\hat{q}_{35}],
\end{aligned}
\tag{A16}
$$

where

$$\begin{aligned}
\hat{q}_{31} &= (1-\lambda_B)^2[H_B + c_{xx}^B \rho m_B^{''}(1-\hat{r})]v_B^{''}[(1-\rho)\theta_G + \rho\theta_{G+}] \\
&-(1-\lambda_G)(\hat{f}\bar{U}\rho c_{xx}^B c_{yy}^G v_G^{''} - (1-\lambda_B)[H_B(v_G^{''} - c_{yy}^G) \\
&+\rho m_B^{''}(1-\hat{r})(v_G^{''} - c_{yy}^G) - c_{yy}^G((1-\rho) + \rho\hat{r}v_B^{''})] \\
&-(1-\lambda_G)^2[c_{xx}^B c_{yy}^G v_G^{''})((1-\rho)\theta_B + \rho\hat{r}\theta_{B+})]. \\
\hat{q}_{32} &= \hat{f}\bar{U}c_{xx}^B v_G^{''} \\
&+(1-\lambda_B)[(c_{xy}^B)^2 + c_{xx}^B(v_B^{''} - c_{yy}^B) + \rho c_{xx}^B(1-\hat{r})(m_B^{''} - v_B^{''})] \\
&+(1-\lambda_G)c_{xx}^B v_G^{''}((1-\rho)\theta_B + \rho\hat{r}\theta_{B+}). \\
\hat{q}_{33} &= (1-\lambda_B)v_B^{''}((1-\rho)\theta_G + \rho\theta_{G+}) + (1-\lambda_G)v_G^{''}. \\
\hat{q}_{34} &= \hat{f}\bar{U} + (1-\lambda_B)(1-\hat{r}) + (1-\lambda_G)(\hat{r}\theta_{B+} - \theta_B). \\
\hat{q}_{35} &= \hat{f}\bar{U}v_G^{'} + (1-\lambda_B)(1-\hat{r})(m_B^{'} - v_B^{'}) + (1-\lambda_G)v'_G(\hat{r}\theta_{B+} - \theta_B).
\end{aligned} \tag{A17}$$

$$\hat{q}_4 = (\theta_G - \theta_{G+})(1-\lambda_B)(\frac{c_{xx}^G v_B^{'}}{1-\lambda_G} - c_{xy}^G)\hat{q}_{41} + c_{xx}^G \hat{q}_{42}[c_{xy}^B \hat{q}_{43} - \frac{c_{xx}^B c_{xx}^G}{1-\lambda_B}\hat{q}_{44}], \tag{A18}$$

where

$$\begin{aligned}
\hat{q}_{41} &= -\frac{1-\lambda_G}{1-\lambda_B}c_{xx}^B v_G^{''}[(1-\rho)\theta_B + \rho\hat{r}\theta_{B+}]. \\
\hat{q}_{42} &= -v_G^{''} - \frac{1-\lambda_B}{1-\lambda_G}v_B^{''}[(1-\rho)\theta_G + \rho\theta_{G+}]. \\
\hat{q}_{43} &= \hat{f}\bar{U} + (1-\lambda_B)(1-\hat{r}) + (1-\lambda_G)(\hat{r}\theta_{B+} - \theta_B). \\
\hat{q}_{44} &= \hat{f}\bar{U}v_G^{'} + (1-\lambda_B)(1-\hat{r})(m_B^{'} - v_B^{'}) + (1-\lambda_G)v_G^{'}(\hat{r}\theta_{B+} - \theta_B).
\end{aligned} \tag{A19}$$

These expressions are extremely complex, and to get some insight, we make the following simplifying assumptions. We assume that the distribution of shocks are uniformly distributed and are the same for the two sexes, so that $F = G$. This implies that $\theta_B = \theta_G = 1$, while $\theta_{B+} = 1$ and $\theta_{G+} = \hat{r}$. Second, we assume a cost function is sufficiently supermodular, so that $c_{xy} > 0$ and is large enough; however, we do not want to make to too large, since we would not have interior solutions for investments by the two sexes. Accordingly, we assume the following cost function:

$$c(x, y) = \frac{\gamma}{3}(x + y)^3 + \frac{1 - \gamma}{2}(x^2 + y^2),$$

with $\gamma = 0.29$ the chosen value.

We also assume quadratic utility for the public good:

$$v_G(y) = v_B(y) = ay - by^2,$$

where $a = 1, b = 0.1$ are the numerical values.

The bargaining power parameters are taken to be $\lambda_B = 0.8$, $\lambda_G = 0.2$. The sex ratio in the market with imbalances, i.e. $\hat{r}$ is set to 0.8. Finally, we take $\hat{f}U = 5$.

Note that we have interior solutions for all feasible values of r, i.e. those in the interval $[\hat{r}, 1]$.

Observe that the investments in the public good for boys are always decreasing in r, i.e. increased marriage market competition induces more investment in the public good, in line with our empirical results. The effects on private good investments by boys are ambiguous—these investments are decreasing in r for r values near 1, but are increasing in r for smaller values of r.

Investments by girls follow the opposite pattern from those by boys. Investments in the public good are consistently increasing in r. Appendix Figure A6 depicts the separate graph of x_G, girls' investments in the private good. These are decreasing in r for r values close to 1, but are increasing in r when the sex ratio is more unbalanced.