

WEALTH, MARRIAGE, AND SEX SELECTION*

Girija Borker[†] Jan Eeckhout[‡] Nancy Luke[§] Shantidani Minz[¶] Kaivan Munshi^{||}
Soumya Swaminathan^{**}

May 12, 2017

Abstract

Sex selection continues to be pervasive in India, despite many decades of economic progress. It is widely believed that large dowries are the main cause of son preference in India. Although dowries and sex selection are positively correlated in our model, the root cause of sex selection is specific frictions in the marriage market. The model delivers the prediction that sex selection is increasing with relative wealth. 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. The main empirical finding is that the probability that a child is a girl is decreasing as we move up the wealth distribution within castes, which define independent marriage markets in India. The variation in sex ratios within castes in a single (unexceptional) district is found to be as large as the variation across all districts in the country. 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.

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.

Given the persistence and the spread of sex selection in India, the question that we ask in this paper is which sub-populations are most affected by the problem and why? 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)). 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 serious 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. The hypothesized relationship between wealth and sex selection, moreover, has not been subjected to rigorous empirical scrutiny.¹

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: (i) marriages occur almost exclusively within castes, (ii) marriages are patrilocal, with women moving into their husbands’ homes, (iii) marriages are arranged, with family wealth often being a major consideration, (iv) marriages involve a payment or dowry to the groom’s family, and (v) the social norm is that all girls must marry. We incorporate these features of the marriage institution in the model and in the empirical analysis that follows.

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, parents must

¹A review of the literature uncovered just three published studies, using data that are now 30 years out of date, that have estimated this relationship; Rosenzweig and Schultz (1982), Krishnaji (1987), and Murthi, Guio, and Drèze (1995). Due to data limitations, these studies are unable to directly estimate the relationship between child sex ratios and household wealth. Moreover, this relationship is not the primary focus of two of the three studies and is only estimated tangentially. As Murthi, Guio, and Drèze (1995) themselves note, “there is a widespread hunch that discrimination against female children is less intense among poorer households [...] but the empirical basis for these observations remains limited.” (p. 755)

use the dowry as a mechanism to share wealth with their daughters. However, 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. 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. Although the marriage market clears instantaneously (without search frictions), it nevertheless generates distortions that result in sex selection at every wealth level. Sex selection in our model is generated by two institutional frictions in the marriage market: (i) the social norm that all girls must marry, and (ii) the seepage in the bequest described above. These frictions leave parents worse off with girls than with boys, resulting in sex selection even when parents do not have an intrinsic preference for sons.

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 solved for 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 wealth with unique data we have recently collected as part of the South India Community Health Study (SICHS),

²This dual role of the dowry as a bequest and a price distinguishes our model from previous models of marriage with dowries. In Botticini and Siow (2003) the dowry served only as a bequest. In Anderson and Bidner (2015), the dowry serves both roles, but two separate instruments are available. The advantage of our specification is that it allows us to solve simultaneously for dowries and sex selection

³To clarify this argument, suppose that the boy-to-girl ratio is 2 at every wealth level. There are 200 boys with wealth 1, 2,..., 99, 100 (two at each level) and 100 girls with wealth 1, 2, . . . , 99, 100 (one at each level). Note that except for the number of boys and girls, the wealth distribution is the same and uniform on [0,100]. Then under positive assortative matching, a boy with wealth 100 marries the girl with wealth 100, another boy with wealth 100 marries the girl with wealth 99, a boy with wealth 99 marries the girl with wealth 98,..., and the last boy to be matched, with wealth 50, marries the girl with wealth 1. The key insight is that except at the top, poorer girls marry richer boys, and that the wealth-gap is larger for poorer girls: at the top (100,100) there is no wealth gap and at the bottom (50,1) the wealth gap is 49.

⁴Previous 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, which is at odds with the data. The additional advantage of our model, in which sex selection is linked to marriage market frictions, is that it is well suited to evaluate the impact of existing and potential policies designed to reduce sex selection.

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 SICHHS; a census of all 298,000 households drawn from 57 castes residing in the study area and a detailed survey of 5,000 representative households. In general, extremely large data sets are needed to detect sex selection with the required level of statistical confidence. The SICHHS 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, which define independent marriage markets in India.

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 emergence of caste-based 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 SICHHS census, is 107, which is clearly biased, but not exceptional for South India or the country as a whole. 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, the worst districts in the country with respect to sex selection have sex ratios around 120. Thus, there is as much variation in sex ratios within castes in a single (unexceptional) district as there is 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 the caste 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. 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.⁵ Although such a policy has not been 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.

In order to evaluate the impact of alternative policy interventions, we structurally estimate the marital sorting model that we laid out above using the SICHHS 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

⁵Although 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.

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 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.

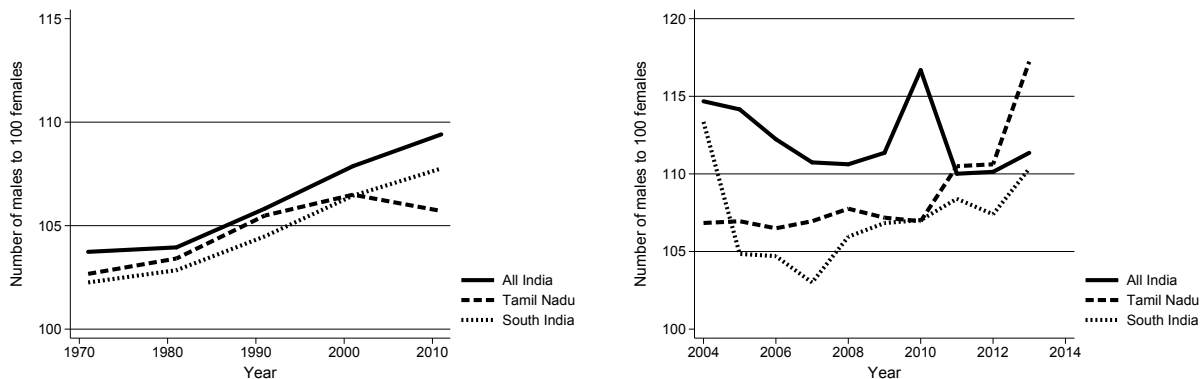
2 Institutional Setting

2.1 Sex Selection

THE PERSISTENCE AND THE SPREAD OF SEX SELECTION. Abnormally high male-female sex ratios have been observed in certain parts of India, particularly the Northwest, since the first British-Indian census in 1871. The consensus in the literature has been that these early findings reflected a broader regional divide, with substantially greater gender bias in the North than in the South. Dyson and Moore (1983) argue in an influential paper that differences in sex ratios, as well as differences in women’s status more generally, can be traced to variation in the kinship structure across the two regions. Women in the North traditionally married outside their natal village into unrelated families (albeit from the same caste). This was seen to result in high dowries, particularly among wealthy landowning castes, which, in turn, resulted in sex selection.

Recent scholarship has challenged the notion of a North-South divide for two reasons. First, there has been evidence of a regional convergence in sex ratios since the 1980’s (Basu (1999), Srinivasan and Bedi (2009)). Second, the differences in the kinship structure that were believed to be responsible for the inter-regional variation in sex ratios have started to disappear. For example, Rahman and Rao (2004) find that the likelihood of girls marrying outside their natal village and the dowry payment are statistically indistinguishable in their sample of villages located in the northern state of Uttar Pradesh and the southern state of Karnataka, respectively. Moreover, numerous studies document a switch from bride-price to groom-price in the South in recent decades and an

associated decline in co-sanguinous marriages; e.g. Caldwell, Reddy, and Caldwell (1983), Kapadia (1993).



(a) Sex ratio (aged 0-6) over time.
(Source: Population Census)

(b) Sex ratio at birth over time.
(Source: Sample Registration System)

Figure 1: Sex Ratios.

Even as sex ratios in the South have been converging towards Northern levels, there has been an increasing bias in these ratios for the country as a whole. This secular trend has been attributed to various factors including reduced fertility; e.g. Basu (1999), improved sex selection technologies; e.g. Arnold, Kishor, and Roy (2002), changes in the economic returns to having boys versus girls; e.g. Foster and Rosenzweig (2001), and dowry inflation; e.g. Das Gupta and Mari Bhat (1997).⁶ Figure 1a examines these trends by plotting the child (aged 0-6) sex ratio over the 1971-2011 period for (i) all Indian states, (ii) southern states, which include Tamil Nadu, Karnataka, Andhra Pradesh, and Karnataka, but exclude Kerala (which is an outlier on many socioeconomic indicators), and (iii) the state of Tamil Nadu where our analysis is situated. The sex ratio is computed as the number of males per 100 females to be consistent with Government of India statistics. The natural sex ratio at birth for this statistic ranges from 103 to 106 (Guilmoto (2009)). It is well known that subsequent mortality favors girls, but the natural sex ratio for 0-6 year olds is not readily available from existing studies. Sex selection appears to have commenced in South India concurrently with the emergence of caste-based 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. It is evident from Figure 1a that there has been a steadily worsening trend in sex ratios over the past four decades, with a mild convergence between the South and the all-India average. Sex selection is evidently a persistent and pervasive problem in India.

⁶Although marriage market frictions generate son preference in our analysis, another popular (potentially co-existing) explanation for son preference 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. Exogenous fertility decline makes the sex ratios worsen earlier, resulting in an overall increase in gender bias.

While the sex ratio worsens relatively rapidly in Tamil Nadu between 1971 and 1991, there is a surprising reversal between 2001 and 2011, which goes against the regional trend. To examine this reversal in greater detail, we turn to the Sample Registration System (SRS), which provides state-level vital statistics over the 2004-2013 period and can thus be used to compare the Sex Ratio at Birth (SRB) in Tamil Nadu with other Indian states. 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. We see in Figure 1b that the Tamil Nadu SRB remains stable from 2004 to 2010, broadly matching the census statistics. However, it worsens sharply in the most recent years, and based on vital registrations, sex selection in Tamil Nadu, and the South Indian region overall, is now around the national average.

SEX SELECTION IN THE CROSS-SECTION. Our objective in this paper is not to explain changes in sex selection over time. Given the persistence of sex selection, we are interested in identifying sub-populations within which this phenomenon is most acute. The general perception is that sex selection is most severe among wealthy households and in upper castes (Miller (1981), Basu (1999)). Wealthy landowning castes, notably the Jats and the Rajputs, have been consistently associated with excessively biased sex ratios since the earliest British-Indian censuses (Jeffery, Jeffery, and Lyon (1984), Krishnaji (1987), Das Gupta (1987)). Although these castes are located in the North, sex selection in the South has also been traditionally associated with specific castes, such as the Gounders and the Kallars (George, Abel, and Miller (1992), Chunkath and Athreya (1997)). However, just as 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)).

Although sex ratios are also believed to worsen with wealth, this relationship has not been subjected to rigorous empirical scrutiny. Our review of the literature uncovered just three published studies, using data that are now nearly 30 years out of date, that explicitly estimate this relationship.

Krishnaji (1987) documents a greater deficit of females in landowning families and in households with higher per capita consumption, using data from the Rural Labour Enquiry, 1963-65 and 1974-75, and the National Sample Survey, 1973-74, respectively. A limitation of her analysis is that it includes children and adults. Adults who have migrated from their rural homes will be missing in her data and Anderson and Ray (2010) document excessive mortality among adult females relative to adult males in India. An alternative explanation for Krishnaji's findings, therefore, is that there is differential male out-migration or differential adult mortality by gender across households with different levels of wealth.

Murthi, Guio, and Drèze (1995) use district-level data from the 1981 population census to estimate the relationship between wealth, measured by the fraction of the population living below the poverty line, and the child (aged 0-6) sex ratio. Conditional on a large set of covariates, sex ratios are found to worsen with wealth. One limitation of their analysis is that it uses a coarse

measure of wealth. A second limitation is that it is based on aggregate, district-level, data.

Rosenzweig and Schultz (1982) overcome the limitations of the preceding studies, using both district-level from the 1961 census and 1971 household-level data to estimate the (conditional) relationship between wealth and child survival by gender. In lieu of a direct measure, wealth is proxied by land ownership and variables associated with agricultural productivity. Although the primary focus of their analysis is on the relationship between female labor force participation and child sex ratios, the wealth variables are included in the estimating equations as covariates. In the household level analysis, landed households are found to have a higher proportion of girls, whereas at the district level, the relationship is reversed. Given the limitations of any district-level analysis, more weight should be placed on the household-level results. In line with these results, our own household-level analysis indicates that absolute household wealth is associated with an increase in the probability that a child is a girl. However, what matters for our theory is relative wealth within the caste, which we expect will reduce the probability that a child is a girl, and that variable has not been included in previous analyses.

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, 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 and demographic indicators. Vellore district is also representative of Tamil Nadu state with respect to child sex ratios. Figure 2 reports child (aged 0-6) sex ratios, by district, in the three most recent population census rounds; 1991, 2001, and 2011. Although the sex ratio in Vellore district is higher than the natural statistic of 102.5, it is fairly stable over time; just under 105 in 1991, between 105 and 106 in 2001, and 106 in 2011.

Vellore is roughly at the median with respect to child sex ratios across districts in Tamil Nadu in all three census rounds. 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 107. None of the castes in Tamil Nadu that have been traditionally associated with 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.

2.2 The Marriage Institution

MARRIAGE STRUCTURE AND ITS EVOLUTION: The following features of marriage in India are relevant for our analysis. First, marriages are endogamous, matching individuals almost exclusively within their caste. Second, marriage is patrilocal, with women moving into their husbands homes, which are often in a different village. Third, marriages are arranged by the parents and relatives

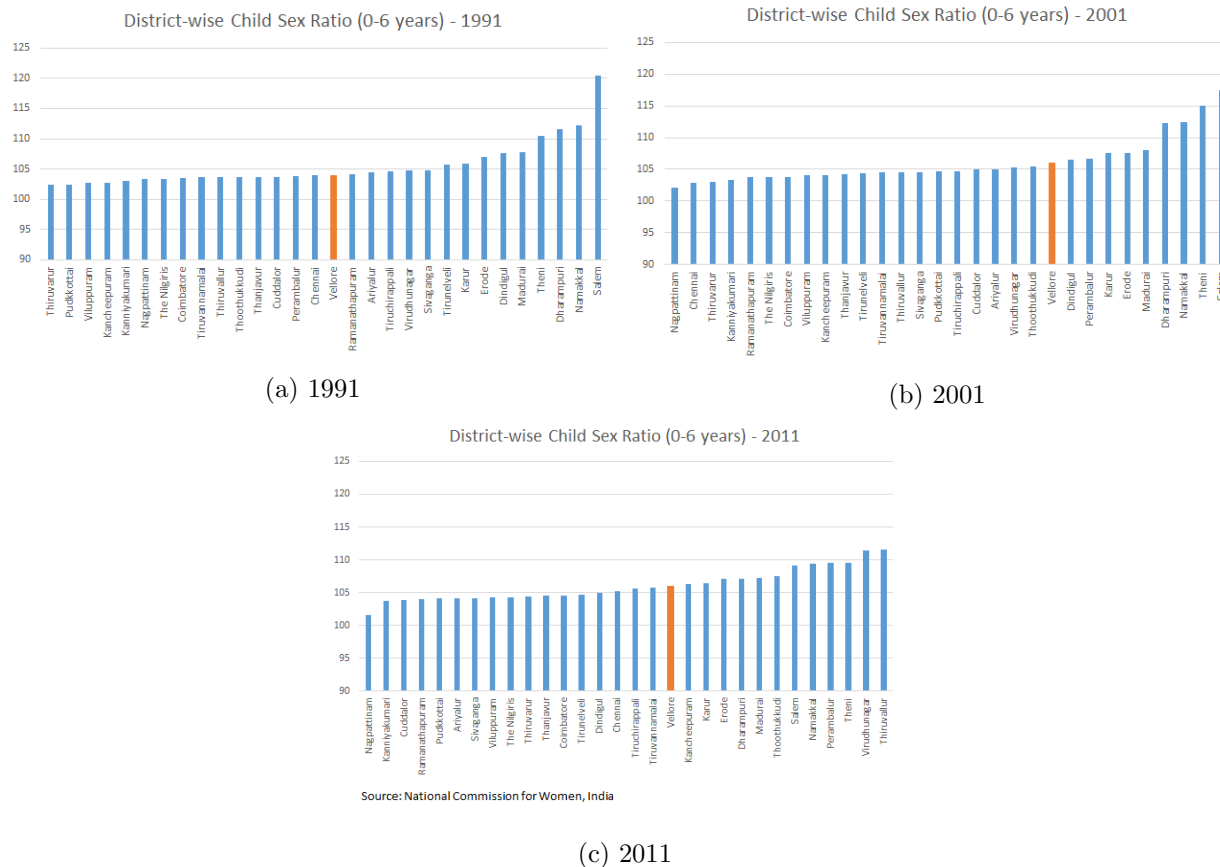


Figure 2: Sex Ratios by District over 3 Census Rounds.

of the groom and bride. Fourth, marriages involve a transfer payment or dowry from the bride’s family to the groom’s family. Fifth, 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 are common to the entire country and have remained stable over time. Other features have traditionally varied between the South and the North but have converged in recent decades.

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 nearly 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)). What

has changed are certain features of marriage in South India. Dyson and Moore (1983)'s regional dichotomy emphasizes two traditional features of marriage in South India: marriage to close kin and marriage within the natal village. As noted, the available evidence indicates that this regional distinction may have largely disappeared.

Perhaps the most visible change in the marriage institution, all over India but especially in the South, is associated with the practice of dowry. Marriages in South India were characterized by transfers from the groom's side to the bride prior to the 1980's (Caldwell, Reddy, and Caldwell (1983), Srinivas (1989)). The direction of the transfers has now been reversed and dowries in South India are 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 now spread across the caste distribution (Bhat and Halli (1999)).⁷

A number of explanations have been proposed for the spread of this practice through the Indian population and for the accompanying dowry inflation. First, in a post-Independence economy that provided greater opportunities for economic and social mobility, the lower castes may have attempted to emulate the upper castes by adopting their traditional practices (Srinivas (1989)). This *Sanskritization* hypothesis does not explain dowry inflation in upper castes where the practice was well established. Second, if men marry younger women, then there will be a shortage of men of marriageable age in a growing population (Caldwell, Reddy, and Caldwell (1983), Rao (1993b), Bhaskar (2011)). This *marriage-squeeze* hypothesis does not explain dowry inflation in states like Tamil Nadu, where fertility rates have been below replacement since the mid-1990's. Third, if there is increased inequality in male incomes within the caste, then the increased competition for eligible men will generate an increase in dowries (Caldwell, Reddy, and Caldwell (1983), Srinivas (1984), Anderson (2003)). This *hypergamy* hypothesis can explain dowry inflation in the dynamic South Indian economy, and is also useful in explaining cross-sectional variation in dowry levels within castes.

MARRIAGE IN THE STUDY AREA. The analysis in this paper makes use of two components of the SICHHS; a census of all households and a detailed survey of 5,000 households who are representative of the castes in the study area.⁸ 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, (iii) whether the girl left her natal village when she moved, and (iv) whether the marriage was arranged. This information was collected from the male primary earner for his own marriage, his parents' marriage, and for the marriages of his children in the five years preceding the survey.

⁷Dowries 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 (Caldwell, Reddy, and Caldwell (1983))

⁸The sampling frame for the household survey included all ever-married men aged 25-60 in the SICHHS census plus 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.

Table 1: Marriage Patterns

Generation	Grandparents	Parents	Children	
	(1)	(2)	Males (3)	Females (4)
Same caste	–	0.98	0.95	0.97
Related	0.49	0.48	0.35	0.38
Same village	0.32	0.25	0.22	0.19
Arranged	–	0.85	0.79	0.87
Observations	2,862	2,862	322	504

Source: SICHHS household survey

Table 1 provides information on marriages over the three generations based on data from the household survey.⁹ In line with nationally representative survey evidence and genetic evidence for the country as a whole, 98% of the parents and 96% of the children married within their caste. The incidence of close-kin marriage declines from 49% in the grandparents' generation and 48% in the parents' generation to 37% in the current generation. Marriage within the natal village also declines from 32% to 25% to 20% over the three successive generations. While some features of the marriage institution have changed, one aspect has stayed the same; approximately 85% of marriages were arranged by parents and relatives in the current generation and the previous generation. Marriage in the study area today can be described as almost always within the caste but not necessarily with close-kin, typically outside the girl's natal village, and usually arranged.

2.3 Wealth and Marriage

Marriages in South India were traditionally between close-kin. The most preferred match for a girl was her mother's younger brother or, if he was unavailable, one of her mother's brothers' sons (Kapadia (1995)). Given that marriage is patrilocal, this meant that girls moved back and forth 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)). Having a girl did not put parents at a disadvantage in this cooperative arrangement and

⁹The 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 our sample.

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. A marriage market consequently emerged to match unrelated families within the caste, with a marriage price or dowry clearing the market. The division of the surplus from marriage in this market apparently favored the boys, resulting in sex selection. Our theoretical model takes as given the presence of a marriage market, but makes precise the connection between the marriage market and sex selection.

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. Thus, while dowries and sex selection will also be correlated in the cross-section, within the caste, they are jointly determined by family wealth in our model. Although the relationship between wealth and sex selection has not been subjected to rigorous scrutiny and will be the focus of our empirical analysis, using the SICHS census data, the positive relationship between wealth and dowries is well established; e.g. Rao (1993a). We use the household survey to corroborate these findings and to look in greater detail at the pattern of dowries and marital matching across households with different levels of wealth.

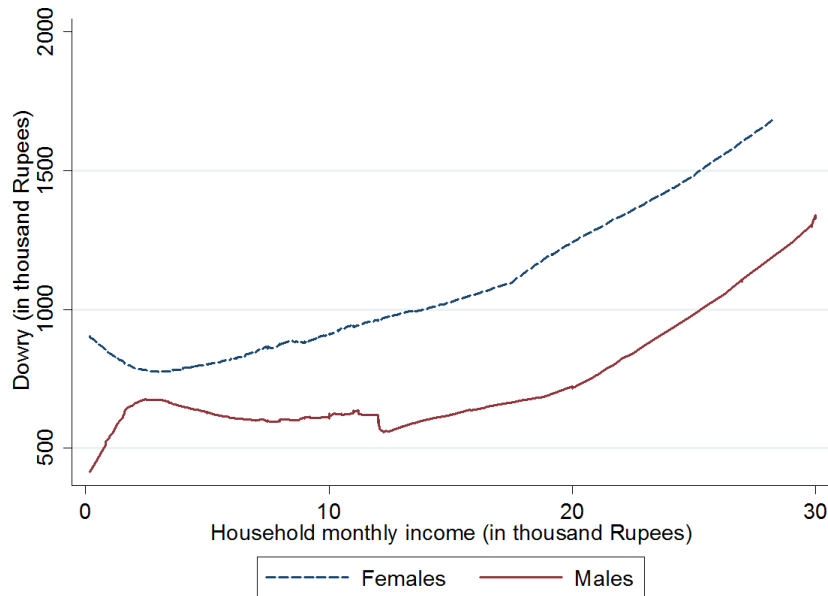


Figure 3: Dowry by Monthly Income.

Figure 3 estimates the relationship between the dowry that was given or received in the recent marriages, by daughters and sons of the primary earner, and the household's wealth. 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.¹⁰ Household wealth is measured by the average monthly income flow, which includes the profit from land in the past year and the wage earnings of all adult members. We see in the figure that dowries are positive at all wealth levels and that they are increasing in wealth. Notice that dowries paid by girls are larger than dowries received by boys.

Table 2: Dowry and Wealth

Dependent variable	Dowry		
	(1)	(2)	(3)
Female	312.0*** (64.00)	306.2*** (70.78)	337.1*** (67.36)
Household wealth	13.25*** (3.384)	–	–
Per capita wealth	–	43.24*** (13.31)	42.27*** (13.98)
Constant	528.5*** (64.62)	549.0*** (72.61)	533.3*** (70.38)
Observations	826	716	716
Caste FE	No	No	Yes

Household wealth measured as income per month; per capita in terms of family members. Dowry and income in thousands of Rupees. Robust standard errors clustered at the panchayat level in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 2 reports the regression results corresponding to Figure 3. Column 1 includes our wealth measure and an indicator variable for the child’s sex as the regressors. Column 2 replaces household wealth with per capita wealth to reflect the resources that are available to each child. Finally, Column 3 includes caste fixed effects as additional regressors. Although the widespread practice of giving a payment to the groom is relatively new in South India, some upper castes have been giving dowries for many generations. The caste fixed effects account for differences in this practice across castes. The robust finding from Table 3 is that dowries are increasing in wealth and that the amount given by girls is significantly higher than the amount received by boys.

One explanation for this gender-gap is reporting bias, 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

¹⁰The 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 sovereigns), and cash (in Rupees).

Table 3: Dowry Statistics

Sex of the child	Males (1)	Females (2)
Median dowry (in thousand Rupees)	447.97	735.97
Median fraction of annual income	3.57	7.78
Partner's parental household		
Wealthier	0.09	0.19
Same wealth	0.62	0.62
Less wealthy	0.28	0.18
Kolmogorov-Smirnov test of equality	P-value = 0.035	
Observations	322	504

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

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 that have been independently associated with sex selection. However, previous studies have failed to make the connection between hypergamy and sex selection.

Table 3 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. Table 3 also reports the level of the dowry, as a fraction of the household's annual income. In line with past studies; e.g. Rao (1993b), Jejeebhoy and Sathar (2001), and Rahman and Rao (2004), the median dowry is 4 to 7 times the household's annual income, which is an enormous sum in an economy where access to market credit is severely restricted.¹¹

¹¹Most 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.

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 $[\underline{z}, \bar{z}]$, with $\Gamma(\bar{z}) = 2$. 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 over consumption, and we assume that families maximize the sum of their members' utilities $U = \log(C_i) + \log(c_i), \forall i = \{x, y\}$.¹² 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 a 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. 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

¹²Equivalently, parents have altruistic preferences over the utility of their children.

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 “Walrasian” schedule of dowries, and the marriage market subsequently clears with a resulting equilibrium price d . Then, the parent of the boy chooses the transfer t .

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.¹³

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 $H(k)$.¹⁴ 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

¹³Following 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)).

¹⁴In 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. In reality there is thus a stochastic element to the cost of sex selection, that we abstract from in our modeling choice.

children) of a married groom-bride pair (x, y) can be written as:

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

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. To derive these equilibrium outcomes, 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

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

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. 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,

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

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

$$v_x + v_u u' = 0. \quad (7)$$

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; $x = \mu(y)$. Applying the Implicit Function theorem to the first order condition,

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

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

Proof. In Appendix. ■

Given that $\alpha \in [1/2, 1)$, the preceding condition will be satisfied if $y \leq x$.

With assortative matching, the market will clear from the top, with the wealthiest available girl matching with the wealthiest available boy as we move down the wealth distribution. 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. 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).

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 1993b). 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).¹⁵ 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.

¹⁵Even with linear preferences and exogenous distributions, the solving multi-dimensional matching problems is challenging. See for example Choo and Siow (2006) and Lindenlaub (2017).

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 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 that $\alpha < 1$, then the total requirement will exceed x .¹⁶ 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; $\bar{y} = \bar{x}$. 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 $\alpha < 1$, leaves them with a greater share of the total surplus from marriage. Parents are thus better off with boys than with girls. Sex selection arises in our model because of frictions in the marriage market and not because parents intrinsically prefer boys to girls or because dowries are “too high.”¹⁷ Note, moreover, that even though the

¹⁶Substituting the expression for t from equation (3) in the expressions for C_x , c_x , it is straightforward to verify that $c_x = \alpha C_x$.

¹⁷Although there is complementarity in match formation due to the structure of the marriage institution, which

payoff to the girls is lower than that of the boys due to the norm, they do have power to determine the equilibrium payoff. This is because girls are on the short side (due to sex selection) and they are able to extract all the rents from the marginal boy's family who is indifferent between marriage and the outside option.

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^* such that

$$k^*(y) = u(y) - v(\mu(y), y; u(\mu(y))). \quad (9)$$

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 $H(k^*(y))$. 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:

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

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

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

or equivalently:

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

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.

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, $H(k^*(y))$. 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 :

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

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

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; $\bar{y} = \bar{x}$, and at the lowest wealth level at which boys match, x^* , where $u(x^*) = 2 \log \left(\frac{x^*}{2}\right)$.

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

1. *at the top $y = \bar{y}$, whenever $d < \frac{\bar{y}}{2}$;*
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.¹⁸

3.3 Numerical Solution and Results

THE ALGORITHM. When solving the model numerically, we assume 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, $u(\bar{x})$, 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 and horizontally (with their own wealth class). As we move down the wealth distribution, the surplus of boys accumulates and it is possible that below a wealth threshold, girls match up exclusively. Our initial guess for the matching ignores this possibility, specifying that girls (boys) match horizontally and one class up (down) over the entire wealth distribution.

Given the initial guess for $u(\bar{x})$ 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 \bar{x} and payoff $u(\bar{x})$, we can solve for v in both wealth classes. Girls' families in the next to

¹⁸Consider 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).

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^* will remain unmatched. A comparison of $u(x^*)$ derived in the first iteration with the outside option, $2 \log(\frac{x^*}{2})$, is used to adjust the guess for $u(\bar{x})$ in the next iteration. Given u and v derived in the first iteration, the level of sex selection $H(k^*(y))$ 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 4a. 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. 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 threshold wealth level, girls match exclusively with wealthier boys, shifting the plot to the right and away from the 45 degree line. On average, girls are matching with wealthier boys, consistent with the hypergamy documented in our survey data.

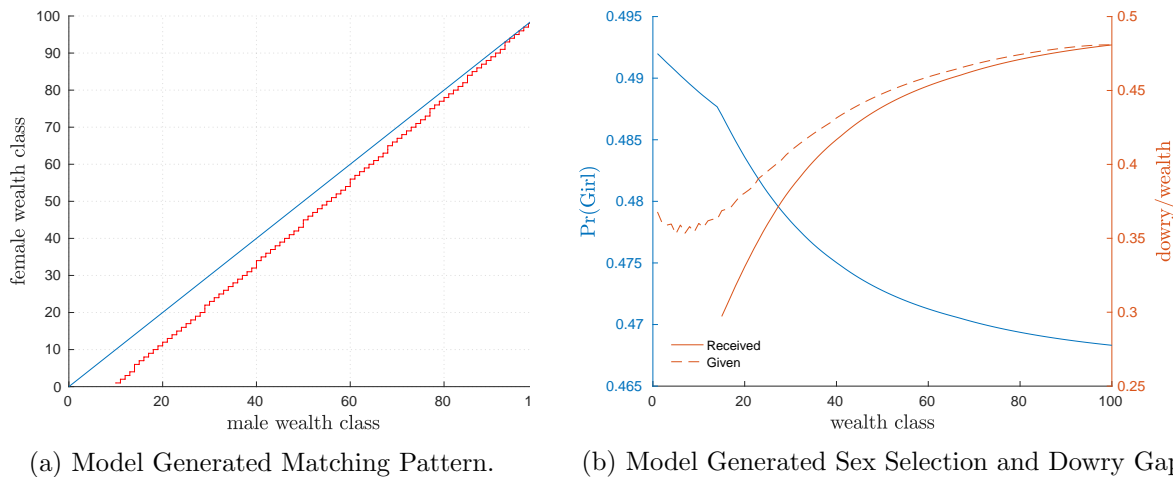


Figure 4: Simulated Model (parameter values $\alpha = 0.6$; $a = 12$).

The dowry received by boys of a given wealth class, x , can be computed directly from $u(x)$ 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 4b, consistent with the dowry gap reported for daughters and sons in the household survey. Notice that the dowry is positive at all wealth levels in the survey data and 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 4b 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.¹⁹ 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 5a 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 5a.²⁰ The direct effect nevertheless dominates, resulting in an increase in sex selection at every wealth level. Figure 5b 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 5b.

The parameter values in Figure 5 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 5a quantifies the impact of this seepage, which is increasing in α , on sex selection.

¹⁹Notice 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.

²⁰If 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.

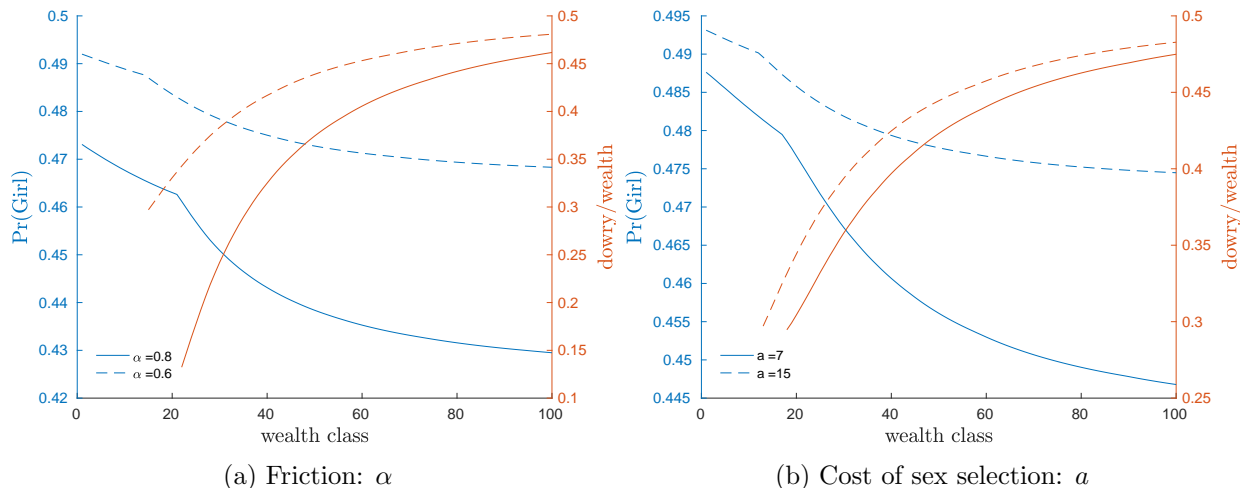


Figure 5: Model Generated Comparative Statics.

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

Figure 5 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. This prediction is inconsistent with the high dowries that are reported in our survey data and by numerous studies in North and South India. In line with the empirical facts, our model is able to generate positive dowries in equilibrium.²¹

POLICY EXPERIMENTS. Our model is able to capture key features of the marriage market 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.

²¹An 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.²² 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,

$$t = \frac{x + \theta d}{2}. \quad (14)$$

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

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

If the dowry d remains fixed, then the boy's family's utility u will decrease.²³ 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,

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

²²In a seemingly related analysis, Bhalotra, Chakravarty, and Gulesci (2016) claim that the price of gold, an important component of the dowry, has a positive effect on sex selection. Their analysis is based on two assumptions. The first assumption is that the amount of gold that is given is unresponsive to the price. This assumption is at odds with the literatures on both dowries and sex selection. If the dowry is treated either as a price or a bequest, its monetary value is the relevant measure for all decisions and the amount of gold will simply adjust to offset the change in the price. The second assumption is that the price of gold at the time of birth is predictive of the price twenty years later at the time of marriage. The gold price at birth will be the best predictor of the future price if it follows a random walk. However, it will still be an extremely weak predictor of prices so far into the future. It is difficult to imagine that parents would make a decision as important as sex selection on the basis of such limited information.

²³The 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.

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

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

If u is fixed, then the most effective scheme will target the family member; i.e. the girl or her parent, 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 (with a population of 193 million) with respect to demographic and socioeconomic characteristics.

Table 4 reports the age distribution, marriage patterns, literacy rates, and labor force participation in the study area, rural Tamil Nadu, and rural South India, respectively.²⁴ 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

²⁴As in Figures 1 and 2, South India includes the states of Tamil Nadu, Andhra Pradesh, Karnataka, and Maharashtra.

Table 4: Comparison of Demographic and Socioeconomic Characteristics

		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 \geq 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
Labour force participation (%; 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 Government 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 ≥ 1 year of education (figures for ≥ 3 years of education are similar, 73.8% for men and 59.5% for women). 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; For Study Area, all statistics based on SICHS Census.

has largely disappeared for children currently enrolled in school (not shown). Labor force participation rates, defined as the fraction of the year for which the individual was employed, 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 2, Vellore district ranks around the median in Tamil Nadu with respect to child (aged 0-6) sex ratios. The child sex ratio in both Vellore district and Tamil Nadu state was 106 in 2011. This is slightly lower than the corresponding statistics of 107 for our study area (from the SICHS census), 108 for South India, and 109 for India as a whole. No castes that are known to be associated with sex selection 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, 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 $\beta < 0$.

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.²⁵

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 model links the family’s rank in the caste wealth distribution to sex selection through the marriage market. However, sex selection could also be determined by other factors. We listed a number of factors that are believed to have contributed to the increase in sex selection over time in India: (i) reduced fertility coupled with the need for at least one son, (ii) improved access to sex selection technology, and (iii) a relative increase in the economic return to boys over girls. The same factors could generate cross-sectional variation in sex selection.

Household wealth and parental education will jointly determine both access to sex selection technology and the labor market returns to sons and daughters (through the educational investments that are made for them). These variables will also determine the mother’s labor force participation, which is one of the most robust determinants of female disadvantage in child survival (Rosenzweig and Schultz (1982), Kishor (1993), Murthi, Guio, and Drèze (1995)). We consequently include household wealth and parental education in an augmented version of equation (18). We measure

²⁵The implicit assumption is that other members of the household, outside the focal 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.

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 imperfections 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 5, Columns 1-3. The coefficient on the wealth rank variable is negative, as predicted, across all specifications. It is not significant at conventional levels in the benchmark specification, but grows steadily larger (in absolute magnitude) and is more precisely estimated as additional regressors are included. The coefficient on household wealth is *positive* and very precisely estimated in Columns 2-3. 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.

Table 5: Wealth and Sex Ratios

Dependent variable	Girl dummy					
	Reported wealth			Predicted wealth		
Wealth measure	(1)	(2)	(3)	(4)	(5)	(6)
Rank in caste wealth distribution	-0.0057 (0.0079)	-0.033** (0.016)	-0.075*** (0.018)	-0.046*** (0.0094)	-0.057*** (0.010)	-0.15*** (0.014)
Wealth	-	0.0010** (0.00046)	0.0020*** (0.00053)	-	0.021*** (0.0053)	0.037*** (0.0067)
Mother's education	-	-0.0004 (0.00051)	-0.0008 (0.00052)	-	0.0003 (0.00053)	0.0000 (0.00053)
Father's education	-	0.00020 (0.00057)	0.00015 (0.00057)	-	0.00035 (0.00060)	0.00030 (0.00059)
Birth order = 1	-	-	0.037*** (0.0056)	-	-	0.080*** (0.0076)
Birth order = 2	-	-	0.029*** (0.0057)	-	-	0.057*** (0.0069)
Constant	0.48*** (0.0029)	0.48*** (0.0046)	0.46*** (0.0059)	0.51*** (0.0058)	0.35*** (0.041)	0.22*** (0.052)
Observations	91,369	91,336	91,336	79,027	78,995	78,995
Caste FE	Yes	Yes	Yes	Yes	Yes	Yes

Sample restricted to children aged 0-6 years. Wealth measured by household income per month divided by the number of family members. Education measured in number of years. Reference category for child birth order is third child or higher. Robust standard errors clustered at the panchayat level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

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 Columns 2-3. However, the effects are not statistically significant, and we will see that both maternal education and paternal education increase the likelihood that the child is a girl with alternative model specifications.

In line with the results from many previous studies, we see in Column 3 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). While these results indicate that the additional motivation for sex selection, which is the need to have at least one son, is also relevant, notice that the coefficient on relative wealth is larger (in absolute magnitude) and more precisely estimated when the birth-order variables are included.

ALTERNATIVE CONSTRUCTION OF HOUSEHOLD 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.²⁶

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

²⁶The 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.

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 are engaged in different occupations, by including caste fixed effects and the interaction of these fixed effects with the historical agricultural productivity variable.²⁷ 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 magnitude. This is indeed what we observe in Table 5, Columns 4-6. The coefficient is also more precisely estimated, and is significant at the 1 percent level, across all specifications.²⁸

ALTERNATIVE SAMPLES. Although we include the birth-order of the child in the augmented specification, this does not fully disentangle the two independent motivations for sex selection. 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. To isolate the marriage market effect, we thus restrict the sample to first-born children. 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. Our data allow us to go further and examine this bias across the wealth distribution within castes; a negative and statistically significant coefficient on the relative wealth variable continues to be obtained in Table 6, Columns 1-2.

As an additional robustness test, we focus on the older group of 7-13 year olds. Mortality is relatively low for children in the 5-20 age-range in all countries and individuals who are younger than 15 rarely leave their homes in India. This implies that the current sex ratio for 7-13 year olds will be very similar to what it was for them when they were 0-6 years old. Sex ratios have been relatively stable over the past decades in Vellore district and so we expect the relationships estimated for the 7-13 year olds to be broadly in line with the estimates for the 0-6 year olds reported above. This is verified in Table 6, Columns 3-5; the probability that the child is a girl is decreasing with relative household wealth among the 7-13 year olds.

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

²⁷Standard errors in this regression, as in equation (18), are clustered at the *panchayat* level.

²⁸Because the new measure of household wealth is a predicted variable, we report bootstrapped standard errors in Columns 4-6. 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).

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 A1. 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.²⁹ This might explain the stronger results that we obtain for that age group in Table 6. As a complementary robustness test, we predicted completed family size for the 0-6 year olds. Results with this alternative measure of per capita wealth, reported in Appendix Table A1, are similar to what we obtain in Tables 5-6.³⁰

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 6 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 for both the benchmark specification and the augmented specification.³¹ 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 6 that the likelihood that a child is a girl is declining continuously with her family’s position in its caste’s wealth distribution. Matching the parametric estimates in Tables 5-6, the decline is steeper with the augmented specification.

The most stringent test of the model is to nonparametrically estimate the relationship between

²⁹Fertility is generally believed to affect the sex ratio through the demand for at least one son. This is captured in our specifications by the birth-order dummies. Note that these dummies will also flexibly capture the direct effect of family size; i.e. the number of children, on sex selection when the sample includes all children in each family. Family size when all children are aged between 5 and 15 can be assumed to be complete and accurately measured because birth-spacing rarely exceeds 5 years in the census data and because children younger than 15 rarely leave their natal homes. Restricting the sample to families where all children are between the age of 5 and 15, and allowing for a more flexible specification that includes four birth-order dummies, relative wealth continues to have a negative and significant effect on the probability that the child is a girl in Appendix Table A1.

³⁰The completed family size is predicted in two steps. First, the Ordered Probit model is used to estimate the relationship between the number of children and predicted 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 sample is restricted to families with 1 to 3 children. The estimated coefficients are used to predict the number of children for all families, 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 or where the predicted number of children is greater than the existing total number of children.

³¹The estimation procedure is implemented in two steps. In the first step, the child’s gender is regressed on a flexible polynomial function of the relative wealth, together with caste fixed effects and, for the augmented specification, with child, parent, and household characteristics. The estimated coefficients are used to partial out the additional regressors. In the second step, the child’s gender, net of the additional regressors, is regressed nonparametrically on the family’s relative wealth.

Table 6: Wealth and Sex Ratios (Robustness Tests)

Dependent variable	Girl dummy				
	(0-6 years)		(7-13 years)		
	First borns		All children		
Birth order	(1)	(2)	(3)	(4)	(5)
Rank in caste wealth distribution	-0.075*** (0.016)	-0.093*** (0.016)	-0.22*** (0.015)	-0.25*** (0.016)	-0.36*** (0.019)
Wealth	-	0.023*** (0.0069)	-	0.043*** (0.0060)	0.063*** (0.0064)
Mother's education	-	0.00069 (0.00085)	-	0.0012** (0.00050)	8.9e-05 (0.00050)
Father's education	-	0.0010 (0.00083)	-	0.0017*** (0.00051)	0.0015*** (0.00051)
Birth order = 1	-	-	-	-	0.12*** (0.0063)
Birth order = 2	-	-	-	-	0.090*** (0.0060)
Constant	0.54*** (0.011)	0.35*** (0.054)	0.59*** (0.0077)	0.25*** (0.044)	0.064 (0.047)
Observations	34,723	34,706	82,648	82,627	82,627
Caste FE	Yes	Yes	Yes	Yes	Yes

Wealth measured by household income per month divided by the number of family members. Education measured in number of years. Reference category for child birth order is third child or higher. Robust standard errors clustered at the panchayat level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

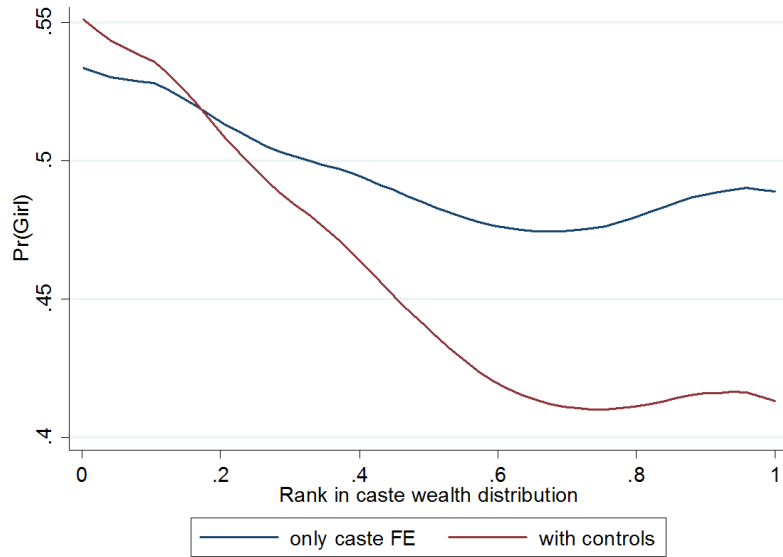


Figure 6: Sex Selection by Percentile in Caste Wealth Distribution (all castes).

relative wealth and sex selection, caste by caste. Figure 7a 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 – Baliya, 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 7b. The relationship is negative at each point in the wealth distribution for all three castes.

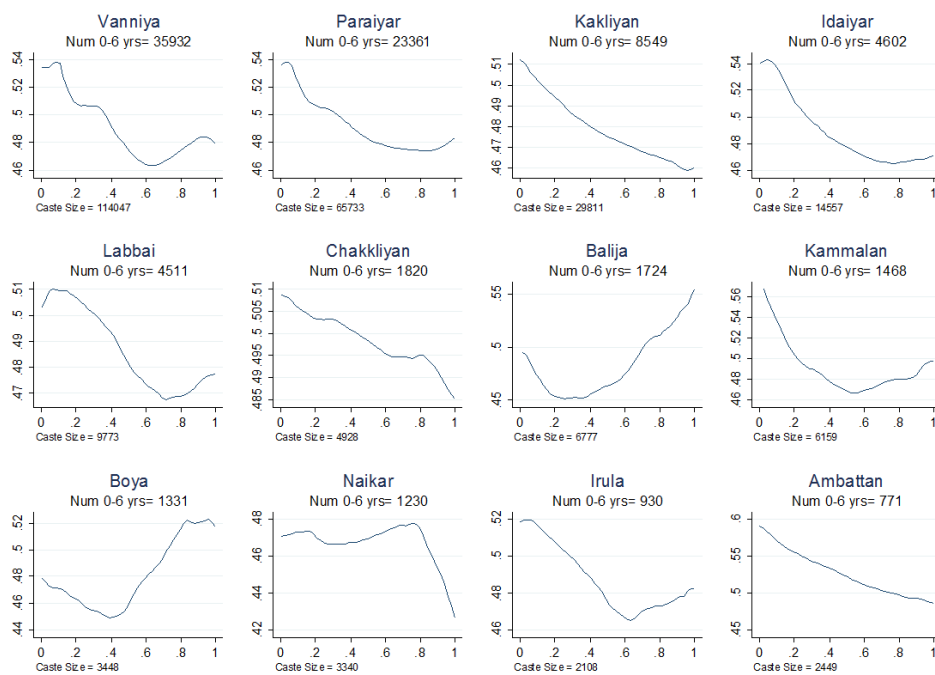
Two castes, the Vanniyas and the Paraiyars, dominate the population in the study area. The Vanniyas are a relatively wealthy landowning caste who have been given the honorific “Gounder” title in Vellore district. Other landowning castes have been given this title in other districts in Tamil Nadu. For example, the Gounders in Salem district, who are believed to be largely responsible for that district’s exceptionally biased sex ratios, belong to the Vellala caste. In Vellore, the Vanniya Gounders are not exceptional with regard to sex selection. The other numerically dominant caste in Vellore, 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, 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 A2 for the benchmark specification with the 0-6 year olds, are very similar to what we obtain for the full sample in Table 5. 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.

4.3 Structural Estimation and Quantification

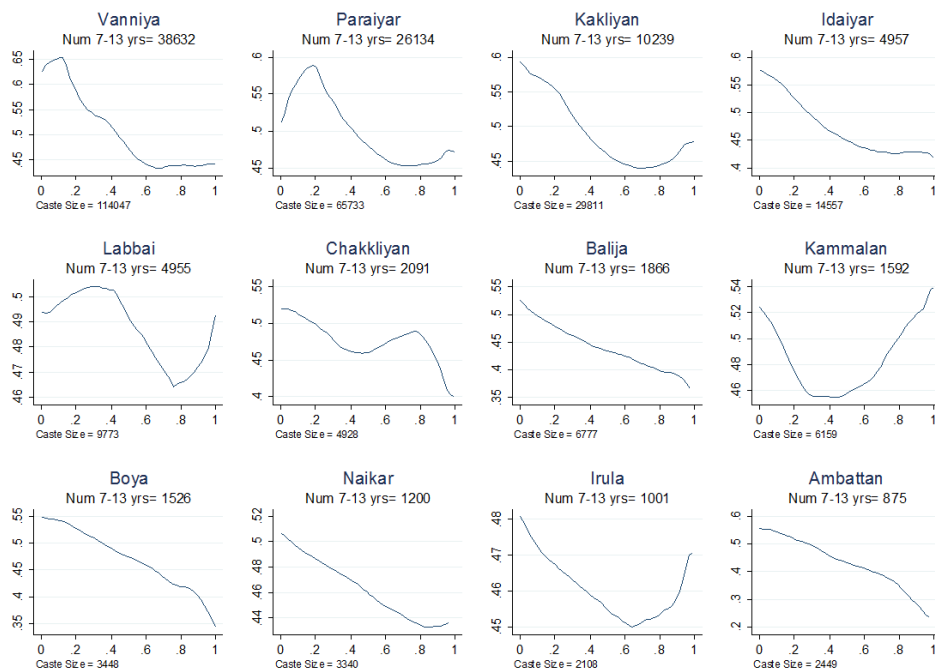
MAGNITUDE OF WITHIN-CASTE VARIATION. A robust finding from the preceding analysis is that the fraction of girls is decreasing as we move up the wealth distribution within castes. To quantify the magnitude of this variation, we partition each caste into eight equally sized wealth classes. The number of classes is chosen by weighting two competing considerations: The larger the number of wealth classes, the closer we can approximate the nonparametric plots. 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 0-6.³²

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

³²In contrast with the preceding specifications, wealth classes are now based on households with children aged 0-6 alone. We will see, nevertheless, that sex selection continues to be increasing with wealth within castes.



(a) Ages 0-6.



(b) Ages 7-13.

Figure 7: Sex Selection by Percentile in Caste Wealth Distribution for 12 Largest Castes.

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 A3). 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 associated with sex selection 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 uncovered by 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. These results are driven by marriage market imperfections that apply to all castes. Sex selection is not restricted to particular castes or particular districts, as is commonly believed, but may, in fact, be a more pervasive problem among relatively wealthy households (within their caste) in India.

STRUCTURAL ESTIMATION AND COUNTER-FACTUAL SIMULATIONS. 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 sex ratio at birth. 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 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 computed in Table 7, 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 7, Column 4 where all castes are sampled with equal probability.

Table 7: Structural Estimates

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

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.

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. 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 8a. 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 complex channels, with potentially unintended consequences.

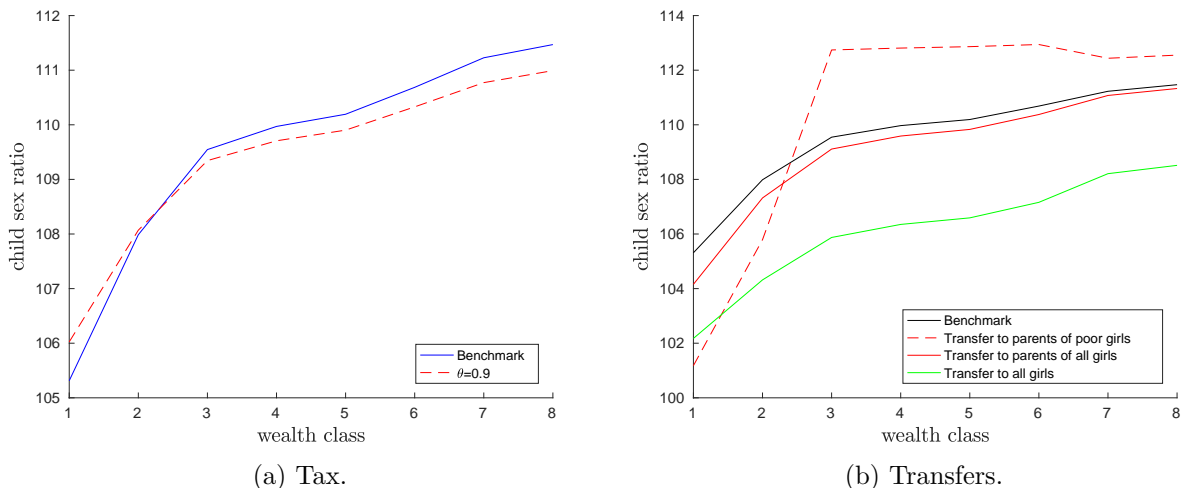


Figure 8: Counterfactual Experiments.

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 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 spillover effects, is perfectly suited to examine the impact of programs with restricted eligibility.

The black solid line in Figure 8b 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.³³ 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 wealth classes). Although the transfer now reduces the sex ratio bias in each wealth class, the magnitude of the effect is small.

The final experiment that we consider is the most promising to fight sex selection with policy intervention. It is the same as the preceding experiment, except that the subsidy goes directly to the adult girls rather than their 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 8b, there is now a substantial increase in the fraction of girls in each wealth class.³⁴ 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. 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 families 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

³³The 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.

³⁴This 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 5a.

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 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.

One limitation of the counter-factual simulations is that they assume that a given policy can always be implemented as designed. In practice, transfers to married women, even if they go directly into their bank accounts, could be appropriated by the husband and his family. 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.³⁵

³⁵Numerous 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 the increased economic returns to having girls reduces sex selection.

Appendix A Robustness Tests

Table A1: Alternative Measures of Family Size

Dependent variable	Girl dummy			
	Household size		Predicted kids	
Family size measure	(1)	(2)	(3)	(4)
Rank in caste wealth distribution	-0.0311** (0.0154)	-0.0744*** (0.0174)	-0.0359*** (0.0138)	-0.111*** (0.0182)
Predicted wealth	–	0.0275* (0.0156)	–	0.0556*** (0.0183)
Mother's education	–	-0.00146 (0.00132)	–	-0.000189 (0.00128)
Father's education	–	0.0000 (0.00144)	–	0.000299 (0.00139)
Birth order = 1	–	0.0655*** (0.0112)	–	0.0995*** (0.0132)
Birth order = 2	–	0.0436*** (0.0116)	–	0.0791*** (0.0141)
Constant	0.502*** (0.00942)	0.273** (0.123)	0.500*** (0.00738)	0.0134 (0.147)
Observations	78,934	78,903	75,997	75,993
Caste FE	Yes	Yes	Yes	Yes

Sample restricted to children aged 0-6 years. Wealth measured by household income per month divided by the number of family members. Education measured in number of years. Reference category for child birth order is third child or higher. Robust standard errors clustered at the panchayat level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A2: Alternative Samples

Dependent variable	Girl dummy			
	Benchmark	Dropping Biggest Caste	Dropping 2 Biggest Castes	Dropping Smaller Castes
Sample	(1)	(2)	(3)	(4)
Rank in caste wealth distribution	-0.0457*** (0.00933)	-0.0408*** (0.0112)	-0.0386*** (0.0120)	-0.0541*** (0.0102)
Observations	79,027	49,522	29,883	69,233
Caste FE	Yes	Yes	Yes	Yes

Sample restricted to children aged 0-6 years. Wealth measured by household income per month divided by the number of family members. Robust standard errors clustered on the panchayat level in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table A3: Within and Between Variation in Sex Ratios

Dependent variable	Fraction of girls			
	12 castes		30 castes	
Sample	(1)	(2)	(3)	(4)
Wealth class				
2	-0.0108 (0.00709)	-0.0110 (0.00781)	-0.0150** (0.00745)	-0.0153* (0.00829)
3	-0.0366*** (0.0111)	-0.0367*** (0.0101)	-0.0386*** (0.0101)	-0.0385*** (0.00917)
4	-0.0444*** (0.00517)	-0.0445*** (0.00488)	-0.0438*** (0.00525)	-0.0443*** (0.00475)
5	-0.0479*** (0.00711)	-0.0480*** (0.00579)	-0.0481*** (0.00675)	-0.0480*** (0.00558)
6	-0.0355*** (0.00605)	-0.0358*** (0.00729)	-0.0344*** (0.00578)	-0.0345*** (0.00672)
7	-0.0267*** (0.00792)	-0.0271*** (0.00940)	-0.0261*** (0.00736)	-0.0268*** (0.00862)
8	-0.0247*** (0.00889)	-0.0250*** (0.00759)	-0.0248*** (0.00851)	-0.0249*** (0.00745)
Constant (wealth class 1)	0.509*** (0.00411)	0.509*** (0.00341)	0.510*** (0.00401)	0.510*** (0.00334)
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. Wealth measured by household income per month divided by the number of family members. 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:

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

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

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

Calculating each of the derivatives, we get:

$$v_x = \frac{1}{x + y - 2\psi(u)} \quad (21)$$

$$v_u = \frac{-2\psi'(u)}{x + y - 2\psi(u)} + \frac{\psi'(u)}{\psi(u)} = \frac{x + y - 4\psi(u)}{2(x + y - 2\psi(u))} \quad (22)$$

$$v_{xy} = \frac{-1}{(x + y - 2\psi(u))^2} \quad (23)$$

$$v_{uy} = \frac{\psi(u)}{(x + y - 2\psi(u))^2}, \quad (24)$$

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:

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

or equivalently

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

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:

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

or

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

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

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

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

For $x = y = \bar{x}$, 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 $k^*(y)$:

$$k^*(y) = u(y) - v(\mu(y), y, u(\mu(y))). \quad (31)$$

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

$$k^{*\prime}(y) = u'(y) - (v_x \mu' + v_y + v_u u' \mu'). \quad (32)$$

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:

$$k^{*\prime}(y) = u'(y) - ((v_x + v_u u') \mu' + v_y) \quad (33)$$

$$= u'(y) - v_y(\mu, y, u(\mu)). \quad (34)$$

This is increasing provided:

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

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 $x = \bar{x}$, under positive sorting we have $\bar{y} = \mu(\bar{x}) = \bar{x}$. Then condition (35) can be written as:

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

or

$$\psi(u) < \frac{3}{4}\bar{x} \quad (37)$$

$$\sqrt{\frac{e^u}{\alpha}} < \frac{3}{4}\bar{x} \quad (38)$$

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

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

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

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

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

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

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

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

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

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.
- 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.
- 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*, forthcoming.
- 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., AND P. MARI BHAT (1997): "Fertility Decline and Increased Manifestation of Sex Bias in India," *Population Studies*, 51(3), 307–315.
- 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.
- 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.