Son Preference, Fertility, and Family Structure.
Evidence from Reproductive Behavior among Nigerian Women *

Annamaria Milazzo †

June 7, 2012
Preliminary draft. Comments are welcome.

Abstract

This paper explores whether and how son preference is reflected in women's fertility behavior and family structure. Using DHS data for Nigeria I find that, compared to women with first-born sons, women with first-born daughters have (and desire) significantly more children; are less likely to use contraceptives; and space births significantly less as the number of daughters after the first increases. The preference for *own* biological sons is also supported by fostering patterns where daughters are substitutes for non-biological girls, while sons are not. Moreover, women with first-born daughters are significantly more likely to end up in a polygynous union, to be divorced, and to be the head of the household. While existing studies for the developing world have mainly documented strong boy-bias in many Asian countries, this paper considers a country in Sub-Saharan Africa and finds that parental gender preferences *do* affect fertility behavior and play a role in shaping complex family structure by interacting with traditional institutions. The results for fertility and birth spacing can partly explain the missing women phenomenon in Sub-Saharan Africa uncovered by Anderson and Ray (2010).

Keywords: gender bias, fertility, Sub-Saharan Africa, child fostering, polygyny
JEL classification: D63, J13, J16, I10

---

*I am grateful to my advisor Eliana La Ferrara for continued guidance and support, and to Maristella Botticini, Éric Brian, Denis Cogneau, Selim Gulesci, Pamela Jakiela, Sylvie Lambert, Andreas Madestam, and Massimo Marinacci for very helpful comments. I thank seminar participants at Bocconi University and the Paris School of Economics for comments and suggestions. All errors are my own.

†Bocconi University. Email: amilazzo14@gmail.com

1 Introduction

There is an extensive literature on parental preferences over a child's gender.[1] For developing countries, son preference has been widely documented in East, South Asia and North Africa, where gender bias is severe especially at young ages. Sen's 1990 article documented high sex ratios (the ratio of males to females) at birth in many Asian countries: he argued that women might be missing because of sex selective abortion, infanticide and neglect of female children. Other authors have studied fertility behavior (son-preferring fertility stopping rules), and shown that subsequent fertility is higher for women who had girls among earlier-born children (Chowdhury and Bairagi, 1990; Clark , 2000; Dreze and Murthi, 2001). Again, most of the evidence comes from Asian countries. Mixed evidence has been found for Western countries (mainly the US) where evidence of increasing son preference has to be considered together with the well documented preference for sex-balance (Ben-Porath and Welch, 1976; Angrist and Evans, 1998). In particular, recent research has documented unusually high sex ratios especially among Asian mothers in the US; this pattern is found to be related to the increasing availability of technologies to determine the sex of the fetus (Abrevaya, 2009; Dahl and Moretti, 2008).

The economic literature on son preference rarely focused on Sub-Saharan Africa, which is often regarded as a continent with low or absent gender preferences. A notable exception is Anderson and Ray (2010), who provide a decomposition of the excess female deaths by age and causes of death in different parts of the world.[2] Surprisingly, they find that many women are missing in Sub-Saharan Africa, despite the relatively low sex ratio at birth in the continent. The vast majority of women are not missing at birth, but throughout the entire age spectrum.[3] Gender bias is likely not to be found at birth in the African context where high fertility is culturally valued and made less costly for families that still rely on the support from the extended family system. This is in contrast with most Asian countries, where fertility is lower and childrearing costs are primarily borne by parents.

Anthropological and demographic evidence emphasize the dominant role of males in traditional patrilineal societies where descent and inheritance are

[1] For a cross-cultural review of literature on parental preferences, see among others Williamson (1976).

[2] Klasen (1996) analyzes anthropometric and mortality data for Sub-Saharan Africa and finds a small and increasing anti-female bias.

[3] Anderson and Ray (2010) find that most women of adult ages die because of HIV infection and maternal mortality.

transmitted through the male line. Furthermore, male children strengthen the relationship between the wife and her husband's kin (by guaranteeing the continuation of his lineage) and secure (the mother's) access to residence and inheritance upon the husband's death.

This paper examines whether the resulting pressure for bearing sons is associated with specific patterns in fertility behavior and family structure. It makes at least two contributions to the existing literature. First, while Anderson and Ray (2010) use aggregate mortality data and find that, contrary to previously-held wisdom, many women are missing in Sub-Saharan Africa, this paper uses individual-level data for Nigeria and explores son preference in fertility decisions in a context where little is known about parental gender preferences. Second, it provides an original explanation for the prevalence of the social institutions of child fostering and polygyny and argues that the preference for biological sons is among the factors motivating these institutions.

I use DHS data for Nigeria and find evidence of son preference in several dimensions. First, compared to women who had a first-born boy, women with a first-born girl exhibit a 2 percent increase in the number of children ever born. While the effect on fertility is stronger when considering the first two or three female births, I focus on the first-born to maintain a causal interpretation of the results. I check for the exogeneity of the sex of the first-born child and discuss about potential sources of bias in the data. The identifying assumption is that the sex of the first-born is uncorrelated with the error term after conditioning for a set of observable characteristics. Second, I find that women's *desired* fertility and the use of contraceptives are affected by the sex of the first born. In particular, women with a first-born girl are 2.4 percentage points more likely to report that they desire to have another child, and 1.1 percentage points less likely to use contraceptives. Thirdly, I investigate whether the pace at which women have children is influenced by the sex of earlier-born children. I exploit the variation in birth spacing for children of the same mother by estimating regressions with mother fixed effects which allow to control for unobserved women's heterogeneity. I find that, among women who had two daughters as first two born children, those who have a third or fourth daughter are 4 percentage points more likely to wait less than 24 months. These findings have implications for maternal and child health that, according to the medical literature, are adversely affected by short spacing (ie, shorter than 24 months).

A second set of findings relates to the institution of child fostering, in which biological children are temporarily sent to live with other families. I focus on the child labor hypothesis as one of the motivations for fostering

according to which children of each gender and age group have specific roles within the household. Therefore, households with an imbalance of biological girls or boys might decide to send or receive a child in order to achieve a balanced gender structure thus maximizing household productivity (Akresh, 2009). If non-biological girls (boys) are substitutes for biological daughters (sons), household should respond symmetrically to the imbalance of girls or sons in fostering decisions. Instead, for the fostering-in decisions, I find that households with an excess of sons are one percentage point more likely to receive a girl, while those with an excess of daughters are no more likely to foster in more boys. This asymmetric response may suggest that, because of the desire to have their own biological son, households lacking sons might want to 'save' and continue bearing children instead of fostering in outside boys. In contrast, households that lack daughters are significantly more likely to receive a girl, as girls are often needed in the house for running domestic chores. This evidence is consistent with the idea that fostered girls are considered as substitutes for own biological daughters in fostering-in decisions, while boys are not substitutes for sons.

Finally, women with first-born daughters are significantly more likely to end up in a polygynous union (ie, that their husband marries another woman), and this effect is specific to first-rank wives. Women with daughters are also more likely to have a nonresident husband, of being divorced, and of being the head of the household.

Therefore, I show that parental gender preferences *do* affect fertility behavior and play a role in shaping complex family structure in Nigeria. Although these findings cannot be generalized to the Sub-Saharan African continent, they might represent a starting point for further research studying the underlying motivations for the existence of specific institutional features and understand their impact on economic development.

This paper is related to the literature on parental gender preferences (and its implications) in different areas of the world, mostly Asian countries and the US (Ben-Porath, and Welch, 1976; Williamson, 1976; Sen, 1990; Dahl and Moretti, 2008; Abrevaya, 2009; Anderson and Ray, 2010; Jayachandran and Kuziemko, 2011). Compared to this literature, this paper investigates if fertility decisions are gender-biased in a context where little is known about parental gender preferences and complements findings of Anderson and Ray (2010).

A second related strand of literature analyzes the motivations and the impact of the institution of child fostering (Isiugo-Abanhie, 1985; Bledsoe, 1990; Ainsworth, 1996; Akresh, 2009; Beck and others, 2011). This paper

links household fostering decisions and fertility and argues that fostering is partly motivated by the desire for sons.

Lastly, this paper relates to the sociological literature on the effects of the gender of child on marital stability, divorce, and parental involvement (Lundberg, 2005 reviews the literature in this field). Among other outcomes, I empirically test qualitative evidence in the anthropological and demographic literature suggesting that infertility and sex composition of earlier-born children are among the factors affecting the husband's decision to marry another woman.

The remainder of the paper is organized as follows: Section 2 provides some insights on the context and cultural background; Section 3 describes the data; Section 4 introduces the empirical methodology used to analyze fertility outcomes and birth spacing; Section 5 presents the results; Section 6 focuses on the interactions between son preference and family structure, by showing how the preference for own biological male child affects child fostering decisions as well as marital outcomes, and Section 7 concludes.

2 Context

As in many parts of Sub-Saharan Africa, the Nigerian society is organized around the extended family which still represents the most basic unit of social organization. Family ties are strong and play an important role in shaping individual behavior, even though there are signs that the extended family system is weakening for some ethnic groups (Wusu and Isiugo-Abanihe, 2006). There is an extraordinary ethnic diversity in Nigeria, but all ethnic groups are predominantly characterized by patrilineality and patrilocality.[4]. The anthropological and demographic literature emphazise the dominant role of males in these traditional patrilineal societies where descent and inheritance are transmitted through the male line. Large progenies are strongly valued because they strengthen a man's family status. The interactions between the principles of social organization, fertility preferences and family structure are documented in several demographic studies, such as Isiugo-Abanihe (1994):

'Childlessness is the most dreaded tragedy for a man or a woman to experience in Nigeria's patrilineal society. (...) The majority of the respondents

[4] The three largest ethnic groups are geographically concentrated in different areas of Nigeria. They are the following: Hausa-Fulani (North) 29%, Yoruba (South-West) 15%, and Igbo (South-East) 14%.

felt that a man without a child, particularly without a son, will not be remembered in the family; his branch of the family will come to an end. For the same reason, a man who has only daughters may acquire a second wife to enhance the chance of having a son. Clearly, in such a patriarchal system, the perpetuation of the family line is a strong motivation for children.' Isiugo-Abanihe (1994, p. 154, from a survey of 3,073 Nigerian couples)

Thus, the desire for many children, especially males, is motivated by descent transmission and the urge to provide continuation to the family lineage. More importantly, these societies are characterized by the fact that only males have the control of family landed property, as documented in Wusu and Isiugo-Abanihe (2006).[5]

'In most cultures only male children are allowed to share in family land holdings within the context of the extended family structure and communal ownership of land. Since farming is central to economic life, the most economically rewarding reproductive goal a couple could pursue is a large family size, ideally with many male children. Against this backdrop couples dread barrenness, and until a 'good' number of male children are born, extended family members exert pressure, which may culminate in the man marrying another wife' Wusu and Isiugo-Abanihe (2006, pp. 141–142)

In this context, sons are highly valued by women as well because they represent the only way through which they can inherit part of their deceased husband's property (see Fapohunda and Todaro, 1988). This consequently generates pressure for bearing male children as a way to protect themselves in the state of widowhood. It may also lead to competition among co-wives as described in Bledsoe et al. (1998).[6]

'In their husbands' compounds, women seek to establish their security and to gain a competitive edge over present and future co-wives and sisters-in-law by bearing a number of children, especially sons, who will retain rights

[5] There are several sources of law in Nigeria: Common law (the *default* Law, applied only when other sources of law are not available), Customary Law, and Sharia (mainly applied among the Muslims). Even though there is some variation in the inheritance rules across states, in most cases only male children are allowed to inherit the father's property. In states where the Sharia is the law, female children are allowed to inherit half of the share of their male siblings, even though there is evidence of noncompliance with the law.

[6] There are contrasting views in the literature on whether there might be competition or cooperation among co-wives in polygynous marriages.

of residence and inheritance in the compound and will eventually take over its leadership roles.' Bledsoe et al. (1998, p. 23)

One of the peculiarity of the West-African context is widespread fostering of children, in which children are temporarily sent to be raised in another family, away from their biological parents. As argued by some demographers, high fertility is also made possible by the social institution of fostering, through which the costs of childrearing are partly shared with the extended family (Bledsoe, 1990; Isiugo-Abanhie, 1985). Fapohunda and Todaro (1988) argue that 'the presence of parental surrogates in the extended family alleviates problems of incompatibility between child care and work and, thereby, lowers the opportunity cost of children.' (p. 572). Smith (2004) argues that the traditional high value of having many children (that allow larger kinship networks) might partly motivate the slow transition to lower fertility in Nigeria, even though families start lamenting the increasing costs of raising children due to modernization. In this paper I study son preference by considering the interactions between fertility behavior and complex family structure related to traditional institutions.

3 Data and descriptive evidence

I use the 2008 Nigeria Demographic and Health Survey (NDHS) which contains individual-level information on birth histories, household composition, biological children living in the household or elsewhere, birth intervals, and other variables. For the analysis of fertility decisions and birth spacing I use the sample of married women age 15–49 in first union (so that the fertility history at time of the survey is relevant for current union) with at least one child ever born.[7] Table A.1 reports summary statistics for this sample.[8]

The 2008 NDHS reports a total fertility rate (TFR) of 5.7 births per woman. This means that, on average, a Nigerian woman will give birth to 5.7 children by the end of her childbearing years. The 1991 Census reports

[7] One of the findings presented in Section 6 is that a first-born girl increases the likelihood of being in a polygynous relationship, of having a nonresident husband, of being divorced and being the (female) head of the household. These effects would lead to underestimation of the effect of a first-born girl on fertility because women with daughters have fewer children than they would have had if their marital status had not changed. Therefore, given that fertility and decisions related to marital status are likely to be simultaneously determined, I consider *currently* married women in first unions to limit the extent of this bias.

[8] Summary statistics for the full sample (including unmarried women, and those not in first union) are reported in Table A.2.

a TFR of 5.9.[9] Knowledge of any contraceptive method is widespread in Nigeria, with 72% of all women and 90% of all men knowing at least one method of contraception, and 29% (41%) of all women (men) who reported ever using a method of contraception at some time[10]. Current use of any contraceptive methods among currently married women is much lower but increasing over time: from 6% in 1990, 13% in 2003, to 15% in 2008 (levels and trend are similar to those for West Africa, WHO 2011).[11,12]

In 2009, infant mortality rates (the probability of dying before reaching age 1) were 92 and 80 deaths per 1,000 live births for male and female children respectively (WHO, 2011).[13] Higher infant mortality rates for male children are consistent with the medical literature documenting that males are biologically weaker than females at birth (Waldron, 1983). Moreover, using the 2008 NDHS, mortality seems to have a U-shaped relationship with birth order, with first and later-born children more likely to die (especially males) than second or third-order children as shown in Appendix Figure A.1. The same pattern is found when using 2003 NDHS data (bottom figure).

Maternal mortality is very high in Nigeria; the WHO estimates that 840 births out of 100,000 live births resulted in maternal death in 2008.[14] Improving maternal health is an enormous challenge in Nigeria, which alone accounts for 14 percent of all maternal deaths worldwide (WHO, UNICEF, UNFPA, and World Bank, 2010). Still, Nigeria is making insufficient progress in the achievement of the fifth Millennium Development Goals (MDGs) by 2015.[15]

[9] The WHO (2011) reports that in 2009, TFR is 5.2 in Nigeria. For comparison, TFR in the same year is 4.9 in the African Region; 2.5 in South-East Asia; 3.4 in Eastern Mediterranean Region; 1.6 in European Region; 2.1 in Region of the Americas; and 1.8 in the Western Pacific Region (WHO region definitions).

[10] Ever use peaks at 36% (52%) among women (men) age 30-39 (age 30-34)

[11] Among the women currently using contraception, about two thirds are using a modern method (mostly male condom, injectables and pills), while one third use traditional methods (rhythm method and withdrawal are the most common).

[12] 20% of currently married women have an unmet need for family planning; 15% for spacing and 5% for limiting; 84% of currently married women age 15-49 who are using a method reported that their husband or partner knows about their use of contraception.

[13] For an international comparison, the male and female mortality rates in 2009 were: 14, 20 in China; 50, 51 in India; 50, 43 in Ghana; 102, 87 in Cameroon; 144, 123 in Afghanistan; 7, 6 in the U.S. For the WHO regions: 85, 74 African Region; 46, 45 South-East Asia (incl. India); 57, 51 Eastern Mediterranean Region; 13, 10 European region; 16, 14 Region of the Americas; and 16, 19 in Western Pacific Region (inlc. China).

[14] For comparison, the WHO estimates for broad world regions are: 620 in the African Region; 240 South-East Asia (incl. India); 320 Eastern Mediterranean Region; 21 European region; 66 Region of the Americas; and 51 in Western Pacific Region (incl. China)

[15] The target of the fifth MDG consists in reducing the maternal mortality ratio (MMR)

Using the 2008 NDHS birth history data, the sex ratio at birth (here defined as the fraction of male births) is 0.512 for all births.[16,17] Figure A.2 shows that there is a clear negative relationship between the sex ratio at birth and the birth year of the child for births occurred between before 1995, while a rather stable (and biologically 'normal', around 0.51-0.515) sex ratio for more recent births. There are several potential explanations for this pattern, including biological conditions, factors related to maternal survival, or misreporting.

As for the biological explanation, Andersson and Bergstrom (1998) and Almond and Mazumder (2011) find evidence that mortality of male fetuses is positively associated with maternal malnutrition. According to this literature, pregnant women exposed to adverse nutritional conditions are more likely to give birth to females.[18] To check for the biological explanation, I first check whether sex ratios are consistently high for births occurred during the period 1975–1994. In absence of census data, I use all the survey rounds of the NDHS (1990, 1999, 2003, and 2008). Figure A.3 shows the sex ratios by the birth year of the child and by survey year. This allows to distinguish between age and cohort-specific effects. The figure shows that the decreasing trend in the sex ratio is common to all the survey rounds, independently of the cohort of birth of the child, and not specific to the period 1975–1994. This finding suggests that there might be other explanations for the systematically higher sex ratios for births that occurred back in time from the time of the survey.

In a recent paper (Milazzo, 2012), I find a similar decreasing pattern

by three quarters between 1990 and 2015, and achieving universal access to reproductive health by 2015.

[16]Based on several data sources, Anderson and Ray (2010) report that sex ratios at birth are about 0.514 in developed countries (including the 'Established Market Economies' as defined by the World Bank: Western Europe, Canada, United States, Australia, New Zealand, and Japan), 0.518 in India, 0.539 in China, and 0.508 in Sub-Saharan Africa. *Overall* sex ratios in India and China are similar, both around 0.514. Several biological, environmental, and genetic factors can partly explain the variation across different areas.

[17]Garenne (2002) analyzed sex ratios at birth in African countries using the DHS surveys. He finds that there is substantial variation within the African region (with Nigeria and Ethiopia being the countries with the highest sex ratio, equal to 0.517) and that the average sex ratio is 0.508. As for Nigeria, Garenne's study does not include the most recent DHS surveys, where the sex ratio is 0.512 (2008) and 0.515 (2003).

[18]Nigeria experienced a severe civil war at the end of the 1960s (that mainly involved the South-East region populated by the Igbos) and several national-level oil shocks during the 1970s and the 1980s. In the empirical strategy, I control for cohort of birth (of the mother) fixed effects and ethnic-specific time trends to control for time-specific events and ethnic-specific trends that might be correlated with fertility and the probability of a female birth.

for the sex ratio using DHS data for India. By comparing the age structure and health indicators of women with first-born daughter and first-born son, I show that son preference is associated with selective maternal mortality through fertility risky behavior.[19] Specifically, women with first-born daughter have higher fertility, tend to space births less and thus are at higher risk of dying because of maternal depletion. A reversal in health indicators (mainly anemia, which responds quickly to reproductive behavior), in which women with first-born daughter are worse-off when young but better-off when older (over age 30) suggests that healthier women from a higher socio-economic background are more likely to survive the birth of daughters. This result is obtained also when controlling for the number of children ever born, suggesting that the effect is mainly due to short spacing. Selective maternal mortality might be happening in Nigeria as well: Figure A.4 uses pooled NDHS data and shows that there are fewer women with first-born girl above the age of 25-30. Figure A.5 shows that only women with a girl as first-born are 'missing', while the sex ratio for higher birth order births is constant for women of all ages. This is because of the association between first-born girl, realized fertility, and spacing. Figure A.6 plots the share of women with first-born daughter for women with no education, and those with at least one year of education. It's only among uneducated women that women with first-born girl above 30 are missing. This is consistent with the idea that women from low socio-economic status (here proxied by the level of education) are more likely to die due to son-preferring fertility behavior. In absence of the anemia variable in NDHS 2008, I compare other health indicators (namely, weight-for-height and body mass index) and find similar results to those found for India: specifically, women above the age of 25-30 with first-born girl exhibit better nutritional conditions.[20] These results suggest that there might be selective maternal mortality in Nigeria. This type of bias would lead to an underestimation of the effect of daughters on fertility because (surviving) women reporting a first-born girl would have fewer children on average (being them better-off).

Finally, being based on self-reported retrospective information, birth histories contained in the NDHS might suffer from misreporting and recall bias (Smith, 1994; Byass et al., 2007). Recall bias requires further investigation only if it is sex-selective (eg, if mothers are more likely to underreport fe-

[19] Using a different dataset for India (with retrospective fertility histories as in the DHS), Rosenblum (2012) similarly finds higher sex ratios for births occurred in the past, but argues that the extent of the bias in the data is small. He also discusses about potential recall and survival bias in the data.

[20] Results available upon request.

Table 1: Summary stats, differences

	firstG	firstB	diff	se diff	n
eduyrs	5.177	4.913	0.264***	(0.088)	18414
partner eduyrs	6.595	6.227	0.368***	(0.093)	18055
husband living in	0.893	0.906	-0.013***	(0.005)	18307
age first birth	19.583	19.430	0.153**	(0.074)	18426
age first marriage	17.992	17.820	0.172**	(0.077)	18426
age	31.587	31.648	-0.061	(0.133)	18426
partner age	41.596	41.724	-0.128	(0.172)	18123
# children ever born	4.286	4.330	-0.044	(0.043)	18426
# children alive	3.546	3.553	-0.007	(0.034)	18426
urban	0.333	0.317	0.016**	(0.008)	18426

Robust standard errors adjusted for clustering at the household level in parentheses. All married women aged 15–49 in first union with at least one child. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels.

male births and deaths than males), which is plausible in countries with son preference. If present, selective recall bias would lead to an underestimation of the effect of first-born girl on fertility, because women from lower socio-economic status (thus, with a higher number of children on average) would be more likely to under report female births. In Milazzo (2012), I show that selective maternal mortality is decreasing over time in India. This is consistent with improved maternal health conditions, but in contrast with selective recall bias because son preference did not show any declining trends in the Indian context (if anything, evidence suggests it is worsening over time).

Table 1 compares some observable characteristics between the group of women with first-born daughter and first-born boy. On average, women with first-born girl are more educated, have a better educated partner, and are more present in urban areas. They also marry and have the first child later than women with first-born boy. These differences are statistically significant. The number of children ever born does not differ between the two groups; this can be explained by differential infant mortality (male children are more likely to die in infancy, thus mothers try to replace the dead child by having more children), and by the fact that mothers with daughters are on average more educated (and have fewer children). Next, I split the sample on the subsamples of women age 15–25, and 26–49. I expect to find smaller differences between women with first-born girl and first-born boy when focusing on the subsample of younger women. This is because women with first-born girl have more children (and tend to space them less) and may have developed differences over time, due to specific reproductive behavior. If son preferring fertility behavior leads to selective

Table 2: Summary stats, differences, subsamples

	Women age 15-25					Women age 26-49				
	firstG	firstB	diff	se diff	n	firstG	firstB	diff	se diff	n
eduyrs	3.897	3.933	-0.036	(0.144)	5418	5.690	5.300	0.390***	(0.107)	12996
partner eduyrs	5.705	5.504	0.201	(0.162)	5289	6.950	6.511	0.439***	(0.113)	12766
husband living in	0.897	0.900	-0.003	(0.009)	5385	0.891	0.908	-0.017***	(0.006)	12922
age first birth	17.608	17.476	0.132*	(0.080)	5423	20.374	20.202	0.172*	(0.093)	13003
age first marriage	16.122	15.976	0.146*	(0.084)	5423	18.743	18.549	0.194**	(0.098)	13003
age	22.009	21.892	0.117	(0.076)	5423	35.427	35.504	-0.077	(0.125)	13003
partnerage	32.544	32.408	0.136	(0.219)	5326	45.210	45.405	-0.195	(0.187)	12797
# children ever born	2.216	2.216	-0.000	(0.035)	5423	5.116	5.165	-0.049	(0.051)	13003
# children alive	1.897	1.888	0.009	(0.031)	5423	4.207	4.211	-0.004	(0.040)	13003
urban	0.248	0.239	0.009	(0.013)	5423	0.366	0.347	0.019**	(0.009)	13003

Robust standard errors adjusted for clustering at the household level in parentheses. All married women in first union with at least one child. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels.

maternal mortality, better off and healthier women would be more likely to survive the birth of girls. As expected, Table 2 shows that the differences are much smaller for younger women, suggesting that selective maternal mortality is a plausible explanation, like in the context of India. Therefore, it is important to control for these observable characteristics in the empirical analysis and examine if the inclusion of these variables affects the coefficients of interest in the expected direction.

Birth spacing Short intervals between births have been found to be related to poor child health and maternal outcomes.[21] Medical research has shown that short birth intervals (<24 months) are associated with poor child and maternal health outcomes (Setty-Venugopal and Upadhyay, 2002). For example, compared with children born less than 2 years after a previous birth, children born 3 to 4 years after a previous birth are 1.5 times more likely to survive the first week of life, and 2.4 times more likely to survive to age five (DHS 2002). With reference to maternal health, compared with mothers who give birth at 9- to 14-month intervals, women who have their babies at 27- to 32-month birth intervals are: 1.3 times more likely to avoid anemia; 1.7 times more likely to avoid third-trimester bleeding; and 2.5 times more likely to survive childbirth (Conde-Agudelo, and Belizan, 2000).[22] Also, there are effects on health of older children (e.g., breastfeeding duration, see Jayachandran, and Kuziemko, 2011). Palloni and Millman

[21] Birth spacing refers to the time interval (# months) between births.

[22] see Conde-Agudelo, Rosas-Bermdez, and Kafury-Goeta (2007) for a review of medical studies on the relationship between birth spacing and maternal health.

Table 3: Child mortality rates by length of previous birth interval (NDHS 2008)

	Neonatal mortality	Infant mortality	Under-five mortality
	< 1 month	< 1 year	< 5 years
<2 years	70	135	252
2 years	37	76	168
3 years	31	59	123
4+ years	23	44	92

Estimates are for deaths per 1,000 live births. Source: NDHS, 2008.

(1986) and Palloni and others (1994) discuss the channels through which short birth spacing and breastfeeding affect early childhood mortality. First of all, breast milk contains the nutritional requirements for infant growth and protects against the formation of bacteria and of malnutrition. Short preceding birth interval does not allow the mother the time to regain the strength between pregnancies, leading to the inability to adequately breastfeed, while short succeeding intervals reduces maternal care for older children.

Table 3 shows some statistics from the 2008 NDHS report. There is a striking negative relationship between the length of the birth interval and mortality. For example, 252 children (out of 1000 live births) die under the age of 5 if the preceding birth interval is shorter than 24 months, while the number of children who die if the interval is 4 years or longer is reduced to 92. In the NDHS dataset, of all births reported by each woman, 7 percent are less than 15 months apart, 12 percent less than 18 months, 32 percent less than two years. The mean birth interval is 32.5 months, and the median is 28. The survival status of the child who opens the birth interval matters: the median interval is 29 months if previous child is alive, 23 months if dead. Moreover, the birth interval increases with birth order (for given # of children).

Given the association between birth spacing and child mortality (which will be explored in detail in section 5.2.1), it is important to understand if the desire for having sons affects the pace at which mothers try to conceive and give birth.

4 Empirical strategy

4.1 Fertility regressions

To investigate the effect of son preference on fertility, I analyze whether the sex of the first-born affects realized fertility, by considering *all children ever*

born to a woman[23]. Given that the sex of the first born can reasonably be considered as random, the empirical strategy should be straightforward: unbiased estimates should be obtained by simply regressing the number of children ever born on the sex of the first-born. However, as discussed in the previous section, there are differences in observables between women with a first-born boy or girl. These differences may have developed over time as a consequence of son-preferring fertility stopping rules (ie, women with a first-born daughter tend to have more children, and higher fertility may influence their health status and ultimately their survival), or may be due to selective recall bias. Even though these types of bias would lead to an underestimation of the effect of first-born girls on fertility, in all regressions I include a set of observable covariates in order to reduce this bias. If couples desire to have male children, women with first-born daughter should exhibit higher fertility than women with first-born son, after controlling for observable characteristics.

In order to understand if the sex of the first-born also affects the *desire* to have more children (which may differ from *realized* fertility), I construct a dummy variable based on the NDHS survey question: *'Would you like to have another child, or would you prefer not to have anymore children?'*. I also examine if the current use of contraceptives differs for women with a first-born daughter and first-born son.

I estimate the following regression (for women with one child or more), which is similar to the empirical strategy in Dahl and Moretti (2004):

$$y_{i,t,r} = \beta_1(firstborngirl)_i + \gamma X_{i,t,r} + \gamma_t + \alpha_r + \epsilon_{i,t,r} \tag{1}$$

with mother i, born in year t, and resident in region r. $y_{i,t,r}$ is the dependent variable, which can alternatively be the # of children ever born, a dummy equal to one if the woman reports the desire to have more children, and a dummy for the current use of any contraceptive methods.

[23] The number of children ever born includes all births to each woman, including dead children and children living elsewhere. Compared to the number of surviving children, this is a more pertinent measure of fertility as it is not affected by differential mortality potentially *caused* by son preference. Clark (2000) uses the number of children ever born in studying fertility stopping rules in India; in that context, poorer treatment of daughters compared to sons has been found to be associated with higher mortality rate for girls. To my knowledge, there is no strong evidence of worse health treatment of daughters in Sub-Saharan Africa; in the NDHS (2008) the infant mortality rates are higher for boys, which is also consistent with the medical evidence that boys are weaker than girls . As boys are more likely to die than girls, the estimated effect of daughters on fertility would be underestimated since mortality is positively associated with fertility due to replacement.

$firstborngirl$ indicates whether the first child ever born is a girl (as opposed to a boy); $X_{i,t,r}$ is the set of covariates including: age (of the mother and her partner) and age squared, seven 5-year age groups, age at first marriage, age at first birth, number of years of education (of the mother and her partner), a wealth index, urban, ethnicity, and religion.[24,25] α_r, γ_t are region, age-group, and cohort of birth fixed effects, respectively. Given that there is differential mortality across genders, in all regressions I control for the survival status of the child (a dummy for whether the first-born child is dead and its interaction with $firstborngirl$). I also control for region and ethnic-specific time-trends, $\delta_r t$ and $\chi_e t$ respectively, to capture to capture region (ethnicity) and cohort-specific effects that may be correlated with the error term (e.g., the evolution of region or ethnic-specific factors that might have affected the sex of first birth and fertility). Identification relies on the assumption that the sex of the first-born is exogenous (uncorrelated with the error term) after conditioning for the observable characteristics. Regression 1 is estimated using OLS when the dependent variable is the # of children ever born, and using a probit model for the probability of desiring more children or of using a contraceptive method.

If women practice son-preferring fertility stopping behavior, I expect to find $\beta_1 > 0$ for the regressions for realized (and desired) fertility, and $\beta_1 < 0$ for the use of contraceptives. The analysis mainly focuses on the sex of the first-born to maintain a causal interpretation (Dahl and Moretti, 2008). Since fertility decisions and decisions related to marital status and living arrangements (as shown in section 6.4) are probably taken simultaneously, the sex of the first two, three, etc, children born cannot be considered as random anymore. However, the effects should be stronger for women who had more children and more daughters among earlier-born children. There is also a potential issue of selection bias when considering the subsamples of women with two, three or more children and comparing realized fertility depending on all children already born. For example, consider evaluating the effect of the first two girls on fertility. In this case, the sample of women with two or more children would be used. While the sex of the first-born

[24] The ethnicity and religion variables are from the woman dataset. Ethnicity include Hausa-Fulani, Igbo, Yoruba, and other minor ethnicities as the omitted category. Religion include Catholic, other Christians, Muslims, and traditionalists (omitted).

[25] The DHS wealth index is a measure of a household's cumulative living standard available in the DHS survey. It's generated using principal component analysis based on data from the households ownership of consumer goods; dwelling characteristics; type of drinking water source; toilet facilities; and other characteristics that are related to a households socio-economic status.

can be considered as random, following children represent a choice that may correlate with other household decisions. Moreover, if the choice to continue having children after a first-born girl or a first-born boy is correlated with unobserved characteristics, the estimated coefficients would be biased. Since it is possible to observe women that progress to next parity only if they have reached the previous one, those who have had children after daughters might be different than women who progressed after sons.[26]

4.2 Birth spacing regressions

If there is pressure on women (or couples) to have male children, mothers who had daughters first might try to conceive another child sooner after the birth of a daughter (compared to those who had sons). To study the association between the length of birth intervals and the sex composition of earlier-born children, I use the sample of all children ever born to each woman (the unit of analysis is the child, and not the woman as in the previous section).[27] I analyze the average succeeding birth interval (expressed in # of months), which indicates how long the mother waits to have another child *after* the realization of the sex of the previous.[28] Given that the NDHS includes information on the interval between each birth for all children, I can use mother fixed effects and exploit the variation in the length of the interval within the fertility history of each woman. This allows to control for all observables and unobservables characteristics that may correlate with the error term. I estimate the following regression:

$$y_{i,j,t} = \beta_1 \; girl_{i,j,t-1} + \sum_k \gamma_k birthorder_{i,j,t-1,k} + \alpha_j + \epsilon_{i,j,t-1} \tag{2}$$

with i child, j mother, t: current child, t 1: preceding child. $y_{i,j,t}$ is the succeeding birth interval in months (the time between the birth of child t 1 and t); α_j are the mother fixed effects; $birthorder_{i,j,t-1,k}$ is a set of k dummies for birth order. I also construct a dummy variable equal to one if

[26] This is a similar empirical issue to that of *dynamic selection bias*, exposed in Cameron and Heckman (1998), where they study schooling decisions over the life cycle.

[27] I exclude twin births and seven mothers with more than 15 children

[28] Fayehun and others (2011) find weak evidence of shorter spacing after the birth of *each* girl in Nigeria. Compared to them, I estimate a model with mother fixed effects to control for unobserved heterogeneity among mothers. Moreover, I consider the subsamples of women with one or more consecutive daughters to see if the increasing pressure for having sons leads to further shortening of the birth intervals.

the succeeding birth interval is shorter than 24 months (which is the critical interval below which maternal and child health are negatively affected).

Given the high fertility context, I expect to find no significant difference in succeeding interval at *each* birth if the previous child is male or female. Therefore, I estimate (2) on the following subsamples: the subsample of women who had a first-born daughter (or, separately, a first-born boy); the subsample of women who had a first and second-born daughter (and all the other combinations); the subsample of women who had three girls (and, separately, those who had three boys, etc) as earlier-born children. Identification with fixed effects requires to consider mothers with at least three, four, five children ever born. The prediction with son preference is that the succeeding birth interval at each birth should be shorter ($\beta_1 < 0$) if a girl is born rather than a boy; given high fertility context, this effect should be found the more girls are born among earlier-born children, because the pressure for having a son is greater.

5 Main results

5.1 Fertility results

Table 4 shows the OLS estimates of the effect of a first-born girl on the number of children ever born. In column (1), the coefficient of $firstborngirl$ is positive but not significant. In column (2), the controls for education and age at first marriage and birth are added, and the coefficient for first-born girl is now bigger in magnitude and significant at 10% level. Adding the controls for mortality in columns (3) and (4) leads to a further increase in the coefficient of interest. Specifically, women with first-born daughter have 0.07 more children (this effect is significant at 1% level), which correspond to a 1.7% increase of the number of children ever born. The effect is 2.1% percent in the subsample of women age 30–49, who have had more children and are closer to completed fertility.[29]

The rising coefficient of $firstborngirl$ across specifications in Table 4 confirms that it is important to control for all the observables characteristics. In particular, more educated women tend to report more first-born girls than uneducated women (probably because of selective maternal mortality discussed in the descriptive section). Since education is negatively

[29]For comparison, Dahl and Moretti (2008) use Census data for the US and find that women age 18–40 with a first-born girl have 0.3 percent more children than women with a first-born boy. Milazzo (2012) uses data for India and finds that a first-born girl predicts an increase in the number of children by 0.247 (equivalent to a 8 percent increase).

Table 4: The effect of first-born girl on the number of children ever born.

	Age 15-49				30-49
	(1)	(2)	(3)	(4)	(5)
	# children ever born				
first-born girl	0.043	0.050**	0.062***	0.073***	0.118***
	(0.029)	(0.024)	(0.024)	(0.025)	(0.043)
age first marriage		-0.011*	-0.013**	-0.013**	-0.008
		(0.006)	(0.006)	(0.006)	(0.007)
age first birth		-0.253***	-0.246***	-0.246***	-0.233***
		(0.006)	(0.006)	(0.006)	(0.007)
mother eduyrs		-0.022***	-0.022***	-0.022***	-0.027***
		(0.004)	(0.004)	(0.004)	(0.006)
partner's eduyrs		-0.002	-0.000	-0.000	-0.002
		(0.003)	(0.003)	(0.003)	(0.005)
first-born girl*first child dead				-0.062	-0.076
				(0.069)	(0.107)
first child dead			0.572***	0.600***	0.722***
			(0.036)	(0.048)	(0.074)
Observations	17589	17589	17589	17589	9777
R-squared	0.557	0.700	0.706	0.706	0.544
Percent effect	1.00	1.15	1.43	1.68	2.09
Mean first-born boy	4.33	4.33	4.33	4.33	5.65

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. All married women aged 15–49 in first union. Using survey weights. Includes 7 five-year age groups dummies, mother and partner's age (and square), six region dummies, birth year fixed effects, ethnic and region-specific time trends, ethnicity, religion, urban dummy, and a wealth index. ***, **, and * indicate significance at 1%, 5% and 10% levels.

correlated with fertility, ignoring to include the control for education leads to an underestimation of the effect of first-born girl. Similarly, not taking into account that boys are biologically weaker than girls at young ages and are therefore more likely to die would generate a downward bias if mortality is associated with higher fertility (because mothers are trying to replace the lost child by having more children).

The evidence found so far suggests that the sex composition of earlier-born children affects subsequent *realized* fertility. The next question is whether it has an effect on mothers' fertility preferences. I estimate a similar regression, where the dependent variable is a dummy equal to one if the mother reports that she wants more children. Table 5 shows the probit estimates. Women with a first-born girl are 2.4 percentage points (significant at 5% level) more likely to report that they want another child.

The results for the use of contraception are shown in column (2) of Table 5. Women with first-born girl are 1.1 percentage points less likely to

Table 5: The effect of first girl on desired fertility and the use of contraception

dep. var:	=1 wants another child	=1 currently using contrac.
	(1)	(2)
first-born girl	0.024**	-0.011**
	(0.010)	(0.005)
age first marriage	0.003**	-0.002
	(0.002)	(0.001)
age first birth	0.019***	0.001
	(0.002)	(0.001)
mother eduyrs	-0.003*	0.005***
	(0.001)	(0.001)
partner's eduyrs	0.004***	0.001
	(0.001)	(0.001)
first-born girl*first child dead	-0.013	0.014
	(0.023)	(0.017)
first child dead	0.111***	-0.032***
	(0.014)	(0.009)
Observations	17499	17589
Pseudo R-squared	0.282	0.224
Percent effect	3.8	6.5
Mean boy	0.64	0.17

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. All married women aged 15–49 in first union. Using survey weights. Includes 7 five-year age groups dummies, mother and partner's age (and square), six region dummies, birth year fixed effects, ethnic and region-specific time trends, urban dummy, ethnicity, religion, # children ever born, and a wealth index. ***, **, and * indicate significance at 1%, 5% and 10% levels.

use any method of contraception. This is an important effect considering that a small fraction of women are currently using contraceptives.

Together with the effects on desired fertility, this evidence suggests that women with daughters *want* to have more children compared to women with sons, probably because they intentionally continue bearing children until they have the desired number of sons. In particular, the effect for the use of contraceptives points to a conscious decision.

5.2 Birth spacing results

I turn to discussing the effects of the sex composition of earlier-born children on the length of the interval between births. Column (1) to (3) of Table 6 show that the birth of a girl has no effect on the birth interval when considering all births. This is probably due to the fact that in a high fertility context the effect on spacing might become negative (and significant) only after the birth of several girls. It is interesting to note that ignoring to include the

indicators for mortality leads to an overestimation of the effect of the birth of a girl. Consistent with the results found in the fertility part, higher mortality of males makes mothers wait less on average after the birth of a boy than that of a girl (being mortality negatively associated with the length of the interval). Therefore, it is important to control for mortality to uncover the real effect on birth spacing. Column (4) to (6) show that the effect of each successive girl birth implies a reduction of the interval for women who already had a first-born girl; however this effect is not significant. Columns (7) to (9) of Table 6 focus on the subsample of women who had at least four children, among which the first two are girls. Column (8) shows that on average mothers wait 1.27 months less after the birth of any successive girl (this effect is significant at 10% level), compared to the birth of any successive boy. The estimates obtained using OLS are similar (column 9). Finally, columns (10) to (12) consider the subsample of women with 5 or more children. Interestingly, for women with the first three born girls, the length of the interval after the birth of a girl (compared to the birth of a boy) is reduced by 3.66 months, and this effect is significant at 1% level. Appendix Table A.3 shows additional results for the subsamples of women with first-born boy (column 3), first and second-born boy (column 5), and mixed gender composition (column 6). As expected, the birth of a girl has no effect on spacing for women who had a first-born boy or mixed gender offspring among earlier-born children.

Next, I estimate the probability that the birth interval is shorter than 24 months which, as indicated in the medical literature, is the critical length below which maternal and child health are negatively affected. The results are shown in Table 7: column (1) shows that, after the birth of any girl, mothers who had two girls as the first two born children are 3 percentage points more likely to wait less than 24 months. Considering that about 33.6% of women wait less than 24 month in this subsample, the effect is important and equal to a 8.9% increase. In column 2, the sample is restricted to the first 4 children (excludes children of birth order 5 and higher), to make sure that the effect on spacing is not exclusively driven by the behavior of women with many children. The result in column 2 suggests that the pace at which mothers have children tends to accelerate especially among earlier female births (the third and/or the fourth girl born) rather than later born. The percent effect on this subsample is a 12% increase of shortly spaced births. This result is quite intuitive: the birth of girls in close succession should have a greater effect on spacing if mothers are trying to have a son. The same check is done in columns 3 to 4 for women with at least five children. Similarly, the likelihood that the next birth is closely spaced increases by 6.9

Table 6: Length of the birth interval (# months), conditional on the gender of earlier-born children

Women with:	3+ children			3+ children			4+ children			5+ children		
	all	all	all	firstG	firstG	firstG	GG	GG	GG	GGG	GGG	GGG
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
girl	0.276	0.112	0.012	-0.042	-0.178	-0.198	-1.034*	-1.271*	-1.344**	-2.790***	-3.664***	-2.970***
	(0.222)	(0.256)	(0.207)	(0.373)	(0.451)	(0.339)	(0.557)	(0.706)	(0.521)	(0.859)	(1.249)	(0.879)
girl*dead child		0.556	0.691*		0.611	0.434		1.238	1.432		2.047	1.507
		(0.485)	(0.392)		(0.706)	(0.570)		(1.032)	(0.889)		(1.496)	(1.302)
dead child		-4.917***	-7.416***		-4.951***	-7.389***		-5.566***	-7.413***		-6.596***	-7.068***
		(0.355)	(0.274)		(0.572)	(0.446)		(0.929)	(0.761)		(1.418)	(1.191)
Birth order dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother FEs	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Observations	51606	51606	51606	24315	24315	24315	11184	11184	11184	4869	4869	4869
R-squared	0.313	0.320	0.025	0.308	0.315	0.026	0.271	0.279	0.030	0.219	0.233	0.038

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. Also includes the interaction girl*first child dead in columns (4)-(12), girl*second child dead in columns (7)-(12), girl*third child dead in columns (10)-(12).

Table 7: Probability of a short birth interval (<24 months), conditional on the gender of earlier-born children

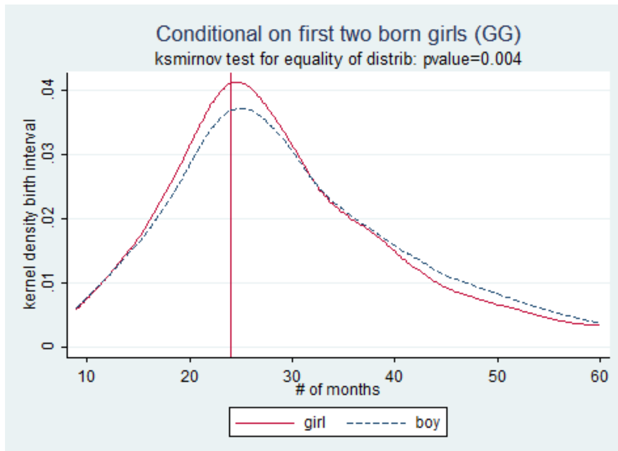
Women with:	4+ children		5+ children	
	GG	GG	GGG	GGG
Subsample	All children	first 4 children	All children	first 5 children
	(1)	(2)	(3)	(4)
girl	0.030*	0.040*	0.049	0.069*
	(0.017)	(0.023)	(0.031)	(0.039)
girl*dead child	-0.036	-0.057	0.007	-0.007
	(0.034)	(0.050)	(0.052)	(0.075)
dead child	0.178***	0.188***	0.140***	0.132*
	(0.028)	(0.044)	(0.045)	(0.068)
Birth order dummies	Yes	Yes	Yes	Yes
Mother FEs	Yes	Yes	Yes	Yes
Observations	11184	8181	4869	3866
R-squared	0.245	0.321	0.220	0.267

OLS estimates, linear probability model. Robust standard errors adjusted for clustering at the household level in parentheses. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. Also includes the interaction girl*first child dead in columns (4)-(12), girl*second child dead in columns (7)-(12), girl*third child dead in columns (10)-(12).

percentage points (equivalent to a 19% increase given that on average 36% of births are spaced less than 24 months apart in this subsample) after the birth of a girl (among the successive two births) for mothers who already had three girls.

Figure 1 plots the distribution of the length of the interval for women with at least three children by the sex of successive births. The continuous (dotted) line represents the distribution after the birth of a girl (boy). Given that the length of the interval increases with birth order, to avoid confounding effects I consider the first four births for each woman (this also allows to exclude the hypothesis that the effect only comes from women who had many children). As expected, compared to the distribution of the length of the interval after the birth of a boy, the whole distribution after the birth of a girl is more skewed to the left. The two distributions are statistically different from each other, as shown by the rejection of the null hypothesis of equality of distributions (Kolmogorov-Smirnov test). Appendix Figure A.7 shows that the distributions in the subsample of women who had two boys or mixed children gender composition are not different from each other. This evidence confirms the idea that the pressure for having sons increases with the number of girls born, and leads to a significant reduction in spacing between births, with implications for the health status of mothers and their children.

Figure 1: Distribution of length of birth interval after the birth of each girl or boy. First four births to each woman.



Subsample of women with at least three children ever born. Excludes births of birth order five or higher. Kolmogorov-Smirnov test for the equality of distribution calculated for range of interval 0-60 months.

5.2.1 Birth spacing and child mortality

In this section I estimate the effects of a short birth interval on child mortality (using mother fixed effects). I focus on the effects of the *preceding* birth interval (shorter than 24 months) on the likelihood of dying before age 2. The preceding birth interval is the birth-to-birth interval that is closed by the child under consideration.[30]

As previously discussed, short birth intervals have been shown to have deleterious effects on child health and mortality. Palloni and Millman (1986) argue that close birth spacing does not allow a woman to regain her physical strength and nutrients required to have a successful following pregnancy. The effects go through retardation of fetal growth as well as the mother's inability to properly breastfeed the child, all leading to increased risk of death of the child. Moreover, competition for resources among siblings born in close succession may aggravate the negative effects on child health.[31]

[30]I focus on the *preceding* interval rather than the *succeeding* because the latter is censored when the child dies: a short succeeding interval may be the consequence (not the cause) of the death of the previous child if mothers try to replace the lost child by having another soon after. In this case, the results would be contaminated by simultaneity bias (Palloni and Millman, 1986).

[31]The effects of short succeeding interval on the health status of the previous (older)

Table 8: The effect of short birth interval on child mortality

	women with at least 3 children			
	all children		children born last 5 years	
	(1)	(2)	(3)	(4)
	fe	ols	fe	ols
preceding interval	0.047***	0.098***	0.027**	0.061***
	(0.004)	(0.004)	(0.012)	(0.006)
breastfeeding duration			-0.016***	-0.010***
			(0.001)	(0.000)
Observations	51606	51606	13796	13796
R-squared	0.303	0.021	0.664	0.079

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. Birth order fixed effects and female dummy included.

Column 1 in Table 8 shows that a child born less than 24 months after the preceding child is 4.7 percentage points more likely to die (at age 2 or before) compared to children born after a longer interval. Column 3 and 4 restrict the sample to children born in the 5 years before the survey was conducted, for whom the duration of breastfeeding is available. In this subsample, both breastfeeding duration (which, as expected, has a negative and significant effect on mortality) and close birth spacing significantly affect mortality. In fact, it is known that one of the channels through which a short birth interval increases mortality is shorter breastfeeding (likely due to maternal depletion). This result, together with the results obtained in the previous section that birth spacing is shorter after the birth of girls, suggests that preference for boys is among the factors affecting child mortality.

6 Son Preference and Family Structure

In this section I provide additional evidence supporting the argument that the preference for biological sons affects several dimensions of people's life, by showing that it correlates with household fostering decisions and women's marital outcomes.

6.1 Child fostering

Child fostering refers to *'the relocation or transfer of children from biological or natal homes to other homes where they are raised and cared for by foster*

child mainly acts through inadequate child care a interrupted breastfeeding.

parents' (Isiugo-Abanihe, 1985, p. 53). It is a widely accepted and practiced social institution in many parts of Sub-Saharan Africa.[32]

The reasons for child fostering have been studied by several researchers (Isiugo-Abanhie, 1985; Bledsoe, 1990; Ainsworth, 1996; Akresh, 2009; Beck and others, 2011). These include insurance against idiosyncratic shocks to family income, education (children in rural households are sent to live with their urban kin members where there are schools), kinship fostering (ie., as part of the obligations between members of the same extended family), and child labor.

Akresh (2009) studies household fostering decisions in Burkina Faso. In particular, he analyzes the child labor hypothesis according to which children participate in household production by performing domestic tasks and children of each gender and age group have their specific roles. According to this idea, households with an excess of female or male children might decide to send or receive a child in order to achieve a balanced gender structure and thus maximize household productivity. I build on Akresh's idea and investigate whether households fostering decisions differ depending on gender of children they have excess of. If non-biological girls (boys) are substitutes for biological daughters (sons), I should find symmetric responses for the imbalance with more girls or boys.[33] If instead there is preference for *own* biological son but not for *own* daughter, responses should not be symmetric.

6.2 Household fostering decisions, data and empirical strategy

The analysis of fostering decisions is done at the household level (as opposed to the fertility part in previous sections where the mother and her children are the unit of analysis). The number of biological children is based on fertility histories of the spouse/s of the head (or of the woman, if the household is female-headed), and includes all daughters and sons alive *before* fostering (summing up resident children and children living elsewhere).[34] The

[32] LLoyd and Desai (1992) analyze data from 16 DHS survey around the world and find that the percentages of children aged 10–14 living away from mother ranges from 0.9 in Tunisia, 4.2 in Sri Lanka, 5.7 in Brazil, 24 in Senegal, 29.4 in Ghana, and 40.9 in Liberia. Isiugo-Abanhie (1985) argues that 'nowhere is it as institutionalized as in parts of West Africa' (pp. 56).

[33] Specifically, if household fostering decisions are influenced by this motivation, the following patterns should be observed. For households with an excess of daughters: (+) foster-in boys; (+) foster-out girls, and for households with an excess of sons: (+) foster-in girls; (+) foster-out boys.

[34] If the household head has more one wife, the sum of children of each spouse is considered.

focus is on children (alive) age 6–14 (included), for which fostering is more common.[35]

First, I empirically examine whether the probability of sending or receiving a child is correlated with household demographic variables, such as the number of biological children (and separately, the number of boys and girls). Second, I construct two variables for the household gender imbalance (less endogenous than the number of children above considered): '*more sons 6–14*', and '*more daughters 6–14*' (similarly to Akresh (2009)).[36] These are dummies equal to one if the number of biological sons age 6–14 (daughters) exceeds the number of daughters (sons), respectively. The omitted category is a dummy equal to one for households in which the number of sons equals the number of daughters. The latter category also include households with no children in the relevant age group, which are the ones that typically receive more children in Nigeria.[37] Compared to Akresh (2009), after having estimated the regressions on the full sample of households, I also restrict the sample to households with at least one biological child age 6–14 (for the fostering in regressions) and separately for households with at least one child, daughter or son (for the corresponding fostering out regression). This allows to focus on fostering decisions for households that have a gender imbalance versus those that are actually balanced because they have the same (strictly positive) number of boys and girls. Furthermore, while Akresh (2009) examines different motivations for fostering within the same regression, I examine only the child labor-gender imbalance hypothesis and also control for a set of additional household-level observable characteristics.

A foster child is defined as being a resident child whose parents are alive and live elsewhere. Accordingly, dummies for whether the household hosts at least one foster child, girl or boy are created (used for fostering-in decisions). In the DHS, the mother also reports whether each biological child is

[35] It is common in the child fostering literature to consider children in this age range. The percentage of households sending out (receiving) at least one child 6–14 is 17% (6%), while it is 3.8% and 1.8% for children 0–5 (2008 NDHS). Children age 15 and older (especially females who are of marriage age very young) are likely to live elsewhere in a separate household.

[36] The issues of simultaneity of household fostering and fertility decisions as well as selection into fostering are not dealt with in this analysis. However, I control for many observable household characteristics and use the variables for the gender imbalance (in addition to the number of children) which are a more exogenous proxy for demographics. On these issues, see Akresh (2004).

[37] In the sample of households considered, the percentage of households that fostered-in at least one child is 6%. This percentage is 8% in the subsample of households with no children in that age group.

living in the household or elsewhere. Based on this information, I construct three dummies which are equal to one if at least one biological child, son or daughter of the household head lives elsewhere (used for fostering-out decisions).

Table A.4 shows some summary statistics. The sample excludes households in which there is no woman eligible for the woman questionnaire (age 15–49); in fact, for these households the number and sex of biological children is not available.[38] 17% of households have at least one child aged 6–14 who is living elsewhere, and 6% of households host at least one child age 6–14.[39] Moreover, female children are more involved in fostering than males. 24% percent of households have an excess of daughters or sons, while 51% are balanced (including households with no children aged 6–14, which represent 40% of the sample). The bottom panel of Table A.4 shows the summary statistics for urban areas only, where fostering out (in) seems to be less (more) frequent than in rural areas. This is consistent with the idea that the quality of the network and the distance to urban areas matters for fostering.

In the empirical part, I estimate a probit model for the probability of fostering in or out a child on the number of biological sons and daughters and the variables for the gender imbalance ('more sons 6–14', and 'more daughters 6–14'). First, if fostering is motivated by the need for child labor, I expect that compared to balanced households, unbalanced households tend to foster more to achieve balance. Second, if there is preference for own biological son, household fostering decisions should differ in response to an excess of male or female children.

[38]Polygynous households in which the senior wife is older than 49 (thus not interviewed) are also not considered in the analysis.

[39]The difference arises mainly from the fact that the sample does not include households in which there are no women eligible for interview and for whom the fertility history is unavailable (eligible women are aged 15–49). However, households with older women tend to host more children (ie, fostering of grandchildren) and at the same time do not have children in the relevant age group (6–14). Among these households with missing information on the number of children, 11% host at least one child (compared to an average of 6% in the main sample). This explains most of the gap between the percentage of households receiving and sending children. Moreover, the definition of fostered (in) children requires *both* parents to be alive and living elsewhere, while there is no information on biological children living elsewhere (in particular, there is no data on the survival status of the father or whether the child is living with the father in a separate compound).

6.3 Household fostering decisions, results

Table 9 shows the results for the probability of fostering-in a child. Column (1) shows that the higher the number of biological sons and daughters, the less likely is that the household receives a child (a Wald test accepts the null of equality of the coefficients on the number of boys and girls). Columns (2) and (3) show the results for the probability of fostering in a girl or a boy. In Columns (4) to (6) the sample is restricted to households with at least one child. The results are similar to those obtained for the full sample but smaller in magnitude. This is consistent with the evidence that the absence of children age 6–14 is a strong motivation for receiving children and is consistent with the child labor hypothesis described above.

This evidence suggests that biological sons and daughters are both substitutes for fostered children in household production. However, from this evidence it is not possible to infer the household response to the gender imbalance, since the coefficients only indicate that fewer children are fostered in if there are more biological children of either sex (by keeping constant the number of children of the opposite sex).

Columns (7) to (12) include the indicators for the gender imbalance and reveal an asymmetric pattern. Households with an imbalanced gender composition are less likely to foster in children if compared to balanced or childless households. Specifically, the excess of daughters is associated with a stronger reduction in the probability of receiving a child (column 7) and in particular a girl (column 8), than the excess of sons (a Wald test rejects the equality of coefficients in columns 7 and 8, while accepts it in column 8). These results (similar to those in Akresh, 2009) probably underestimate the real effect of gender imbalances on fostering in decisions because the omitted category (same number of boys and girls) includes childless households (the ones more involved in fostering in). When restricting the sample to households with at least one child I can focus on the effect of the imbalance. Column (11) shows that households in which there are more sons than daughters are 1 percentage point more likely to receive a girl (significant at 5% level and different from the coefficient on the excess of daughters by a Wald test). By contrast, households with an excess of daughters do not foster in more boys (column 12). This asymmetric response may suggest that households lacking sons do not foster in boys because of the desire of having their own biological son. This motivation may lead them to 'save' (by avoiding to host boys in the household) to continue bearing children. Instead, households that lack daughters are significantly more likely to receive a girl, as girls are often needed in the house for performing domestic chores

Table 9: Household fostering-in decision: foster child, girl, or boy age 6–14

	full sample			subsamples			full sample			subsamples		
Households with:				at least one child 6–14						at least one child 6–14		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
dep. var:	foster child	foster girl	foster boy	foster child	foster girl	foster boy	foster child	foster girl	foster boy	foster child	foster girl	foster boy
# biol. sons	-0.015***	-0.011***	-0.004***	-0.007***	-0.005***	-0.001						
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)						
# biol. daughters	-0.020***	-0.014***	-0.006***	-0.011***	-0.007***	-0.004***						
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)						
more biol. sons							-0.017***	-0.013***	-0.005*	0.013**	0.010**	0.003
							(0.004)	(0.003)	(0.003)	(0.005)	(0.005)	(0.003)
more biol. daughters							-0.025***	-0.019***	-0.008***	0.005	0.004	-0.000
							(0.003)	(0.003)	(0.002)	(0.005)	(0.005)	(0.003)
Observations	20578	20578	20578	12477	12477	12477	20578	20578	20578	12477	12477	12477
Pseudo R-squared	0.054	0.056	0.043	0.042	0.042	0.046	0.044	0.046	0.039	0.037	0.038	0.044

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the village level in parentheses. Using survey weights. All controls and fixed effects included. ***, **, and * indicate significance at 1%, 5% and 10% levels. The dependent variables are dummies for whether the household hosts at least one foster child, girl or boy, where a foster child is defined as being a resident child whose parents are alive and live elsewhere. Other controls include region fixed effects, urban dummy, wife (or woman head if female headed household) religion and ethnicity, years of education and age of the male head and his wife/ves (or woman head), a wealth index, household land ownership, a dummy for polygynous household head and female headed household, # children ever born, # of daughters and sons alive age 0–5, # of household members older than 15 (male and female).

(Isiugo-Abanhie, 1985; Ainsworth 1996). This evidence is constistent with the idea that fostered girls are considered as substitutes for own biological daughters in fostering-in decisions, while boys are not substitutes for sons.

Table 10 reports the results for the fostering-out decision. Column (1) shows that the probability of sending away a child increases with the number of children, which is again consistent with the child labor motivation for fostering. Moreover, having many daughters increases this probability *more* than having many sons (the Wald test rejects the null of equality of coefficients). This suggests that biological sons are fostered out less than daughters, probably because parents tend to privilege having sons under their direct control more than daughters (consistent with son preference). Sons and daughters are not substitutes for each other in fostering out, as shown in columns 2 and 3 (in contrast with the results for fostering in). The higher the number of biological daughters (sons), and the more daughters (sons) are sent away: this is consistent with the idea of gender-specific roles within the household. Columns 4 to 6 focus on the subsample of households with at least one child, daughter or son, with similar results. Column (7) shows that unbalanced households send out children more than balanced or childless ones, even though a Wald test indicates that the coefficients of the variables for the excess of sons and daughters are not significantly different from each other. Columns (8)-(9) and (11)-(12) suggest that households tend to respond symmetrically to the imbalance of children of either gender.

The asymmetries found for the household fostering decisions seem to indicate that biological sons are preferred to non-biological boys, while this is not the case for female children. Alternatively, this is also partly consistent with the possibility that girls are in general more 'productive' or have a greater weight in the household production function.[40] Consistent with this idea and considering that in rural areas boys should have a greater weight in the household production function if they work in the field, the estimates obtained above should differ in urban and rural areas (eg. in fostering-in regressions, if there is excess of daughters, should foster in boys more in rural areas). Appendix Table A.5 shows that this is not the case, even though the asymmetries are stronger in rural areas (this results is in agreement with stronger son preference in more traditional rural areas). It is important to consider that the motivations for fostering are widely different in urban and rural areas and that these results deserve further examination.

[40] This alternative interpretation is not fully consistent with the results for the fostering in decisions according to which the more children of either sex and the fewer boys and girls are fostered in (when using the number of sons and daughters as independent variables). These results suggested that daughters and sons can be substitute for each other.

Table 10: Household fostering-out decision: biological child, daughter or son age 6-14 living elsewhere

	full sample			subsamples			full sample			subsamples		
Households with:				at least one child	at least one dau	at least one son				at least one child	at least one dau	at least one son
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
dep. var:	child elsewhere	daughter elsewhere	son elsewhere	child elsewhere	daughter elsewhere	son elsewhere	child elsewhere	daughter elsewhere	son elsewhere	child elsewhere	daughter elsewhere	son elsewhere
# biol. sons	0.062***	0.003	0.060***	0.048***	-0.007	0.058***						
	(0.003)	(0.002)	(0.002)	(0.005)	(0.005)	(0.006)						
# biol. daughters	0.071***	0.068***	0.003*	0.062***	0.078***	-0.007						
	(0.003)	(0.003)	(0.002)	(0.005)	(0.006)	(0.005)						
more biol. sons							0.150***	-0.021***	0.124***	-0.015	-0.043***	0.054***
							(0.009)	(0.005)	(0.008)	(0.012)	(0.015)	(0.011)
more biol. daughters							0.162***	0.135***	-0.021***	-0.000	0.064***	-0.044***
							(0.009)	(0.008)	(0.005)	(0.012)	(0.012)	(0.014)
Observations	20578	20578	20578	12477	9156	9246	20578	20578	20578	12477	9156	9246
Pseudo R-squared	0.218	0.257	0.233	0.083	0.072	0.063	0.193	0.213	0.201	0.068	0.061	0.058

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the village level in parentheses. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. The dependent variables are three dummies equal to one if at least one biological child, son or daughter of the household head lives elsewhere. Other controls include region fixed effects, urban dummy, wife (or woman head if female headed household) religion and ethnicity, years of education and age of the male head and his wife/ves (or woman head), a wealth index, household land ownership, a dummy for polygynous household head and female headed household, # children ever born, # of daughters and sons alive age 0–5, # of household members older than 15 (male and female).

6.4 Marital outcomes

The anthropological and demographic evidence introduced in Section 2 suggests that infertility and sex composition of earlier-born children are among the factors affecting the husband's decision to marry another woman. Dahl and Moretti (2004) analyze Census data for Kenya and find that, among all married women, those with daughters among earlier born children are more likely to be in a polygamous relationship compared to women with sons.[41] They interpret this as evidence that the desire for boys leads some husbands to marry another woman if the first wife delivers a girl. This could be due to the fact that women who had only girls are discriminated by their husband's kin and the society or because husbands might perceive that the probability of giving birth to boys at next births is lower.[42] Another interpretation could be that taking an additional wife, possibly younger, maximizes the probability of having many sons, given that the fertile period of the first wife is reduced (as it is reduced the possible number of male children that she can possibly bear considering that she already had daughters). Being in a polygynous union has important consequences on women's and children's welfare. In particular, even though there are contrasting views in the literature on the desirability of being in a polygynous union by women, several papers have documented that women and children in polygynous households exhibit poorer health outcomes and suffer because of competition for resources (see Bove and Valeggia, 2009, for a review of the empirical evidence on the relationship between polygyny and women's health in Sub-Saharan Africa).[43]

Female headship has often, but not always, been found to be associated with poverty and insecurity (see Buvinic and Rao Gupta 1997 for a review of the literature). Horrell and Krishnan (2007) find that female-headed households in rural Zimbabwe are no different from male-headed ones in

[41] In their empirical analysis for Kenya, they consider *resident* children rather than all children ever born (cohabitation is a very imperfect proxy given the flexibility of household structure in the African context). Moreover, they do not distinguish between wives according to their rank (likely because of lack of this information in the data). For the US, they find that women with first-born daughters are less likely to marry, more likely to be divorced, and thus more likely to have non-resident husband.

[42] Using data for the US, Ben-Porath and Welch (1976) find weak evidence that the sex realization of successive births is influenced by the sex of previous children. There is contrasting evidence in the biology literature that this might be the case (James, 2009; Stansfield and Carlton, 2007).

[43] Boserup (1970) and Goody (1976) argue that the prevalence of polygyny is positively associated with the degree of female involvement in agriculture. Jacoby (1995) empirically links polygyny to women's productivity using data for Cote d'Ivoire.

terms of income poverty, but they lack assets for agricultural production and are thus constrained in the capacity to improve productivity. In this section I empirically explore whether the sex of the first-born child affects marital status, which in turns affects children's and women's well-being.

6.5 Data and results

I use the sample of all married (or ever-married) women and their full fertility history to explore whether polygyny and other marital outcomes are associated with the sex of the first-born child. 70% of women 15–49 in the NDHS are married. Among married women, 35% have a polygynous husband, and 90% have a resident husband, while among ever-married women, 10% are divorced, and 12% are head of the household. Among polygynous married women, 41.2% are ranked first, 48.7% second, 8% third, 2% higher rank (rank assigned based on the duration of marriage). The probability of being in first union is lower for higher-rank wives (89% of rank1, 71% rank2, 58% rank3; while 92% for monogamous).[44]

I analyze the effect of first-born daughter on four outcomes: that the husband marries another woman after the first wife (he is polygynous), that he resides in the household, that the woman is divorced or head of the household. Table 11 shows the results. The empirical strategy is similar to that employed in the analysis of fertility outcomes illustrated in section 4.1.[45] Again, the identification assumption is that the sex of the first-born is exogenous after conditioning on the controls. Each regression is run separately for all married or ever-married women, and on the subsample of older women age 30–49, who are more likely to have experienced the outcomes under consideration.

I exploit the information on the rank of each wife in a polygynous union (available in the NDHS) and run the regressions on the subsample of monogamous married women and first-rank wives in a polygynous union (excluding wives of higher order rank). This is because the outcome of interest is whether the husband marries another woman as a consequence of the fact that the first wife had daughters instead of sons. Thus, I compare monogamous women to women who are currently first-rank wives but entered the union with their current husband as monogamous. Column (1) shows that women with a first-born daughter are 1.3 percentage points (significant at

[44] Levirate marriage (traditional practice in which the wife of a deceased man is obliged to marry the husband's brother) is not uncommon in Nigeria.

[45] Additional controls include the number of children ever born and household size which may be correlated both with the outcomes of interest and with the sex of the first-born.

Table 11: Sex first-born child, marital outcomes and living arrangements

Dep. var.:	=1 polygyn husband		=1 husband resident		=1 female head		=1 divorced	
Sample:	married women				ever-married women			
AGE	all	30-49	all	30-49	all	30-49	all	30-49
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
first-born girl	0.013**	0.027***	-0.009**	-0.014**	0.007*	0.018***	0.004	0.014*
	(0.006)	(0.010)	(0.004)	(0.006)	(0.004)	(0.006)	(0.005)	(0.007)
age first marr	-0.006***	-0.005***	0.001	0.001	-0.000	0.001	-0.007***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
age first birth	-0.003***	-0.005***	0.001	0.000	-0.001	-0.001	-0.002**	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
eduyrs	-0.007***	-0.009***	-0.002***	-0.002**	0.004***	0.003***	-0.003***	-0.003***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
husbeduyrs	0.002***	0.003**	-0.005***	-0.005***				
	(0.001)	(0.001)	(0.001)	(0.001)				
urban	-0.027***	-0.038***	0.007	0.011	-0.003	0.001	0.018**	0.017*
	(0.008)	(0.012)	(0.006)	(0.007)	(0.006)	(0.008)	(0.007)	(0.010)
firstchdead	0.058***	0.074***	-0.002	0.001	-0.008	-0.006	0.044***	0.061***
	(0.011)	(0.016)	(0.008)	(0.009)	(0.007)	(0.010)	(0.009)	(0.012)
first-born girl*	-0.027**	-0.046***	0.002	0.001	-0.008	-0.009	0.002	-0.014
first child dead	(0.011)	(0.017)	(0.011)	(0.013)	(0.010)	(0.013)	(0.010)	(0.013)
Observations	16314	9395	20312	11747	22812	13321	21805	12558
Pseudo R-squared	0.296	0.301	0.09	0.121	0.277	0.338	0.156	0.134

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. In Columns (1) and (2) the sample is restricted to married women who are monogamous or rank 1 (excluding women of rank 2 or higher); columns (3) and (4) include all married women; columns (5)-(8) the sample of ever married women; and columns (7)-(8) additionally excludes widowed women and those currently not living with the partner (because none of them reports to be divorced). Other controls include region fixed effects, birth year fixed effects, ethnicity and region-specific time trends, mother religion and ethnicity, age of the mother and her partner, 5-year age group dummies, a wealth index, # of children ever born, hh size. Columns 5-6 include dummies for whether women are married and, among unmarried women, if they are living with a partner, widowed, divorced, not living with a parner. Columns 7-8 include dummies for whether women are married and, among unmarried women, if they are living with a partner.

5% level) more likely to be first-rank wives (as opposed to monogamous) compared to women with a first-born son. As a placebo test, I run a regression in which I compared monogamous women and last-rank wives (of rank higher than the first). If husbands choose to marry another woman because the first had a daughter, the effect should be stronger for first-rank wives (compared to monogamous), while no effect (or a smaller effect, since the husband can always marry more women) should be found for higher rank wives. As expected, the probability of being in a polygynous union is not higher for last-rank wives who have had a first-born daughter rather than a

first-born boy (the coefficient of first-born girl is 0.000 with a standard error equal to 0.006).

Column (4) shows that the husband of women with first-born daughters is 1.4 percentage points less likely to live with his wife. Column (6) and (8) show that ever married women with first-born girl are also 1.8 (1.4) percentage points more likely to be the head of the household (divorced).[46] This evidence clearly suggests the preferences for the sex of children affects marriage outcomes.

6.6 Polygyny and the sex of children born at the time of husband's re-marriage

In the previous section, I considered the sample of *all* women, independently of their relationship to the head of the household (spouses, daughters, other relatives, etc...) and irrespectively of whether or not they co-reside with their husband or not. Given that, in this context, many women with a polygynous husband do not co-reside with him, it is impossible to know the number and the composition of children born to every wife of the same husband. However, it would be interesting to know the number and sex of children (if any) born *before* the husband decides to marry a second wife.

In order to check the *exact* sex composition of children born at the time of the husband's re-marriage, I should focus on a subsample of women with the following characteristics: resident spouses of the head in their first unions (because only the date of *first* marriage is available, and I need to have the date of marriage of the first and second wife). Therefore, I end up with a sample of 1621 women rank1 and 2 (3242 women in total), and very few of higher order. With this dataset, I compare monogamous women with women ranked first in a polygynous relationship (who were monogamous until the date of the husband's marriage with the second wife). The fraction of polygynous women in this sample is 9.5%, which is significantly lower than the fraction of polygynous women in the full sample due to the restrictions made in the dataset.

By analyzing realized fertility data for rank-one wives, I obtain the following descriptive statistics: the median number of children born before re-marriage is 2 (it's 3 after excluding women who did not have children); 19% of married women ranked first (305 out of 1621) did not have any child before re-marriage; the median number of years between 1st and 2nd marriage is 7 years (only 3 years if the first wife did not have any child).

[46]The sample of ever-married women includes, in addition to currently married women, also those who have been married in the past but are currently unmarried.

Table 12: Polygynous husband (timing of his re-marriage)

Dep. var.:=1 husband polygynous	(1)	(2)	(3)
did not have children	0.154***		
	(0.017)		
share of sons		-0.008**	
(children born before husband's re-marriage)		(0.003)	
first-born girl			0.005**
(born before husband's re-marriage)			(0.002)
Observations	14827	13266	13266
Pseudo R2	0.454	0.417	0.416

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. All married women aged 15–49 in first union. Using survey weights. Other controls include region fixed effects, mother religion and ethnicity, age and education of the mother and her partner, mother's age at first marriage and birth, urban dummy, dummies for mortality of first daughter or boy born, 5-year age group dummies, a wealth index. # of children ever born, hh size are included only in columns (2) and (3). ***, **, and * indicate significance at 1%, 5% and 10% levels.

Turning to the results, column 1 in table 12 shows that infertility is strongly associated with polygyny. Specifically, women without children are 15.4 percentage points more likely to have an husband who is polygynous. Column 2 instead shows that, among women who have at least one child, those who had all sons (compared to all daughters) are 0.8 percentage points (significant at 5% level) less likely to end up in a polygynous relationship. Consistently, women those with a first-born girl are 0.5 percentage points (significant at 5% level) more likely to have a polygynous husband. These results are consistent with the findings of the previous section.

7 Conclusions

The economic literature has long studied the prevalence and the consequences of parental son preference in developed and developing countries. In the latter context, most studies have focused on Asian and North African countries, where sex ratios at birth are considerably high. As noted by Anderson and Ray (2010), 'normal' sex ratio (at birth) in Sub-Saharan countries might be one of the reason leading to the underestimation of the phenomenon of missing women in this context. They show that women are indeed missing in Sub-Saharan Africa but at later ages because they are more vulnerable to certain diseases than men. Using individual-level data for Nigeria, this paper shows that women (or couples) make fertility decisions that seem to be motivated by the desire for giving birth to sons. In-

deed, male children can strengthen the relationship between the wife and her husband's kin (by guaranteeing the continuation of his lineage) and secure access to residence and inheritance after the husband's death. I find that son preference affects fertility behavior and has potential negative consequences for women's and child health through birth spacing. Interestingly, findings also suggest that the preference for *own* biological sons interacts with the way people participate in traditional institutions, such as child fostering and polygyny.

A Appendix

Figure A.1: Mortality, by birth order and gender, NDHS 2008 and 2003

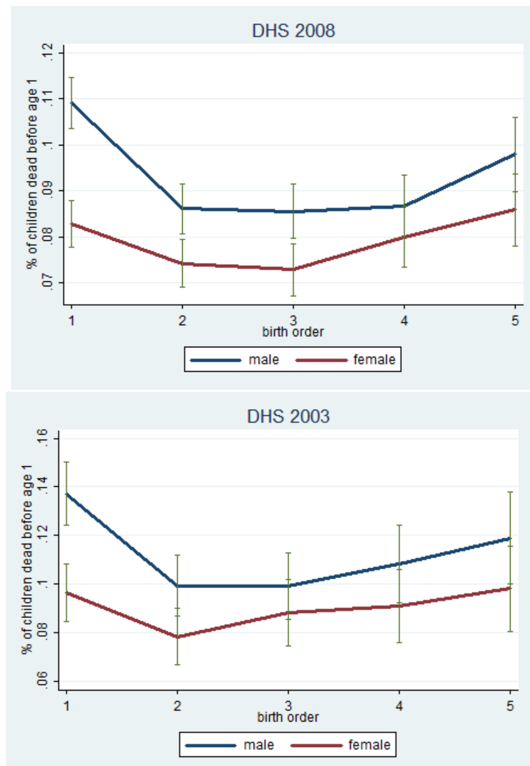


Figure A.2: Male births (out of total births), by birth year of the child, NDHS 2008

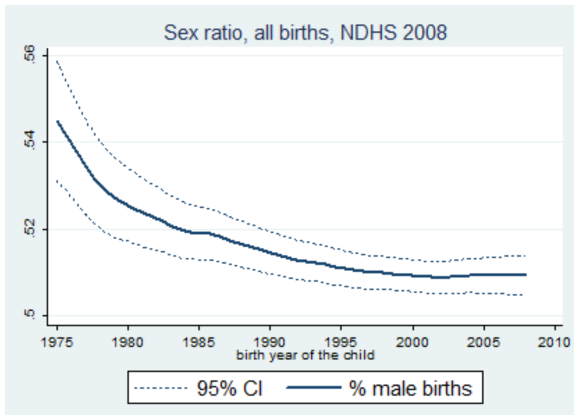


Figure A.3: Male births (out of total births), by birth year of the child, by NDHS survey

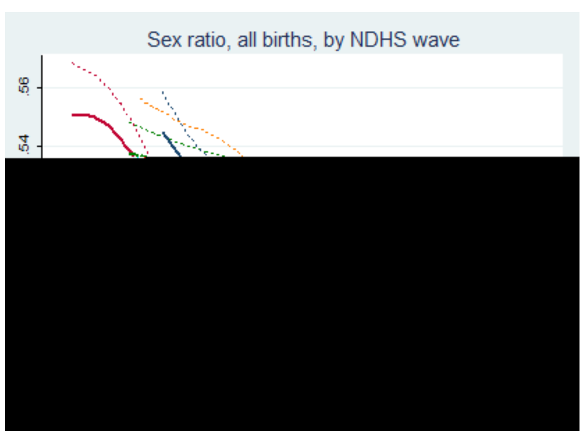


Figure A.4: Share of women with first-born girl, by age of the mother, Pooled NDHS 1990, 1999, 2003, and 2008

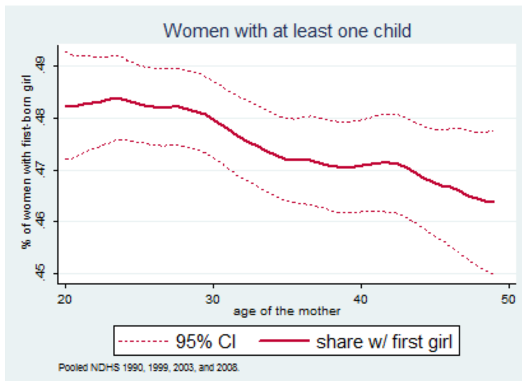


Figure A.5: Share of women with first-born girl, by age of the mother and birth order of the child, Pooled NDHS 1990, 1999, 2003, and 2008

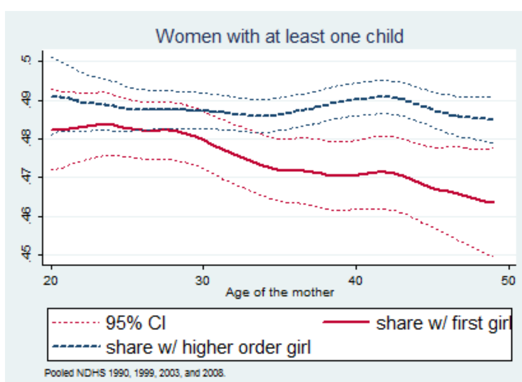


Figure A.6: Share of women with first-born girl, by age of the mother and education, Pooled NDHS 1990, 1999, 2003, and 2008

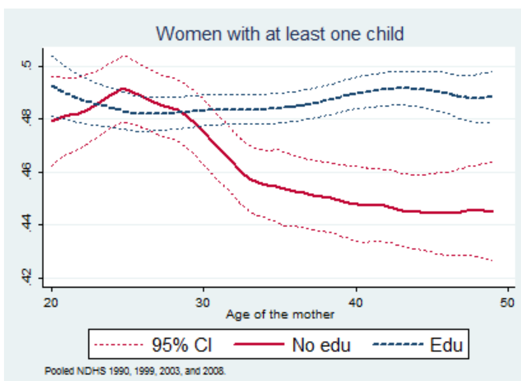


Table A.1: Summary statistics, sample of married women 15–49 in first union with at least one child ever born

	mean	p50	sd	min	max	N
# children ever born	4.31	4	2.70	1	18	18426
# daughters ever born	2.10	2	1.68	0	11	18426
# sons ever born	2.21	2	1.74	0	12	18426
# children alive	3.55	3	2.13	0	15	18426
# daughters alive	1.74	1	1.41	0	10	18426
# sons alive	1.81	2	1.44	0	11	18426
first child dead	0.17	0	0.38	0	1	18426
first daughter dead	0.08	0	0.26	0	1	18426
first son dead	0.10	0	0.30	0	1	18426
first-born girl	0.48	0	0.50	0	1	18426
first-born boy	0.52	1	0.50	0	1	18426
married	1.00	1	0.00	1	1	18426
only union	1.00	1	0.00	1	1	18426
polygynous husband	0.30	0	0.46	0	1	18346
female head	0.09	0	0.29	0	1	18426
divorce	0.00	0	0.00	0	0	18426
widowed	0.00	0	0.00	0	0	18426
husband living in	0.90	1	0.30	0	1	18307
age	31.62	30	8.39	15	49	18426
age first marriage	17.90	17	4.66	2	45	18426
age first birth	19.50	19	4.48	9	44	18426
partner age	41.66	40	10.89	17	96	18123
contraceptive use	0.17	0	0.37	0	1	18426
wants another child	0.65	1	0.48	0	1	18314
wants no more children	0.21	0	0.41	0	1	18314
urban	0.32	0	0.47	0	1	18426
education yrs	5.03	5	5.36	0	22	18414
husband eduyrs	6.40	6	5.82	0	21	18055
catholic	0.10	0	0.30	0	1	18286
other christian	0.36	0	0.48	0	1	18286
muslim	0.52	1	0.50	0	1	18286
hausa-fulani	0.34	0	0.48	0	1	18318
igbo	0.13	0	0.34	0	1	18318
yoruba	0.18	0	0.38	0	1	18318

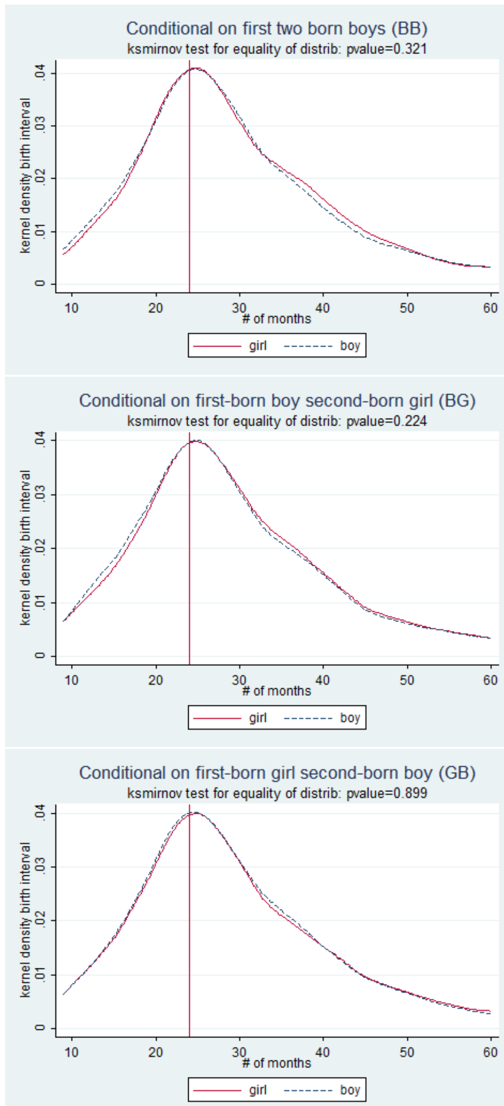
2008 NDHS. Using sample weights.

Table A.2: Summary statistics, sample of women 15–49 with at least one child ever born

	mean	p50	sd	min	max	N
# children ever born	4.35	4	2.76	1	18	23751
# daughters ever born	2.12	2	1.72	0	12	23751
# sons ever born	2.23	2	1.76	0	12	23751
# children alive	3.53	3	2.15	0	15	23751
# daughters alive	1.74	1	1.42	0	10	23751
# sons alive	1.79	2	1.45	0	11	23751
first child dead	0.18	0	0.39	0	1	23751
first daughter dead	0.08	0	0.27	0	1	23751
first son dead	0.10	0	0.31	0	1	23751
first-born girl	0.48	0	0.50	0	1	23751
first-born boy	0.52	1	0.50	0	1	23751
married	0.90	1	0.30	0	1	23750
only union	0.86	1	0.34	0	1	23050
polygynous husband	0.34	0	0.47	0	1	21690
female head	0.14	0	0.34	0	1	23751
divorce	0.10	0	0.30	0	1	23751
widowed	0.05	0	0.22	0	1	23751
husband living in	0.90	1	0.30	0	1	21652
age	32.13	31	8.61	15	49	23751
age first marriage	17.69	17	4.63	2	45	23163
age first birth	19.36	19	4.44	9	44	23751
partner age	42.49	40	11.29	15	96	21457
contraceptive use	0.16	0	0.36	0	1	23751
wants another child	0.63	1	0.48	0	1	23598
wants no more children	0.22	0	0.42	0	1	23598
urban	0.32	0	0.47	0	1	23751
education yrs	4.91	4	5.26	0	22	23735
husband eduyrs	6.16	6	5.79	0	21	22657
catholic	0.10	0	0.30	0	1	23575
other christian	0.37	0	0.48	0	1	23575
muslim	0.51	1	0.50	0	1	23575
hausa-fulani	0.34	0	0.47	0	1	23617
igbo	0.12	0	0.33	0	1	23617
yoruba	0.17	0	0.37	0	1	23617

2008 NDHS. Using sample weights.

Figure A.7: Distribution of length of birth interval after the birth of each girl or boy. First 4 children (or less) ever born to a woman



Kolmogorov-Smirnov test for the equality of distribution calculated for range of interval 0-60 months.

Table A.3: Succeeding birth interval (# months), conditional on the gender of earlier-born children

	all	firstG	firstB	G,G	B,B	mixed	G,G,G	B,B,B	mixed
Women with:	3+ children	4+ children		5+ children			6+ children		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
girl	0.112	-0.178	-0.203	-1.271*	-0.113	0.139	-3.664***	0.314	-0.195
	(0.256)	(0.451)	(0.420)	(0.706)	(0.628)	(0.406)	(1.249)	(1.007)	(0.350)
girl*dead child	0.556	0.611	0.329	1.238	-2.047**	1.023	2.047	-1.945	0.404
	(0.485)	(0.706)	(0.745)	(1.032)	(1.014)	(0.794)	(1.496)	(1.571)	(0.653)
dead child	-4.917***	-4.951***	-4.711***	-5.566***	-4.022***	-4.960***	-6.596***	-3.743***	-4.930***
	(0.355)	(0.572)	(0.484)	(0.929)	(0.592)	(0.546)	(1.418)	(0.797)	(0.463)
B.order dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	51606	24315	27291	11184	13085	21947	4869	5862	28378
R-squared	0.320	0.315	0.325	0.279	0.277	0.269	0.233	0.249	0.229

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. Using survey weights. Also includes the interaction girl*first child dead in columns (2)-(9), girl*second child dead in columns (4)-(9), and girl*third child dead in columns (7)-(9). ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table A.4: Fostering, summary statistics

Sample of households (urban and rural)

	mean	sd	min	max	N
at least one child 6–14 living elsewhere	0.172	0.377	0	1	21501
at least one daughter 6–14 living elsewhere	0.106	0.308	0	1	21501
at least one son 6–14 living elsewhere	0.098	0.298	0	1	21501
at least one foster child 6–14 in the hh	0.061	0.239	0	1	21501
at least one foster girl 6–14 in the hh	0.041	0.198	0	1	21501
at least one foster boy 6–14 in the hh	0.025	0.156	0	1	21501
more daughters 6–14 (# daughters > #sons)	0.240	0.427	0	1	21501
more sons 6–14 (# daughters < #sons)	0.246	0.431	0	1	21501
same number (# daughters = #sons)	0.515	0.500	0	1	21501
no biol. children age 6–14	0.398	0.490	0	1	21501
# biological children 6–14	1.417	1.562	0	14	21501
# biological daughters 6–14	0.702	0.977	0	12	21501
# biological sons 6–14	0.715	0.987	0	10	21501
Urban areas only					
at least one child 6–14 living elsewhere	0.133	0.340	0	1	6572
at least one daughter 6–14 living elsewhere	0.078	0.269	0	1	6572
at least one son 6–14 living elsewhere	0.077	0.266	0	1	6572
at least one foster child 6–14 in the hh	0.069	0.254	0	1	6572
at least one foster girl 6–14 in the hh	0.047	0.212	0	1	6572
at least one foster boy 6–14 in the hh	0.027	0.162	0	1	6572
# biological children 6–14	1.187	1.398	0	10	6572
# biological daughters 6–14	0.589	0.879	0	8	6572
# biological sons 6–14	0.598	0.888	0	6	6572

2008 NDHS. Using sample weights. Excludes households in which there is no woman eligible for the woman questionnaire (age 15–49) for whom the number and sex of children is not available.

Table A.5: Household fostering decisions, urban and rural areas

	Fostering-in			Fostering-out		
	foster child	foster girl	foster boy	biol. child elsewhere	biol. girl elsewhere	biol. boy elsewhere
	(1)	(2)	(3)	(4)	(5)	(6)
more biol. sons	0.015**	0.015***	0.001	-0.028**	-0.051***	0.050***
	(0.006)	(0.006)	(0.004)	(0.014)	(0.016)	(0.013)
more biol. daughters	0.007	0.006	0.000	0.000	0.062***	-0.034**
	(0.006)	(0.006)	(0.004)	(0.014)	(0.013)	(0.016)
more biol. sons*urban	-0.005	-0.011	0.009	0.045	0.030	0.012
	(0.011)	(0.007)	(0.009)	(0.029)	(0.040)	(0.026)
more biol. daughters*urban	-0.006	-0.006	-0.002	-0.001	0.006	-0.041
	(0.010)	(0.008)	(0.006)	(0.027)	(0.026)	(0.029)
urban	-0.002	0.004	-0.003	-0.006	-0.016	0.019
	(0.010)	(0.008)	(0.006)	(0.024)	(0.025)	(0.024)
Observations	12477	12477	12477	12477	9156	9246

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the village level in parentheses. Using survey weights. All controls and fixed effects included. ***, **, and * indicate significance at 1%, 5% and 10% levels. Other controls include region fixed effects, urban dummy, wife (or woman head if female headed household) religion and ethnicity, years of education and age of the male head and his wife/ves (or woman head), a wealth index, household land ownership, a dummy for polygynous household head and female headed household, # children ever born, # of daughters and sons alive age 0–5, # of household members older than 15 (male and female). Regressions in columns (1) to (4) are for the subsample of households with at least one child alive age 6–14; and the subsamples of hhs with at least one daughter and son age 6–14 in columns (5) and (6), respectively.

References

[1] Abrevaya, Jason. 2009. "Are There Missing Girls in the United States? Evidence from Birth Data" *American Economic Journal: Applied Economics*, Vol. 1, No. 2, pp. 1-34.

[2] Ainsworth, Martha. 1996. "Economic Aspects of Child Fostering in Côte d'Ivoire." In *Research in Population Economics*, ed. Schultz T. Paul, pages 25-62. Greenwich: JAI Press Inc.

[3] Akresh, Richard. 2009. "Flexibility of Household Structure: Child Fostering Decisions in Burkina Faso" *Journal of Human Resources*, Fall 2009, 44(4), pp. 976–997.

[4] Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2011. "Fertility and the Plough." *American Economic Review: Papers & Proceedings 2011*, 101:3, pp-499-503.

[5] Almond, Douglas, and Bhashkar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance During Pregnancy." *American Economic Journal: Applied Economics*. Vol. 3, No. 4, pp. 56–85.

[6] Anderson, Siwan and Debraj Ray. 2010. "Missing Women: Age and Disease." *Review of Economic Studies*, Vol. 77, pp. 1262-1300.

[7] Andersson, Roland, and Staffan Bergstrom. 1998. "Is Maternal Malnutrition Associated with a Low Sex Ratio at Birth?" *Human Biology*. Vol. 70, No. 6, pp. 1101-1106.

[8] Angrist, J. D., and W. N. Evans. 1998. "Children and their Parents Labor Supply: Evidence from Exogenous Variation in Family Size" *American Economic Review*. Vol. 88, No. 3, pp. 450–477.

[9] Simon Beck, Philippe De Vreyer, Sylvie Lambert, Karine Marazyan and Abla Safir. 2011 "Child Fostering in Senegal" working paper, Paris School of Economics.

[10] Becker, Gary. 1981. "A Treatise on the Family." Cambridge, MA: Harvard University Press, 1981.

[11] Becker, Gary and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*

[15] Behrman, Jere R., Robert A. Pollak and Paul Taubman 1997. "Do Parents Favor Boys?" *International Economic Review*, Vol. 27, No. 1, pp. 33–54.

[16] Ben-Porath, Yoram and Finis Welch. 1976. "Do Sex Preferences Really Matter?." *The Quarterly Journal of Economics*, Vol. 90, No. 2, pp. 285–307.

[17] Bergstrom, Theodore C. 1997. "A Survey of Theories of the Family." *Handbook of Population and Family Economics*, 1(1), pp. 21–80.

[18] Bledsoe, Caroline. 1990. "Transformations in Sub-Saharan African Marriage and Fertility." *Annals of the American Academy of Political and Social Science*, Vol. 510, pp. 115–125.

[19] Bledsoe, Caroline, Fatoumatta Banja, and Allan G. Hill. 1998. "Reproductive Mishaps and Western Contraception: An African Challenge to Fertility Theory." *Population and Development Review*, Vol. 24, No. 1, pp. 15–57.

[20] Boserup, Ester. 1970. "Womans role in economic development." London: Allen and Unwin.

[21] Bove, R.M. and Valeggia, C.R. 2008. "Polygyny and Womens Health in Sub-Saharan Africa." *Social Science and Medicine* Vol. 68, No. 1, pp. 21-29.

[22] Buvinic, Mayra, and Geeta Rao Gupta. 1997. "Female-Headed Households and Female-Maintained Families: Are They Worth Targeting to Reduce Poverty in Developing Countries?." *Economic Development and Cultural Change*, Vol. 45, No. 2, pp. 259–280.

[23] Byass, P., A. Worku, A. Emmelin, and Y. Berhane. 2007. "DSS and DHS: Longitudinal and Cross-sectional Viewpoints on Child and Adolescent Mortality in Ethiopia." *Population Health Metrics*, 5(12).

[24] Chowdhury, Mridul K. and Radheshyam Bairagi. 1990. "Son Preference and Fertility in Bangladesh." *Population and Development Review*, 16(4), pp. 749–757.

[25] Clark, Shelley. 2000. "Son Preference and Sex Composition of Children: Evidence from India." *Demography*, 37(1), pp. 95–108.

[26] Conde-Agudelo, A., Anyeli Rosas-Bermdez, and Ana C. Kafury-Goeta. 2007. "Effects of birth spacing on maternal health: a systematic review" *American Journal of Obstetrics and Gynecology*, Volume 196(4), pp. 297–308.

[27] Conde-Agudelo, A., and Belizan, J.M. 2000. "Maternal morbidity and mortality associated with interpregnancy interval: Cross sectional study" *British Medical Journal*, 321(7271): 1255-1259. Nov. 18, 2000.

[28] Dahl, Gordon B. and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*, Volume 75(4), pp. 1085–1120.

[29] Dreze, Jean and Mamta Murthi. 2001. "Fertility, Education, and Development: Evidence from India." *Population and Development Review*,

[30] Fapohunda, Eleanor R., and Michael P. Todaro. 1988. "Family Structure, Implicit Contracts, and the Demand for Children in Southern Nigeria." *Population and Development Review*, Volume 14(4), pp. 571–594.

[31] Fayehun O.A, Omololu O.O., Isiugo-Abanihe U.C. 2011. "Sex of preceding child and birth spacing among Nigerian ethnic groups." *African Journal of Reproductive Health*, Volume 15(2), pp. 79–89.

[32] Garenne, Michel. 2002. "Sex Ratios at Birth in African Populations: A Review of Survey Data", *Human Biology*, Vol. 74, No. 6, pp. 889-900.

[33] Garenne, Michel. 2003. "Sex Differences in Health Indicators Among Children in African DHS Surveys", *Journal of Biosocial Science*, Vol. 35, pp. 601-614.

[34] Garenne, Michel. 2009. "Sex ratio at birth and family composition in sub-saharan Africa: inter-couple variations", *Journal of Biosocial Science*, Vol. 41, No. 3, pp. 399–407.

[35] Goody, Jack. 1976. "Production and Reproduction: A Comparative Study of the Domestic Domain." Cambridge: Cambridge University Press, 1976.

[36] Isiugo-Abanihe, Uche C. 1985. "Child Fosterage in West Africa." *Population and Development Review*, Vol. 11, No. 1, pp. 53–73.

[37] Isiugo-Abanihe, Uche C. 1994. "Reproductive Motivation and Family-Size Preferences among Nigerian Men." *Studies in Family Planning*, Vol. 25, No. 3, pp. 149–161.

[38] Jacoby, Hanan G. 1995. "The Economics of Polygyny in Sub-Saharan Africa: Female Productivity and the Demand for Wives in Côte d'Ivoire" *The Journal of Political Economy*, Vol. 103, No. 5, pp. 938–971.

[39] Jayachandran, Seema and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less Than Boys: Evidence and Implications for Child Health in India." *The Quarterly Journal of Economics*, Vol. 126, No. 3, pp. 1485–1538.

[40] James, William H. 2009. "Variation of the Probability of a Male Birth Within and Between Sibships" *Human Biology*, Vol. 81, No. 1, pp. 13–22.

[41] Jensen, Robert. 2003. "Equal Treatment, Unequal Outcomes? Generating Sex Inequality Through Fertility Behavior" Mimeo, Harvard University.

[42] Klasen, Stephan. 1996. "Nutrition, Health, and Mortality in Sub Saharan Africa: Is There a Gender Bias" *Journal of Development Studies*, Vol. 32, pp. 913-933.

[43] Lloyd, Cynthia B., and Sonalde Desai. 1992. "Childrens Living Arrangements in Developing Countries." *Population Research and Policy Review*, Vol. 11, No. 3, pp. 193-216.

[44] Lundberg, Shelly J. 2005. "Sons, Daughters, and Parental Behaviour." *Oxford Review of Economic Policy*, Vol. 21, No. 3, pp. 340–356.

[45] Milazzo, Annamaria. 2012. "Why Are Older Women Missing? Son Preference and Maternal Survival in India." Bocconi University, working paper.

[46] Morduch, Jonathan. 2000. "Sibling Rivalry in Africa." *The American Economic Review*, Vol. 90, No. 2, Papers and Proceedings of the One Hundred Twelfth Annual Meeting of the American Economic Association (May, 2000), pp. 405–409.

[47] Palloni, Alberto, Guido Pinto Aguirre, and Santiago Lastiri. 1994. "The Effects of Breast-feeding and the Pace of Childbearing on Early Childhood Mortality in Mexico." *Bullettin of PAHO*, Vol. 28, No. 2, pp. 93–110.

[48] Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Income on Sex Imbalance." *The Quarterly Journal of Economics*, Vol. 123, No. 3, pp. 1251–1285.

[49] Rosenblum, Daniel. 2012. "The Effect of Fertility Decisions on Excess Female Mortality in India." working paper, Dalhousie University.

[50] Rosenzweig, Mark R. and T. Paul Schultz. 1982. "Market Opportunities, Genetic Endowments, and Intrafamily Resource Distribution: Child Survival in Rural India." *The American Economic Review*, Vol. 72, No. 4, pp. 803–815.

[51] Sen, Amartya. 1990. "More Than 100 Million Women Are Missing." *The New York Review of Books*, 37(20).

[52] Setty-Venugopal, V and U.D. Upadhyay. 2002. "Birth Spacing: Three to Five Saves Lives." *Population Reports*, Series L, No. 13. Baltimore: Johns Hopkins Bloomberg School of Public Health, Population Information Program.

[53] Smith, Herbert L. 1994. "Nonreporting of Births or Nonreporting of Pregnancies? Some Evidence from Four Rural Counties in North China." *Demography*, 31(3), pp. 481–486.

[54] Smith, Daniel Jordan. 2004. "Contradictions in Nigeria's Fertility Transition: The Burdens and Benefits of Having People." *Population and Development Review*, 30(2), pp. 221–238.

[55] van de Walle, Dominique. 2011. "Lasting Welfare Effects of Widowhood in a Poor Country." *World Bank Policy Research Working Paper 5734* .

[56] Waldron, Ingrid. 1983. "Sex differences in human mortality: The role of genetic factors." *Social Science & Medicine.*, Vol. 17, Issue 6, 1983, pp. 321-333.

[57] Williamson, Nancy E. 1976. "Sons or daughters: A cross-cultural survey of parental preferences" SAGE Library of Social Research Volume 31, Beverly Hills: Sage Publications.

[58] World Health Organization (WHO). 2011. "World Health Statistics" World Health Organization, Geneva.

[59] WHO/UNICEF/UNFPA/World Bank. 2010. "Trends in Maternal Mortality: 1990 to 2008" World Health Organization, 2010.

[60] Wusu Onipede, and Uche C. Isiugo-Abanihe. 2006. "Interconnections among changing family structure, childrearing and fertility behaviour among the Ogu,

Southwestern Nigeria: A qualitative study" *Demographic Research*, Vol. 14, No. 8, pp. 139–156.