Wealth, Marriage, and Sex Selection*

Girija Borker[†] Jan Eeckhout[‡] Nancy Luke[§] Shantidani Minz[¶] Kaivan Munshi[‖]
Soumya Swaminathan[**]

July 5, 2017

Abstract

Sex selection continues to be a serious problem in India, despite many decades of economic progress. It is widely believed that large marriage payments to the groom's family or dowries are the main cause of sex selection in India, especially among the wealthy. Our theoretical model clarifies this argument, showing that the root cause of sex selection is not dowries but specific frictions in the marriage market, which arise because of the structure of the marriage institution. The model predicts that *relatively* wealthy households within castes, which define independent marriage markets in India, will be more likely to practice sex selection. This prediction is tested with unique data we have collected, covering the entire population of 1.1 million individuals residing in half a rural district in South India. We find that the variation in sex ratios within castes in this single district is comparable to the variation across all districts in the country. Given that the marriage market is organized the same way in all castes, sex selection may be a more pervasive problem than is currently believed. Estimation of the model's structural parameters allows us to quantify the impact of alternative policies, which operate through the marriage market, to reduce sex selection. We find that cash transfers to adult women, which forward-looking parents will take into account when making sex selection decisions, are substantially more effective than transfers to their parents when they are children.

Keywords. Sex Selection. Marriage Market. Caste. Assortative Matching. Wealth Distribution. Inequality.
JEL. J12. J16. D31. I3.

*Akhil Lohia provided exceptional research assistance, especially with the computation. We are grateful to Debraj Ray for detailed notes, which helped us better understand the model. Research support from the National Institutes of Health through grant R01-HD046940, the Keynes Fund at the University of Cambridge, Cambridge-INET, ERC Grant 339186, and grant ECO2015-67655-P is gratefully acknowledged. We are responsible for any errors that may remain.

[†]Brown University
[‡]University College London and UPF-ICREA-GSE
[§]Pennsylvania State University
[¶]Christian Medical College, Vellore
[‖]University of Cambridge
[**]Indian Council of Medical Research

1 Introduction

Sex selection through female feticide, infanticide, or neglect is an extreme manifestation of gender discrimination. Amartya Sen brought sex selection to public attention over 25 years ago when he famously claimed that 100 million women were "missing" in Asia (Sen (1990)). Since that time, countries like China and India have made tremendous economic progress, and we would expect this progress to be accompanied by greater gender equality. However, sex ratios at birth and in childhood remain skewed in favor of boys and, in many settings, have worsened. In India, a problem that was historically confined to landowning castes in the North has spread throughout the country, down to the lowest castes.

It is widely believed that large marriage payments to the groom's family, or dowries, are the main cause of son preference in India (e.g. Das Gupta (1987), Basu (1999), Diamond-Smith, Luke, and McGarvey (2008)).[1] There is also a common perception that the wealthy are more likely to practice sex selection because they must pay higher dowries (e.g. Murthi, Guio, and Drèze (1995)). However, these arguments have limitations. Wealthy girls match with wealthy boys who provide them with greater resources for consumption during marriage. If girls' parents internalize these benefits, then it is not obvious that having a girl is especially disadvantageous for wealthy parents, despite the fact that they must pay higher dowries.

In this paper, we show how the marriage institution in India contributes to sex selection. In particular, we develop a model that links wealth to sex selection through the marriage market. Dowries and sex selection are endogenously determined in the model, and although these variables are positively correlated in equilibrium, dowries do not causally determine sex selection. The root cause of sex selection is specific frictions in the marriage market, which arise because of the structure of the marriage institution in India. These frictions do not affect all households equally. In particular, sex selection is shown to be increasing with *relative* wealth within the relevant marriage market.

Parents are perfectly altruistic in the model and want to share their wealth equally with children, regardless of their gender. Because girls leave their natal homes when they marry in India, parents must use the dowry as a mechanism to share wealth with their daughters. However, the dowry is given to the in-laws and so the daughter receives only a fraction of its value, by way of the transfer she receives from her husband's parents. This seepage in the bequest is one friction in the marriage market in our model. The second friction is the social norm that all girls must marry, which lowers their outside options and divides the surplus from marriage in favor of the boy's side. Although the marriage market clears instantaneously (without search frictions), these institutional frictions leave parents worse off with girls than with boys, resulting in sex selection at every wealth level,

[1] Another common explanation for sex selection, which we account for in the empirical analysis, is that parents want at least one boy.

even when parents do not have an intrinsic preference for sons.

In addition to its role as a bequest, the dowry also functions as a price to clear the marriage market.[2] This results in Positive Assortative Matching on wealth in equilibrium because wealthier parents are willing to pay a higher dowry to get their daughters married into wealthier families, where they will consume at a higher level. Once there is sex selection, the distribution of wealth on the two sides of the market becomes endogenous. The sex selection decision, the wealth distribution (which determines the pattern of matching) and the marriage price or dowry must be determined simultaneously. This is a challenging problem, which has not been previously solved in the matching literature. We are nevertheless able to show that sex selection is increasing with wealth. Although analytically deriving the actual solution is complicated because of the endogenous wealth distribution, the intuition for this result is straightforward. Suppose that sex selection and the accompanying shortage of girls is constant at every wealth level. The wealthiest boys still match with the wealthiest girls, but because of the shortage of girls all other boys marry less wealthy girls. Given assortative matching, this wealth-gap increases as we move down the wealth distribution, making it more attractive for poorer parents to have a girl versus a boy.[3] As a result, sex selection will decline endogenously in equilibrium as we move down the wealth distribution.[4]

We test the key theoretical prediction that sex selection is increasing with relative wealth with unique data we have recently collected as part of the South India Community Health Study (SICHS), which covers the entire rural population of 1.1 million individuals residing in half of Vellore district in the South Indian state of Tamil Nadu. The analysis makes use of two components of the SICHS; a census of all 298,000 households drawn from 57 castes residing in the study area, completed in 2014, and a detailed survey of 5,000 representative households, completed in 2016. In general, extremely large data sets are needed to detect sex selection with the required level of statistical confidence. The SICHS census is the only data set we are aware of that is large enough to estimate the relationship between family wealth and sex selection *within* castes or *jatis*, which define independent marriage

[2]This dual role of the dowry as a bequest and a price distinguishes our model from previous models of marriage with dowries; e.g. Botticini and Siow (2003), Anderson and Bidner (2015). The advantage of our specification is that it explains the coexistence of sex selection and dowries (despite the fact that girls are on the short side of the market).

[3]To clarify this argument, suppose that there are 100 wealth levels and two boys and one girl at each wealth level. Except for the number of boys and girls, the wealth distribution is the same and uniform on [0,100]. Then under positive assortative matching, one of the boys with wealth 100 marries the girl with wealth 100 and the other boy marries the girl with wealth 99, the boys with wealth 99 marry the girls with wealth 98 and 97, and so on, until the last boy to be matched, with wealth 50, marries the girl with wealth 1. The key insight is that the wealth-gap is increasing as we move down the wealth distribution: at the top (100,100) there is no wealth gap and at the bottom (50,1) the wealth gap is 49.

[4]Previous models (e.g. Edlund (1999) and Bhaskar (2011)) also generate the result that sex selection is increasing with wealth. However, these models, in which son preference is exogenously determined, generate the prediction that there will be bride-price in equilibrium; i.e. payments from boys to girls because girls are on the short side of the market. While this prediction is consistent with large payments from the groom's side to the bride's side in countries like China, it is at odds with the payments in the opposite direction that are observed in India. Groom-price or dowries arise naturally in our model, in which sex selection is generated by particular features of the marriage institution, despite the fact that there is a shortage of girls.

markets in India.[5]

Our main empirical finding is that the probability that a child (aged 0-6) is a girl is decreasing as we move up the wealth distribution within castes. This result is robust to alternative specifications, samples, and measures of wealth. It is also obtained caste by caste, across the social spectrum. Sex selection is a relatively recent phenomenon in South India, coinciding with the decline in close-kin marriage and the emergence of marriage markets in the early 1980's. Based on population census statistics prior to that period, the natural child (aged 0-6) sex ratio in South India is 102.5 boys per 100 girls. The current child sex ratio in the study area, obtained from the SICHS census, is 108, which is clearly biased, but not exceptional for South India or the country as a whole today. Across wealth classes *within* castes, however, the sex ratio ranges widely, from a slight surplus of girls; i.e. below 100, to as high as 117. To put these numbers in perspective, a sex ratio of 120 is considered to be severely biased and only a handful of districts in the country have sex ratios exceeding that level. Thus, the magnitude of the variation in sex ratios within castes in a single (unexceptional) district is comparable to the variation across all districts in the country.

The marriage market is organized in the same way in all castes. The positive relationship between relative wealth within castes and sex selection that we have uncovered is thus likely to apply more widely. Given these new findings on the severity and the extent of the sex selection problem, the design of policies to address it becomes especially important. In order to evaluate the impact of alternative policy interventions, we structurally estimate the marital sorting model that we laid out above using the SICHS data. The exogenous determinants of the equilibrium allocation are (i) the parameter governing the seepage in the bequest from the girl's parents, which arises because part of the dowry is siphoned off by the boy's family, and (ii) the parameter governing the cost of sex selection. For a given pattern of marital matching, which is defined by the wealth-gap between girls at every wealth level and their partners, the seepage in the bequest together with the social norm that all girls must marry will determine the utility differential between having a girl and a boy. The cost of sex selection parameter maps this utility differential into the sex ratio, which in turn determines the distribution of wealth for boys and girls and the associated pattern of marital matching, to close the model. The two structural parameters are estimated by matching the sex ratios predicted by the model at each wealth level in each caste to the actual sex ratios.

Given these parameter estimates, we can use the model to evaluate and interpret the impact of different policy interventions, which will work through the marriage market equilibrium to change patterns of sex selection. The dowry is effectively the price for a boy, and so one policy lever to reduce the (excess) demand for boys would be to tax dowries.[6] Although such a policy has not been

[5]The basic rule in Indian society is that individuals must marry within their caste (or its non-Hindu equivalent kinship group). Recent genetic evidence, discussed in the next section, indicates that Indians have been marrying in this way for over 2,000 years.

[6]Although dowries are illegal in India, parents circumvent the law by claiming that the dowry is a gift to the newly married couple. A tax on the dowry could thus be implemented as a gift tax.

implemented, the central government as well as many state governments have introduced welfare schemes with the stated objective of reducing sex selection. These schemes typically consist of a cash transfer to parents, conditional on having a girl, or a direct transfer to the girl (when she becomes an adult). In some cases, the schemes are restricted to households below an income ceiling. The main findings of the counter-factual policy simulations are (i) that a tax on the dowry will have both positive and negative effects at different points in the wealth distribution, (ii) that interventions targeted at specific (low income) households will have unintended pecuniary externalities on other caste members by changing the equilibrium marriage price, possibly *increasing* the overall bias in sex ratios, and (iii) that direct transfers to girls when they are adults are much more effective than transfers to their parents when they are children. The latter result suggests a promising way forward, which we examine in greater detail in the concluding section.

2 Institutional Setting

2.1 Marriage and Sex Selection in India

Abnormally high male-female sex ratios were observed in certain parts of India, particularly the Northwest, as far back as the first British-Indian census in 1871. This problem appears to have intensified with economic development, spreading throughout the country in recent decades (Basu (1999), Srinivasan and Bedi (2009)). To quantify the magnitude of the problem, it is useful to examine changes in sex ratios over time and across regions of India. In this paper, we define South India by the states of Tamil Nadu, Karnataka, Andhra Pradesh, and Maharashtra, excluding Kerala which is an outlier on many socioeconomic indicators. The natural child (aged 0-6) sex ratio is 102.5 males per 100 females.[7] The child sex ratio for the country as a whole, based on the 2011 population census is substantially higher, at 109. Starting with virtually no sex selection as recently as the 1970's, South India is not far behind at 108. More recent data from the Sample Registration System (SRS) indicate that sex ratios in South India may now have converged with the all-India average.[8]

The emergence of sex selection in South India in the 1980's is especially useful in helping us understand why sex selection is linked to marriage in India and why this problem appears to intensify with economic development. Although Indians have married within their castes or *jatis* for centuries, marriages in South India were, in addition, traditionally between close-kin (Dyson and Moore (1983); The most preferred match for a girl was her mother's younger brother or, if

[7]The natural sex ratio at birth ranges from 103 to 106 (Guilmoto (2009)). Subsequent mortality favors girls, but the natural sex ratio for 0-6 year olds is not directly available from existing studies. As discussed below, sex selection appears to have commenced in South India concurrently with the emergence of marriage markets and associated dowries in the 1980's. Assuming that the sex ratios for South India in the 1961 and 1971 censuses are thus unbiased, the natural child sex ratio is 102.5.

[8]The Sample Registration System provides the Sex Ratio at Birth (SRB) at the state level over the 2004-2013 period. Most sex selection at this point in time in India occurs prior to birth (Srinivasan and Bedi (2009)) and so the SRB is a good proxy for the child sex ratio.

he was unavailable, one of her mother's brothers' sons (Kapadia (1995)). Given that marriage in India is patrilocal, with girls moving into their husbands' homes, this meant that girls were traded between families (or dynasties) that were related through marriage from one generation to the next. Family sizes were historically large and so each dynasty maintained multiple long-term marital relationships.

Upon marriage, the girl typically moved into her grandfather's or a maternal uncle's home. Given the extremely close pre-existing relationship between the girl's natal family and her husband's family, the two families effectively functioned as a cooperative unit. There were no major payments at the time of marriage, just a ritual gift or *stridhan* from the groom's side to the girl (Anderson (2007), Srinivas (1989)). Having a girl did not put parents at a disadvantage in this cooperative arrangement and thus there was no sex selection.

Caldwell, Reddy, and Caldwell (1983) and Srinivas (1984) attribute the demise of this system to economic development and the resulting changes in wealth within castes. Families that had traded girls over many generations were no longer matched with respect to wealth. Close-kin marriage declined (Caldwell, Reddy, and Caldwell (1983), Kapadia (1993)) and a marriage market consequently emerged to match unrelated families within the caste, with a marriage price or dowry clearing the market.[9] Marriage markets have always existed within castes in North India, with dowries being used to clear these markets. Although this is a relatively recent phenomenon in South India, dowries in that region are now as high as they are in the North (Caldwell, Reddy, and Caldwell (1983), Srinivas (1984), Rahman and Rao (2004), Anderson (2007)). A practice that was formerly confined to the upper castes has also spread across the caste distribution (Bhat and Halli (1999)).[10]

The widespread emergence of dowries in South India in the early 1980's coincided with the onset of sex selection. Although it is tempting to conclude from the temporal correlation that dowries caused the sex selection, the root cause of sex selection in our model is frictions in the marriage market that emerged with the dowry, which are, in turn, determined by particular features of the marriage institution in India.

2.2 Marriage and Sex Selection in the Study Area

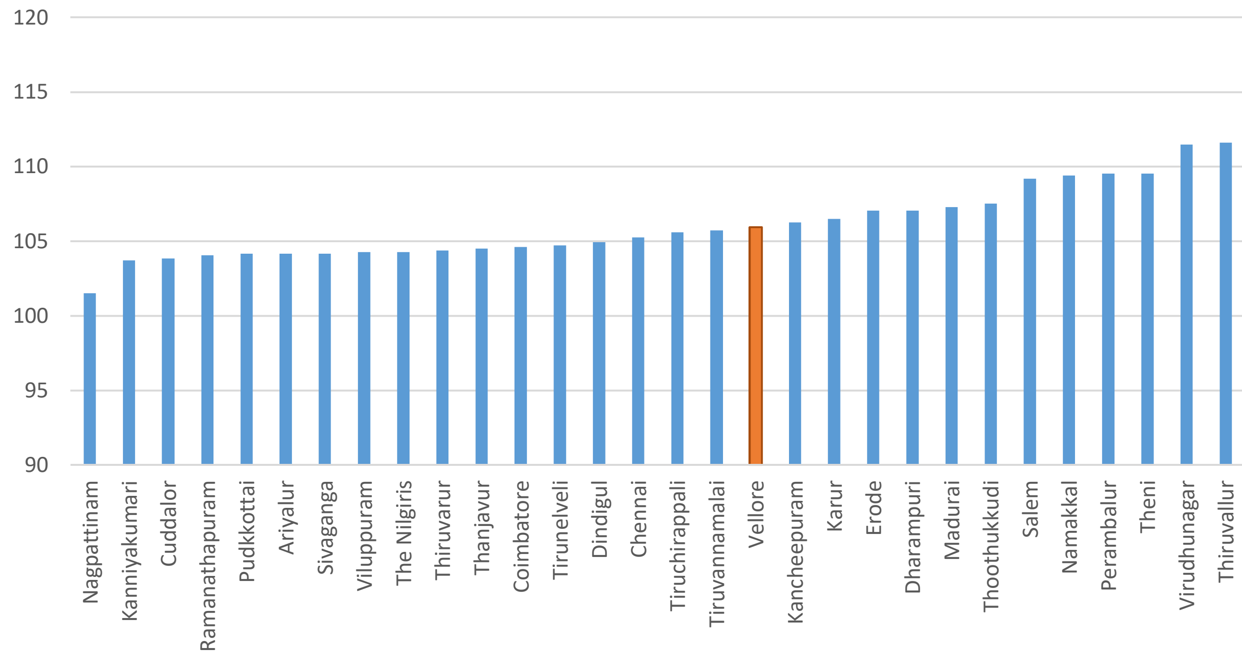
The data that we use for the analysis in this paper were collected as part of the South India Community Health Study (SICHS), which covers a rural population residing in half of Vellore district in the South Indian state of Tamil Nadu. We will see below that the SICHS study area is broadly representative of rural Tamil Nadu and rural South India with respect to key socioeconomic

[9] In related research, Anderson (2003) links dowry inflation to economic development and increased income inequality on the male side of the marriage market.

[10] Dowries were first observed among Brahmins in Madras in the 1930's. They had spread to provincial towns, but continued to be restricted to the Brahmin caste, by the 1960's (Srinivas (1984)). By the 1980's, the practice of dowry was observed even among the lower castes in South India (Caldwell, Reddy, and Caldwell (1983)).

and demographic indicators. Vellore district is also representative of Tamil Nadu state with respect to child sex ratios. Figure 1 reports child (aged 0-6) sex ratios, by district, in the most recent 2011 population census round. The sex ratio in Vellore district at 106 boys per 100 girls is just about the Tamil Nadu average of 106 and slightly lower than the South Indian average of 108.

Figure 1: **Sex Ratio (Aged 0-6) by District in 2011**



Source: National Commission for Women, India

There is nothing exceptional about Vellore district with respect to either socioeconomic characteristics or sex ratios. This is also true of our study area within the district where, based on the SICHS census, the sex ratio is 108. None of the castes in Tamil Nadu that have been traditionally associated with severe sex selection are present in the study area. Nevertheless, we uncover a strong and robust relationship between household wealth and sex selection, within castes. These results are obtained for castes across the social spectrum and are tied in our analysis to frictions in the marriage market that operates in the same way within each caste.

The following features of the marriage institution in India are relevant for our analysis: (i) marriages are endogamous, matching individuals almost exclusively within their caste or *jati*, (ii) marriages are patrilocal, with women moving into their husbands' homes, (iii) marriages are arranged by the parents and relatives of the groom and bride, with family wealth being a major consideration, (iv) marriages involve a transfer payment or dowry from the bride's family to the groom's family, and (v) despite the cost (largely associated with the dowry) of marrying a daughter, the social norm is that all girls must marry.

Some of these features of the marriage institution have remained stable over time. For example, recent evidence from nationally representative surveys such as the 1999 Rural Economic Development Survey (REDS) and the 2005 India Human Development Survey (IHDS) establishes that over 95% of Indians marry within their caste. Complementary genetic evidence indicates that these patterns of endogamous marriage have been in place for over 2,000 years (Moorjani, Thangaraj, Patterson, Lipson, Loh, Govindaraj, Berger, Reich, and Singh (2013)). The Indian population today is divided into 4,000 distinct genetic groups, each of which is a caste (or its non-Hindu equivalent kinship group) and within which an independent marriage market is organized. Marriages continue to be arranged, and early and universal marriage remains the norm for females, especially in rural India (Caldwell, Reddy, and Caldwell (1983), Arnold, Choe, and Roy (1998), Bhat and Halli (1999), Basu (1999), Das Gupta, Zhenghua, Bohua, Zhenming, Chung, and Hwa-Ok (2003)).[11] What has changed are certain features of marriage in South India; marriage to close-kin has declined and, in parallel, a marriage market within each caste has emerged, with dowries being used to clear the market.

The analysis in this paper makes use of two components of the SICHS; a census of all households and a detailed survey of 5,000 households who are representative of the castes in the study area.[12] The survey collected information on key aspects of the marriage institution: (i) whether marriage was within the caste, (ii) whether marriage was between close-kin, and (iii) whether the marriage was arranged. This information was collected from the male primary earner for his own marriage and for the marriages of his children in the five years preceding the survey.

Table 1: **Marriage Patterns**

Generation	Parents	Children	
		Males	Females
	(1)	(2)	(3)
Same caste	0.97	0.95	0.95
Related	0.48	0.35	0.35
Arranged	0.86	0.80	0.88
Observations	3,524	421	611

Source: SICHS household survey

Table 1 provides information on marriages over the two generations based on data from the household survey.[13] In line with nationally representative survey evidence and genetic evidence for

[11] According to the 2005-06 DHS, less than 2% of women in Tamil Nadu are never married by age 35.

[12] The sampling frame for the household survey included all ever-married men aged 25-60 in the SICHS census plus divorced or widowed women with "missing" husbands who would have been aged 25-60, based on the average age-gap between husbands and wives. The sample was subsequently drawn to be representative of each caste in the study area, excluding castes with less than 100 households in the census.

[13] The larger number of marriages for daughters versus sons in the last 5 years is because girls marry younger than boys in India. Given that the fathers are aged 25-60, there are more girls of marriageable age in the households in

the country as a whole, 97% of the parents and 95% of the children married within their caste. The incidence of close-kin marriage declines, in line with the general trend in South India, from 48% in the parents' generation to 35% in the current generation. However, most marriages continue to be arranged.

Table 2: **Dowries and Marital Matching**

Sex of the child	Males (1)	Females (2)
Mean dowry (in thousand Rupees)	147.78	199.72
Mean fraction of annual income	3.16	4.15
Partner's parental household		
Wealthier	0.09	0.18
Same wealth	0.62	0.64
Less wealthy	0.29	0.17
Kolmogorov-Smirnov test of equality	P-value = 0.001	
Observations	421	611

Source: SICHS household survey; sample: marriages of children in the last 5 years.

Table 2 reports the dowry in levels and as a fraction of the household's annual income, for the marriages of the children in the household that took place in the last five years. A dowry was paid by the girl's side to the boy's side in all marriages. The dowry amount is computed by summing up the monetary value of gifts, such as household items, vehicles, and gold, as well as the cost of the wedding celebration.[14] The annual income is measured by the profit in the past year from land owned, leased, or rented plus profit from livestock plus the wage earnings of all adult members. In line with past studies; e.g. Rao (1993b), Jejeebhoy and Sathar (2001), and Rahman and Rao (2004), the dowry is three to four times the household's annual income on average, which is an enormous sum in an economy where access to market credit is severely restricted.[15]

Notice that dowries paid by girls are larger than dowries received by boys. As reported in Appendix Table A1, this gender-gap in the dowry is statistically significant, conditional on household income and even with household fixed effects. One explanation for the gender-gap is reporting bias,

our sample.

[14]The list of items for the dowry include bed, bureau, kitchen utensils (bronze and stainless steel), grinder, mixer, refrigerator, TV, microwave, washing machine, silk saris, groceries, motorcycle, bicycle, car, gold jewelry (in grams), and cash (in Rupees).

[15]Most households will receive support from their close relatives and other caste members to pay the dowry. Munshi and Rosenzweig (2016) use data from the Rural Economic and Development Survey (REDS) to document that gifts and loans within the caste are the primary source of support for meeting major contingencies, including marriage, in rural India. Note that the amount of money that must be raised is large, even by the standards of a developed economy. For example, the maximum amount that banks lend in developed economies for the purchase of a home is typically 2.5 to 3 times the annual household income.

with respondents inflating the amount they gave and under-reporting the amount they received. A second explanation is hypergamy; i.e. that girls marry wealthier boys on average. If there is an equal number of boys and girls and an equal distribution of wealth on both sides of the marriage market, then girls cannot marry up on average. However, hypergamy emerges naturally in a marriage market with sorting on wealth, as a consequence of sex selection; given the shortage of girls, wealthier boys are forced to match with poorer girls. Hypergamy has a long tradition in North India, especially among the upper castes (Bhat and Halli (1999)). Hypergamy has also been associated with the emergence of dowry in South India (Caldwell, Reddy, and Caldwell (1983), Srinivas (1984)). These are all settings with sex selection. However, previous studies have failed to make the connection between hypergamy and sex selection.

Table 2 provides evidence supportive of hypergamy based on our household survey data. The survey respondents were asked whether their child's spouse's family had the same wealth, more wealth, or less wealth than their own. These are coarse categories and the majority of marriages, for sons and daughters, are reported to be with families of equal wealth. However, the respondents are more likely to report that their daughters married up in wealth than their sons. Conversely, they are more likely to report that their sons married down in wealth than their daughters. The Kolmagorov-Smirnov test easily rejects the null hypothesis that the distribution of responses is equal for sons and daughters.[16]

3 A Theory of Wealth, Marriage, and Sex Selection

3.1 Model Set Up

Population. Consider a population of families with measure 2. We assume that a family consists of one parent and one child. The gender of the parent is irrelevant. The gender of the child is the purpose of our analysis. Under natural circumstances, without sex selection, a child is born a boy or a girl with equal probability, and the distribution of children would each have measure one. Families are indexed by their wealth z which is distributed according to the measure $\Gamma(z)$ on $[\underline{z}, \overline{z}]$, with $\Gamma(\overline{z}) = 2$. Denote the boy's family wealth by x and the girl's by y. The measure of families with boys and with girls will be endogenous, as will be the distribution of wealth. We denote the wealth distribution of families with boys by $F(x)$ and with girls by $G(y)$. Under natural circumstances without selection and with equal probability of having a boy or a girl, the wealth distribution of boys is identical to that of girls: $F(\cdot) = G(\cdot) = \frac{1}{2}\Gamma(\cdot)$.

Preferences, Payoffs and Consumption. Denote the wealth-contingent consumption of parents by C_x, C_y and that of the children by c_x, c_y. All individuals have logarithmic preferences

[16]Multinomial logit estimates, with same wealth as the reference category, yield the same result in Appendix Table A2; girls are significantly more (less) likely to marry up (down) than boys. The child's education significantly increases the probability that the spouse's family is wealthier, but does not change the matching on wealth (by gender).

over consumption, and we assume that families maximize the sum of their members' utilities $U = \log(C_i) + \log(c_i), \forall i = \{x, y\}$.[17] Denote the maximized utility of the groom's family with income x marrying a bride with income y by $u(x, y)$ and the associated utility for the bride's family by $v(x, y)$.

The Marriage Institution. The model incorporates the key features of the marriage institution listed above. Castes form independent marriage markets and we can think of the model as describing one such market. Marriages are arranged, with family wealth being the major consideration when forming a match. The additional institutional feature that is especially relevant for the model is that marriage in India is patrilocal; i.e. women move into their husbands' homes. Patrilocal marriage has benefits and costs. The benefits are that if a girl lands a well off boy, she will get to consume a fraction of the wealth her future husband receives as a transfer (or bequest) from his parent. We denote the transfer by t. While parents are altruistic towards their own children, they are not towards their children's marriage partners. Therefore, the boy's parent would like to give nothing to the bride, but given that the husband and wife live together, the groom obtains a fraction $\alpha \geq 1/2$ of the transfer, and the bride cannot be stopped from consuming a fraction $1 - \alpha$ of what her husband receives.[18] The groom's (or the bride's) parent cannot earmark what fraction of the transfer goes to the bride, and thus α is exogenously given. The cost of patrilocal marriage to the bride is that the boy's parent is only willing to accept the match if the girl's parent pays a dowry d.

When a match is arranged between a boy with family wealth x and a girl with family wealth y, the dowry must be large enough that the maximized utility of the boy's family, $u(x)$, exceeds its outside option, which is for the boy to remain single. If the boy stays single, the family wealth is divided equally between the parent and the child, given altruistic preferences, resulting in a level of utility, $2\log\left(\frac{x}{2}\right)$. Non-participation in the marriage market is not an option for girls, however, given the social norm that all girls must marry and the resulting extreme disutility from staying single.

Matching in this marriage market is frictionless with the transfer between the bride and the groom's family d determined competitively. We denote the equilibrium allocation by $\mu(y)$, i.e. the family wealth of the groom who is married to a bride with family wealth y is $x = \mu(y)$. The timing of the decisions is, first, participants in the marriage market choose their best partner given a

[17] Equivalently, parents have altruistic preferences over the utility of their children. The assumption that preferences are logarithmic is broadly consistent with Euler equation estimates of the inter-temporal elasticity of substitution; e.g. Attanasio and Weber (1993), Blundell, Browning, and Meghir (1994). Although there are only two generations in the model, we can interpret the weight on the child's consumption utility as reflecting the cummulative (discounted) weight on all future generations. It is possible that the parent's and the child's consumption would then no longer receive equal weight, but this extension to the model would not change the results that follow.

[18] If future generations are accounted for in the parent's utility function, then the boy's parent could place positive weight on his daughter-in-law's consumption utility (to the extent that this affects the utility of her offspring). However, as long as he places more weight on his son's utility than his daughter-in-law's utility, the results that follow will remain unchanged.

"Walrasian" schedule of dowries, and the marriage market clears with a resulting equilibrium price d. Then, the parent of the boy chooses the transfer t.[19]

Observe that in this setup, the dowry plays a dual role: it serves as a price in the marriage market (Becker (1973)) as well as a bequest from the girl's parent to her (Botticini (1999)). By paying a higher dowry, the bride's parent is able to arrange a match with a wealthier groom, who will receive a larger transfer from his parent, resulting in higher consumption for his wife. This dual role for the dowry distinguishes our model from existing models of marriage with dowries. In Botticini and Siow (2003) the marriage market clears by wealth matching between brides and grooms, and dowries serve only as a bequest. In Anderson and Bidner (2015), the dowry serves both roles, but two separate instruments are available. The advantage of our parsimonious specification, with a single marriage payment instrument, is that it will allow us to solve simultaneously for dowries and sex selection.[20] The bequest motive also explains why marriage payments flow from the girl's side to the boy's side, despite the fact that with sex selection there is a shortage of girls in equilibrium.

THE SEX SELECTION TECHNOLOGY. We assume that parents who are expecting a girl can replace her with a boy (with probability one) at a utility cost k, which is distributed according to the cumulative density function $H(k)$.[21] k incorporates the monetary cost, which is relatively small, and the more important ethical cost of sex selection. We assume that k is uncorrelated with wealth and is bounded below at zero.

CONSUMPTION. Given the setup described above, the consumption of all agents (parents and children) of a married groom-bride pair (x, y) can be written as:

$$\begin{aligned} c_x &= \alpha t \\ C_x &= x - t + d \\ c_y &= (1-\alpha)t \\ C_y &= y - d. \end{aligned} \tag{1}$$

[19]The sequential nature of the model highlights a key imperfection in the marriage market, which is that the boy's parent cannot credibly commit to transferring a specific amount, which is agreed upon *ex ante* with the girl's parent, directly to his daughter-in-law. The resulting seepage in the dowry exacerbates sex selection in our model. Note that seepage would occur even if dowries and transfers were determined simultaneously.

[20]Following common convention, we refer to the marriage transfer as the "dowry" throughout the paper. The technically more accurate terminology is that the price component of the transfer is the groom-price and the bequest component is the dowry (Anderson (2007)).

[21]In reality, the decision is more subtle. First, if parents use sex selective abortion rather than infanticide or neglect to eliminate unwanted girls, then all parents who anticipate that they will make this decision must bear the *ex ante* cost of sex determination. Second, even if parents do eliminate a girl, there is no guarantee that the next pregnancy will result in a boy. There is thus a stochastic element to the cost of sex selection that we abstract from in our modeling choice.

3.2 Analytical Solution and Results

MATCHING. We solve the model backwards. For any match between a girl's family with wealth y and a boy's family with wealth x, and given a dowry d, the boy's parent will choose a transfer t that maximizes his family's utility. The maximized utility of the boy's family can be written as:

$$u(x, d) = \max_t \{\log(x - t + d) + \log(\alpha t)\}, \tag{2}$$

which is independent of y, conditional on the dowry d. The first order condition from this maximization problem implies that:

$$t = \frac{x + d}{2}. \tag{3}$$

Given the sequence of decisions, the dowry d is determined competitively in the marriage market, together with the equilibrium matching pattern, $x = \mu(y)$, taking the preceding optimal transfer t as given. In competitive equilibrium, the allocation must be optimal for each agent and the market must clear. In the marriage market, we derive conditions for optimality on the girl's side, taking as given the maximized utility on the boy's side, $u(x)$, for each wealth level. Notice that u is now a function of the boy's family wealth x alone because once the marriage price d has been determined in equilibrium, there will be a distinct price for each wealth level.

To derive the optimality condition, it will be convenient to express d and t as functions of u. Substituting the expression for t from equation (3) in the boy's family's utility function, we obtain:

$$u = \log\left(\frac{x + d}{2}\right) + \log\left(\alpha\left(\frac{x + d}{2}\right)\right) = \log\left(\alpha\left(\frac{x + d}{2}\right)^2\right). \tag{4}$$

This permits us to write the dowry and the transfer as a function ψ of the utility u obtained by the boy's family, $t = \psi(u)$ and $d = 2\psi(u) - x$, where

$$\frac{x + d}{2} = \sqrt{\frac{e^u}{\alpha}} = \psi(u). \tag{5}$$

A girl's family with wealth y will take the hedonic Walrasian price schedule, $u(x)$, as given when choosing the partner with wealth x that maximizes its utility,

$$v(x, y, u) = \log(x + y - 2\psi(u)) + \log((1 - \alpha)\psi(u)). \tag{6}$$

This is a matching problem with Imperfectly Transferable Utility (ITU) first analyzed in Legros and Newman (2007). The first order condition to this problem satisfies

$$v_x + v_u u' = 0. \tag{7}$$

Having established the condition for optimality, the remaining condition to be satisfied for a competitive equilibrium is market clearing. To establish market clearing, we must first determine the pattern of matching; $x = \mu(y)$.

Lemma 1 *There is Positive Assortative Matching on wealth, i.e., $\mu'(y) > 0$ if*

$$y \leq x\left(\frac{2}{\sqrt{\alpha}} - 1\right). \tag{8}$$

Proof. In Appendix. ■

Observe that given that $\alpha \in [1/2, 1)$, the preceding condition will be satisfied if $y \leq x$. With positive assortative matching, the market will clear from the top, with the wealthiest available girl matching with the wealthiest available boy. To establish the existence of an equilibrium with assortative matching, Lemma 1 indicates that we must derive the matching pattern when the market clears from the top and then verify *ex post* that $y \leq x$ at every wealth level.

Observe that there is no technological complementarity between the boy's and the girl's wealth in our model. The complementarity that gives rise to positive sorting is derived from the structure of the marriage institution in conjunction with the parents' preferences to leave a bequest. Wealthy parents are willing to pay a higher dowry to secure a wealthy match, which will ensure higher consumption for their daughters (because their husbands receive a larger bequest). The first order condition, equation (7), ensures that the hedonic price, u, and, hence, the dowry is increasing sufficiently steeply in x so that the matching on wealth is stable.[22]

Our assumption that families match on wealth alone is consistent with the empirical evidence that household characteristics, especially land wealth, matter more for matching than individual characteristics in rural India (Rao (1993a)). This is also consistent with our own results, reported in Appendix Table A2, that the pattern of matching by households on wealth varies by gender but is unaffected by the child's education. We could add individual characteristics to the model, but then the matching problem becomes a multi-dimensional allocation problem, which is analytically intractable once the wealth distribution is allowed to be endogenous (on account of sex selection).[23] Individual characteristics are thus omitted from the model, although we condition for parental education as an independent determinant of sex selection in the empirical analysis.

Sex Selection. Without sex selection, a child is born a boy or girl with equal probability. This implies that the wealth distribution on either side of the market is the same. It follows that girls and

[22] If the dowry was chosen on the basis of the bequest motive alone, then for a given x, a girl's family with wealth y would choose d to maximize $v(x, y, d) = log(y - d) + log((1 - \alpha)t)$. Substituting the expression for t from equation (3), the first order condition for this maximization problem implies that $d = (y - x)/2$. The maximized utility for the girl's family, $v(x, y) = 2log\left(\frac{x+y}{2}\right) + log\left(\frac{1-\alpha}{2}\right)$ is monotonically increasing in x. All girls'parents would want them to match with the wealthiest boys and the market would not clear.

[23] Multi-dimensional matching problems are difficult to solve even with exogenous distributions and linear preferences. See for example Choo and Siow (2006) and Lindenlaub (2017).

boys of equal wealth will match with each other; $y = x$, when the market clears from the top. The *ex post* condition required to establish the existence of an equilibrium with assortative matching, $y \leq x$, is satisfied at every wealth level.

With sex selection, the solution is more complicated because the wealth distribution is now endogenous and is no longer the same for boys and girls. We can, nevertheless, derive the following result without placing any additional restrictions on the model:

Proposition 1 *In equilibrium there is sex selection at every wealth level and Positive Assortative Matching on wealth, which implies hypergamy.*

Proof. We first establish that there is sex selection at every wealth level. Given that the cost of sex selection k is bounded below at zero, the required condition is $u(y) > v(y)$; i.e. families receive higher utility if they have a boy rather than a girl.

Let girls with family wealth y match with boys with family wealth x in equilibrium, so that the total wealth available for consumption for the two families is $x + y$. Given the outside option of remaining single, the boy's family must receive at least $2\log\left(\frac{x}{2}\right)$. The minimum total wealth that it needs to achieve this utility is x, but this requires that the boy and his parent consume the same amount. If their consumption levels differ, as they must given α, then the total requirement will exceed x.[24] In that case, the girl's family is left with less than y and so the maximum utility it can achieve (with equal consumption across generations) is less than $2\log\left(\frac{y}{2}\right)$. It follows that $v(y) < 2\log\left(\frac{y}{2}\right) \leq u(y)$.

To establish that there is assortative matching, we derive the matching pattern when the market clears from the top. At the very top of the distribution, girls and boys of equal wealth match with each other; $\overline{y} = \overline{x}$. Given that there is a shortage of girls at every wealth level, $y < x$, and there is hypergamy, at all other wealth levels. The *ex post* condition required from Lemma 1 to establish the existence of an equilibrium with assortative matching, $y \leq x$, is thus satisfied at every wealth level. ■

The root cause of sex selection in our model is the social norm that all girls must marry, which results in extreme disutility if a girl stays single. The positive outside option for the boys, together with the distortion in their consumption because of α, leaves them with a greater share of the total surplus from marriage. Parents are thus better off with boys than with girls, which results in sex selection.[25] Although the payoff to the girls is lower than that to the boys due to the norm, they do

[24] Substituting the expression for t from equation (3) in the expressions for C_x, c_x, it is straightforward to verify that $c_x = \alpha C_x$.

[25] Although the complementarity in match formation that is generated by the structure of the marriage institution results in assortative matching in equilibrium, the total consumption utility for both families is still less than what would be obtained if they consumed independently. The reason why marriages take place is because of a social norm which makes remaining single for girls extremely costly. Although this is outside the scope of our model, one way to motivate the presence of this norm is that there are positive social externalities from marriage which are not internalized by girls' families. This could be because marriages create links in a larger (caste-based) economic network, which have a multiplier effect, or because there is a social value to procreation.

have some bargaining power. This is because the boys are on the long side of the market, and thus the lowest wealth boy to marry is pushed to his outside option. This pins down the boy's payoff at that wealth level and, through the resulting hedonic price schedule, the division of the surplus from marriage across the wealth distribution.

Wealth and sex selection. Proposition 1 establishes that there will be sex selection at every wealth level. However, it does not tell us how sex selection will vary across the wealth distribution. With positive assortative matching, girls with family wealth y match with boys with family wealth $\mu(y)$, where $d\mu(y)/dy > 0$. When a family with wealth y that is expecting a girl decides to have a boy instead, it will receive utility $u(y) - k$. Note that the boy will then match with a poorer girl with family wealth $\mu^{-1}(y)$. If the family had chosen instead to keep the girl, it would have received $v(\mu(y), y; u(\mu(y)))$, which we know from Proposition 1 is less than $u(y)$. Thus, the family will proceed with sex selection if its cost $k < u(y) - v(\mu(y), y; u(\mu(y)))$. In general, for families with wealth y there is a critical cutoff $k^\star$ such that

$$k^\star(y) = u(y) - v(\mu(y), y; u(\mu(y))). \tag{9}$$

Given that the cost of sex selection, k, is distributed according to the cumulative density function, $H(k)$, the fraction of families with wealth y that choose sex selection is thus $H(k^\star(y))$. For most of what follows, we assume H uniform on $[0, a]$. The pattern of sex selection at every wealth level generates an endogenous and distinct distribution of wealth for girls and boys. The economy-wide distribution of wealth z is $\Gamma(z)$. The measure of families with boys whose wealth exceeds x and the measure of families with girls whose wealth exceeds y can thus be described as follows:

$$F(x) = \int_x^{\overline{x}} (1 + H(k^\star(z)) d\Gamma(z)/2 \quad \text{and} \quad G(y) = \int_y^{\overline{y}} (1 - H(k^\star(z)) d\Gamma(z)/2, \tag{10}$$

where $\overline{x} = \overline{y} = \overline{z}$. With Positive Assortative Matching, the market clearing condition is

$$\int_{\mu(y)}^{\overline{x}} dF(x) = \int_y^{\overline{y}} dG(y) \tag{11}$$

or equivalently:

$$\int_{\mu(y)}^{\overline{x}} (1 + H(k^\star(z)) d\Gamma(z)/2 = \int_y^{\overline{y}} (1 - H(k^\star(z)) d\Gamma(z)/2. \tag{12}$$

Sex selection determines the distribution of wealth for boys and girls, which, in turn, determines the pattern of matching in equation (12). The pattern of matching determines sex selection in equation (9). Sex selection and matching must thus be solved simultaneously.

If we knew the payoff $u(x)$ at every level of wealth x on the boys' side, then we could solve for sex selection and matching recursively, starting at the top of the wealth distribution and moving down. We would know $\mu(y)$ at any wealth level y on the girls' side, given the pattern of sex selection at higher wealth levels, and so would be able to compute $u(y) - v(\mu(y), y; u(\mu(y)))$ and, hence, $H(k^\star(y))$. However, the hedonic price schedule $u(x)$ must also be derived endogenously in the model. To do this we integrate the first order condition in equation (7), $v_x + v_u u' = 0$, which implies $u' = -\frac{v_x}{v_u}$, with respect to x:

$$u(x) = \int_{x^\star}^{x} -\frac{v_x(x, \mu^{-1}(x); u(x))}{v_u(x, \mu^{-1}(x); u(x))} dx + u(x^\star) \tag{13}$$

where the denominator is negative, and where $x^\star$ is the lowest wealth boy who is matched. From the outside option, we know that $u(x^\star) = 2 \log \frac{x^\star}{2}$.

The equilibrium is fully defined by the sex selection condition, the matching condition, and the payoff condition, as specified in equations (9), (12), and (13). This system of equations must be solved simultaneously. The additional consideration is that the payoff condition holds a fixed point because $u(x)$ appears on both sides of equation (13). We cannot solve the system of equations analytically to determine sex selection at each wealth level. However, the model can be solved numerically. Analytical results can also be obtained at the very top of the wealth distribution where the matching pattern is exogenously determined; $\overline{y} = \overline{x}$, and at the lowest wealth level at which boys match, $x^\star$, where $u(x^\star) = 2 \log \left(\frac{x^\star}{2}\right)$.

Proposition 2 *Sex selection is increasing in wealth (i.e., $\frac{dk^\star(y)}{dy} > 0$):*

1. *at the top $y = \overline{y}$, whenever $d < \frac{\overline{y}}{2}$;*

2. *at the bottom $y = \underline{y}$, whenever $\alpha \geq \tilde{\alpha}$ for some $\tilde{\alpha}$.*

Proof. In Appendix. ∎

If the result in Proposition 2 holds over the entire wealth distribution, then this implies that sex selection worsens monotonically with wealth. The intuition for this conjecture, which is validated by the numerical results, is that the shortage of girls grows as we move down the wealth distribution because more and more boys are left unmatched above them. This implies that poorer girls match with relatively wealthy boys; i.e. there is an increase in hypergamy, making the switch (through sex selection) to a boy relatively unattractive. The girls at the bottom of the wealth distribution benefit the most from the sex selection above them and thus the sex ratio is most favorable at the bottom. The combination of sex selection and Positive Assortative Matching also implies that boys below a threshold wealth level remain unmatched.

The implicit assumption in the model is that boys and girls of the same cohort match with each other. We could reduce some of the male surplus in the marriage market by allowing boys to match with younger girls, but this cannot be a steady-state solution. The age gap between husbands and wives will widen over successive age cohorts until the girls that are needed to match with the surplus boys are too young to marry.[26]

3.3 Numerical Solution and Results

THE ALGORITHM. The numerical solution of the model assumes that there is a finite number of wealth classes. This implies that boys and girls in a given wealth class could potentially match across multiple wealth classes. The matching allocation then looks like a step function instead of a smooth curve. With a continuum of wealth classes, the first order condition in equation (7), $\frac{dv}{dx} = 0$, ensures that the allocation and transfers are optimal for girls' families in each wealth class. With a finite number of wealth classes, the equivalent condition is that girls' families in a given wealth class will obtain the same utility across all the wealth classes that they match with. Given that the equilibrium payoff for the boys' families, $u(x)$, is a function of their wealth alone, the symmetric condition is that boys' families in a given wealth class receive the same utility across all the wealth classes that they match with.

The solution to the model must satisfy the sex selection condition, the measure preserving allocation or matching condition, and the payoff condition simultaneously. The algorithm that we use to solve the model numerically begins with an initial guess for the payoff at the top of the wealth distribution, $u(\overline{x})$, and for the pattern of matching. We know from Proposition 1 that there will be sex selection at every wealth level. This implies that there will be a shortage of girls in the highest wealth class and so girls in the next to highest wealth class will match up (with boys one wealth level higher than themselves) and horizontally (with boys in their own wealth class). As we move down the wealth distribution, the excess of boys accumulates and it is possible that below some wealth level, girls match exclusively with wealthier boys.

Given any initial guess for $u(\overline{x})$ and the matching pattern, we can solve for $u(x)$ and $v(y)$ in each wealth class. $v(y)$ is a function of y, x, and $u(x)$, as specified in equation (6). Given that girls in the highest and the next to highest wealth class match with the wealthiest boys, with family wealth $\overline{x}$ and payoff $u(\overline{x})$, we can solve for v in both wealth classes. Girls' families in the next to highest wealth class must receive the same utility, v, from matching with the wealthiest boys and boys in their own wealth class. This allows us to solve for u in the next to highest wealth class. We continue to solve recursively in this way down the wealth distribution.

With sex selection, boys below a wealth level $x^\star$ will remain unmatched. A comparison of $u(x^\star)$ derived in the first iteration with the outside option, $2\log\left(\frac{x^\star}{2}\right)$, is used to adjust the guess for $u(\overline{x})$ in the next iteration. Given u and v derived in the first iteration, the level of sex selection $H(k^*(y))$

[26]Consider the following thought experiment. Suppose that we are out of steady state and the number of boys is double the number of girls, and all boys marry at the age of 25. Then the first cohort of boys will marry girls aged 25 and 24, the second cohort will marry girls aged 24 and 23, and so on. Eventually the girls will be too young and some boys must remain unmarried. This is independent of the sex ratio as long as it is different from one. See also Bhaskar (2011).

can be determined in each wealth class y. The pattern of matching implied by this sex selection is used as the starting point for the next iteration. This iterative process continues until there is convergence. The numerical solution thus simultaneously satisfies the sex selection condition, the matching condition, and the payoff condition.

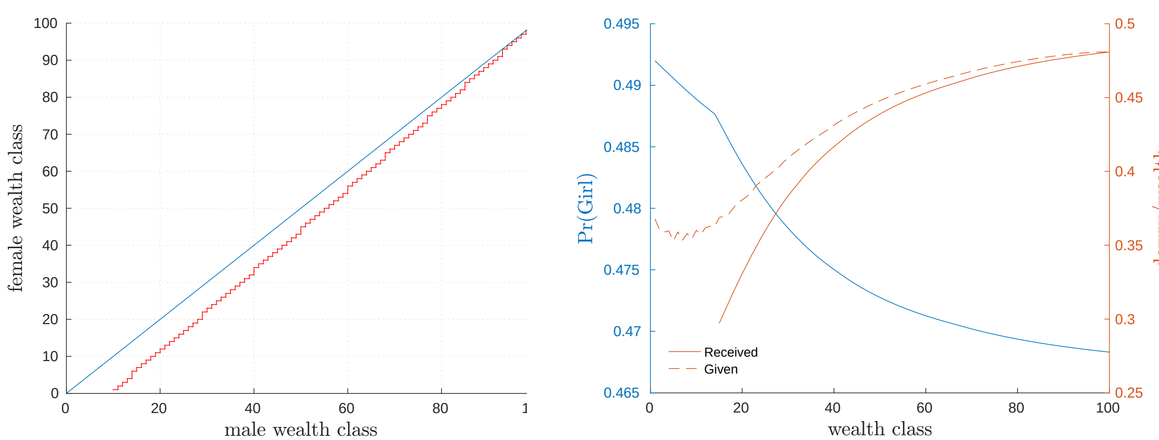
Numerical simulations. The wealth distribution is assumed to be log-normal in the numerical simulations, with the parameters selected to match the census data (within castes). The wealth distribution is divided into 100 classes for the simulations. As in Proposition 2, $k \sim U[0, a]$, which implies that there are two parameters in the model: α and a. We select values for these parameters: $\alpha = 0.6$ and $a = 12$ that are close to the values estimated below.

The matching pattern generated by the model is reported in Figure 2a. Notice that the plot is not a smooth function and has small steps. This is due to the discreteness of the wealth distribution, which results in each wealth class matching with multiple wealth classes of the opposite sex as described above. At higher wealth classes, girls and boys match horizontally as well as up and down, respectively. This is why the plot touches the 45 degree line at those wealth levels. However, below a certain wealth level, girls match exclusively with wealthier boys, shifting the plot down and away from the 45 degree line. On average, girls are matching with wealthier boys, consistent with the hypergamy documented in our survey data.

Figure 2: **Simulated Model (parameter values $\alpha = 0.6; a = 12$)**

(a) Model Generated Matching Pattern

(b) Model Generated Sex Selection and Dowry



The dowry received by boys of a given wealth class, x, can we computed directly from the value of $u(x)$ derived for that wealth class. While this is the same, regardless of the wealth of the girls' families that they match with, the dowry paid by girls in a given wealth class will vary with the wealth of the families that they match with. It is thus necessary to take account of the matching pattern in each wealth class when computing the average dowry paid by girls' families over the wealth distribution. Given that girls are matching up on average, it must nevertheless be the case that the dowry given is greater than the dowry received at a given wealth level. This is what we observe in Figure 2b, consistent with the dowry gap reported for daughters and sons in the household survey. The dowry is positive at all wealth levels in the survey data and this is also true in our numerical simulation.

We have shown analytically that there will be sex selection at every wealth level (Proposition 1) and that sex selection is increasing in wealth at the bottom and the top of the distribution (Proposition 2). However, we cannot analytically characterize the relationship between wealth and sex selection over the entire distribution. Figure 2b reports this relationship, based on the numerical solution to the model. The proportion of girls declines monotonically as we move up the wealth distribution, with an accompanying increase in the dowry.[27] In line with conventional wisdom, dowries and sex selection are positively correlated. However, as shown in Proposition 1, the root cause of sex selection is not dowry *per se*, but the social norm that all girls must marry.

Figure 3a reports the relationship between wealth and both dowries and sex selection for different values of α. The direct effect of an increase in α is to make boys' families better off at the expense of girls' families. However, this effect is partially offset by the decrease in the equilibrium dowry observed in Figure 3a.[28] The direct effect nevertheless dominates, resulting in an increase in sex selection at every wealth level. Figure 3b reports the same plots for different values of a, the sex selection parameter. The direct effect of a reduction in the cost of sex selection is a decline in the fraction of girls at every wealth level. This is partially (but not completely) offset by the reduced competition for boys, which shifts dowries down in Figure 3b.

The parameter values in Figure 3 are chosen to cover a relatively wide parameter space. Nevertheless, the robust finding is that the dowry is always positive and increasing in wealth, while the proportion of girls is decreasing in wealth. An increase in α reduces the proportion of girls, while a decrease in a, which implies a reduction in the cost of sex selection, also has the same effect. In our model, sex selection is generated by two frictions in the marriage market: (i) the social norm that all girls must marry, and (ii) the seepage in the bequest from the girl's parent because part of the dowry is siphoned off by the boy's family. Figure 3a quantifies the impact of this seepage, which is increasing in α, on sex selection.

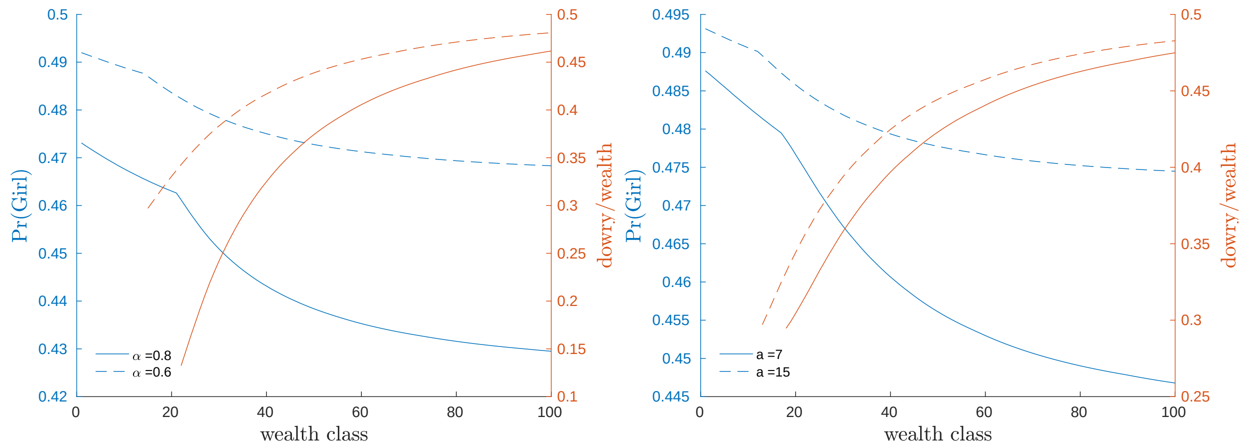
The α parameter will, in general, reflect the woman's status and bargaining power in her marital home. Although α is treated as exogenously determined and fixed in our model and the cross-sectional empirical analysis, this parameter could change over the course of the development process. Anderson and Bidner (2015) show theoretically that women's status could decline with economic development. The resulting increase in α, in the context of our model, would worsen sex selection,

[27]Notice that the dowry is less than half the family's wealth at the top of the wealth distribution. This satisfies the condition for sex selection to be increasing in wealth at that point in the distribution, as derived in Proposition 2.

[28]If there was sufficient curvature in the utility function, then altruistic parents would compensate for the decline in their daughters' consumption by increasing the dowry (at the cost of their own consumption). This condition is evidently not satisfied with logarithmic preferences. The effect of an increase in α is to reduce competition for boys in the marriage market, shifting down the dowry.

Figure 3: **Model Generated Comparative Statics**

(a) Friction: α (b) Cost of Sex Selection: a



providing one explanation for the worsening sex ratios in the the dynamic South Indian economy over the past decades.

Figure 3 distinguishes our model, in which son preference is generated endogenously by frictions in the marriage market, from the models in Edlund (1999) and Bhaskar (2011), where son preference is exogenously determined. All three models generate the prediction that the proportion of girls is decreasing with wealth, which we verify below. However, Bhaskar's and Edlund's models, in which marriage payments clear the market *ex post*, generate the prediction that there will be bride-price in equilibrium because there is a shortage of girls.[29] While this prediction is consistent with the bride-prices that are observed in countries like China, where sex selection is determined by the demand for a son, it is inconsistent with the marriage transfers from the girl's side to the boy's side that are universally reported in India. In line with the Indian data, our model, which incorporates relevant features of the marriage institution in that country, can generate both sex selection and positive dowries in equilibrium, despite the fact that girls are on the short side of the market.[30]

POLICY EXPERIMENTS. Our model is able to capture key features of the marriage institution and, based on these features, to generate sex selection in equilibrium. The model is also well suited to evaluate the effect of policies designed to address the sex selection problem because most policies will either directly or indirectly work through the marriage market.

[29] Bhaskar's model will also generate a groom-price, due to a net shortage of men, if men marry younger women and the population is growing sufficiently fast. However, his model cannot explain why a groom-price is observed in states like Tamil Nadu, the setting for our research, where fertility rates have been below replacement since the mid-1990's and where there is a surplus of men due to sex selection.

[30] An additional prediction of Edlund (1999)'s model is that the imbalance in the number of girls and boys will decrease with the availability of a more certain sex selection technology. As noted by Bhaskar (2011), this prediction is inconsistent with the findings from many studies that sex ratios worsened with access to sex selective abortion in India, China, and Korea. Our model does not suffer from this limitation of Edlund (1999)'s model. There is no uncertainty in the sex selection technology. Nevertheless, sex ratios are biased in equilibrium, at every wealth level.

Given that the dowry is effectively the price for a boy, one obvious policy lever to reduce the demand for boys would be to tax the dowry.[31] A dowry tax is introduced in our model by assuming that the boy's parent receives an amount θd, where $\theta < 1$. This directly affects the transfer from the boy's parent to his son and equation (3) can be rewritten as,

$$t = \frac{x + \theta d}{2}. \tag{14}$$

Substituting as before, $t = \psi(u)$ and $d = \frac{2\psi(u) - x}{\theta}$, where $\psi(u) = \sqrt{\frac{e^u}{\alpha}}$. This allows us to write the maximized utility for the girl's family as

$$v = \log\left(y - \frac{2\psi(u) - x}{\theta}\right) + \log((1 - \alpha)\psi(u)). \tag{15}$$

If the dowry d remains fixed, then the boy's family's utility u will decrease.[32] The girl's family's utility will also decrease; although her parent's utility is unchanged, the girl's utility declines with the decline in t. Thus, the net effect on sex selection is ambiguous. This ambiguity is compounded by the fact that the dowry d will shift up in response to the dowry tax. The advantage of our model is that it can be solved numerically, using the modified expressions for u and v, to derive the equilibrium price response to the dowry tax and the accompanying consequences for sex selection at each wealth level.

The preceding discussion highlights the value of our model in analyzing the complex equilibrium response to external interventions. Although a tax on the dowry has yet to be implemented, many schemes have already been introduced with the specific objective of reducing the bias in child sex ratios. In the framework of our model, these schemes either provide a wealth transfer to girls' parents, conditional on having a girl, or a direct transfer to the girl. These policies will change the maximized utility of the girl's family in the following ways:

(a) If the wealth transfer, w, is to the girl's parents,

$$v = \log(y + w + x - 2\psi(u)) + \log((1 - \alpha)\psi(u)). \tag{16}$$

(b) If the wealth transfer is directly to the girl,

$$v = \log(y + x - 2\psi(u)) + \log((1 - \alpha)\psi(u) + w). \tag{17}$$

If u is fixed, then the most effective scheme will target the family member; i.e. the girl or her parent,

[31] In a related exercise, Bhalotra, Chakravarty, and Gulesci (2016) estimate the relationship between the price of gold at the time of birth, which is assumed to determine the monetary cost of marriage to girls' parents many years later, and sex selection.

[32] The transfer t declines, from the expression above. Given that $t = \psi(u)$ and that ψ is an increasing function of u, it follows that u will decrease as well.

who has a lower level of consumption in equilibrium. However, the effect of the wealth transfer is more complex than that because it will change the equilibrium marriage price and, hence, matching and sex selection over the entire wealth distribution. This is especially important when evaluating existing transfer schemes that are targeted at less wealthy parents. While the targeted families may be induced to have more girls, there will be spillover effects through the equilibrium price that could *increase* sex selection at other points in the wealth distribution. Once the structural parameters are estimated, they will be used to quantify the impact of different policies, taking into account these equilibrium price effects.

4 Wealth and Sex Selection: Empirical Results

4.1 The Setting

The South India Community Health Study (SICHS) covers half of Vellore district in the state of Tamil Nadu. There are 298,000 households drawn from 57 castes in the study area. The study area (with a population of 1.1 million) is representative of rural Tamil Nadu (with a population of 37 million) and rural South India (comprising Tamil Nadu, Andhra Pradesh, Karnataka, and Maharashtra with a total population of 193 million) with respect to demographic and socioeconomic characteristics.

Table 3 reports the age distribution, marriage patterns, literacy rates, and labor force participation in the study area, rural Tamil Nadu, and rural South India, respectively. Statistics for Tamil Nadu and South India are based on official Government of India data, while the corresponding statistics for the study area are derived from the SICHS census. The age distribution and marriage patterns are combined in a composite statistic that measures the number of married individuals in 5-year age categories as a fraction of the total population, separately for men and women. If this statistic is the same across two populations, then it follows that both the age distribution and marriage rates must be the same in these populations. Based on the Kolmogorov-Smirnov test, we are unable to reject the null hypothesis that the age distribution of married individuals is equal, for men and for women, between the study area and both rural Tamil Nadu and rural South India.

Literacy rates and labor force participation rates, for men and for women, are similarly comparable between the study area and both rural Tamil Nadu and rural South India. Notice that literacy rates are much higher for men than for women, 80% versus 60%, although this gender gap has largely disappeared for children currently enrolled in school (see Appendix Table A3). Labor force participation rates match the patterns for literacy; 80% for men versus 40% for women. The extremely low female labor force participation rates will be relevant later when we discuss potential policies to reduce sex selection.

The study area was purposefully selected to be representative of the South Indian region with respect to key demographic and socioeconomic indicators. As observed in Figure 1, Vellore district

Table 3: **Comparison of Demographic and Socioeconomic Characteristics**

Region	Men			Women		
	South India	Tamil Nadu	Study Area	South India	Tamil Nadu	Study Area
Age distribution						
married (%) <10Yrs	0.0	0.0	0.0	0.0	0.0	0.0
10-14	0.0	0.0	0.0	0.0	0.0	0.0
15-19	0.1	0.0	0.0	1.6	0.6	0.8
20-24	2.2	1.0	1.0	7.0	5.4	5.9
25-29	5.4	4.4	4.8	7.4	7.8	8.1
30-34	6.9	6.8	6.5	7.4	7.4	6.6
35-39	6.5	6.6	7.1	6.2	6.2	7.7
40-44	6.2	6.3	3.7	5.9	6.2	3.1
45-49	5.4	5.8	6.8	4.5	4.8	5.8
50-54	4.6	5.0	4.8	3.2	3.7	3.5
55-59	3.4	4.0	4.1	3.5	3.9	3.3
60-64	2.8	3.1	3.8	1.8	2.0	1.9
65-69	2.1	2.4	2.4	1.1	1.1	1.1
70-74	1.3	1.4	1.6	0.6	0.6	0.3
75-79	0.8	0.9	0.8	0.2	0.2	0.1
80-84	0.3	0.3	0.4	0.1	0.1	0.0
85>Yrs	0.2	0.2	0.2	0.0	0.0	0.0
Kolmogorov-Smirnov test of equality (p-value)	1.00	1.00	–	1.00	0.75	–
Literacy rate (%)	79.2	82.1	76.9	61.1	65.5	62.4
Labor force participation rate (%, 15-59 years)	79.8	81.1	81.0	44.9	42.6	40.0

Notes: % married measures the number of married individuals in each age category as a fraction of the total population, seperately for men and women. South India includes Maharashtra, Karnataka, Andra Pradesh, and Tamil Nadu. Literacy defined by the Goverment of India as those aged 7+ who can, with understanding, read and write a short, simple statement on their everyday life; SICHS Census definition is those aged 7+ with $\geq$ 1 year of education (figures for $\geq$ 3 years of education are similar, 73.8% for men and 59.5% for women). Labor force participation defined as the proportion of 15-59 year old persons of the total 15-59 years population who are either employed or seeking or available for employment.
Sources: For Tamil Nadu and South India, age/marriage data from Ministry of Home Affairs, GOI, and literacy data from 2011 census reported by Office of the Registrar General and Census Commissioner, GOI; Labor force participation: Ministry of Labor and Employment, Government of India, 2009-10. For Study Area, all statistics based on SICHS Census.

ranks around the median in Tamil Nadu with respect to child (aged 0-6) sex ratios. The child sex ratio in the study area (from the SICHS census) is 108, which is just around the average for South India and slightly lower than the all-India average of 109. No castes that are traditionally associated with severely biased sex ratios are present in the study area. In this apparently innocuous setting we will, nevertheless, uncover substantial variation in sex ratios *within* castes. Given the representativeness of the study area and the fact that the marriage market is organized the same way in all castes, it is very possible that the same findings would be obtained elsewhere in the country.

4.2 Estimation

Specification. The benchmark equation that we estimate has the following specification:

$$Pr(G_{ij} = 1) = \beta R_{ij} + \delta_j + \epsilon_{ij}. \tag{18}$$

G_{ij} is an indicator variable, which takes the value one if child i belonging to caste j is a girl and zero if it is a boy. The sample is restricted to children aged 0-6 to be comparable with Government of India statistics on child sex ratios. R_{ij} is the child's family's rank in its caste's wealth distribution. δ_j is a caste fixed effect and ϵ_{ij} is a mean-zero disturbance term. The caste fixed effect will capture all caste-level variation in sex selection, including variation in norms that condemn this practice and common household characteristics, such as land ownership, that are believed to be associated with boy preference. The relative wealth variable, R_{ij}, reflects the marriage market effect. Based on the model, we expect $\beta < 0$.[33]

Although each family consists of a parent and a child in the model, there is substantial heterogeneity in family size and structure in the data. We account for this heterogeneity in the empirical analysis by restricting attention to "single family" households and by ranking families in the caste wealth distribution on the basis of their per capita wealth. Single family households, consisting of one couple and their children, but possibly including other adults (typically a grandparent) account for 96.2% of all households with children in the census. By dividing household wealth by the number of family members, we obtain a measure of the wealth that is potentially available to each child.[34]

The household's wealth is measured by its associated income flow; i.e. the sum of the total profit from land owned, leased, or rented and the total labor income of all members, including those that have temporarily migrated to work. The total profit is measured in the year prior to the census and the labor income is measured in the prior month, which allows us to compute average monthly household income. Although the household's average monthly income in the year prior to the census will be positively correlated with its wealth, it also includes a transitory income component. The resulting measurement error will bias the relative wealth coefficient towards zero. We will see that the results get even stronger with an alternative measure of relative wealth that purges this measurement error.

The increasingly biased sex ratios over time in India have been attributed to a number of factors including: (i) improved sex selection technologies; e.g. Arnold, Kishor, and Roy (2002), (ii) changes in the economic returns to having boys versus girls; e.g. Foster and Rosenzweig (2001), and (iii)

[33] A review of the literature uncovered just three published studies, using data that are now 30 years out of date, that have estimated the relationship between wealth and sex selection; Rosenzweig and Schultz (1982), Krishnaji (1987), and Murthi, Guio, and Drèze (1995). No previous research has examined the theoretically relevant relationship between sex selection and *relative* wealth within the marriage market, which is defined by the caste in India.

[34] The implicit assumption is that other members of the household, outside the nuclear family, do not receive a share of the wealth. For example, a grandparent living with the son and his family would have already distributed his wealth among his children.

reduced fertility coupled with a desire to have at least one son; e.g. Basu (1999).[35] The same factors could generate cross-sectional variation in sex selection.

Modern technologies have substantially reduced the monetary cost of sex selection. These technologies have been in use for many decades and are well known in the general population. However, household wealth and parental education could jointly determine the willingness of families to use these technologies. These variables will also determine the labor market returns to sons and daughters, through the educational investments that are made for them, as well as the mother's labor force participation, which has been shown to increase the survival of female children (Rosenzweig and Schultz (1982), Kishor (1993), Murthi, Guio, and Drèze (1995)).[36] We consequently include household wealth and parental education in an augmented version of equation (18). We measure the household's wealth by its average monthly income in the year preceding the census and parental education by years of schooling.

While we focus on marriage market frictions as the primary determinant of sex selection, a desire for at least one son can also bias the sex ratio. With this motivation, the sex ratio will worsen with the birth-order of the child. We account for this alternative channel by including the birth-order of the child in the augmented specification. There has been below replacement fertility in Tamil Nadu since the mid-1990's, and few families have more than three children. The reference category when constructing the birth-order measure is thus the third child or above, with separate indicator variables for the first-born child and the second-born child.

Results. The estimation results with the benchmark specification and the augmented specification are reported in Table 4, Columns 1-2. The coefficient on the wealth rank variable is negative, as predicted. It is not significant at conventional levels in the benchmark specification, but is larger (in absolute magnitude) and is more precisely estimated once additional regressors are included. The coefficient on household wealth is *positive* and very precisely estimated in Column 2.[37] This result emphasizes the distinction between relative wealth, which our model predicts will reduce the probability that a child is a girl, and absolute wealth. There are many channels through which absolute wealth could determine sex selection and, hence, the sign of its effect is ambiguous.

The findings from previous studies, summarized by Murthi, Guio, and Drèze (1995), indicate that the effect of parental education on sex selection is similarly ambiguous. In our data, maternal education is associated with an increase in sex selection, while paternal education works in the opposite direction in Column 2. However, the effects are not statistically significant, and we will

[35] Although marriage market frictions generate sex selection in our model, another popular (potentially co-existing) explanation for sex selection is that parents want at least one boy. It is straightforward to verify that this implies that sex ratios will worsen at higher birth-orders, as documented in previous studies. Exogenous fertility decline makes the sex ratios worsen earlier, resulting in an overall increase in sex bias.

[36] As shown in Appendix Table A3, using SICHS census data, household wealth and parental education determine both secondary school enrollment, by gender, among 14-17 year olds, as well as female labor force participation.

[37] This is consistent with Rosenzweig and Schultz (1982) who find, using household data, that land ownership is associated with an increase in the probability that a child is a girl.

Table 4: **Wealth and Sex Ratios**

Dependent variable	Girl dummy					
Wealth measure	Reported wealth		Predicted wealth			
Rank measure based on	all households in the caste				households with at least one child aged 0-6 years	
	(1)	(2)	(3)	(4)	(5)	(6)
Rank in caste per capita wealth distribution	-0.0057 (0.0079)	-0.075*** (0.018)	-0.046*** (0.0068)	-0.15*** (0.010)	-0.0409*** (0.0064)	-0.135*** (0.0097)
Wealth	–	0.0020*** (0.00053)	–	0.037*** (0.0051)	–	0.0360*** (0.0051)
Mother's education	–	-0.0008 (0.00052)	–	0.0000 (0.00041)	–	-0.0001 (0.00041)
Father's education	–	0.00015 (0.00057)	–	0.00030 (0.00042)	–	0.00027 (0.00042)
Birth order = 1	–	0.037*** (0.0056)	–	0.080*** (0.0060)	–	0.0778*** (0.0059)
Birth order = 2	–	0.029*** (0.0057)	–	0.057*** (0.0049)	–	0.0552*** (0.0049)
Sample mean of dependent variable	0.479		0.480		0.480	
Observations	91,369	91,336	79,027	78,995	79,068	79,036
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Sample restricted to children aged 0-6 years. Reference category for child birth order is third child or higher. Standard errors (in parentheses) are calculated on the basis of 500 bootstrap replications for Columns 3-6. All standard errors are clustered at the panchayat level, *** p<0.01, ** p<0.05, * p<0.1

see that both maternal education and paternal education increase the likelihood that the child is a girl with alternative model specifications (although the effects remain insignificant). In line with the results from many previous studies, we also see in Column 2 that the first-born child is most likely to be a girl, followed by the second-born child, who is also significantly more likely to be a girl relative to the reference category (third-born and higher children). These results indicate that the additional motivation for sex selection, which is a desire to have at least one son, is also relevant.

Alternative construction of relative wealth. We measure household wealth by an income flow, which includes a transitory component. The resulting measurement error will bias the estimated relative wealth coefficient towards zero. If we observed the household's income realizations over many years, then the income shocks could be purged by using the average income over time. Because we have a single income realization from the census, we purge the income shocks by predicting current wealth with historical wealth.

There are 377 *panchayats* or village governments in the study area. These *panchayats* were historically single villages, which over time sometimes divided or added new habitations. The *panchayat* as a whole, which often consists of multiple modern villages, can thus be linked back to a single historical village. The agricultural revenue tax, per acre of cultivated land, that was collected from these villages by the British colonial government is available in 1871. The colonial government carefully measured soil quality, irrigation, and other growing conditions in each village. The revenue tax was designed to reflect the potential output per acre, which, based on the detailed data collected, would have been highly correlated with actual agricultural productivity.

Soil quality is a fixed characteristic, which will continue to determine productivity and agricultural income today. Irrigation in the study area in the nineteenth century was almost entirely provided by surface tanks. A relatively small number of villages, which had access to tank irrigation, could grow rice, which increased their income substantially. Tank irrigation has been largely replaced by well irrigation, which is less geographically constrained. However, historical advantage could have persisted in an economy with imperfect credit markets by allowing households in historically wealthy villages to make profit-enhancing investments over time.[38]

We test the hypothesis that a household's current wealth is determined by historical agricultural productivity in its village by estimating the relationship between household income, obtained from the SICHS census, and historical agricultural productivity, measured by the tax revenue per acre in 1871. We allow for the possibility that this relationship could vary across castes, given that they have historically been engaged in different occupations, by including caste fixed effects and the interaction of these fixed effects with the historical agricultural productivity variable.[39] Historical agricultural productivity strongly predicts current household income, with the F-statistic measuring joint significance of the productivity coefficient and the productivity-caste interactions as high as 20.4. Predicted current household income can thus be used in place of the reported income as the measure of household wealth in equation (18).

Once the measurement error has been purged, we expect the estimated relative wealth coefficient in equation (18) to be larger in (absolute) magnitude. This is indeed what we observe in Table 4, Columns 3-4. The coefficient is also more precisely estimated, and is significant at the 1 percent level, for both the benchmark specification and the augmented specification.[40]

Although there is a single cohort in the model, in practice the age-gap between partners can vary across marriages. In lieu of a clear partition of age cohorts into independent marriage markets within the caste, we compute the household's relative wealth with respect to the entire caste. The implicit assumption with this measure is that the distribution of wealth within the caste is stable across age

[38]The implicit assumption underlying the historical persistence is that households, or dynasties, have remained in the same village over many generations. This assumption is supported by recent evidence that permanent migration from rural to urban areas is extremely low in India (Munshi and Rosenzweig (2016)). The 1871 population census provides the caste composition of each historical village in the study area. This allows us to construct the population share of each caste in each village in 1871 and the corresponding statistic today, based on the SICHS census. The correlation between these statistics is as high as 0.8.

[39]Standard errors in this regression, as in equation (18), are clustered at the *panchayat* level.

[40]Because the new measure of household wealth is a predicted variable, we report bootstrapped standard errors in Columns 3-4. A fresh sample of households is drawn, with replacement, for each iteration and this sample is used for the predicting equation and to estimate equation (18). The same procedure is followed for all specifications in which household wealth is a predicted variable.

cohorts. Our predictor of household wealth, based on historical productivity, is independent of the age composition of the household by construction. However, per capita wealth, which is used to determine relative wealth, depends on family size, which could vary with the age of the child. We thus construct an alternative measure of relative wealth, based only on households in the caste with children aged 0-6. The implicit assumption when constructing this measure is that the 0-6 year olds in each caste will form an independent marriage market in the future. The estimated coefficients in Table 4, Columns 5-6 are very close, both in magnitude and significance, to the corresponding coefficients in Columns 3-4. We thus continue to use the full caste to construct the household's relative wealth (unless noted) in the robustness tests that follow.

Alternative samples. When sex selection is determined by the demand for at least one son, the sex of the first-born child will be unbiased, with an increasing bias in higher-order births (Jayachandran (2017)). When sex selection is generated by imperfections in the marriage market, the bias will be observed at all births, with the additional prediction that the sex ratio varies with relative wealth within the caste. Consistent with the marriage market channel, Jha, Kumar, Vasa, Dhingra, Thiruchelvam, and Moineddin (2006) use data from the Sample Registration System (SRS) to document that first-births are biased in favor of males in India. Arnold, Kishor, and Roy (2002) obtain the same result with DHS data. Our results indicate that the probability that a child is a girl decreases with birth order, so both the marriage market channel and the demand for a son likely play a role in determining sex selection in the study area. To isolate the marriage market effect, we restrict the sample to first-born children in Table 5, Columns 1-2. A negative and statistically significant coefficient on the relative wealth variable continues to be obtained.

As an additional robustness test, we focus on the older group of 7-13 year olds. The coefficient on relative wealth is negative and significant in Table 5, Columns 3-4, and is larger in absolute magnitude than the corresponding coefficient for the 0-6 year olds in Table 4, Columns 3-4. This increase in absolute magnitude could be a cohort-specific effect; there is an unexplained improvement in the child (aged 0-6) sex ratio in the 2011 population census in Tamil Nadu, which, based on more recent data from the Sample Registration System, appears to have reversed. The larger relative wealth effect for the 7-13 year olds could also be because the forces driving sex selection in our model persist past the age of six, consistent with Anderson and Ray (2010) who find that sex selection continues into adulthood in India.

When constructing per capita wealth, we divide household wealth by the number of family members; i.e. the two parents plus the number of children. The implicit assumption is that other members of the household, typically single grandparents, do not share (inherit) the wealth. The alternative assumption is that the wealth is divided among all household members. The results are robust to this alternative measure of per capita wealth, as reported in Appendix Table A4. A second concern with the per capita wealth variable is that family size will not be complete, especially for the youngest children in the sample. A small fraction of families in Tamil Nadu have more than three children and there is typically a short spacing between births. The completed family size and, hence, per capita wealth is thus more accurately measured for children in the 7-13 age group. This might explain, in part, the stronger relationship between relative wealth and the sex ratio that we observed for that age group, relative to the 0-6 year olds, in Table 5. In a more stringent test, we restrict the sample to households with all children aged 5-15. The spacing between births rarely exceeds five years in the census data and children rarely leave their natal homes before the age of 15. The family size in this restricted sample is thus reliably complete. Using the restricted sample to construct the relative wealth measure, a negative and significant relationship between relative wealth and the probability that the child is a girl continues to be obtained in Table 5, Columns 5-6. Given the overlap in the age range, the relative wealth coefficient in Columns 5-6 is similar in magnitude to the corresponding coefficient for the 7-13 year olds in Columns 3-4.[41]

Table 5: **Wealth and Sex Ratios (Robustness Tests)**

Dependent variable	Girl dummy					
Sample age range	(0-6 years)		(7-13 years)		(5-15 years)	
Birth order	First borns		All children		All children	
	(1)	(2)	(3)	(4)	(5)	(6)
Rank in caste per capita wealth distribution	-0.0739***	-0.0917***	-0.218***	-0.363***	-0.156***	-0.244***
	(0.0115)	(0.0121)	(0.0066)	(0.0093)	(0.0062)	(0.0086)
Wealth	–	0.0234***	–	0.0635***	–	0.0627***
		(0.0065)		(0.0072)		(0.00734)
Mother's education	–	0.000718	–	0.0001	–	-0.00062
		(0.00063)		(0.0004)		(0.00038)
Father's education	–	0.00110	–	0.00146***	–	0.000199
		(0.00063)		(0.00038)		(0.00038)
Birth order = 1	–	–	–	0.116***	–	0.132***
				(0.0044)		(0.0136)
Birth order = 2	–	–	–	0.0909***	–	0.115***
				(0.0040)		(0.0133)
Birth order = 3	–	–	–	–	–	0.0524***
						(0.0135)
Sample mean of dependent variable	0.485		0.483		0.467	
Observations	34,723	34,706	82,648	82,627	79,040	79,018
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Reference category for child birth order is third child or higher in Columns 3-4, and fourth child or higher in Columns 5-6. In Columns 5-6, the estimation sample is households with all children aged 5-15 years. Standard errors (in parentheses) are calculated on the basis of 500 bootstrap replications. All standard errors are clustered at the panchayat level, *** p<0.01, ** p<0.05, * p<0.1

As noted, a decline in fertility, coupled with a desire to have at least one son, will increase sex selection. Variation in fertility across families is accounted for in our specifications by the birth-order dummies. These dummies will flexibly capture both the birth-order effect and the direct

[41] In an independent robustness test, we predicted completed family size for the 0-6 year olds. Results with this alternative measure of relative wealth, reported in Appendix Table A4, are similar to what we obtain for the 0-6 year olds in Table 4.

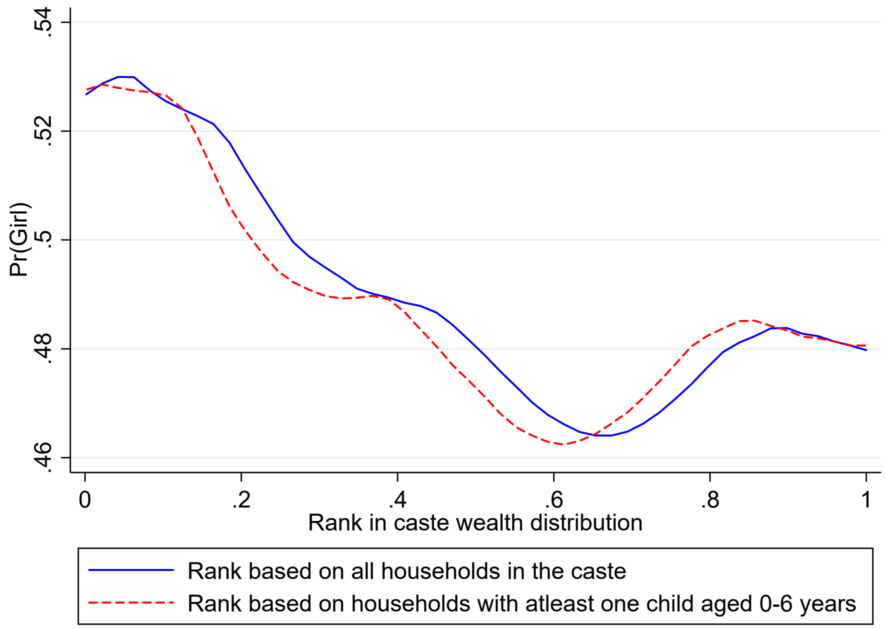
effect of family size; i.e. the number of children, on sex selection when the sample includes all children in the family. The restricted sample of families with all children aged 5-15 satisfies this requirement. Because this is a sample of completed families, the number of children is larger on average than it is for the families of the 0-6 year olds or the 7-13 year olds. We thus include three birth-order dummies in Table 5, Column 6. The reference category is now families with four or more children, but almost no families have more than four children. The birth-order dummies continue to be positive and precisely estimated with this specification and are declining in magnitude with increasing birth-orders. Although we cannot disentangle the birth-order effect from the family size effect, what is important is that the coefficient on relative wealth retains its sign and statistical significance.

Nonparametric estimation. The linear specification in equation (18) allows us to determine the sign of the overall relationship between relative wealth and sex selection. However, the model generates the stronger prediction that sex selection should be monotonically increasing as we move up the wealth distribution. Figure 4 examines this prediction by nonparametrically estimating the relationship between the sex of the child and the family's position in its caste's wealth distribution. This relationship is reported with two measures of relative wealth: (i) based on all households in the caste, and (ii) based on households in the caste with at least one child in the 0-6 age range. Large samples are needed to identify sex selection with the requisite level of statistical confidence. This problem is exacerbated with the nonparametric estimates because we are attempting to identify sex selection at different points in the wealth distribution. Nevertheless, we see in Figure 4 that the likelihood that a child is a girl is declining continuously with her family's position in its caste's wealth distribution. The flattening out at the top of the distribution matches the results from the numerical simulations in Figure 3.

The most stringent test of the model is to nonparametrically estimate the relationship between relative wealth and sex selection, caste by caste. Figure 5a reports this test for the 0-6 age group for the 12 largest castes, which account for 82% of the children in this age group. The probability that a child is a girl is decreasing with relative wealth for 9 of the 12 castes. For the three castes that it is not – Balija, Boya, and Naikar – the number of observations is relatively small (less than 2,000 children per caste). It is possible that the anomalous pattern for these castes is simply a consequence of the small sample size, which makes the estimated relationship unstable. To examine this possibility, we report the relationship between relative wealth and the probability that the child is a girl for the 7-13 year olds in Figure 5b. The relationship is negative at each point in the wealth distribution for all three castes.

Wealthy landowning castes, located primarily in the North, have been associated with excessively biased sex ratios since the earliest British-Indian censuses (Miller (1981), Krishnaji (1987), Das Gupta (1987)). Severe sex selection in the South has also been traditionally associated with specific castes (George, Abel, and Miller (1992), Chunkath and Athreya (1997)). However, just as

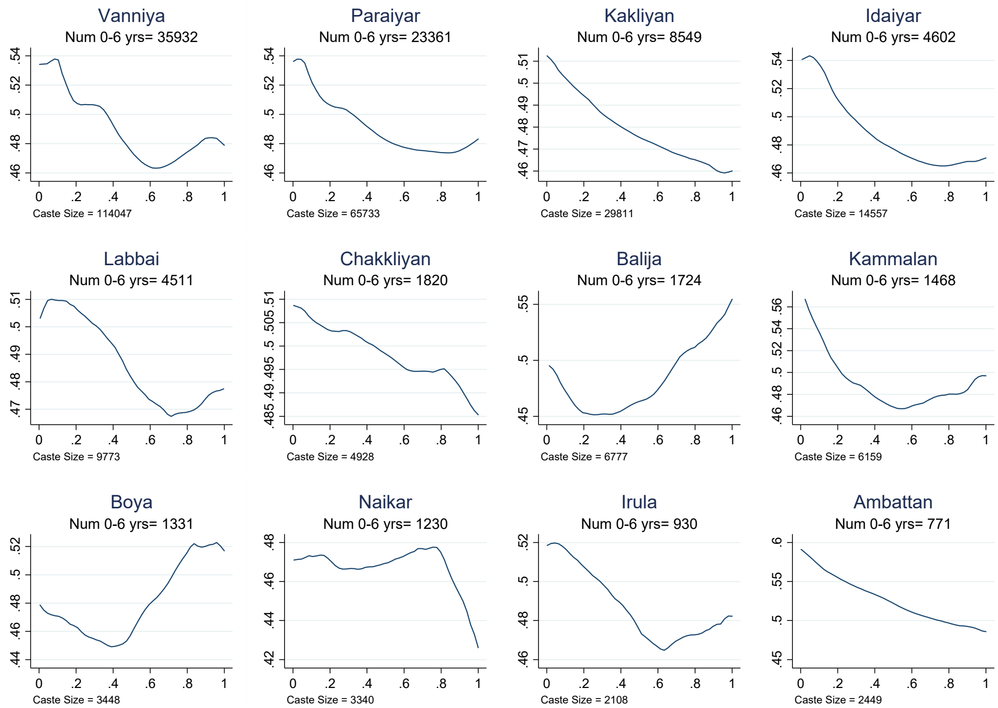
Figure 4: **Sex Selection by Percentile in Caste Wealth Distribution (all castes)**



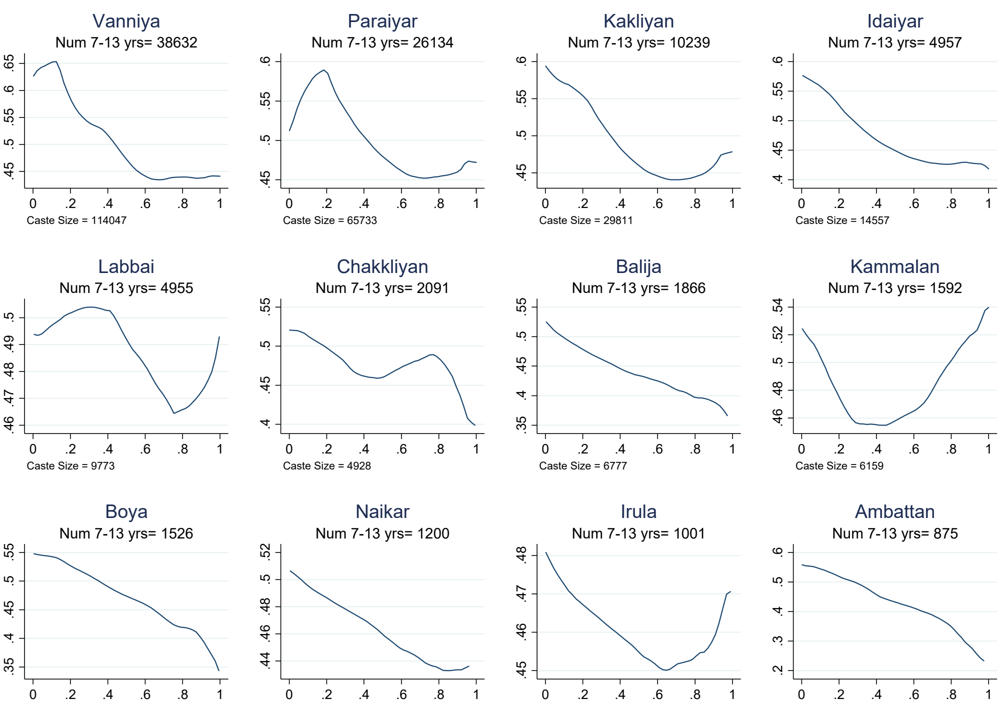
dowries have spread across the entire caste distribution in recent decades, there has been, in parallel, a trend towards sex selection even among the lower castes (Jeffery, Jeffery, and Lyon (1984), Sudha and Rajan (1999)). Two castes, the Vanniyas and the Paraiyars, dominate the population in the study area. The Vanniyas are a relatively wealthy landowning caste who have been given the honorific "Gounder" title in Vellore district. Other landowning castes have been given this title in other districts in Tamil Nadu. For example, the Gounders in Salem district, who are believed to be largely responsible for that district's exceptionally biased sex ratios in the past, belong to the Vellala caste. In Vellore, the Vanniya Gounders are not exceptional with regard to sex selection. The other numerically dominant caste in the study area, the Paraiyars, lie at the very bottom of the social hierarchy (the English word "pariah" is derived from this caste name). Despite their social differences, and in line with the current view that sex selection has spread across the caste distribution, the probability that a child is a girl is decreasing with wealth within each of these castes. As a final robustness test, we estimated equation (18), (i) without the Vanniyas, (ii) without the Vanniyas and the Paraiyars, and (iii) with just the 12 largest castes. The estimates with these different samples, reported in Appendix Table A5 for the benchmark specification with the 0-6 year olds, are very similar to what we obtain for the full sample in Table 4. There is a robust negative relationship between a family's position in its caste's wealth distribution and the probability that a child will be a girl.

Figure 5: **Sex Selection by Percentile in Caste Wealth Distribution for 12 Largest Castes**

(a) Ages 0-6



(b) Ages 7-13



4.3 Structural Estimation and Quantification

Magnitude of Within-Caste Variation. A robust finding from the preceding analysis is that the fraction of girls is decreasing as we move up the wealth distribution within castes. To quantify the magnitude of this variation, we partition the households with children aged 0-6 in each caste into eight equally sized wealth classes. The number of classes is chosen by weighting two competing considerations: The larger the number of wealth classes, the closer we can approximate the corresponding nonparametric plots where relative wealth in each caste is based on households with children aged 0-6. However, this comes at the cost of less precise estimates of the sex ratio within wealth classes, especially for castes with just a couple thousand children aged 0-6.

The benchmark equation that we use to quantify the magnitude of the within-caste variation in sex ratios has the fraction of girls in each wealth class in the caste as the dependent variable and a full set of wealth-class dummies as regressors. The R^2 in this regression, which indicates how much of the overall variation in sex ratios can be explained by relative wealth within the caste is 0.23 when the sample is restricted to the 30 largest castes and 0.39 when the sample is restricted even further to the 12 largest castes. To compare the magnitude of the within-caste and between-caste variation in sex ratios, we estimate an augmented equation that incorporates both sources of variation by including caste fixed effects. The R^2 with this specification increases to 0.33 for the sample with 30 castes and 0.45 for the sample with 12 castes (the coefficient estimates for all specifications are reported in Appendix Table A6). This implies that within-caste variation accounts for 70% of the explained variation with the 30-caste sample and as much as 87% of the explained variation with the 12-caste sample. No caste that is known to be traditionally associated with severely biased sex ratios is present in the study area. It is possible that in other districts where such castes are present, the between-caste variation will be more substantial. Nevertheless, these results highlight the importance of the within-caste variation that is the focus of our analysis.

A second approach to quantify the magnitude of the within-caste variation would be to measure the range of sex ratios across the eight wealth classes. Converting the fraction of girls to the number of boys per 100 girls, to be consistent with Government of India statistics once again, the sex ratio ranges from 97 to 117. To put these statistics in perspective, a sex ratio of 120 is considered to be severely biased and only a handful of districts in the country have sex ratios in excess of that level.[42] The variation in sex selection within castes that we uncover is driven by marriage market frictions that apply to all castes. Sex selection may thus be more pervasive than is commonly believed, affecting relatively wealthy households (within their caste) throughout the country.

[42] Over the last three census rounds, the worst sex ratio reported in the state of Tamil Nadu was by Salem district, which recorded a sex ratio of 120 in 1991. Salem had the most biased sex ratio in the entire country in that census round.

Table 6: **Structural Estimates**

	Fraction of girls			
	Weighted		Unweighted	
	Actual	Predicted	Actual	Predicted
	(1)	(2)	(3)	(4)
Wealth class				
1	0.509	0.487	0.522	0.489
		[0.483, 0.491]		[0.484, 0.493]
2	0.499	0.481	0.494	0.483
		[0.475, 0.487]		[0.476, 0.490]
3	0.472	0.477	0.471	0.480
		[0.470, 0.484]		[0.472, 0.488]
4	0.465	0.476	0.475	0.479
		[0.469, 0.483]		[0.471, 0.488]
5	0.460	0.476	0.464	0.479
		[0.468, 0.483]		[0.470, 0.488]
6	0.468	0.475	0.474	0.478
		[0.467, 0.482]		[0.469, 0.487]
7	0.483	0.473	0.488	0.477
		[0.465, 0.481]		[0.468, 0.486]
8	0.477	0.473	0.471	0.476
		[0.465, 0.481]		[0.467, 0.486]
Parameters				
a	–	12.975	–	16.367
		[9.053, 16.897]		[8.009, 24.726]
α	–	0.613	–	0.620
		[0.589, 0.637]		[0.552, 0.688]

Weighted estimates sample castes in proportion to their size. Unweighted estimates sample castes with equal probability. The 12 largest castes are used for the structural estimation. 95% confidence intervals in brackets.

Structural estimation. While our analysis provides new evidence on the extent of the sex selection problem, the problem itself is well known and widely discussed in academic and policy circles. Many states and the central government have responded to the problem by introducing Conditional Cash Transfer schemes with the stated objective of improving the survival and the welfare of girls and reversing the bias in the child sex ratio. Once the structural parameters have been estimated, counter-factual simulations with our model can be used to assess the impact of different schemes.

The estimation of the structural parameters, α and a, is straightforward. The algorithm we used to solve the model for given values of α and a was described above. To estimate the parameters, we search over all combinations of α and a to find the combination for which the predicted fraction of girls across the eight equal-sized wealth classes matches most closely with the actual fractions; i.e. for which the sum of squared errors is minimized. We use the 12 largest castes for the structural estimation. Bootstrapped means and confidence intervals for the parameters are reported in Table 6, Column 2 by drawing repeated samples with replacement, where the probability that a particular caste is drawn is proportional to its size. Given the numerical dominance of a small number of castes, the bootstrapped estimates will be largely determined by those castes. To assess the robustness of our results to alternative sampling procedures, we also report bootstrapped estimates in Table 6, Column 4 where all castes are sampled with equal probability.

The estimated α parameter is just over 0.6 with both sampling procedures, indicating that boys consume a greater fraction of the transfers from their parents than their wives. The a parameter is more sensitive to the sampling procedure, but the point estimate with each procedure nevertheless lies within the 95% confidence interval generated by the other. This is also true for the predicted fraction of girls in each wealth class. Despite the fact that the confidence intervals for the predicted fraction of girls are very narrow, the actual fraction (reported in Column 1 and Column 3) lies within the confidence interval for 3 of the 8 wealth classes, and just outside for the remainder.

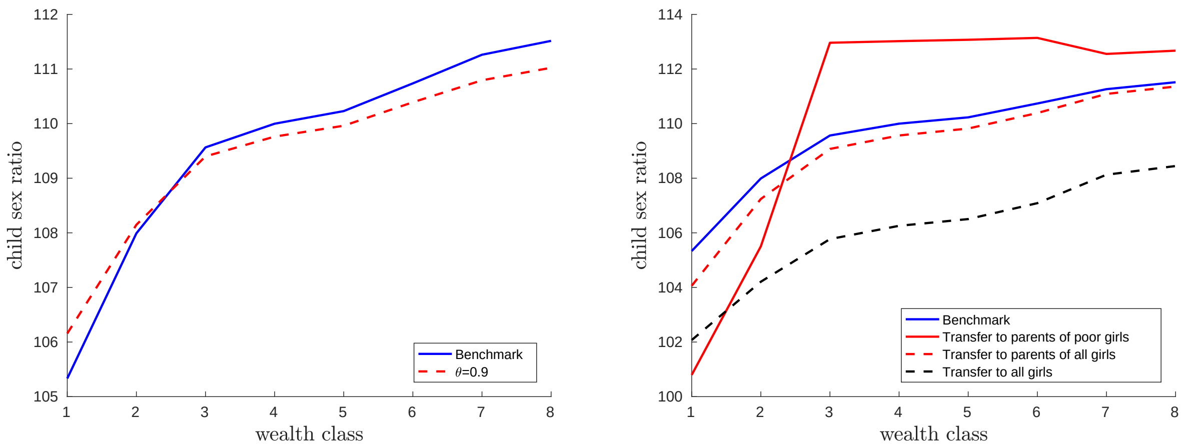
COUNTER-FACTUAL SIMULATIONS. Although dowries have been illegal in India since 1961, families can easily circumvent the law by claiming that the dowry is a gift. Given that the dowry is effectively the price for a boy, one potential policy instrument to reduce the (excess) demand for boys would be a gift tax on the dowry. As described above, such a tax would affect the welfare of both boys' and girls' families in ways that are difficult to determine analytically. To quantify the impact of a gift tax, we consider a policy that levies a 10% tax on the dowry; i.e. the θ parameter in the model is equal to 0.9. While the dowry tax does reduce the fraction of boys in the upper wealth classes, it has the opposite effect in the lower wealth classes in Figure 6a. One reason why this result is obtained is that as sex selection at the top of the wealth distribution decreases, the advantage of marrying up at the bottom is mitigated, and as a result sex selection there increases. This result emphasizes the fact that any policy intervention will affect sex ratios through the marriage market equilibrium channel, with potentially unintended consequences.

The second set of policies that we consider are based on the Conditional Cash Transfer schemes that are currently in place. Sekher (2010) evaluates 15 such schemes. These schemes have a number of common features. Parents receive a cash transfer when (i) the birth of a female child is registered, (ii) she receives the requisite immunizations, and (iii) she achieves specific educational milestones. In addition, an insurance cover is provided, which matures when the girl turns 18 or 20. In the framework of our model, a transfer to the girl's parent is equivalent to an exogenous increase in the wealth of the girl's family. Although governmental transfers when she is young go directly to her parents, the insurance payment, when it matures, goes directly into a bank account that is set up for the girl. Even though the money is in a bank account in the girl's name, however, it is not clear whether the insurance payment should be seen as a wealth transfer to the girl's family or a direct

Figure 6: **Counterfactual Experiments**

(a) Tax

(b) Transfers



transfer to her. We allow for the latter possibility by examining a policy that provides a direct transfer to the married daughter (in addition to her share of the transfer that her husband receives from his parent). Although some of the welfare schemes are available to all families with girls, many are restricted to families below the poverty line. While a Conditional Cash Transfer program will encourage eligible families to have a girl, it will, in addition, affect all families in a caste by shifting the equilibrium marriage price (dowry). Our model, which allows for these pecuniary externality effects, is perfectly suited to examine the impact of programs with restricted eligibility.

The blue solid line in Figure 6b is the benchmark sex ratio (the number of boys per 100 girls) predicted by the model in each wealth class. The first counter-factual policy experiment that we consider is a 20% wealth transfer to families in the bottom two classes with girls. This experiment is meant to reflect the wealth eligibility requirement in many existing schemes. The sex ratio declines substantially in each of the two treated wealth classes. This increase in the number of girls at the bottom of the wealth distribution will shift the entire equilibrium price (dowry) schedule and we see in the figure that this results in an increasingly biased sex ratio in the upper six wealth classes.[43] Combining all wealth classes, the net effect of this scheme is to *worsen* the overall sex ratio.

The next policy experiment that we consider provides the wealth transfer to all girls' parents. To be comparable with the first experiment, the amount of the per family transfer is divided by four (because the beneficiaries are now in 8 rather than 2 of the equal-sized wealth classes). Although the transfer now reduces the sex ratio bias in each wealth class, the magnitude of the effect is small.

The final policy experiment that we consider has the most promise. It is the same as the preceding experiment, except that the subsidy goes directly to the adult girls rather than their

[43]The wealth increase of the girls at the bottom pushes up the dowry for them, but also for those who do not receive the subsidy, since they compete for the same boys. The higher equilibrium dowry leads to more sex selection.

parents. Crucially, the transfer should not be given until the girl is married and it cannot be used as a dowry payment. As we can see in Figure 6b, there is now a substantial increase in the fraction of girls in each wealth class.[44] This is because the (optimal) bequest that must be transferred to the girl through the inefficient dowry mechanism will decline. With less seepage, it is less costly to have a girl. Policies that give resources directly to girls when they are adults, as opposed to their parents when they are children, may thus be especially effective in reducing the bias in child sex ratios in India.

5 Conclusion

Sex selection continues to be pervasive in India; indeed, it has spread and intensified, despite many decades of economic progress. It is widely believed that large marriage payments to the groom's family or dowries are the main cause of son preference in India, especially among the wealthy. We build on this idea to develop a model that links wealth to sex selection through the marriage market. Although dowries and sex selection are positively correlated in the model, sex selection is generated by specific frictions in the marriage market, which arise because of the structure of the marriage institution in India.

The model predicts that *relatively* wealthy households within the relevant marriage market will be more likely to practice sex selection. Using unique data we have collected from rural South India, we find that the probability that a child (aged 0-6) is a girl is decreasing as we move up the wealth distribution within castes, which define independent marriage markets in India. The estimates indicate that the variation in sex ratios in a single (unexceptional) district is comparable in magnitude to the variation across all districts in the country. Aggregate district-level statistics will mask this underlying variation in sex ratios. Based on our findings, sex selection may be more serious and more pervasive than currently believed.

The design of policies to reduce the sex selection problem assumes special significance in light of our findings. One class of policy interventions finds ways to reduce sex selection, taking as given the marriage market frictions that generate the problem. A gift tax on the groom-price or dowry is a seemingly obvious solution to the problem by reducing the demand for grooms (boys). However, any intervention will work through the marriage market affecting the entire equilibrium price (dowry) schedule. Our counter-factual simulations indicate that a gift tax on the dowry would have mixed effects, reducing sex selection higher in the wealth distribution, but increasing it lower down.

A second set of policy interventions in this broad class, which have been implemented by the central government and many states, effectively shift the wealth distribution on the girls' side of the marriage market. Once again, these interventions affect the equilibrium prices in the marriage

[44] This policy experiment is conducted holding constant the α parameter. It is possible that the girl's bargaining power will increase when she has direct control of the resources she brings into the marriage. The resulting decline in α will further increase the fraction of girls from Figure 3a.

market, and the effects are sometimes surprising. Cash transfers to less wealthy parents, conditional on having a girl, do generate an increase in the fraction of girls. However, there is a negative pecuniary externality on wealthier households in their caste, through the accompanying change in the dowry. This results in an overall worsening of the sex ratio. A transfer to all parents, conditional on having a girl, reduces sex selection across the wealth distribution. But the effects are small because the marriage market frictions dampen the incentives of parents to change their behavior. Based on the counter-factual simulations, by far the most effective policy is to give wealth transfers directly to girls when they are adults; forward-looking altruistic parents will take these transfers into account and the resulting equilibrium price schedule leads to a substantial reduction in sex selection.

Ultimately, the most effective and the most sustainable class of policies would address the root causes of the sex selection problem: (i) the social norm that all girls must marry, and (ii) the seepage in the bequest to the girl through the dowry. The seepage is determined by the woman's bargaining power in her marital home, which, in turn, determines the α parameter in the model. One way to simultaneously shift the social norm and increase the woman's bargaining power would be to increase female labor force participation. With economic independence, marriage is no longer the only option, and the woman's bargaining power will thus also increase. Female labor force participation remains extremely low, even in South India, despite large increases in female education. Our analysis indicates that policies that increase female labor force participation would not only generate economic growth, but also improve the sex ratio, through a new channel that has not been previously identified in the literature.[45]

[45]Numerous studies, going back to Rosenzweig and Schultz (1982), have documented the positive effect of female labor force participation on sex ratios. Their interpretation of these findings is that female labor force participation is positively correlated across generations and that the accompanying increased economic returns to having girls reduces sex selection.

Appendix A Tables

Table A1: **Dowry by Gender**

Dependent variable	Dowry	
	(1)	(2)
Female	57.80***	74.44**
	(9.549)	(36.65)
Household wealth	1.573*	–
	(0.934)	
Household wealth squared	0.00180	–
	(0.0129)	
F-test of joint significance of wealth variables (p-value)	0.002	–
Mean of dependent variable	176.58	178.53
Observations	978	1,032
Caste FE	Yes	–
Household FE	–	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A2: **Evidence on Hypergamy**

Partner's family	Wealthier	Less wealthy	Wealthier	Less wealthy
	(1)	(2)	(3)	(4)
Female	0.684***	-0.556***	0.640***	-0.565***
	(0.206)	(0.156)	(0.208)	(0.158)
Child's education	–	–	0.0536*	0.0133
			(0.0302)	(0.0244)
Constant	-1.946***	-0.761***	-2.514***	-0.899***
	(0.176)	(0.110)	(0.371)	(0.276)
Observations	1,025	1,025	1,018	1,018

The reference category is "Same wealth". Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A3: **Female Labour Force Participation and School Enrollment**

Dependent variable	Mother's labor force participation	Girls higher secondary enrollment	Boys higher secondary enrollment
Age range	≤59	14-17	14-17
	(1)	(2)	(3)
Mother's education	-0.0162***	0.0127***	0.0102***
	(0.00029)	(0.00051)	(0.00048)
Father's education	-0.0181***	0.0116***	0.0113***
	(0.00030)	(0.00053)	(0.00049)
Wealth	-0.0476***	0.00979***	0.0131***
	(0.00202)	(0.0023)	(0.0019)
Constant	1.033***	0.626***	0.615***
	(0.0154)	(0.0181)	(0.0151)
Observations	291,151	21,938	23,282

Labor force participation indicates whether the mother is employed. Higher secondary enrollment indicates whether the child is enrolled in school. The lower bound for the age range is set at 14 because most children in rural Tamil Nadu study till the 8^{th} grade (age 13). The upper bound is set at 17 because girls start to marry (and leave their parental homes) by the age of 18. Standard errors (in parentheses) are calculated on the basis of 500 bootstrap replications, *** p<0.01, ** p<0.05, * p<0.1

Table A4: **Alternative Measures of Family Size**

Dependent variable	Girl dummy			
Family size measure	Household size		Predicted kids	
	(1)	(2)	(3)	(4)
Rank in caste per capita wealth distribution	-0.0369***	-0.0696***	-0.0450***	-0.149***
	(0.0063)	(0.0069)	(0.00675)	(0.0103)
Wealth	–	0.0201***	–	0.0368***
		(0.0044)		(0.0051)
Mother's education	–	-0.00060	–	-0.00005
		(0.0004)		(0.00041)
Father's education	–	-0.00005	–	0.00029
		(0.0004)		(0.00042)
Birth order = 1	–	0.0410***	–	0.0819***
		(0.0045)		(0.00600)
Birth order = 2	–	0.0337***	–	0.0588***
		(0.0043)		(0.00492)
Sample mean of dependent variable	0.480		0.480	
Observations	79,027	78,995	78,999	78,995
Caste FE	Yes	Yes	Yes	Yes

Sample restricted to children aged 0-6 years. Reference category for child birth order is third child or higher. The completed family size in Columns 3-4 is predicted in two steps. First, the Ordered Probit model is used to estimate the relationship between the number of children and household wealth, mother's age, mother's age squared, mother's education, father's education. This regression is run for families with all children between the age of 5 and 15, where the family size can be assumed to be complete and accurately measured. The estimated coefficients are used to predict the number of children for all families in the caste, based on their characteristics. In the second step, we replace the actual number of children with the predicted number of children in families where the youngest child is less than 5 years old. Standard errors (in parentheses) are calculated on the basis of 500 bootstrap replications. All standard errors are clustered at the panchayat level, *** p<0.01, ** p<0.05, * p<0.1

Table A5: **Alternative Samples**

Dependent variable	Girl dummy (0-6 years)			
Sample	Benchmark	Dropping biggest caste	Dropping 2 biggest castes	Dropping smaller castes
	(1)	(2)	(3)	(4)
Rank in caste per capita wealth distribution	-0.0457***	-0.0408***	-0.0386***	-0.0541***
	(0.0068)	(0.0068)	(0.0104)	(0.0071)
Sample mean of dependent variable	0.480	0.482	0.481	0.480
Observations	79,027	49,522	29,883	69,233
Caste FE	Yes	Yes	Yes	Yes

Standard errors (in parentheses) are calculated on the basis of 500 bootstrap replications. All standard errors are clustered at the panchayat level, *** p<0.01, ** p<0.05, * p<0.1

Table A6: **Within and Between Variation in Sex Ratios**

Dependent variable	Fraction of girls			
Sample	12 castes		30 castes	
	(1)	(2)	(3)	(4)
Wealth class				
1	0.0247***	0.0250***	0.0248***	0.0249***
	(0.00889)	(0.00759)	(0.00851)	(0.00745)
2	0.0139	0.0140	0.00973	0.00964
	(0.00978)	(0.00977)	(0.00979)	(0.0101)
3	-0.0119	-0.0117	-0.0138	-0.0136
	(0.0129)	(0.0117)	(0.0119)	(0.0109)
4	-0.0197**	-0.0195**	-0.0190**	-0.0194**
	(0.00848)	(0.00767)	(0.00824)	(0.00750)
5	-0.0232**	-0.0230***	-0.0233**	-0.0231***
	(0.00979)	(0.00824)	(0.00927)	(0.00802)
6	-0.0108	-0.0108	-0.00958	-0.00957
	(0.00905)	(0.00936)	(0.00859)	(0.00886)
7	-0.00204	-0.00208	-0.00129	-0.00187
	(0.0104)	(0.0111)	(0.00972)	(0.0104)
Constant (wealth class 8)	0.484***	0.484***	0.485***	0.485***
	(0.00789)	(0.00680)	(0.00751)	(0.00666)
Observations	96	96	237	237
R-squared	0.392	0.453	0.231	0.332
Caste FE	No	Yes	No	Yes

Sample restricted to children aged 0-6 years. Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Appendix B Omitted Proofs

B.1 Proof of Lemma 1

Proof. The solution to the girl's family maximization problem is to choose the optimal x that maximizes the family's utility. The first order condition is:

$$v_x + v_u u' = 0 \quad \Rightarrow \quad u' = -\frac{v_x}{v_u}. \tag{19}$$

Then the surplus is supermodular and the allocation will be PAM (see Legros and Newman (2007) and Chade, Eeckhout, and Smith (2017)) provided:

$$\frac{\partial^2 v(x,y,u)}{\partial x \partial y} = v_{xy} + v_{uy}u' > 0 \quad \Rightarrow \quad v_{xy} > \frac{v_x}{v_u} v_{uy}. \tag{20}$$

Calculating each of the derivatives, we get:

$$\begin{aligned} v_x &= \frac{1}{x+y-2\psi(u)} && (21)\\ v_u &= \frac{-2\psi'(u)}{x+y-2\psi(u)} + \frac{\psi'(u)}{\psi(u)} = \frac{x+y-4\psi(u)}{2(x+y-2\psi(u))} && (22)\\ v_{xy} &= \frac{-1}{(x+y-2\psi(u))^2} && (23)\\ v_{uy} &= \frac{\psi(u)}{(x+y-2\psi(u))^2}, && (24) \end{aligned}$$

where we have used the fact that $\psi'(u) = \frac{e^{u/2}}{2\sqrt{\alpha}}$ and $\frac{\psi'(u)}{\psi(u)} = \frac{1}{2}$. Then there is PAM provided:

$$\frac{-1}{(x+y-2\psi(u))^2} > \frac{\frac{1}{x+y-2\psi(u)}}{\frac{x+y-4\psi(u)}{2(x+y-2\psi(u))}} \times \frac{\psi(u)}{(x+y-2\psi(u))^2} \tag{25}$$

or equivalently

$$\frac{x+y-2\psi(u)}{x+y-4\psi(u)} < 0 \tag{26}$$

The numerator is equivalently to $y-d$ (and the denominator to $(y-d)-(x+d)$). Clearly the girl's family cannot pay a dowry more than their wealth so the numerator here is positive. Hence, for PAM we must have $x+y < 4\psi(u)$, or the net wealth of the girls' family after giving dowry must be less than that of the boys after the transaction ($y-d < x+d$). This is true when:

$$x + y < 4\sqrt{\frac{e^u}{\alpha}} \tag{27}$$

or

$$u > 2\log\left(\frac{\sqrt{\alpha}}{4}(x+y)\right). \tag{28}$$

A sufficient condition for this to be satisfied is that the equilibrium payoff exceeds the outside option: $u(x) \geq 2\log\left(\frac{x}{2}\right)$ or

$$\frac{\sqrt{\alpha}}{4}(x+y) \leq \frac{x}{2} \tag{29}$$

$$\sqrt{\alpha}(x+y) \leq 2x \tag{30}$$

For $x = y = \overline{x}$, this is true when $\sqrt{\alpha} \leq 1$. Since $\alpha \in [0,1]$, we have PAM at the top for any α. For all other x, y the inequality is satisfied when $y < x\left(\frac{2}{\sqrt{\alpha}} - 1\right)$. This establishes the proof. ■

B.2 Proof of Proposition 2

Proof. The extent of sex selection is given by $k^\star(y)$:

$$k^\star(y) = u(y) - v(\mu(y), y, u(\mu(y))). \tag{31}$$

We need to show that $k^\star(y)$ is increasing in y or

$$k^{\star'}(y) = u'(y) - \left(v_x\mu' + v_y + v_u u'\mu'\right). \tag{32}$$

From the first order condition (7), along the equilibrium matching $\mu(y)$, it must be that $v_x + v_u u' = 0$, so the derivative can be written as:

$$k^{\star'}(y) = u'(y) - \left((v_x + v_u u')\mu' + v_y\right) \tag{33}$$

$$= u'(y) - v_y(\mu, y, u(\mu)). \tag{34}$$

This is increasing provided:

$$\frac{-2}{y + \mu^{-1}(y) - 4\psi(u(y))} - \frac{1}{\mu(y) + y - 2\psi(u(\mu(y)))} > 0, \tag{35}$$

since $u' = -\frac{v_x}{v_u}$ from the First-Order Condition (19) and substituting for v_x and v_u from equations (21) and (22), and v_y obtains from partially differentiating expression (6).

1. At the top of the wealth distribution. At $x = \overline{x}$, under positive sorting we have $\overline{y} = \mu(\overline{x}) = \overline{x}$. Then condition (35) can be written as:

$$\frac{-2}{2\overline{y} - 4\psi(u(\overline{y}))} - \frac{1}{2\overline{y} - 2\psi(u(\overline{y}))} > 0 \tag{36}$$

or

$$\psi(u) < \frac{3}{4}\overline{x} \tag{37}$$

$$\sqrt{\frac{e^u}{\alpha}} < \frac{3}{4}\overline{x} \tag{38}$$

$$u(\overline{x}) < 2\log\frac{3\overline{x}\sqrt{\alpha}}{4}. \tag{39}$$

Now we know that $u(\overline{x}) = \log\left(\frac{(\overline{x}+d)\sqrt{\alpha}}{2}\right)^2$. Therefore a sufficient condition for $k^\star$ increasing at $\overline{x}$ is:

$$\frac{(\overline{x}+d)\sqrt{\alpha}}{2} < \frac{3\overline{x}\sqrt{\alpha}}{4} \tag{40}$$

or $d < \frac{\overline{x}}{2}$.

2. At the bottom of the wealth distribution. We verify the condition at the lower bound $\underline{y}$, where $\mu(\underline{y}) = x^\star$. Now $\mu^{-1}(\underline{y})$ is not defined as it outside of the range of matches. But we know that for the unmatched men, the utility is $u(x) = 2\log\frac{x}{2}$ for all $x \in [\underline{x}, x^\star]$. Therefore $u'(x) = \frac{4}{x}$ in that range. At an income level $\underline{y}$, we can then write condition (34) as:

$$\frac{4}{\underline{y}} - \frac{1}{x^\star + \underline{y} - 2\psi(u(x^\star))} > 0. \tag{41}$$

Observe that the marginal man $x^\star$ is indifferent between the outside opinion and being matched, so that $u(x^\star) = 2\log\frac{x^\star}{2}$. After rearranging, we therefore obtain:

$$\alpha > \left(\frac{4x^\star}{4x^\star + 3\underline{y}}\right)^2. \tag{42}$$

This is satisfied for large enough α. $x^\star$ is a function of the distribution Γ and the parameters α and a, and $x^\star \in [\underline{y}, \overline{y}]$. So there is a critical $\tilde{\alpha} \in \left(0.327, \left(\frac{4\overline{y}}{4\overline{y}+3\underline{y}}\right)^2\right)$ such that for all $\alpha > \tilde{\alpha}$ sex selection is increasing at the bottom of the distribution. To see this, at $x^\star = \underline{y}$ this condition implies $\tilde{\alpha} = \left(\frac{4}{7}\right)^2 = 0.327$, and if $x^\star = \overline{y}$ then $\tilde{\alpha} = \left(\frac{4\overline{y}}{4\overline{y}+3\underline{y}}\right)^2$. ■

References

ANDERSON, S. (2003): "Why Dowry Payments Declined with Modernization in Europe but are Rising in India," *Journal of Political Economy*, 111(2), 269–310.

——— (2007): "The Economics of Dowry and Brideprice," *The Journal of Economic Perspectives*, 21(4), 151–174.

ANDERSON, S., AND C. BIDNER (2015): "Property Rights over Marital Transfers," *The Quarterly Journal of Economics*, 130(3), 1421–1484.

ANDERSON, S., AND D. RAY (2010): "Missing Women: Age and Disease," *The Review of Economic Studies*, 77(4), 1262–1300.

ARNOLD, F., M. K. CHOE, AND T. K. ROY (1998): "Son Preference, the Family-Building Process and Child Mortality in India," *Population Studies*, 52(3), 301–315.

ARNOLD, F., S. KISHOR, AND T. ROY (2002): "Sex-Selective Abortions in India," *Population and Development Review*, 28(4), 759–785.

ATTANASIO, O. P., AND G. WEBER (1993): "Consumption Growth, the Interest Rate and Aggregation," *The Review of Economic Studies*, 60(3), 631–649.

BASU, A. M. (1999): "Fertility Decline and Increasing Gender Imbalance in India, Including a Possible South Indian Turnaround," *Development and Change*, 30(2), 237–263.

BECKER, G. S. (1973): "A theory of Marriage: Part I," *Journal of Political economy*, 81(4), 813–846.

BHALOTRA, S. R., A. CHAKRAVARTY, AND S. GULESCI (2016): "The Price of Gold: Dowry and Death in India," Essex Working paper.

BHASKAR, V. (2011): "Sex Selection and Gender Balance," *American Economic Journal: Microeconomics*, 3(2), 252–253.

BHAT, P. M., AND S. S. HALLI (1999): "Demography of Brideprice and Dowry: Causes and Consequences of the Indian Marriage Squeeze," *Population Studies*, 53(2), 129–148.

BLUNDELL, R., M. BROWNING, AND C. MEGHIR (1994): "Consumer Demand and the Life-Cycle Allocation of Household Expenditures," *The Review of Economic Studies*, 61(1), 57–80.

BOTTICINI, M. (1999): "A Loveless Economy? Intergenerational Altruism and the Marriage Market in a Tuscan Town, 1415–1436," *The Journal of Economic History*, 59(01), 104–121.

Botticini, M., and A. Siow (2003): "Why Dowries?," *American Economic Review*, 93(4), 1385–1398.

Caldwell, J. C., P. H. Reddy, and P. Caldwell (1983): "The Causes of Marriage Change in South India," *Population Studies*, 37(3), 343–361.

Chade, H., J. Eeckhout, and L. Smith (2017): "Sorting through Search and Matching Models in Economics," *Journal of Economic Literature*, 55(2), 493–544.

Choo, E., and A. Siow (2006): "Who marries whom and why," *Journal of political Economy*, 114(1), 175–201.

Chunkath, S. R., and V. B. Athreya (1997): "Female Infanticide in Tamil Nadu: Some Evidence," *Economic and Political Weekly*, pp. WS21–WS28.

Das Gupta, M. (1987): "Selective Discrimination against Female Children in Rural Punjab, India," *Population and Development Review*, pp. 77–100.

Das Gupta, M., J. Zhenghua, L. Bohua, X. Zhenming, W. Chung, and B. Hwa-Ok (2003): "Why is Son Preference so Persistent in East and South Asia? A Cross-Country Study of China, India and the Republic of Korea," *The Journal of Development Studies*, 40(2), 153–187.

Diamond-Smith, N., N. Luke, and S. McGarvey (2008): "Too many girls, too much dowry: Son Preference and Daughter Aversion in Rural Tamil Nadu, India," *Culture, Health and Sexuality*, 10(7), 697–708.

Dyson, T., and M. Moore (1983): "On kinship Structure, Female Autonomy, and Demographic Behavior in India," *Population and Development Review*, pp. 35–60.

Edlund, L. (1999): "Son Preference, Sex Ratios, and Marriage Patterns," *Journal of Political Economy*, 107(6), 1275–1304.

Foster, A., and M. Rosenzweig (2001): "Missing women, the marriage market and economic growth," Brown mimeo.

George, S., R. Abel, and B. D. Miller (1992): "Female Infanticide in Rural South India," *Economic and Political Weekly*, pp. 1153–1156.

Guilmoto, C. Z. (2009): "The Sex Ratio Transition in Asia," *Population and Development Review*, 35(3), 519–549.

Jayachandran, S. (2017): "Fertility Decline and Missing Women," *American Economic Journal: Applied Economics*, 9(1), 118–139.

JEFFERY, R., P. JEFFERY, AND A. LYON (1984): "Female Infanticide and Amniocentesis," *Social Science & Medicine*, 19(11), 1207–1212.

JEJEEBHOY, S. J., AND Z. A. SATHAR (2001): "Women's Autonomy in India and Pakistan: the Influence of Religion and Region," *Population and Development Review*, 27(4), 687–712.

JHA, P., R. KUMAR, P. VASA, N. DHINGRA, D. THIRUCHELVAM, AND R. MOINEDDIN (2006): "Low Male-to-Female (sic) Sex Ratio of Children Born in India: National Survey of 1· 1 million Households," *The Lancet*, 367(9506), 211–218.

KAPADIA, K. (1993): "Marrying Money: Changing Preference and Practice in Tamil Marriage," *Contributions to Indian Sociology*, 27(1), 25–51.

——— (1995): *Siva and her Sisters: Gender, Caste, and Class in Rural South India.* Westview Press, Inc.

KISHOR, S. (1993): "" May God Give Sons to All": Gender and Child Mortality in India," *American Sociological Review*, pp. 247–265.

KRISHNAJI, N. (1987): "Poverty and Sex Ratio: Some Data and Speculations," *Economic and Political Weekly*, pp. 892–897.

LEGROS, P., AND A. F. NEWMAN (2007): "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities," *Econometrica*, 75(4), 1073–1102.

LINDENLAUB, I. (2017): "Sorting multidimensional types: Theory and application," *Review of Economic Studies*, forthcoming.

MILLER, B. D. (1981): *The Endangered Sex: Neglect of Female Children in Rural North India.* Cornell University Press.

MOORJANI, P., K. THANGARAJ, N. PATTERSON, M. LIPSON, P.-R. LOH, P. GOVINDARAJ, B. BERGER, D. REICH, AND L. SINGH (2013): "Genetic evidence for Recent Population Mixture in India," *The American Journal of Human Genetics*, 93(3), 422–438.

MUNSHI, K., AND M. ROSENZWEIG (2016): "Networks and Misallocation: Insurance, Migration, and the Rural-Urban Wage Gap," *The American Economic Review*, 106(1), 46–98.

MURTHI, M., A.-C. GUIO, AND J. DRÈZE (1995): "Mortality, fertility, and gender bias in India: A district-level analysis," *Population and Development Review*, pp. 745–782.

RAHMAN, L., AND V. RAO (2004): "The Determinants of Gender Equity in India: Examining Dyson and Moore's Thesis with New Data," *Population and Development Review*, 30(2), 239–268.

RAO, V. (1993a): "Dowry 'Inflation'in Rural India: A Statistical Investigation," *Population Studies*, 47(2), 283–293.

——— (1993b): "The Rising Price of Husbands: a Hedonic Analysis of Dowry Increases in Rural India," *Journal of Political Economy*, 101(4), 666–677.

ROSENZWEIG, M. R., AND T. P. SCHULTZ (1982): "Market opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India," *The American Economic Review*, 72(4), 803–815.

SEKHER, T. (2010): "Special Financial Incentive Schemes for the Girl Child in India: a Review of Select Schemes," *United Nations Population Fund*.

SEN, A. (1990): "More than 100 million women are missing," *The New York Review of Books*.

SRINIVAS, M. N. (1984): *Some Reflections on Dowry*. Published for the Centre for Women's Development Studies, New Delhi, by Oxford University Press.

——— (1989): *The Cohesive Role of Sanskritization and other essays*. Oxford University Press, USA.

SRINIVASAN, S., AND A. S. BEDI (2009): "Girl Child Protection Scheme in Tamil Nadu: an Appraisal," *Economic and Political Weekly*, pp. 10–12.

SUDHA, S., AND S. I. RAJAN (1999): "Female Demographic Disadvantage in India 1981–1991: Sex Selective Abortions and Female Infanticide," *Development and Change*, 30(3), 585–618.