Gender discrimination and excess female under-five mortality in India: a new perspective using mixed-sex twins

Ridhi Kashyap (i) and Julia Behrman (ii)

(i) University of Oxford, UK, Ridhi.Kashyap@nuffield.ox.ac.uk
(ii)Northwestern University, IL, USA, Julia.Behrman@northwestern.edu

Abstract

Son preference has been linked to excess female under-five mortality in India and considerable literature explores whether parents invest more resources in sons relative to daughters—which we refer to as *explicit discrimination*— leading to girls' poorer health status and consequently, higher mortality. However, this literature does not adequately control for the *implicit* discrimination processes that sort girls into different types of families (e.g. larger) and at earlier parities. To better address the endogeneity associated with implicit discrimination processes, we explore the association between child sex and post-neonatal under-five mortality using a sample of mixed-sex twins from four waves of the Indian National Family Health Survey. Mixed-sex twins provide a natural experiment that exogenously assigns a boy and a girl to families at the same time, thus controlling for selectivity into having an unwanted female child. We document a sizeable impact of explicit discrimination on girls' excess mortality in India, particularly compared to a placebo analysis in Africa where girls have a survival advantage. We also show that while explicit discrimination weakened for birth cohorts after the mid-1990s especially in northern India, further weakening has stalled since the mid-2000s, thus contributing to understandings of how the micro- processes underlying the female mortality disadvantage have changed over time.

Keywords: son preference, under-five mortality, excess female child mortality, India, fixed effects, twins

Acknowledgments We are grateful to Kieron Barclay, Jere Behrman, Sonia Bhalotra, Monica Caudillo, Seema Jayachandran, Stephanie Koning, Felix Tropf, and Abigail Weitzman for helpful comments on earlier versions of this manuscript. We gratefully acknowledge support for this article through the Global Family Change (GFC) Project (http://web.sas.upenn.edu/gfc), which is a collaboration between the University of

Pennsylvania, University of Oxford (Nuffield College), Bocconi University, and the Centro de Estudios Demograficos (CED) at the Universitat Autonoma de Barcelona. Funding for the GFC Project is provided through NSF Grant 1729185 (PIs Kohler & Furstenberg), ERC Grant 694262 (PI Billari), ERC Grant 681546 (PI Monden), the Population Studies Center and the University Foundation at the University of Pennsylvania, and the John Fell Fund and Nuffield College at the University of Oxford. RK also acknowledges support from the Leverhulme Trust within the Leverhulme Centre for Demographic Science.

Introduction

Son preference continues to be a defining feature of family life in India, shaping the well-being of Indian women and girls throughout the life course. One of the most striking demographic manifestations of son preference in India is the persistence of excess female infant and child mortality. Despite declining levels of overall under-five mortality, India continues to experience one of the highest levels of excess female under-five mortality in the world (Alkema et al. 2014, Guilmoto et al. 2018, Kashyap 2019). The term 'excess' implies that girls experience higher than biologically expected levels of mortality relative to boys, which, as famously characterized by Amartya Sen, results in women and girls being 'missing' from population structures (Sen 1990).[1]

The dominant demographic explanation for the female mortality disadvantage in India has been that parents invest more resources (e.g. healthcare, nutrition, immunization) in sons relative to daughters—a set of processes which we refer to as *explicit discrimination*— leading to girls' poorer health status and consequently, higher mortality (Miller 1981, Caldwell, Reddy and Caldwell 1982, Das Gupta 1987, Caldwell and Caldwell 1990). However, research linking explicit discrimination to excess female child mortality in India confronts a major measurement issue: the sex[2] composition of families is not random in India and the literature typically does not adequately control for the passive son-preferring fertility behaviors—which we call *implicit discrimination*—that sort girls into different types of families and at earlier parities within families. For example, son-preferring fertility stopping rules imply that families may continue

---

[1] The male/female ratio (hereafter, sex ratio) of infant mortality is generally higher than 1, indicative of higher male mortality compared with female mortality. The same pattern also exists in early childhood (between ages 1 to 4 years) although sex ratios are less masculine in most populations in this age group compared with infancy (first 12 months of life). In the first year of life, newborn girls have a biological advantage over boys due to their lower vulnerability to perinatal conditions, congenital abnormalities, and certain infectious diseases such as intestinal and lower respiratory infections (Drevenstedt et al. 2008).

[2] Throughout the paper to be consistent with demographic convention (e.g. sex ratios, sex-specific mortality) and the variable in the National Family Health Survey (NFHS), we use sex of the child. We recognize nevertheless that the female mortality disadvantage is the product of social—rather than biological—processes related to *gender*-based discrimination.

to have children until their desired number of sons is reached, which results in girls being born into larger families (Yamaguchi 1989, Clark 2000, Filmer, Friedman and Schady 2009, Rosenblum 2013), and at earlier parities relative to boys (Basu and De Jong 2010). Furthermore, prenatal sex selection in the form of sex-selective abortion allows some families to “opt out” of having daughters (Jha et al 2006, Bhalotra and Cochrane 2010, Hu and Schlosser 2015, Anukriti, Bhalotra and Tam 2018, Kashyap 2019). If technology access to practice sex selection is concentrated in urban, better-off families, then girls may be disproportionately sorted into disadvantaged households. The conceptual implication of different implicit discrimination processes is that the population-level female mortality disadvantage could emerge from girls being disproportionately born into larger and/or poorer families where overall mortality is higher (i.e. implicit discrimination), rather than differential resource allocation within the same family (i.e explicit discrimination).

It is empirically difficult to measure explicit discrimination net of implicit discrimination because prenatal sex selection remains unobserved at the family level and other forms of differential selection, such as family size or birth order, are endogenous to mortality and parental son preference. It is further complicated to explore how explicit discrimination has changed over time given that the implicit processes that might sort girls into qualitatively different households have also changed with diffusion of ultrasound technology, fertility declines, improvements in women’s education, and other social and economic changes. Nonetheless, it is fundamental for our understanding of son preference in India to understand if explicit discrimination underpins observed mortality disadvantage for girls. Furthermore, accurately documenting explicit discrimination is essential from a policy perspective because the policy responses implied by parents’ differential resource allocation to boys versus girls within families are different than if girls’ mortality disadvantage accrue primarily from selection into different families. If there is intra-household discrimination against girls, then policies targeting health and welfare investments in girls specifically within families will be necessary. However, if girls do worse on average because of selection into larger and/or poorer families, then policies addressing generalized poverty reduction will be better motivated.

To better address the endogeneity of implicit discrimination processes we use a large sample of mixed-sex twins to investigate the association between child sex and post-neonatal under-five mortality with data from four waves of the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016). Mixed-sex twins provide a natural experiment that exogenously assigns both a boy and a girl to families at the same time, thus allowing us to control for differential family selectivity into having an unwanted female child and other implicit discrimination processes. To validate our estimates of explicit discrimination we conduct a placebo analysis using a large sample of twins from sub-Saharan Africa, a region that does not have a history of female mortality disadvantage. We also explore heterogeneity in explicit discrimination by stratifying by region and number of older female siblings. Finally, we investigate whether there have been declines in explicit discrimination over subsequent birth cohorts, thus providing important insight into how the micro- processes that contribute to the female mortality disadvantage have changed over time in contemporary India.

Explicit and implicit discrimination processes that contribute to the female under-five mortality disadvantage in India

Patterns of excess female infant and child mortality arising from a strong preference for male offspring have been long noted in India. Sex ratios (male/female) of infant and child mortality in India have remained under 1.00, and thus indicative of a female mortality disadvantage that is attributable to social (gender) discrimination processes because biologically males are more vulnerable to mortality in infancy, and to a lesser extent, in early childhood. In particular in the age group 1—4 years, India has the most anomalous levels of excess female mortality in the world (Alkema et al 2014, Guilmoto et al. 2018, Kashyap 2019). In what follows we highlight both explicit and implicit discrimination processes that contribute to the female under-five mortality disadvantage, with discussion of how these processes may have changed over time.

Excess female infant and child mortality has long been thought to arise from explicit postnatal discrimination against girls in the allocation of resources such as healthcare (e.g. immunization, medical treatment) or nutrition (e.g. food, breastfeeding) that are relevant for survival (Miller 1981, Das Gupta 1987, Arnold et al 1998, Pande 2003, Mishra et al 2004). Nonetheless, the empirical evidence for sex-differential allocation of resources has often been mixed or inconclusive (Barcellos, Carvalho and Lleras-Muney 2014). While several studies have found that girls were less likely to receive healthcare and vaccinations (Ganatra and Hirve 1994, Pande 2003, Mishra et al 2004, Borooah 2004, Corsi et al. 2009, Rajan and Morgan 2018), others have found similar vaccination rates for boys and girls (Deaton 2003, Barcellos, Carvalho and Lleras-Muney 2014). Evidence on sex-differentials in children's diet is similarly mixed (Basu 1998, Fledderjohann et al. 2014). Population-level studies of anthropometric measures such as malnutrition and stunting also show an ambiguous picture with no clear female disadvantage in these measures, and in some cases a male disadvantage in the first 24 months of life but with the gap closing at later ages (Mishra et al 2004, Corsi et al. 2015). Studies, however, have found a clear female disadvantage with respect to duration of breastfeeding (Jayachandran and Kuziemko 2011, Fledderjohann et al 2014, Barcellos, Carvalho and Lleras-Muney 2014). Although in some cases son preference may actually disfavor boys, who may be exclusively breastfed longer in the ages between 6-9 months when breastfeeding alone is not sufficient to meet an infant's energy needs (Mishra et al 2004). Parents in India have also been shown to make differential prenatal investments in male versus female fetuses (Bharadwaj and Lakdawala 2013).

In seeking to reconcile these mixed findings, which on one hand show a female disadvantage in mortality linked to son preference but weaker evidence for sex-differentials in resources and anthropometric measures, studies have argued that the female disadvantage is not generalized but concentrated among certain subsets of girls, particularly those born at later parities and without brothers. These birth order and sibling composition effects have been found both in terms of health inputs and outcomes (Pande 2003, Mishra et al 2004) as well as in mortality (Muhuri and Preston 1991, Arnold et al 1998). According to this perspective, son preference does not imply that all girls are unwanted but daughters deemed to be "redundant" are more likely to be discriminated against. In contrast to these perspectives, Rajan and Morgan (2018) have argued that generalized discrimination against girls – that affects all daughters at a given parity, rather only than those with older sisters and no brothers – provides a better description of patterns of female disadvantage observed in India for outcomes such as immunization, treatment for respiratory illness and breastfeeding after 17 months.

Most of the abovementioned studies have estimated the effect of being female on a mortality outcome or particular investment (e.g. immunization, breastfeeding), and compared boys and girls between different families usually in a regression framework. However, in a context where son preference shapes fertility behaviors and family sex composition is not random, girls are likely to be selected into different types of families and this differential selection—which we term *implicit discrimination* —may also affect the observed aggregate-level disadvantage in girls' outcomes. As we describe later, controlling for different forms of implicit discrimination, which may be changing over time, is important to detect if there is gender discrimination within families that leads to excess female mortality, but is not always straightforward.

For example, studies have found that son-preferring fertility rules result in girls being born into larger families, that is having larger siblingship sizes relative to boys (Yamaguchi

1989, Clark 2000, Basu and De Jong 2010, Rosenblum 2013). Girls as a result of this are likely to share resources in the same family with larger sibling cohorts, and thus will be worse off than boys when comparing boys versus girls between families if fewer resources per member are available in larger families. Conversely, if there are returns to scale for certain resources in large families then girls might not be worse off but could benefit from these instead (Barcellos, Carvalho and Lleras-Muney 2014). Rosenblum (2013) has found that these son-preferring stopping rules that result in girls being born into larger families exacerbate excess female child mortality in India. Couples with first-born boys had fewer total children born and a higher proportion of males in their families. Using reduced form estimates, Rosenblum found that second- and higher-order girls in families with first-born sons had 25% higher mortality than boys, while those in first-born girl households had 38% higher mortality than boys.

Another implication of son-preferring fertility behavior is that girls are more likely to be born at earlier parities within families (Basu and De Jong 2010). The implication of this form of selection for mortality is a priori ambiguous. While some studies find a J-shaped relationship with infant mortality, with first-borns showing the highest risks, others have found a linearly increasing risk from earlier-borns to later-borns, which could protect earlier-born girls (Mishra et al. 2018). On the other hand, Basu and De Jong (2010) hypothesize that in a context with son preference, earlier-born daughters may experience negative consequences for their wellbeing as a result of having to assist in the care of later-born children.

Yet another form of selection that may affect the family conditions into which girls are born is prenatal sex selection, most commonly in the form of sex-selective abortion. Prenatal sex selection became practiced in India starting the early 1990s, as indicated by distorted sex ratios at birth, especially in northern Indian states of Punjab, Haryana, Delhi and others (Jha et al. 2006, Guilmoto and Tove 2015, Hu and Schlosser 2015). Whether prenatal sex selection works to protect or worsen girls' mortality outcomes depends on which families are able to access it. If sex-selective abortion enables households with the strongest son preference, and those that might have otherwise resorted to explicit discrimination the option to avoid having unwanted daughter(s) and reduce family size, this form of implicit discrimination may work to protect girls (Goodkind 1996, Hu and Schlosser 2015, Kashyap 2019). However, uneven access to technology enabling sex-selection may imply that wealthier families are better able to avoid unwanted female births, and girls may be differentially sorted into households with overall fewer resources because these households cannot afford to opt out of having daughters even if sons are preferred (Hu and Schlosser 2015, Kashyap 2019). This may worsen the population-level disadvantage experienced by girls. Studies from India have found that sex ratios at birth are most distorted among wealthier, urban and more educated families (Jha et al 2006), which has generally been interpreted as a sign of better access to ultrasound technology among these groups (Guilmoto and Tove 2015).

Aggregate indicators—such as the sex-disaggregated under-five mortality rate —mask both explicit and implicit discrimination processes, and cannot adequately capture if the micro-level processes of explicit discrimination have changed over time. Explicit discrimination could have changed over time in India through different channels. Some scholars suggest that diffusion of ultrasound technology could lead to a "substitution" whereby postnatal discrimination in the allocation of nutrition and health resources between male and female children (e.g. explicit discrimination) is weakened as a result of the uptake of prenatal discrimination via sex-selective abortion (e.g. implicit discrimination) (Goodkind 1996, Sen 2003, Kashyap 2019). Evidence for this hypothesis in the Indian context has so far been mixed. Whereas Hu and Schlosser (2015) did not find faster reductions in girls' mortality relative to boys for cohorts that witnessed prenatal sex selection, Anukriti, Bhalotra and Tam (2018) report evidence for faster reductions in girls' mortality in the period in which ultrasound became widely available in India (after 1995). Disentangling the effects of weakening son preference from the practice of sex selection is empirically challenging, however, and son preference can be weakening even as sex ratios at birth become more masculine (Kashyap and Villavicencio 2016). Sex-selective abortion may enable families to reconcile son preference with a small family size, and thus also facilitate fertility decline in contexts with son preference (Kashyap and Villavicencio 2016, Jayachandran 2017). On the other hand, explicit discrimination could also decline as a result of weakening son preference. There is some indication that son preference, as measured by different indicators of ideal sex composition, may be weakening over time in India with wider processes of socioeconomic development and fertility decline (Bhat and Zavier 2003, Retherford and Roy 2003, Bongaarts 2013, Kashyap and Villavicencio 2017).

Accounting for implicit discrimination in measurement of explicit discrimination

A standard approach for measuring the mortality attributable to explicit discrimination would be to estimate the association between child sex and mortality controlling for birth order, family size, family SES and other measures related to the implicit discrimination processes that sort girls and boys into different families. However, with this approach it would be difficult to appropriately control for implicit discrimination processes because the intensity of son preference and prenatal sex selection are unobserved at the family level, and variables such as completed family size and birth order are endogenous to mortality and parental son preference. Controlling for variables such as family size or sex composition would allow us to compare outcomes of boys and girls in families of the same size. However, as Barcellos, Carvalho and Lleras-Muney (2014) note, in the presence of son-preferring stopping rules if we compare girls and boys in families of the same size, girls are, on average, in families that desire fewer sons (than the family of the average child). In other words, even conditional on family size and sex composition, child sex is not exogenous but it is correlated with parental preferences for the sex composition of children.

One approach for addressing the endogeneity of family size, as followed by Rosenblum (2013), has been to capture exogenous variation related to son-preferring stopping rules by using the sex of the first child as a natural experiment.[3] While this approach demonstrates how sex-differential stopping rules exacerbate mortality outcomes, it is unable to estimate an effect of explicit discrimination, net of implicit discrimination, for girls. An alternative approach, used by Barcellos, Carvalho and Lleras-Muney (2014) to examine if boys and girls receive differential resources, has been to focus on the youngest child in the family – when they are young enough and the next birth has not yet occurred – to measure boy-girl differences in this sample. In the absence of sex-selective abortion, the sex of the child in this sample can be assumed to be exogenous, and consequently, their study focuses on births that occurred before the 1990s, after which sex-selective abortion became practiced in India. Both strategies, by Rosenblum (2013) and Barcellos, Carvalho and Lleras-Muney (2014), face limitations in a context where sex-selective abortion is practiced, and cannot adequately address how explicit discrimination is changing over time.

We propose a novel strategy to better address the endogeneity associated with implicit discrimination processes by leveraging mixed-sex twins as a natural experiment in which both a boy and a girl is assigned to a family at the same time, thus allowing us to control for implicit discrimination processes such as differential family selectivity into having an unwanted female child. Mixed-sex twins are exposed to the same prenatal environment and are born at the same time and thus exposed to the same family environment (e.g. wealth at birth). Since the principal difference in mixed sex twins is child sex—and not family size, birth order, maternal age,

[3] Sex selection at first birth in northern India has been noted in recent work (e.g. Anukriti 2018) that challenges this assumption.

family wealth at birth etc.— elevated female mortality among mixed-sex twins should be more readily attributable to differential parental behaviors based on child sex (e.g. explicit discrimination). This is particularly the case since biologically male children are more vulnerable to death in infancy and early childhood (Drevenstedt et al 2008), and this has also been shown to hold for twin populations (Pongou 2013, Ahrenfeldt et al 2017). Thus, in a context without differential treatment of male and female children we would actually expect a female survival advantage.

Data & Sample

We use pooled standardized data on mixed-sex twins from the 1992-1993, 1998-1999, 2005-2006, and 2015-2016 India National Family Health Survey (NFHS). The NFHS is cross-sectional micro-data on key demographic and health outcomes that is nationally representative of all women ages 15-49 and follows the format and structure of the Demographic and Health Surveys (DHS). The NFHS is collected by the Indian Ministry of Health and Family Welfare and International Institute for Population Sciences, with input from MACRO (ORC).

We identify a sample of mixed-sex twins using the NFHS birth recode, which provides detailed information on all children born to women in the sample. Respondents are queried about whether each child is still alive, and if not at what age in months the child died. Respondents are also asked whether each birth is a multiple birth (e.g. twin birth), birth order, and whether each birth was male or female. Combining this information allows us to identify which births are mixed sex-twins. We exclude households with multiple sets of twins and households with triplets and quadruplets due to the exceptional nature of these events, which suggests these household might be categorically different from others in the sample (in particular they may be genetically predisposed to twinning). Our analytical sample includes 6,200 mixed-sex twins from 3,100 families.

Empirical Approach

Measures

Mortality: Our main outcome is a dichotomous indicator of whether the birth resulted in death in infancy or early childhood (e.g. between 1 and 59 months). We exclude mortality in the first month of life to because we are interested in capturing social rather than biological processes that impact mortality and previous literature on son preference has shown that its mortality manifestations in under-five mortality are most apparent in the post-neonatal ages (Das Gupta 1987, Arnold et al. 1998, Arokiasamy 2004). All infant and child deaths are self-reported by mothers and thus are subject to reporting bias. Nonetheless, the death of an offspring is a rare and important event, thus it is reasonable to believe that mothers would accurately remember the age of offspring death.

Child sex: Throughout the models the main treatment outcome is a dichotomous indicator of whether the birth was female.

First born twin: We include a control for which twin was born first because on average first born twins are heavier than second born twins, which may have implications for parental investment and later life outcomes (Pongou 2013).

Estimation Strategy

We use a within-twin fixed effects model that allows us to compare boy-girl differences in mortality within twin pairs born into the same family. The fixed effect (in eq. 1, α_i) captures all observed and unobserved factors (e.g. family socioeconomic status and environment, prenatal inputs) shared between the twin pair. For an individual *j* in twin-pair *i,* the within-twin fixed effect model of sex of the child on mortality can be expressed as:

$$Mortality_{ij} = \beta_0 + \beta_1 FirstBorn_j + \tau Female_j + \alpha_i + \varepsilon_{ij} \quad (1)$$

First, we use the within-twin fixed effects model in eq. 1 to estimate the effect of being female (τ) on the probability of post-neonatal under-five mortality pooling across the four NFHS survey waves. We interpret τ as a measure of the female mortality disadvantage that is attributable to explicit discrimination. We use linear probability models throughout analyses for the ease of interpretation and comparability of coefficients across models. Furthermore, we assess whether there is evidence of changes in explicit discrimination by running models across three different birth cohorts of mixed sex twins: (1) twins born prior to 1995; (2) twins born between 1995 and 2005; (3) and twins born after 2005. Our birth cohorts roughly correspond with those suggested by Anukriti, Bhalotra and Tam (2018) regarding different periods of diffusion of ultrasound technology in India, whereby the first period represents the early diffusion period when ultrasound was still new and less common, the second period represents a period by which ultrasound use is widespread, and the third period corresponds with more recent history in India.[4] Due to the diffusion of ultrasound, after the mid-1990s implicit discrimination associated with sex-selective abortion – in addition to sex-differential stopping behaviour – became possible. The measurement problems introduced by sex-selective abortion do not affect the within mixed-sex twin estimated precisely because sex-selective abortion is not a viable option for mixed-sex twins. Nevertheless, this form of implicit discrimination shapes the broader context of the families into which mixed-sex twins get randomly assigned over this time period (e.g. parents after the mid-1990s could use sex-selective abortion to implement their son preference prior to having twins, or also could have for future children).

To validate our measure, we also conduct a placebo analysis using a large sample of sub-Saharan African (SSA) twins. The Africa data comes from the Demographic and Health

[4] Unlike Anukriti, Bhalotra and Tam (2018), we also use the most recent wave of the NFHS-4, and our sample covers more recent births as well (i.e. after 2005).

Surveys (which are standardized with the NFHS) with multiple survey waves occurring over the same period as the NFHS (see Appendix Table 1 for further detail on the African sample). Because SSA does not show patterns of son-preferring fertility behaviours (Basu and De Jong 2010, Jayachandran and Pande 2017), and aggregate-level under-five female mortality disadvantage (Alkema et al 2014), we hypothesize that the females in mixed-sex twin pairs in this region should not experience a mortality disadvantage. Thus, if there is a female mortality disadvantage among mixed-sex twins in India, but not SSA (i.e. a higher magnitude of τ in India compared with SSA), this is further evidence of social—as opposed to biological—processes leading to the under-five mortality disadvantage.

Given evidence that suggests heterogeneity in son preference in India—including by region and by family structure—we also re-run the within-twin fixed effects models stratifying by region and number of older sisters. Regional variations in son preference and its manifestations have long been noted in India, with son preference notably stronger in northern states (Dyson and Moore 1983, Arnold et al. 1998, Bhat and Zavier 2003, Arokiasamy 2004).[5] As a further extension, we also stratify by both region and birth cohort to see if there are changes over birth cohorts at the regional level. Finally, we re-run our mixed-sex twin fixed effects models stratifying by number of older sisters because explicit discrimination may be more common in families where there are multiple older sisters (Arnold et al 1998, Pande 2003, Mishra et al 2004).

Results

Descriptive summary

Table 1 presents weighted proportions (and means for continuous variables) for descriptive characteristics of our twin sample, including how twin characteristics have changed

[5] The states included in the northern category include Bihar, Delhi, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Madhya Pradesh, Maharashtra, Punjab, Rajasthan, and Uttar Pradesh.

over subsequent birth cohorts. On average, about 9% of the twins in our pooled sample died before the age of five, although survival improves over time. For example, 17% of twins born before 1995 died before the age of five, compared to 7% and 5% of twins born between 1995 and 2005 and after 2005 respectively.

These declines in mortality likely correspond with a number of other important changes in fertility and family life that also occurred over subsequent birth cohorts. For example, between the oldest (e.g. born before 1995) and youngest (e.g. born after 2005) birth cohorts, twins were increasingly born into smaller families at earlier birth orders, which correspond with overall fertility declines in India in recent history. Mothers in the sample are increasingly better educated and have children at older ages, which also makes sense given that these are key correlates of fertility declines. Mother's stated ideal number of boys and girls also both significantly declined over subsequent birth cohorts of twins, which may reflect preference for increasingly smaller family sizes. There are also significant declines over birth cohorts in mother's ideal sex ratio of boys to girls (ideal boys divided by ideal number of children), which does suggest some lessening of stated son preference over time. Nonetheless, even in the most recent birth cohort, mother's ideal sex ratio is still skewed towards boys (0.54).

Exploring explicit discrimination and the female post-neonatal under-five mortality disadvantage using within-twin fixed effects

To better estimate the effects of explicit discrimination, we start by using a pooled sample of mixed-sex twins to conduct a within-twin fixed effects analysis of the association between child sex and post-neonatal under-five (1—59 months) mortality. Unlike past estimates of the female mortality disadvantage, this specification allows us to control for implicit discrimination processes by accounting for unobserved twin-level confounders that do not vary between twins (e.g. prenatal conditions, family SES, family size, birth timing and order). We find females are associated with a 2-percentage point higher probability of post-neonatal under-five mortality than males ($p<0.001$) (Table 2, Panel A, col 1), which corresponds to a 26% higher mortality for girls relative to the mortality probability for boys of the pooled sample (0.075). These results provide strong evidence of explicit discrimination playing an important role in the female mortality disadvantage observed in our data, net of implicit discrimination processes. The substantive findings from Table 2 are robust to re-specification as logistic regression fixed effects and Cox proportional hazard models where the outcome is age (in months) at death (Appendix Table 2, Panels A-B). Results are also robust to limiting to children born within 10 years of the survey to minimize recall bias in the reporting of children's age at death (Appendix Table 2, Panel C).

Over the timespan covered in our study, son preference and forms of implicit discrimination changed with fertility declines, socioeconomic development, and diffusion of ultrasound technology to facilitate sex-selective abortion. Our next step is to explore how mortality attributable to explicit discrimination changed over subsequent birth cohorts. As Table 2 Panel B shows, among mixed-sex twins born before 1995 (e.g. before the widespread diffusion of ultrasound technology and uptake of prenatal sex selection), females are associated with a 5.3-percentage point higher probability of post-neonatal under-five mortality compared to males ($p<0.001$) (Table 2, Panel A, col 2), or a 45% higher mortality relative to male mortality of this period. On the other hand, females are associated with 0.6 percentage point (or 9%) and 0.7 percentage point (or 15.3%) higher post-neonatal under-five mortality compared to males respectively (neither of these coefficients is statistically significantly different from zero at $p<0.05$). The female coefficient in the earliest birth cohort (e.g. before 1995) is significantly higher than the female coefficients in the latter two birth cohorts, thus providing evidence of the female mortality disadvantage attributable to explicit discrimination weakening between the first cohort and the two subsequent ones. The coefficient between the second and third cohort are similar in magnitude, and indicate a stagnation of improvements after the mid-2000s. Re-specifying the birth cohorts by decades (e.g. born in or before the 1980s; born in the 1990s; born in the 2000s in Appendix Table 2, Panel C) yields substantively similar results, with a significantly elevated female mortality before the 1990s, both in absolute and relative measures, compared with the two successive cohorts. The changes that we document over subsequent birth cohorts are important given the difficulties of empirically assessing whether the micro-level processes that underlie the female under-five mortality disadvantage have changed over time in the context of changing forms of implicit discrimination.

Within mixed-sex twin placebo analysis using data from sub-Saharan African twins

To validate our within-twin measure, we conduct a placebo analysis using data on mixed-sex twins from sub-Saharan Africa (for an overview of the SSA sample see Appendix Table 3, Panel A). In the SSA analysis, we find females are associated with a 1.6 percentage point lower probability of infant and child mortality compared to males ($p<0.001$) (Table 2, Panel B, col 1), which corresponds to an 8.6% female advantage relative to male mortality in the sample. This finding is consistent with literature suggesting that males are biologically more vulnerable to infant and child mortality than females, which means that in a context without sex-differentials in allocation of resources there should actually be a female under-five mortality advantage (Drevenstedt et al 2008, Pongou 2013). Upon disaggregating by birth cohort, we find no evidence of statistically significant differences in the female coefficient across the three birth cohorts among mixed-sex twins in Africa (Table 2, Panel B, cols 2-4), unlike patterns of change in India. The fact that the SSA results are very different than those presented in India, provides further support that the elevated female mortality observed in our India twin sample captures explicit discrimination behaviors towards daughters. The India and

Africa comparative results are robust to a pooled, difference-in-difference analysis (Appendix Table 3, Panel B). The female x India interaction in Appendix Table 3, Panel B indicates that female mortality is about 3.5 percentage point higher in India relative to Africa, which corresponds to a 22% relative excess female mortality compared with male mortality in the pooled sample.

Exploring explicit discrimination and in infant and child mortality separately using within-twin fixed effects

We disaggregate our main outcome into post-neonatal infant mortality (1-11 months) and child mortality (12-59 months). Figure 1 presents results of mixed-sex within twin FE with post-neonatal infant and child mortality as separate outcomes and shows that mortality disadvantage for females is particularly pronounced between 12-59 months in the first time period (e.g. before 1995). Figure 2 further highlights how the survival gap between boys and girls widens after infancy. While girls experience a 2-percentage point (or 22.8%) higher mortality probability than boys between 1 and 11 months in the first time period ($p < 0.1$), the size of this effect for 12-59 months is 3.6 percentage points (150% higher) ($p < 0.5$).

Heterogeneity in explicit discrimination and the female mortality disadvantage by region and family composition using within-twin fixed effects

We re-run the within-twin FE models stratifying by region because we would expect to see more evidence of explicit discrimination in regions of the country where son preference has historically been the strongest such as in northern India. Consistent with this hypothesis we find that females experience a 3.6 percentage point (or 45%) higher probability of post-neonatal under-five mortality than males in northern India ($p<0.001$) (Table 3, Panel A, col 1). In contrast, the female coefficient in the pooled sample of other regions is much smaller in magnitude (slightly negative) (Table 3, Panel B), suggesting stronger explicit discrimination in the north relative to other regions. It is important to note that these findings do not mean that there is no evidence of female mortality disadvantage in other regions of the country, given that we would actually expect a female mortality advantage in a population with no son preference (see the Africa placebo in Table 2, Panel B).

Disaggregating by birth cohort, among mixed-sex twins born before 1995 (e.g. the earliest birth cohort), we find a sizeable impact of explicit discrimination in the northern region, with girls experiencing a 9.1 percentage point (or 76.4%) higher probability of mortality than males ($p<0.001$) (Table 3, Panel A, col 2). In the latter two birth cohorts, the size of this mortality disadvantage attributable to explicit discrimination weakens, to 0.6 and 1.9 percentage point higher probabilities of mortality compared to boys (neither of these coefficients is statistically significantly different from zero at $p<0.05$). Post-estimation tests of significance indicate the female coefficient in the earliest birth cohort in the northern region is significantly different from the female coefficients in the latter two birth cohorts in the northern region, which is indicative of weakening explicit discrimination over time in the region where son preference was historically strongest. However, we see a stalling of improvements in the cohorts after the mid-2000s compared with the mid-1990s in the northern region. Here, the female coefficient increases slightly between the second and third cohort, although the differences between the two cohorts are not statistically significant at $p < 0.05$. On the other hand, among cohorts born after the mid-2000s in the other regions there is indication of the emergence of a female mortality advantage as the coefficient becomes negative, although the female coefficients in the other regions are not statistically significantly different by birth cohorts (Table 3, Panel B). While we see a convergence in the female coefficient for cohorts in 1995-2005 between the northern and other regions, in the post-2005 cohorts we see a divergence between the regions, with the other region tending towards reductions in the female disadvantage whilst the northern region experiences stagnation.

Next, we explore heterogeneity by family composition given literature to assess if gender discrimination is selective by birth order and sibling composition. We re-run the within-twin FE models stratifying by number of older sisters and find females experience a 7.7 (96%) and 8.5 (126%) percentage point higher probability than males of post-neonatal under-five mortality among twins with two and three older sisters respectively ($p<0.001$) (Table 4, Panel A). Post-estimation tests of significance indicate the female coefficient in the model with no older sisters is significantly different from the female coefficients in the models with one, two, and three or more older sisters. This suggests that explicit discrimination is experienced particularly by later-born girls who have one or more sisters in the family already, rather than generalized to all girls within a family. In contrast, when we conduct a placebo analysis using data from Africa we find that there are no significant differences in the magnitude of the female coefficient across different sister combinations (Table 4, Panel B).

Supplementary Analyses

Although twin studies have widely been used to account for unobserved heterogeneity in demographic research (Guo and Tong 2006, Li et al. 2008, Marteleto and de Souza 2012, Pongou 2013, Nisen et al. 2013, Tropf and Mandemakers 2017), the external validity of estimates generated from twin data could be a concern. This might be particularly the case if the likelihood of having twins is not random, either because of genetic disposition for twins or because use of assisted reproductive technology (ART) that leads to higher rates of twinning (for dizygotic twins). [6] To partially account for the first possibility, we exclude families with

[6] ART has predominantly been associated with increased rates of dizygotic twinning. Dizygotic twins may be mixed-sex or same-sex, but all monozygotic twins are same-sex (Pison, Monden and Smits 2015).

multiple twins in the family (a sign of genetic predisposition for twinning). To partially account for the second possibility, we show the average birth year and other demographic characteristics for both mixed-sex and same-sex twins are very similar (see Appendix Table 4, panels B and C for same-sex twins). This is suggestive that ART is not driving the composition of mixed-sex twins, which would affect the mixed-sex twins' sample (entirely dizygotic) more than the same-sex sample (both monozygotic and dizygotic).[7] Although ART has increased over time in India, it is still concentrated to a relatively small urban elite, and thus it is highly unlikely that those who practice ART would be a large enough group to change the trend for the whole country over the long period of consideration (Smits and Monden 2011). Our samples show consistent known maternal factors associated with spontaneous twinning of maternal age and birth order (Hoekstra et al 2007), and in our sample both same and mixed-sex twins have mothers who are older, and likely in relation to this age pattern, have somewhat higher education than singletons. Perhaps the most striking difference between twin-births and singleton-births is the prevalence of infant and child mortality—a finding that holds for both same and mixed sex twins. This corresponds with literature suggesting that twins might be more vulnerable to mortality in infancy or childhood (Monden and Smits 2017).

As Table 1 shows, over time mixed-sex twins are increasingly born into different types of households (more educated, smaller sibship size etc.), but Appendix Table 4 shows that this is true for singletons and same-sex twins as well, thus indicating that the household characteristics of *both* singletons and twins are changing. This is further confirmed in Appendix Table 5 that shows that the interaction terms between family characteristics and birth cohort in predicting mixed-sex twin births are not changing significantly differently for twins rather than singletons. The sole exception is that there is a significant positive interaction between the last birth cohort and mother's tertiary education. This positive interaction only appears when the

[7] We are unable to measure zygosity in the NFHS data.

last three years (2014-2016) of births in our sample are included. Re-estimating the main models shown in Table 2 with these years excluded yields substantively similar results (see Appendix Table 7).

We also conduct an OLS regression analysis where we regress mortality on child sex using a large sample of singleton children and conduct our analyses using within family fixed effects (Appendix Table 6, Panels A and B). In the OLS and within family fixed effects models, we find females experience 0.7 and 0.6 percentage point (about 20%) higher probability of post-neonatal under-five mortality than males ($p<0.001$), net of controls for family and child characteristics. While we are hesitant to directly compare these OLS results to those generated by the within-twin FE, we cautiously note that the female disadvantage in mortality, both in absolute and relative measures, indicated by the magnitude of the coefficient on the pooled OLS models is smaller than those on the within-twin FE models. This difference is especially so for cohorts born prior to 1995. Although they use a different estimation strategy than ours to exploit random variation in the sex of the child, Barcellos, Carvalho and Lleras-Muney (2014) also find that estimates of gender discrimination in health investments and breastfeeding indicated in their "experimental" sample is higher than in standard OLS estimates before 1992. This suggests that in this period OLS estimates likely overstate the role of implicit discrimination processes when estimating the female mortality disadvantage as completed family size and birth order variables are likely to capture some of the excluded explicit discrimination processes as well. Interestingly, the opposite pattern is visible for the cohorts born 1995-2005, for which the relative female disadvantage in mortality indicated by the within-twin FE model is smaller than that OLS estimates, suggesting that actually explicit discrimination is weaker than OLS estimates would suggest.

In the main analysis, we do not include controls for birthweight due to the very high missing values for this variable (e.g. over 80% of respondents in the pooled sample are missing this information), but our indicator of first-born twin likely captures whether the twin was higher or lower birth weight due to the correlation between twin birth order and birth weight. We nevertheless perform sensitivity tests with imputed birth weight and re-run our analyses; results are unchanged (Appendix Table 2, Panel E).

Limitations

One potential limitation of the within-twin fixed effects approach is if there are unobserved confounders that vary across mixed-sex twins. Unlike identical (or monozygotic) twins who share 100% of their genetic material, mixed sex twins are fraternal (or dizygotic) and share 50% of their genetic material (about the same amount of genetic material that non-twin siblings share). Thus, it is plausible that there are unobservable genetic differences between twins that lead to differential parental care and attention, net of child sex. Another biological mechanism that may operate distinctively among mixed-sex twins is intrauterine hormonal exposure (Tapp et al. 2011, Ahrenfeldt et al 2019). According to this, girls in mixed-sex twins may experience high levels of prenatal testosterone exposure and consequently may be more “masculinized” in their development than singleton or same-sex twin girls. Conversely, boys exposed to a female co-twin may be more “feminized” due to estrogen exposure. Evidence for this hypothesis in relation to early life mortality outcomes is limited, and when available mixed and inconclusive (Pongou 2013, Ahrenfeldt et al 2017). Nevertheless, throughout our analysis, by comparing mixed-sex twins to each other, and conduct our placebo analysis also on African mixed-sex twins, our results are unaffected even if sex differences in mixed-sex twin populations may be less pronounced than in singleton populations.

It is plausible that twins may be a greater negative shock than singleton births, leading twins to receive differential treatment than other types of children. If this is the case, we could interpret our estimates to be upper bound estimates of explicit discrimination, given as, described before, twins are not that dissimilar to other children on most observable characteristics. Even if overall mortality outcomes are worse for twins, in the absence of son preference we should not expect a female disadvantage in mortality among females in mixed-sex twins. Ultimately, the exceptional nature of twin births is what makes them so interesting for our experimental design by allowing us to control for differential family selectivity into having a less desired female child.

A final limitation is that our analysis provides a useful way to measure the mortality impact of explicit discrimination net of implicit discrimination, but it does not shed much insight into the specific mechanisms through which explicit discrimination operates. Although we hypothesize that differential allocation of resources is the main mechanism through which the explicit discrimination that leads to elevated female mortality among mixed-sex twins operates, it is difficult to test this directly using NFHS data and the within-twin FE approach due to significant data limitations. The NFHS collects early childhood health and nutrition measures for children born only in the last five years, leaving us with a small sample of twins born in the last five years with full nutrition and health information. Even for children born in the last five years, there is no information on key measures (e.g. height-for-age, stunting) for deceased children. Furthermore, children who have died will likely be different in key characteristics (e.g. they might have been breastfed less, have lower probability of vaccination etc.) than children who survived through early childhood, and would not be possible to know whether these differences led to death (e.g. they died because they were not immunized), or whether early death led to these differences (e.g. they would have been immunized if they had survived longer). Finally, if a child died, their surviving twin may have received better treatment because of the death (e.g. parents may dote on the offspring who survives to compensate for the loss ex-post facto). Ultimately, mortality is the strongest benchmark of discrimination, one that is least prone to recall bias and it is what our data best allow us to measure, and thus, we have focused on it.

Discussion

One of the most striking demographic manifestations of son preference in India is the persistence of excess female under-five mortality. While considerable literature attributes the under-five female mortality disadvantage to parents differentially investing more resources in boys versus girls within families, which we term *explicit discrimination,* this literature does not adequately control for what we term *implicit* discrimination processes that sort girls into different types of families (e.g. larger, poorer, or with varying son preference) and at different birth orders than boys. It is conceptually important to recognise these two distinctive micro-level mechanisms of discrimination as the family-level processes implied by each are different, as are the policy responses required by them. To better address the endogeneity associated with implicit discrimination processes, we explored the association between child sex and post-neonatal under-five mortality using a sample of mixed-sex twins. We argued that mixed-sex twins provided a natural experiment that exogenously assigned a boy and a girl to families at the same time, thus controlling for family selectivity into having an unwanted female child, birth order, and other implicit discrimination processes.

Our within-twin fixed effects models showed that female children experienced significantly higher probability of post-neonatal under-five mortality. This provided strong evidence of explicit discrimination playing an important role in the female mortality disadvantage observed in our data because our models controlled for implicit discrimination processes that resulted in boys and girls being differentially sorted into different kinds of families and birth orders. The Indian estimates for explicit discrimination were particularly striking when compared to a placebo analysis conducted in sub-Saharan Africa where female twins actually had a survival advantage, which corresponded with literature showing that males have a biological disadvantage in early life.

Using our novel measure, we found that that the role of explicit discrimination underlying the female mortality disadvantage weakened for cohorts born after the mid-1990s relative to those born prior to mid-1990s. Subsequent analyses also showed that our temporal results were largely driven by northern India, and explicit discrimination declined over time in this region that has historically been characterized by high son preference. Nevertheless, our results do not indicate the disappearance of a mortality disadvantage attributable to explicit discrimination in India, and indeed the cohorts born after the mid-2000s in northern India appeared to show stalling improvements.

Although we are not able to test them directly, we anticipate a combination of contextual changes – weakening son preference, policy initiatives aimed at improving girls' status, fertility decline, as well as the ability to realize son preference at lower parities due to the practice of sex-selective abortion –underpin weakening explicit discrimination in India since the mid-1990s. Compared with the pre-1995 period, we find that later cohorts of mixed sex twins are born into smaller families with lower indicators of stated son preference. While these results refer to stated preference indicators, it is plausible that son preference has weakened through socioeconomic development (Chung and Das Gupta 2007, Kashyap and Villavicencio 2016) via channels such as improved educational and economic opportunities for women (Murthi et al 1995, Bhat and Zavier 2003, Pande and Astone 2007, Luke and Munshi 2011), and media exposure (Jensen and Oster 2009, Lin and Adsera 2013, Ting et al. 2014). Furthermore, since the 1990s, several states across India launched financial incentive policies targeted at encouraging investments in daughters' health and education. While implementation of these policies has been irregular and their systematic review limited (Sekher 2012), available evidence suggests improvements in postnatal outcomes after their implementation (Sinha and Joong 2009, Srinivasan and Bedi 2009, Sekher and Ram 2015).

Existing research has found that weakening son preference is correlated with fertility decline (Bhat and Zavier 2003). With reductions in overall family size, differences in resource allocations may become less pronounced, particularly as we find that it is girls at higher birth orders with existing sisters who are most vulnerable to explicit discrimination. Explicit discrimination could have also partly weakened not because son preference weakened but because parents were able to realize son preference with their desire for smaller families through sex-selective abortion (e.g. Anukriti, Bhalotra and Tam 2018, Kashyap 2019). In a context with strong son preference such as India, fertility decline may be enabled by sex-selective abortion. While the mechanism of "opting out" of having unwanted daughters does not apply to mixed-sex twins in the same way as it does to other births, which precisely makes our strategy better at controlling for this form of implicit discrimination, parents with mixed-sex twins could opt out of having other children and get their preferred sex ratio with fewer overall children.

The discussion above suggests that targeting son preference solely through policies that ban prenatal sex selection may be counterproductive for reducing explicit discrimination. The uneven adoption of sex-selective abortion among more advantaged households however could imply that girls are increasingly selected into the most disadvantaged households, which may underpin continued explicit discrimination. The fact that improvements appear to have stalled among more recent cohorts since the mid-2000s suggests that policy responses that target health investments in girls specifically such as through financial incentive schemes, rather than generalized poverty policies that are target for all children, are still necessary. These policies, combined with indirect measures to weaken son preference through media advocacy as well as measures to improve women's outcomes in political, legal and economic domains, are needed to further reduce explicit discrimination.

While the approach presented in our paper does not provide an all-encompassing measure of son preference (indeed it is one of several ways to explore son preference), we provided a conceptual contribution by distinguishing between explicit and implicit discrimination processes, and demonstrated a quasi-experimental approach to better estimate explicit discrimination effects using a novel sample. It is fundamental for our understanding of son preference if the effects of son preference arise due to parents' actively investing more in boys over girls in families or whether they accrue due to changing patterns of implicit discrimination. A significant contribution of our approach was that it allowed for the temporal analysis of the impact of explicit discrimination on the female mortality disadvantage in India over four decades.

Although our analysis was focused on India, the points we raise about the different processes behind both explicit and implicit discrimination can also be applied to other high son preference contexts in South, East, and Central Asia. This distinction may become particularly important as wider changes such as fertility decline and technology diffusion continue across countries with historically strong son preference, implying that changing family selectivity — as opposed to differential resource allocation within families—could become a particularly important mechanism through which the mortality manifestations of son preference emerge. Future research should explore the mechanisms of both explicit and implicit discrimination, and better understand the interrelationship between changing son preference, fertility decline, and excess female mortality.

References

Ahrenfeldt, Linda Juel, Kaare Christensen, Nancy L. Segal, and Yoon-Mi Hur. "Opposite-sex and same-sex twin studies of physiological, cognitive and behavioral traits." *Neuroscience & Biobehavioral Reviews* (2019).

Ahrenfeldt, Linda Juel, Lisbeth Aagaard Larsen, Rune Lindahl-Jacobsen, Axel Skytthe, Jacob vB Hjelmborg, Sören Möller, and Kaare Christensen. "Early-life mortality risks in opposite-sex and same-sex twins: a Danish cohort study of the twin testosterone transfer hypothesis." *Annals of epidemiology* 27, no. 2 (2017): 115-120.

Alkema, Leontine, Fengqing Chao, Danzhen You, Jon Pedersen, and Cheryl C. Sawyer. "National, regional, and global sex ratios of infant, child, and under-5 mortality and identification of countries with outlying ratios: a systematic assessment." *The Lancet Global Health* 2, no. 9 (2014): e521-e530.

Anukriti, S. "Financial incentives and the fertility-sex ratio trade-off." *American Economic Journal: Applied Economics*10, no. 2 (2018): 27-57.

Anukriti, S., Sonia Bhalotra, and Hiu Tam. *On the Quantity and Quality of Girls: Fertility, Parental Investments, and Mortality*. No. 950. Boston College Department of Economics, 2018.

Arnold, Fred, Minja Kim Choe, and Tarun K. Roy. "Son preference, the family-building process and child mortality in India." *Population studies* 52, no. 3 (1998): 301-315.

Arokiasamy, Perianayagam. "Regional patterns of sex bias and excess female child mortality in India." *Population* 59, no. 6 (2004): 831-863.

Barcellos, Silvia Helena, Leandro S. Carvalho, and Adriana Lleras-Muney. "Child gender and parental investments in India: Are boys and girls treated differently?" *American Economic Journal: Applied Economics* 6, no. 1 (2014): 157-89.

Basu, Alaka Malwade. "Is discrimination in food really necessary for explaining sex differentials in childhood mortality?" *Population studies* 43, no. 2 (1989): 193-210.

Basu, Deepankar, and Robert De Jong. "Son targeting fertility behavior: some consequences and determinants." *Demography* 47, no. 2 (2010): 521-536.

Bhalotra, Sonia. R., & Cochrane, Tom. "Where have all the young girls gone? Identification of sex selection in India" *Institute for Labor Economics (IZA) Discussion Paper 5381* (2010).

Bharadwaj, Prashant, and Leah K. Lakdawala. "Discrimination begins in the womb: Evidence of sex-selective prenatal investments." *Journal of Human Resources* 48, no. 1 (2013): 71-113.

Bhat, PN Mari, and AJ Francis Zavier. "Fertility decline and gender bias in northern India." *Demography* 40, no. 4 (2003): 637-657.

Bongaarts, John. "The implementation of preferences for male offspring." *Population and Development Review* 39, no. 2 (2013): 185-208.

Borooah, Vani K. "Gender bias among children in India in their diet and immunisation against disease." *Social science & medicine* 58, no. 9 (2004): 1719-1731.

Caldwell, John C., P. Hemachandra Reddy, and Pat Caldwell. "The causes of demographic change in rural South India: A micro approach." *Population and development review* (1982): 689-727.

Caldwell, Pat. and John C. Caldwell. 1990. *Gender Implications for Survival in South Asia*. Health Transition Working Paper No. 7. Canberra: National Centre for Epidemiology and Population Health, Australian National University.

Chung, Woojin, and Monica Das Gupta. "The decline of son preference in South Korea: The roles of development and public policy." *Population and development review* 33, no. 4 (2007): 757-783.

Clark, Shelley. "Son preference and sex composition of children: Evidence from India." *Demography* 37, no. 1 (2000): 95-108.

Corsi, Daniel J., Diego G. Bassani, Rajesh Kumar, Shally Awasthi, Raju Jotkar, Navkiran Kaur, and Prabhat Jha. "Gender inequity and age-appropriate immunization coverage in India from 1992 to 2006." *BMC international health and human rights* 9, no. 1 (2009): S3.

Corsi, Daniel J., Michelle F. Gaffey, Diego G. Bassani, and S. V. Subramanian. "No Female Disadvantage in Anthropometric Status among Children in India: Analysis of the 1992–1993 and 2005–2006 Indian National Family Health Surveys." *Journal of South Asian Development* 10, no. 2 (2015): 119-147.

Das Gupta, Monica. "Selective discrimination against female children in rural Punjab, India." *Population and development review* (1987): 77-100.

Deaton, Angus. "Health, inequality, and economic development." *Journal of economic literature* 41, no. 1 (2003): 113-158.

Drevenstedt, Greg L., Eileen M. Crimmins, Sarinnapha Vasunilashorn, and Caleb E. Finch. "The rise and fall of excess male infant mortality." *Proceedings of the National Academy of Sciences* 105, no. 13 (2008): 5016-5021.

Dyson, Tim, and Mick Moore. "On kinship structure, female autonomy, and demographic behavior in India." *Population and development review* (1983): 35-60.

Filmer, Deon, Jed Friedman, and Norbert Schady. "Development, modernization, and childbearing: The role of family sex composition." *The World Bank Economic Review* 23, no. 3 (2009): 371-398.

Fledderjohann, Jasmine, Sutapa Agrawal, Sukumar Vellakkal, Sanjay Basu, Oona Campbell, Pat Doyle, Shah Ebrahim, and David Stuckler. "Do girls have a nutritional disadvantage compared with boys? Statistical models of breastfeeding and food consumption inequalities among Indian siblings." *PLoS one*, no. 9 (2014): e107172.

Ganatra, Bela, and Siddhivinayak Hirve. "Male bias in health care utilization for under-fives in a rural community in western India." *Bulletin of the World Health Organization* 72, no. 1 (1994): 101.

Goodkind, Daniel. "On substituting sex preference strategies in East Asia: Does prenatal sex selection reduce postnatal discrimination?" *Population and Development Review* (1996): 111-125.

Guilmoto, Christophe Z., Nandita Saikia, Vandana Tamrakar, and Jayanta Kumar Bora. "Excess under-5 female mortality across India: a spatial analysis using 2011 census data." *The Lancet Global Health* 6, no. 6 (2018): e650-e658.

Guilmoto, Christophe Z., and James Tove. "The masculinization of births. Overview and current knowledge." *Population* 70, no. 2 (2015): 185-243.

Guo, Guang, and Yuying Tong. "Age at first sexual intercourse, genes, and social context: Evidence from twins and the dopamine D4 receptor gene." *Demography* 43, no. 4 (2006): 747-769.

Hu, Luojia, and Analía Schlosser. "Prenatal Sex Selection and Girls' Well-Being: Evidence from India." *The Economic Journal* 125, no. 587 (2015): 1227-1261.

Jayachandran, Seema. "Fertility decline and missing women." *American Economic Journal: Applied Economics* 9, no. 1 (2017): 118-39.

Jayachandran, Seema, and Ilyana Kuziemko. "Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India." *The Quarterly journal of economics* 126, no. 3 (2011): 1485-1538.

Jayachandran, Seema, and Rohini Pande. "Why are Indian children so short? The role of birth order and son preference." *American Economic Review* 107, no. 9 (2017): 2600-2629.

Jensen, Robert, and Emily Oster. "The power of TV: Cable television and women's status in India." *The Quarterly Journal of Economics* 124, no. 3 (2009): 1057-1094.

Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin. "Low male-to-female sex ratio of children born in India: national survey of 1· 1 million households." *The Lancet* 367, no. 9506 (2006): 211-218.

Kashyap, Ridhi. "Is prenatal sex selection associated with lower female child mortality?" *Population Studies* 73, no. 1 (2019): 57-78.

Kashyap, Ridhi, and Francisco Villavicencio. "The dynamics of son preference, technology diffusion, and fertility decline underlying distorted sex ratios at birth: A simulation approach." *Demography* 53, no. 5 (2016): 1261-1281.

Kashyap, Ridhi, and Francisco Villavicencio. "An agent-based model of sex ratio at birth distortions." In *Agent-based modelling in population studies*, pp. 343-367. Springer, Cham, 2017.

Li, Hongbin, Junsen Zhang, and Yi Zhu. "The quantity-quality trade-off of children in a developing country: Identification using Chinese twins." *Demography* 45, no. 1 (2008): 223-243.

Lin, Tin-chi, and Alícia Adserà. "Son preference and children's housework: the case of India." *Population research and policy review* 32, no. 4 (2013): 553-584.

Luke, Nancy, and Kaivan Munshi. "Women as agents of change: Female income and mobility in India." *Journal of Development Economics* 94, no. 1 (2011): 1-17.

Marteleto, Letícia J., and Laetícia R. de Souza. "The changing impact of family size on adolescents' schooling: Assessing the exogenous variation in fertility using twins in Brazil." *Demography* 49, no. 4 (2012): 1453-1477.

Miller, Barbara D. *The endangered sex: neglect of female children in rural North India*. Cornell University Press, 1981.

Mishra, Vinod, Tarun K. Roy, and Robert D. Retherford. "Sex differentials in childhood feeding, health care, and nutritional status in India." *Population and development review* 30, no. 2 (2004): 269-295.

Mishra, S. K., Bali Ram, Abhishek Singh, and Awdhesh Yadav. "Birth order, stage of infancy and infant mortality in India." *Journal of biosocial science* 50, no. 5 (2018): 604-625.

Monden, Christiaan WS, and Jeroen Smits. "Mortality among twins and singletons in sub-Saharan Africa between 1995 and 2014: a pooled analysis of data from 90 Demographic and Health Surveys in 30 countries." *The Lancet Global Health* 5, no. 7 (2017): e673-e679.

Muhuri, Pradip K., and Samuel H. Preston. "Effects of family composition on mortality differentials by sex among children in Matlab, Bangladesh." *Population and Development Review* (1991): 415-434.

Murthi, Mamta, Anne-Catherine Guio, and Jean Dreze. "Mortality, fertility, and gender bias in India: A district-level analysis." *Population and development review* (1995): 745-782.

Nisén, Jessica, Pekka Martikainen, Jaakko Kaprio, and Karri Silventoinen. "Educational differences in completed fertility: A behavioral genetic study of Finnish male and female twins." *Demography* 50, no. 4 (2013): 1399-1420.

Pande, Rohini P. "Selective gender differences in childhood nutrition and immunization in rural India: the role of siblings." *Demography* 40, no. 3 (2003): 395-418.

Pande, Rohini P., and Nan Marie Astone. "Explaining son preference in rural India: the independent role of structural versus individual factors." *Population Research and Policy Review* 26, no. 1 (2007): 1-29.

Pison, Gilles, Christiaan Monden, and Jeroen Smits. "Twinning rates in developed countries: trends and explanations." *Population and Development Review* 41, no. 4 (2015): 629-649.

Pongou, Roland. "Why is infant mortality higher in boys than in girls? A new hypothesis based on preconception environment and evidence from a large sample of twins." *Demography* 50, no. 2 (2013): 421-444.

Rajan, Sowmya, and S. Philip Morgan. "Selective Versus Generalized Gender Bias in Childhood Health and Nutrition: Evidence from India." *Population and Development Review* (2018).

Retherford, Robert D., and T. K. Roy. "Factors affecting sex-selective abortion in India and 17 major states." Mumbai: International Institute for Population Sciences" (2003) (National Family Health Survey Subject Reports No. 21).

Rosenblum, Daniel. "The effect of fertility decisions on excess female mortality in India." *Journal of Population Economics* 26, no. 1 (2013): 147-180.

Sekher, T. V. "Ladlis and Lakshmis: financial incentive schemes for the girl child." *Economic and Political Weekly* (2012): 58-65.

Sekher, T. V., and F. Ram. "Conditional cash transfers for girls in India: Assessment of a girl child promotion scheme from beneficiary perspective." *International Institute for Population Sciences* (2015).

Sen, Amartya. "More than 100 million women are missing." *New York Review of Books* (1990): 61-66.

Sen, Amartya. "Missing women revisited." *BMJ* (2003): 327.

Sinha, Nistha, and Joanne Yoong. *Long-term financial incentives and investment in daughters: Evidence from conditional cash transfers in North India*. The World Bank, 2009.

Smits, Jeroen, and Christiaan Monden. "Twinning across the developing world." *PLoS One* 6, no. 9 (2011): e25239.

Srinivasan, Sharada, and Arjun S. Bedi. "Girl child protection scheme in Tamil Nadu: an appraisal." *Economic and political weekly* (2009): 10-12.

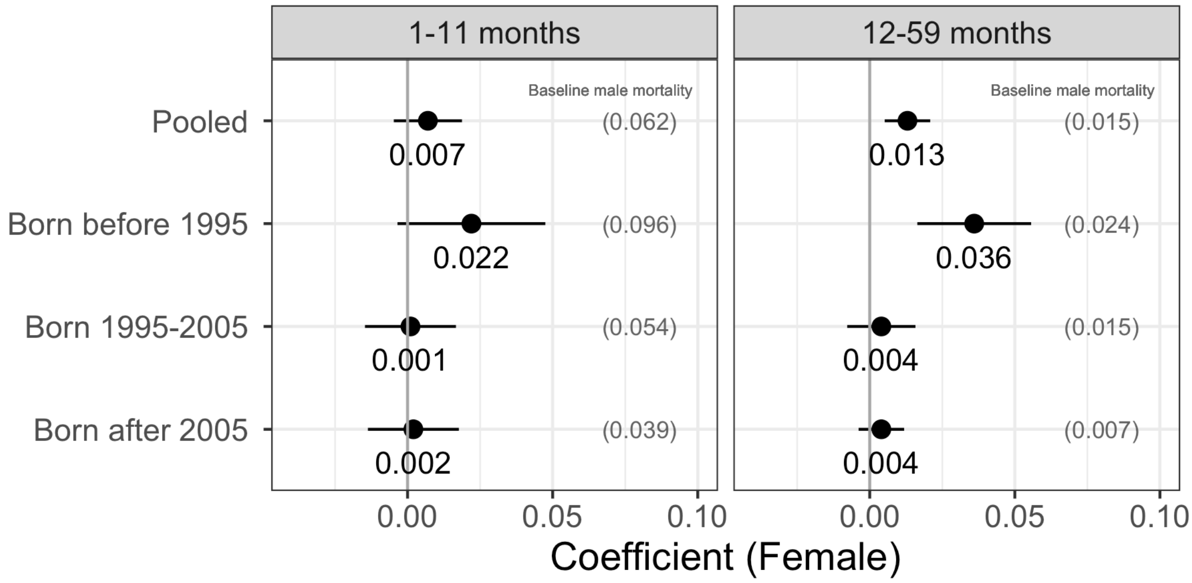
Tapp, Aimee L., Murray T. Maybery, and Andrew JO Whitehouse. "Evaluating the twin testosterone transfer hypothesis: a review of the empirical evidence." *Hormones and behavior* 60, no. 5 (2011): 713-722.

Ting, Hsin-Lan, Chon-Kit Ao, and Ming-Jen Lin. "Television on women's empowerment in India." *The Journal of Development Studies* 50, no. 11 (2014): 1523-1537.

Tropf, Felix C., and Jornt J. Mandemakers. "Is the association between education and fertility postponement causal? The role of family background factors." *Demography* 54, no. 1 (2017): 71-91.

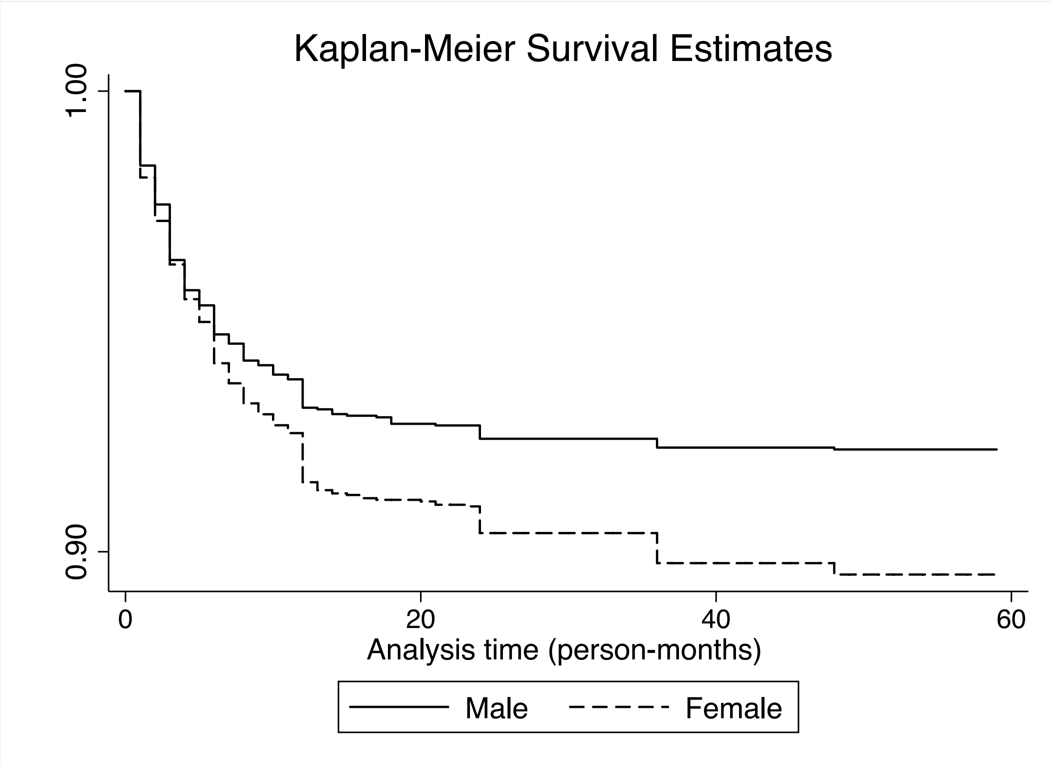
Yamaguchi, Kazuo. "A formal theory for male-preferring stopping rules of childbearing: Sex differences in birth order and in the number of siblings." *Demography* 26, no. 3 (1989): 451-465.

Figure 1. Results of within mixed-sex twin fixed effects analyses of the effect of child sex on infant mortality (1-11 months) (left panel) and within mixed-sex twin fixed effects analyses of the effect of child sex on child mortality (12-59 months) (right panel). *Note: Baseline male mortality probability of the relevant sample is shown in parentheses.*



Source: Created by the authors using data from the NFHS

Figure 2. Kaplan-Meier Survival estimates of time to death in person-months for male and female twins in our sample.



Note: Children enter our sample at 1 month and are censored at age at survey end or 59 months, whichever comes first. Neonatal mortality excluded.

Source: Created by the authors using data from the NFHS

Table 1. Descriptive summary of background characteristics of mixed-sex twins, including tests for significant difference between the first birth cohort and the two subsequent cohorts. All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016) and use sampling weights provided by the NFHS.

	(1) *Pooled*	(2) *Born before 1995*	(3) *Born 1995-2005*	(4) *Born after 2005*
Mortality (1-59 mos.)	0.09	0.17	**0.07**	**0.05**
Female	0.50	0.50	0.50	0.50
Birth year	2000	1986	**2000**	**2011**
Birth order	3.26	3.68	**3.18**	**2.98**
Rural	0.68	0.69	0.66	0.68
Northern region	0.60	0.61	0.58	0.60
Hindu	0.79	0.77	0.80	0.79
Poorest 40%	0.40	0.36	0.38	**0.46**
Mother no school	0.45	0.58	**0.44**	**0.33**
Mother primary school	0.16	0.20	0.17	**0.12**
Mother secondary school	0.30	0.19	**0.32**	**0.39**
Mother tertiary	0.09	0.04	**0.07**	**0.16**
Mother age at birth	25.32	24.68	24.92	**26.34**
Total children born to mother	4.32	5.22	**4.19**	**3.69**
Mother's ideal number of boys	1.36	1.56	**1.31**	**1.26**
Mother's ideal number of girls	1.08	1.14	**1.07**	**1.05**
Mother's ideal sex ratio	0.55	0.57	**0.55**	**0.54**
N	6200	1868	2278	2054

Notes: All measures are dichotomous except birth year (ranges from 1958 to 2016), birth order (ranges from 1 to 11), mother age at birth (ranges from 12 to 47), total children born (ranges from 2 to 13), mother's ideal boys (ranges from 0 to 7), mother's ideal girls (ranges from 0 to 6), and ideal sex ratio (ranges from 0 to 1 and excludes women who desire zero children). Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Two-sample t tests for all continuous outcomes, and chi-square tests for all dichotomous outcomes.

Table 2. Within mixed sex twin fixed effects models of the association between child sex and infant and child mortality (1-59 months) in India (Panel A) and Africa (Panel B). Estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016) and the Demographic and Health Surveys in Africa (See Appendix Table 1 for full list of African countries and survey waves). Analysis conducted in STATA 15.

Panel A. India	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.020**	0.053***	**0.006**	**0.007**
	(0.006)	(0.015)	**(0.010)**	**(0.009)**
First born twin	-0.039***	-0.023	-0.047***	-0.042***
	(0.006)	(0.015)	(0.010)	(0.009)
Constant	0.097***	0.131***	0.093***	0.069***
	(0.006)	(0.013)	(0.009)	(0.008)
Observations	6,200	1,868	2,278	2,054
R-squared	0.017	0.019	0.021	0.023
Number of families	3,100	934	1,139	1,027
Baseline male mortality	0.075	0.118	0.068	0.046

Panel B. Africa	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	-0.016***	-0.024**	-0.008	-0.013
	(0.005)	(0.008)	(0.007)	(0.009)
First born twin	-0.042***	-0.066***	-0.037***	0.001
	(0.005)	(0.008)	(0.007)	(0.009)
Constant	0.209***	0.280***	0.189***	0.098***
	(0.004)	(0.007)	(0.007)	(0.008)
Observations	17,963	7,124	7,533	3,306
R-squared	0.009	0.018	0.007	0.001
Number of families	8,988	3,566	3,767	1,655
Baseline male mortality	0.185	0.243	0.169	0.099

*** p<0.001, ** p<0.01, * p<0.05

Notes: Bold numbers indicate statistically significant (p<0.05) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995).

Table 3. Within mixed sex twin fixed effects models of the association between child sex and infant and child mortality (1-59 months) for Northern regions (Panel A) and other regions (Panel B). All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016). Analysis conducted in STATA 15.

Panel A. Northern regions	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.036***	0.091***	**0.006**	**0.019**
	(0.009)	(0.020)	**(0.014)**	**(0.013)**
First born twin	-0.047***	-0.013	-0.060***	-0.063***
	(0.009)	(0.020)	(0.014)	(0.013)
Constant	0.108***	0.127***	0.113***	0.082***
	(0.008)	(0.018)	(0.013)	(0.012)
Observations	3,482	1,074	1,248	1,160
R-squared	0.028	0.042	0.030	0.047
Number of families	1,741	537	624	580
Baseline male mortality	0.080	0.119	0.079	0.047

Panel B. Other regions	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	-0.001	0.004	0.003	-0.010
	(0.009)	(0.021)	(0.013)	(0.011)
First born twin	-0.026**	-0.032	-0.031*	-0.014
	(0.009)	(0.021)	(0.013)	(0.011)
Constant	0.083***	0.134***	0.070***	0.052***
	(0.008)	(0.019)	(0.011)	(0.010)
Observations	2,718	794	1,030	894
R-squared	0.006	0.006	0.011	0.005
Number of families	1,359	397	515	447
Baseline male mortality	0.069	0.116	0.054	0.045

*** $p<0.001$, ** $p<0.01$, * $p<0.05$

Notes: Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995).

Table 4. Within mixed sex twin fixed effects models of the association between child sex and infant and child mortality (1-59 months) disaggregated by number of older sisters for India (Panel A) and Africa (Panel B). All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016) and the Demographic and Health Surveys in Africa (See Appendix Table 1 for full list of African countries and survey waves). Analysis conducted in STATA 15.

Panel A. India	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	No older sisters	One older sister	Two older sisters	Three older sisters
Female	-0.009	**0.025**	**0.077******	**0.085******
	(0.008)	**(0.013)**	**(0.019)**	**(0.024)**
First born twin	-0.029***	-0.049***	-0.054**	-0.062*
	(0.008)	(0.013)	(0.019)	(0.024)
Constant	0.088***	0.110***	0.108***	0.098***
	(0.007)	(0.011)	(0.017)	(0.021)
Observations	3,026	1,740	930	504
R-squared	0.009	0.021	0.051	0.070
Number of families	1,513	870	465	252
Baseline male mortality	0.071	0.084	0.080	0.067

Panel B. Africa	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	No older sisters	One older sister	Two older sisters	Three older sisters
Female	-0.026**	-0.003	-0.014	-0.024
	(0.008)	(0.009)	(0.011)	(0.012)
First born twin	-0.045***	-0.043***	-0.039***	-0.031*
	(0.008)	(0.009)	(0.011)	(0.012)
Constant	0.217***	0.209***	0.201***	0.200***
	(0.007)	(0.008)	(0.010)	(0.011)
Observations	6,253	5,662	3,384	2,664
R-squared	0.012	0.009	0.008	0.007
Number of families	3,129	2,835	1,693	1,332
Baseline male mortality	0.192	0.184	0.178	0.183

*** p<0.001, ** p<0.01, * p<0.05

Notes: Bold numbers indicate statistically significant (p<0.05) difference between the sister category in question and no older sisters.

Appendix Table 1. Overview of Sub-Saharan Africa DHS survey year and samples used

Country	**Year**	**Year**	**Year**	**Year**	**Year**	**Year**
Benin	1996	2001	2006	2011-12		
Burkina Faso	1993	1998-99	2003	2010		
Cameroon	1991	1998	2004	2011		
Ghana	1993	1998	2003	2008	2014	
Kenya	1993	1998-99	2003	2008	2014	
Madagascar	1992	1997	2003-04	2008-09		
Malawi	1992	2000	2004	2010		
Mali	1995-96	2001	2006	2012-13		
Namibia	1992	2000	2006-07	2013		
Niger	1992	1998	2006	2012		
Nigeria	1990	2003	2008	2013		
Rwanda	1992	2000	2005	2010	2014-15	
Senegal	1992	1997	2005	2010-11	2012-13	2014
Tanzania	1991-92	1996	1999	2004-05	2010	
Uganda	1995	2000-01	2006	2011		
Zambia	1992	1996	2001-02	2006	2013-14	
Zimbabwe	1994	1999	2005-06	2010-11		

Appendix Table 2. Alternative specifications of within mixed sex twin fixed effects models of the association between child sex and infant and child mortality (1-59 months) in India; Panel A uses logistic regression fixed effects with results presented as odds ratios; Panel B uses Cox proportional hazard models with fixed effects with results presented as hazard ratios; Panel C uses linear regression fixed effects, but limits to births in the 10 years before the survey; Panel D uses linear regression fixed effects, but uses an alternative specification of birth cohorts; Panel E uses linear regression fixed effects, with a control for birth weight created through multiple imputation. Estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016). Analysis conducted in STATA 15.

Panel A. Logistic regression fixed effects	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	1.468***	1.891***	1.088	0.980
	(0.160)	(0.301)	(0.219)	(0.263)
First born twin	0.514***	0.680*	0.397***	0.311***
	(0.056)	(0.108)	(0.080)	(0.083)
Observations	782	366	248	168
Number of families	391	183	124	84

Panel B. Cox proportional hazards	(1)	(2)	(3)	(4)
	Age at death (months)	Age at death (months)	Age at death (months)	Age at death (months)
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	1.427***	2.006***	1.094	0.685
	(0.146)	(0.290)	(0.207)	(0.181)
First born twin	0.420***	0.716*	0.280***	0.146***
	(0.044)	(0.103)	(0.056)	(0.042)
Observations	6,190	1,866	2,278	2,046
Number of families	3095	933	1139	1023

Panel C. Limiting to births in last 10 years	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.022*	0.102***	**-0.006**	**0.007**
	(0.009)	(0.027)	**(0.022)**	**(0.009)**
First born twin	-0.019*	0.043	-0.001	-0.042***
	(0.009)	(0.027)	(0.022)	(0.009)

Constant	0.077***	0.076**	0.100***	0.069***
	(0.008)	(0.025)	(0.020)	(0.008)
Observations	3,298	600	644	2,054
R-squared	0.007	0.049	0.000	0.023
Number of families	1,649	300	322	1,027
Baseline male mortality	0.066	0.100	0.099	0.046

Panel D. Alternative birth cohorts	(1)	(2)	(3)	(4)
	Mortality 1-59 mos. Pooled	Mortality 1-59 mos. Born before or in 1980s	Mortality 1-59 mos. Born 1990s	Mortality 1-59 mos. Born in 2000s
Female	0.020**	0.063**	**0.014**	**0.009**
	(0.006)	(0.020)	**(0.013)**	**(0.007)**
First born twin	-0.039***	-0.014	-0.047***	-0.042***
	(0.006)	(0.020)	(0.013)	(0.007)
Constant	0.097***	0.145***	0.112***	0.073***
	(0.006)	(0.018)	(0.012)	(0.006)
Observations	6,200	1,142	1,676	3,382
R-squared	0.017	0.020	0.019	0.022
Number of families	3,100	571	838	1,691
Baseline male mortality	0.075	0.137	0.086	0.050

Panel E. Including birth weight	(1)	(2)	(3)	(4)
	Mortality 1-59 mos. Pooled	Mortality 1-59 mos. Born before 1995	Mortality 1-59 mos. Born 1995-2005	Mortality 1-59 mos. Born after 2005
Female	0.020**	0.053***	**0.006**	**0.007**
	(0.006)	(0.015)	**(0.010)**	**(0.009)**
First born twin	-0.039***	-0.022	-0.047***	-0.042***
	(0.006)	(0.015)	(0.010)	(0.009)
Birth weight	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.103***	0.131**	0.106***	0.076*
	(0.021)	(0.048)	(0.032)	(0.032)
Observations	6,200	1,868	2,278	2,054
Number of families	3,100	934	1,139	1,027
Baseline male mortality	0.075	0.118	0.068	0.046

*** p<0.001, ** p<0.01, * p<0.05

Notes: Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). In Panel B, respondents are censored at age at survey or 59 months (whichever comes first).

Appendix Table 3. Additional information on sub-Saharan African sample; Panel A presents descriptive summary of background characteristics of mixed-sex twins for sub-Saharan Africa sample, including tests for significant difference between the first birth cohort and the two subsequent cohorts (estimates use sampling weights provided by the DHS); Panel B presents difference-in-difference analysis using pooled Africa and India data and an interaction between female and India.

Panel A.	(1)	(2)	(3)	(4)
	Pooled	*Born before 1995*	*Born 1995-2005*	*Born after 2005*
Mortality (1-59 mos.)	0.18	0.24	**0.17**	**0.09**
Female	0.50	0.50	0.50	0.50
Birth year	1996	1987	**2000**	**2009**
Birth order	4.46	4.39	**4.53**	4.45
Rural	0.72	0.74	0.73	**0.67**
Mother no school	0.47	0.51	**0.46**	**0.42**
Mother primary school	0.36	0.37	0.36	**0.33**
Mother secondary school	0.15	0.11	**0.16**	**0.21**
Mother tertiary	0.02	0.01	**0.02**	**0.04**
Mother age at birth	27.66	26.36	**28.17**	**29.23**
Total children born to mother	6.52	7.39	**6.26**	**5.27**
Mother's ideal number of boys	2.49	2.59	**2.50**	**2.32**
Mother's ideal number of girls	2.34	2.47	**2.35**	**2.12**
Mother's ideal sex ratio	0.51	0.51	**0.52**	**0.52**
N	17,963	7,124	7,533	3,306

Notes: All measures are dichotomous except birth year (ranges from 1960 to 2014), birth order (ranges from 1 to 14), mother age at birth (ranges from 10 to 47), total children born (ranges from 1 to 16), mother's ideal boys (ranges from 0 to 20), mother's ideal girls (ranges from 0 to 15), and ideal sex ratio (ranges from 0 to 1 and excludes women who desire zero children). Bold numbers indicate statistically significant (p<0.05) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Two-sample t tests for all continuous outcomes, and chi-square tests for all dichotomous outcomes. Singleton sample all live in households with no other twins.

Panel B.	(1)
	Mortality 1-59 mos.
Female	-0.016**
	(0.005)
First born twin	-0.040***
	(0.005)
India	-0.094***
	(0.008)
Female*India	0.035***
	(0.011)
Born 1995-2005 (ref=before 1995)	-0.071***
	(0.005)
Born after 2005 (ref=before 1995)	-0.130***
	(0.006)
Constant	0.264***
	(0.005)

Observations	24,163
R-squared	0.033
Baseline level of male mortality	0.157

*** p<0.001, ** p<0.01, * p<0.05

Appendix Table 4. Descriptive summary of background characteristics of singletons (Panel A); female same-sex twins (Panel B); and male same-sex twins (Panel C) including tests for significant difference between the first birth cohort and the two subsequent cohorts. All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016) and use sampling weights provided by the NFHS.

Panel A.	Singleton sample:			
	(1)	(2)	(3)	(4)
	Pooled	*Born before 1995*	*Born 1995-2005*	*Born after 2005*
Mortality (1-59 mos.)	0.04	0.07	**0.03**	**0.02**
Female	0.48	0.48	0.48	0.48
Birth year	1996	1985	**2000**	**2010**
Birth order	2.43	2.48	2.48	**2.29**
Rural	0.72	0.72	**0.71**	**0.71**
Northern region	0.60	0.59	**0.60**	**0.60**
Hindu	0.80	0.81	**0.79**	**0.78**
Poorest 40%	0.44	0.40	**0.46**	**0.48**
Mother no school	0.52	0.63	**0.51**	**0.35**
Mother primary school	0.16	0.16	**0.15**	**0.15**
Mother secondary school	0.27	0.18	**0.29**	**0.42**
Mother tertiary	0.05	0.02	**0.05**	**0.09**
Mother age at birth	23.29	22.14	**23.75**	**24.61**
Total children born to mother	3.83	4.49	**3.70**	**2.86**
Mother's ideal number of boys	1.40	1.54	**1.35**	**1.24**
Mother's ideal number of girls	1.06	1.10	**1.04**	**1.00**
Mother's ideal sex ratio	0.57	0.58	**0.56**	**0.55**
N	2004684	803490	692572	508622

Notes: All measures are dichotomous except birth year (ranges from 1954 to 2016), birth order (ranges from 1 to 18), mother age at birth (ranges from 10 to 50), total children born (ranges from 1 to 18), mother's ideal boys (ranges from 0 to 20), mother's ideal girls (ranges from 0 to 11), and ideal sex ratio (ranges from 0 to 1 and excludes women who desire zero children). Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Two-sample t tests for all continuous outcomes, and chi-square tests for all dichotomous outcomes.

Panel B.	Female same sex twin sample:			
	(1)	(2)	(3)	(4)
	Pooled	*Born before 1995*	*Born 1995-2005*	*Born after 2005*
Mortality (1-59 mos.)	0.12	0.20	**0.09**	**0.06**
Female	1.00	1.00	1.00	1.00
Birth year	1999	1986	**2000**	**2011**
Birth order	3.04	3.27	**3.03**	**2.82**
Rural	0.66	0.68	0.64	0.67
Northern region	0.55	0.51	**0.55**	**0.60**

Hindu	0.76	0.77	0.77	0.74
Poorest 40%	0.40	0.33	**0.38**	**0.49**
Mother no school	0.45	0.57	**0.41**	**0.36**
Mother primary school	0.15	0.16	0.18	**0.12**
Mother secondary school	0.33	0.23	**0.35**	**0.40**
Mother tertiary	0.07	0.04	0.06	**0.12**
Mother age at birth	24.43	23.26	**24.65**	**25.36**
Total children born to mother	4.48	5.15	**4.53**	**3.74**
Mother's ideal number of boys	1.34	1.46	**1.31**	**1.27**
Mother's ideal number of girls	1.14	1.18	1.18	**1.06**
Mother's ideal sex ratio	0.54	0.55	**0.53**	0.54
N	6149	1991	2134	2024

Notes: All measures are dichotomous except birth year (ranges from 1962 to 2016), birth order (ranges from 1 to 13), mother age at birth (ranges from 12 to 48), total children born (ranges from 2 to 14), mother's ideal boys (ranges from 0 to 8), mother's ideal girls (ranges from 0 to 6), and ideal sex ratio (ranges from 0 to 1 and excludes women who desire zero children). Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Two-sample t tests for all continuous outcomes, and chi-square tests for all dichotomous outcomes.

Panel C.	Male same sex twin sample:			
	(1)	(2)	(3)	(4)
	Pooled	*Born before 1995*	*Born 1995-2005*	*Born after 2005*
Mortality (1-59 mos.)	0.09	0.17	**0.07**	**0.05**
Female	0.00	0.00	0.00	0.00
Birth year	1999	1986	**2000**	**2011**
Birth order	3.19	3.39	3.29	**2.86**
Rural	0.67	0.66	0.67	0.69
Northern region	0.54	0.53	0.55	0.53
Hindu	0.77	0.81	0.78	**0.73**
Poorest 40%	0.36	0.33	**0.39**	0.37
Mother no school	0.45	0.58	**0.48**	**0.27**
Mother primary school	0.14	0.16	0.13	**0.13**
Mother secondary school	0.33	0.19	**0.31**	**0.49**
Mother tertiary	0.08	0.06	0.07	**0.11**
Mother age at birth	24.95	23.69	**25.34**	**25.79**
Total children born to mother	4.33	5.09	**4.31**	**3.56**
Mother's ideal number of boys	1.42	1.54	**1.44**	**1.29**
Mother's ideal number of girls	1.02	1.04	1.03	**0.97**
Mother's ideal sex ratio	0.58	0.59	0.58	**0.57**
N	6469	2074	2310	2085

Notes: All measures are dichotomous except birth year (ranges from 1961 to 2016), birth order (ranges from 1 to 13), mother age at birth (ranges from 10 to 44), total children born (ranges from 2 to 15), mother's ideal boys (ranges from 0 to 9), mother's ideal girls (ranges from 0 to 8), and ideal sex ratio (ranges from 0 to 1 and excludes women who desire zero children). Bold numbers indicate statistically significant ($p<0.05$) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Two-sample t tests for all continuous outcomes, and chi-square tests for all dichotomous outcomes.

Appendix Table 5. Analysis of the factors predicting having a mixed sex twin birth (compared to singleton birth) using linear probability models and including interactions between family characteristics and birth period. All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016) and use sampling weights provided by the NFHS.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mixed-sex twin birth	Mixed-sex twin birth	Mixed-sex twin birth	Mixed-sex twin birth	Mixed-sex twin birth	Mixed-sex twin birth	Mixed-sex twin birth
Born 1995-2005 (ref=before 1995)	0.001***	0.001**	0.001*	0.000	0.001*	0.001**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Born after 2005 (ref=before 1995)	0.002***	0.001	0.001	0.000	0.001	0.000	0.001
	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Rural		0.000	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Hindu		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Northern India		0.000	0.000	0.000	0.000	0.000	0.001
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Poorest 40%		0.000	0.000	0.000	0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother primary (ref=no school)		0.002***	0.002***	0.002***	0.002***	0.002***	0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mother secondary (ref=no school)		-0.000*	-0.000*	-0.000*	-0.000*	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Mother tertiary (ref=no school)		-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Born 1995-2005*Rural			-0.000				
			(0.000)				
Born after 2005*Rural			-0.000				
			(0.001)				
Born 1995-2005*Hindu				0.000			
				(0.000)			
Born after 2005*Hindu				0.001			
				(0.001)			
Born 1995-2005*North					-0.000		
					(0.000)		
Born after 2005*North					0.000		
					(0.000)		
Born 1995-2005*Poorest						-0.001	
						(0.000)	
Born after 2005*Poorest						0.000	
						(0.000)	
Born 1995-2005*Mother primary							0.000
							(0.001)
Born after 2005*Mother primary							0.000
							(0.000)
Born 1995-2005*Mother secondary							0.001

							(0.001)
Born after 2005*Mother secondary							-0.001
							(0.001)
Born 1995-2005*Mother tertiary							-0.000
							(0.000)
Born after 2005*Mother tertiary							0.003*
							(0.001)
Observations	2,010,880	2,010,880	2,010,880	2,010,880	2,010,880	2,010,880	2,010,880
R-squared	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Robust standard errors in parentheses clustered at the family level

*** p<0.001, ** p<0.01, * p<0.05

Notes: Models 2-7 include controls for birth year, female, and survey round (not shown). 4 observations from the twin sample are dropped due to missing information on background covariates.

Appendix Table 6. OLS regression models of the association between child sex and infant and child mortality (1-59 months) using a sample of singleton children (Panel A); within family fixed effects estimates of the association between child sex and infant and child mortality (1-59 months) (Panel B). All estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016). Analysis conducted in STATA 15.

Panel A. OLS	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.007***	0.011***	**0.005*****	**0.002*****
	(0.000)	(0.001)	**(0.000)**	**(0.000)**
Background controls	YES	YES	YES	YES
Observations	2,004,684	803,490	692,572	508,622
R-squared	0.021	0.023	0.008	0.005
Baseline male mortality	0.036	0.056	0.027	0.019

*** p<0.001, ** p<0.01, * p<0.05
Notes: Background controls include rural, northern region, mother's school, poorest 40%, birth year, birth order, and survey round. Bold numbers indicate statistically significant (p<0.05) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Robust standard errors clustered at the family level. Neonatal mortality excluded to correspond with within twin FE estimates.

Panel B. Within family FE	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.006***	0.012***	**0.005*****	**0.001**
	(0.000)	(0.001)	**(0.001)**	**(0.001)**
Background controls	YES	YES	YES	YES
Observations	1,971,714	787,717	681,111	502,886
R-squared	0.007	0.008	0.002	0.004
Number of families	704,348	293,432	343,664	289,148
Baseline male mortality	0.036	0.065	0.034	0.021

*** p<0.001, ** p<0.01, * p<0.05
Notes: Background controls include birth order and year of birth. Bold numbers indicate statistically significant (p<0.05) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995). Excludes families with twins. Neonatal mortality excluded to correspond with within twin FE estimates.

Appendix Table 7. Within mixed sex twin fixed effects models of the association between child sex and infant and child mortality (1-59 months) in India, excluding twins born in 2014-2016. Estimates use pooled data from the Indian National Family Health Survey (1992-1993, 1998-1998, 2005-2006, 2015-2016). Analysis conducted in STATA 15.

	(1)	(2)	(3)	(4)
	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.	Mortality 1-59 mos.
	Pooled	Born before 1995	Born 1995-2005	Born after 2005
Female	0.022***	0.053***	0.006	0.011
	(0.007)	(0.015)	(0.010)	(0.010)
First born twin	-0.039***	-0.023	-0.047***	-0.046***
	(0.007)	(0.015)	(0.010)	(0.010)
Constant	0.100***	0.131***	0.093***	0.071***
	(0.006)	(0.013)	(0.009)	(0.009)
Observations	5,800	1,868	2,278	1,654
R-squared	0.017	0.019	0.021	0.027
Number of families	2,900	934	1,139	827
Baseline male mortality	0.078	0.118	0.068	0.046

*** p<0.001, ** p<0.01, * p<0.05

Notes: Bold numbers indicate statistically significant (p<0.05) difference between the birth cohort in question and the first birth cohort (e.g. born before 1995).