DECLARATORIA SULLA TESI DI DOTTORATO

Il/la sottoscritto/a

COGNOME	Milazzo
NOME	Annamaria
Matricola di iscrizione al dottorato	1370055

Titolo della Tesi:

Gender, Family and Traditional Norms in Developing Countries

Dottorato di ricerca in	Economics
Ciclo	XXIV
Tutor del dottorando	Eliana La Ferrara
Anno di discussione	2013

DICHIARA

sotto la sua responsabilità di essere a conoscenza:

1) che, ai sensi del D.P.R. 28.12.2000, N. 445, le dichiarazioni mendaci, la falsità negli atti e l'uso di atti falsi sono puniti ai sensi del codice penale e delle Leggi speciali in materia, e che nel caso ricorressero dette ipotesi, decade fin dall'inizio e senza necessità di nessuna formalità dai benefici previsti dalla presente declaratoria e da quella sull'embargo;

2) che l'Università ha l'obbligo, ai sensi dell'art. 6, comma 11, del Decreto Ministeriale 30 aprile 1999 prot. n. 224/1999, di curare il deposito di copia della tesi finale presso le Biblioteche Nazionali Centrali di Roma e Firenze, dove sarà consentita la consultabilità, fatto salvo l'eventuale embargo legato alla necessità di tutelare i diritti di enti esterni terzi e di sfruttamento industriale/commerciale dei contenuti della tesi;

3) che il Servizio Biblioteca Bocconi archivierà la tesi nel proprio Archivio istituzionale ad Accesso Aperto e che consentirà unicamente la consultabilità on-line del testo completo (fatto salvo l'eventuale embargo);

4) che per l'archiviazione presso la Biblioteca Bocconi, l'Università richiede che la tesi sia consegnata dal dottorando alla Società NORMADEC (operante in nome e

per conto dell'Università) tramite procedura on-line con contenuto non modificabile e che la Società Normadec indicherà in ogni piè di pagina le seguenti informazioni:

- tesi di dottorato (titolo *tesi):* "Gender, Family and Traditional Norms in Developing Countries";
- di *(cognome e nome del dottorando*) Milazzo Annamaria;
- discussa presso l'Università commerciale Luigi Bocconi – Milano nell'anno (*anno di discussione) 2013 ;*
- La tesi è tutelata dalla normativa sul diritto d'autore (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche). Sono comunque fatti salvi i diritti dell'Università Commerciale Luigi Bocconi di riproduzione per scopi di ricerca e didattici, con citazione della fonte;
- **solo nel caso sia stata sottoscritta apposita altra dichiarazione con richiesta di embargo**: La tesi è soggetta ad embargo della durata di …..................... mesi (indicare durata embargo);

5) che la copia della tesi depositata presso la NORMADEC tramite procedura online è del tutto identica a quelle consegnate/inviate ai Commissari e a qualsiasi altra copia depositata negli Uffici dell'Ateneo in forma cartacea o digitale e che di conseguenza va esclusa qualsiasi responsabilità dell'Ateneo stesso per quanto riguarda eventuali errori, imprecisioni o omissioni nei contenuti della tesi;

6) che il contenuto e l'organizzazione della tesi è opera originale realizzata dal sottoscritto e non compromette in alcun modo i diritti di terzi (legge 22 aprile 1941, n.633 e successive integrazioni e modifiche), ivi compresi quelli relativi alla sicurezza dei dati personali; che pertanto l'Università è in ogni caso esente da responsabilità di qualsivoglia natura, civile, amministrativa o penale e sarà dal sottoscritto tenuta indenne da qualsiasi richiesta o rivendicazione da parte di terzi;

7) **scegliere l'ipotesi 7a o 7b indicate di seguito**:

7a) che la tesi di dottorato non è il risultato di attività rientranti nella normativa sulla proprietà industriale, non è stata prodotta nell'ambito di progetti finanziati da soggetti pubblici o privati con vincoli alla divulgazione dei risultati; non è oggetto di eventuali registrazioni di tipo brevettale o di tutela, e quindi non è soggetta a embargo;

Oppure

7b) che la tesi di Dottorato rientra in una delle ipotesi di embargo previste nell'apposita dichiarazione "**RICHIESTA DI EMBARGO DELLA TESI DI DOTTORATO**" sottoscritta a parte.

Data _______________

COGNOME	Milazzo	NOME	Annamaria

Contents

Acknowledgements

Writing this Thesis has been a challenging experience. I am grateful to all those who supported me during the last four years as a PhD student.

First I would like to thank my advisor, Eliana La Ferrara, who has always been (and is!) a rich source of inspiration and advice. Also, I wish to thank Eliana for believing in me and giving me the unique opportunity to work with her on a joint project which is now the third chapter of this dissertation. I have been fascinated by the readings and research conducted for this project which inspired me in writing the other two chapters of this Thesis.

I am deeply grateful to Maristella Botticini for helping me shape and develop my research ideas, always with enthusiasm and a positive attitude. I have also benefited greatly from discussions with many faculty members at Bocconi University, in particular with Alberto Alesina, Selim Gulesci, Andreas Madestam, and Katja Kaufmann.

Sylvie Lambert, Denis Cogneau, and Marc Gurgand have been welcoming and supportive while I was visiting researcher at the Paris School of Economics. In particular, I want to thank Sylvie for her invaluable suggestions and for sharing her rich knowledge of West Africa.

I thank all seminar and conference participants at Bocconi University, the Paris School of Economics, CRED at the University of Namur, CSAE Conference in Oxford, NEUDC Conference at Dartmouth College, CEPR/AMID Conference in Milan, AEL Conference in Bonn, and several other conferences where my coauthor presented our joint work for very helpful comments and discussions.

I am grateful to all my friends and PhD colleagues, and in particular to Greta, Annalisa, Marianna, Marinella, Ksenia, Patrice, Davide, Emiliano, Francesco, Diana, Giovanna, Guilhem, Lore, Maëlys, who have been supportive and kept me smiling through the ups and downs of the PhD process.

I thank my parents Oriana and Stefano, my brother Paolo, my grandmother Rina, Lucia, Filippo and Salvatore for always being there.

To my husband Paolo, without you this dissertation would not have been possible. I find it difficult to find the right words to express my gratitude for your love and constant presence in my life. I shared every moment of this entire journey with you, and I am more than happy to dedicate this dissertation to you.

Preface

This dissertation explores the consequences of parental gender preferences and traditional norms on different economic outcomes in developing countries. These includes fertility, health and educational outcomes. It is composed by three essays. The first essay analyzes how the preference for male children, or 'son preference', affects fertility behavior and traditional family structure in Nigeria, a context where little is known about parental gender preferences. The second essay explores the 'unintended' consequences of son preference for maternal health in India where the phenomenon of missing women at birth is well-known, while the newly uncovered problem of women missing of adult and reproductive ages is still widely unexplored. While the first two essays focus on gender preferences, the third chapter analyzes how different customary norms can be a source of distortion in parents' human capital investment decisions for their children by exploiting a policy change in Ghana that represents an exogenous shock to the strength of the social norms.

The literature on parental preferences over a child's gender is extensive. It has mainly focused on gender bias at birth or at very young ages in many South and East Asian countries where there is considerable evidence of infanticide, sex selective abortion and neglect of female children within the household. Yet, Anderson and Ray (2010) recently provide a decomposition of the excess female deaths by age and causes of death using aggregate mortality data for different parts of the world and unveil several new findings.[1] While they still find strong evidence of severe gender bias at young ages in India and China, they also find that, as a proportion of the total female population, the number of missing women is largest in Sub-Saharan Africa (mostly women of reproductive ages), and that the majority of missing women in India are of adult ages. The first two chapters of this Thesis attempt to make a step forward in understanding the causes of excess female mortality among women of reproductive and adult ages. This is done by using individual-level data and analyzing how reproductive behavior (i.e., the number of children ever born or desired, the time interval between each birth, and the use of contraceptives) is affected by the sex of earlier-born children. High fertility and especially short birth spacing are in fact associated with poor maternal (and child) health outcomes. As the focus is on women of reproductive ages, the two countries under study, India and Nigeria, are chosen as the two countries with the largest number of maternal deaths, accounting for 18% and 14% of all deaths for worldwide, respectively.[2] The empirical strategy used in the first two chapters mainly relies on the exogeneity of the sex of the first-born child. Therefore, 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. While the effects are stronger when considering the first two female births, I

[1] Anderson, Siwan and Debraj Ray. 2010. "Missing Women: Age and Disease." *Review of Economic Studies*, Vol. 77, No. 4, pp. 1262-1300.

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

focus on the sex of the first-born to maintain a causal interpretation of the results. The effect of the sex of earlier-born children on birth spacing is estimated using mother fixed effects to control for unobserved mother's heterogeneity.

The first chapter, titled *'Son Preference, Fertility, and Family Structure. Evidence from Reproductive Behavior among Nigerian Women'*, explores whether and how son preference affects women's fertility behavior and complex family structure. By using individual-level data for Nigeria, the empirical analysis shows evidence of son preference in several dimensions. Specifically, 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 are more likely to reduce spacing between following births the more daughters they had among earlier-born children. 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 second chapter, titled *'Why Are Adult Women Missing? Son Preference and Maternal Survival in India'*, explores the phenomenon of missing women of reproductive and adult ages in India. The empirical analysis uses individual-level data and provides a comparison of the age structure and health indicators of women by the sex of their first-born to uncover several new findings. First, among mothers with at least one child, the share of those with first-born daughter decreases with women's age (observed at the time of the survey). Second, women with first-born daughter are significantly more likely to be anemic when young (especially under age 30) while they show better anthropometric indicators and no different incidence of anemia as they get older. Interestingly, differences in anemia prevalence arise only some time after the birth of the first child. Third, women with first-born daughter have significantly higher (realized and desired) fertility and are less likely to use contraceptives. Moreover, a female birth is associated with shorter spacing which might put mothers at higher risk of dying for complications during or after childbirth and maternal depletion. I argue that all these findings are consistent with a selection effect in which healthier women of higher socioeconomic status are more likely to survive the birth of daughters. To my knowledge, this is the first paper arguing that women's excess mortality can (at least partly) be an unintended consequence of son preference.

Parental investment in human capital has been shown to have important consequences for children's well being and is a crucial input into their social mobility prospects. Available evidence mostly comes from industrialized countries, where parents are generally free to pass their wealth on to their offspring according to their will. In many developing countries, however, bequests respond not only to parental decisions, but to a series of claims by extended family and lineages that are enforced through customary norms. The presence of these norms is a potential source of distortions in parents' allocation decisions, the extent of which is not yet fully understood.

The third and last chapter of this dissertation, titled *'Customary Norms, Inheritance, and Human Capital: Evidence from a Reform of the Matrilineal System in Ghana'*, a joint paper with Eliana La Ferrara, examines the effects of traditional inheritance rules on parental investments in their children's human capital. In a context where parents are constrained in the possibility of passing land on to their children (e.g., because it is considered property of the extended family, or clan), investment in their children's human capital might not be optimal. The paper focuses on matrilineal inheritance rules and exploits a policy experiment in Ghana, the introduction of the 1985 Intestate Succession Law. The Law introduced minimum quotas for the land that parents can devolve on their children through intestacy, substantially reducing the share going to the matriclan. In this setting, the Law allowed parents to move closer to the unconstrained optimum. The empirical strategy is a difference-in-difference strategy that exploits differences across cohorts and ethnic groups separately for males and females. We find evidence that, compared to the other patrilineal ethnic groups in Ghana (unaffected by the Law), children in matrilineal groups exposed to the reform received significantly less education. This effect is specific to males for whom the matrilineal constraint was binding, while there is no effect for females. This evidence suggests that before the reform matrilineal groups in Ghana invested more in their children's education to substitute for land inheritance and more generally that traditional norms are important in determining the intergenerational accumulation of property and human capital investments.

To conclude, each chapter provides an original contribution to the literature on parental gender preferences, traditional norms, and economic outcomes. The first two chapters focus on two aspects that have been widely unexplored in the literature on son preference. These include how gender preferences interact with fertility behavior and traditional social institutions in an African country, and the consequences of son preference for maternal health in India. Finally, by exploiting a natural experiment, the last chapter allows to make some progress in establishing a causal link going from social constraints to economic choices.

Chapter 1

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

This paper explores whether and how son preference is reflected in women's reproductive behavior and family structure. Using individual-level 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 are more likely to reduce spacing between following births the more daughters they had among earlier-born children. The preference for *own* biological sons is also supported by child fostering patterns in which daughters are found to be substitutes for non-biological girls, while the same does not hold true for sons and non-biological boys. 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 social institutions. These results can partly explain the missing women phenomenon in Sub-Saharan Africa uncovered by Anderson and Ray (2010).

1.1 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 South and East Asia, 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, and shown that subsequent fertility is higher for women who had girls among earlier-born children (son-preferring fertility stopping rules) (Chowdhury and Bairagi, 1990; Clark, 2000; Dreze and Murthi, 2001). It has been argued that the 'try until you have a son' fertility rule may 'passively' lead to a lower amount of resources being allocated to female children within the household. This may happen either as a consequence of shorter weaning time for girls because parents want to 'try again' for a son if they had a girl first (with reduced health benefits from breastfeeding for female children), or because girls tend to be born in larger households (Jayachandran and Kuziemko, 2011; Jensen, 2003).[2] Again, most of the evidence comes from Asian countries. Recently, in addition to the well-documented preference for sex-balance (Ben-Porath and Welch, 1976; Angrist and Evans, 1998), new evidence of son preference has been reported also for the United States. In particular, recent research has documented unusually high sex ratios

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

[2] Barcellos *et al.* (2012) find that, accounting for the effect of son-preferring stopping rules on family size, boys still receive more parental investment (including childcare time, breastfeeding, vaccinations, and vitamin supplementation) than girls in India.

especially among Asian mothers in the United States: 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.[3] In a recent study, Anderson and Ray (2010) provide a decomposition of the excess female deaths by age and causes of death in different parts of the world. They find that many women are indeed 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 and especially during their reproductive years.[4] 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.

The reasons given for son preference are various and include cultural, kinship, economic and institutional factors that typically differ across different societies (Das Gupta *et al.*, 2003; Almond *et al.*,

[3]Two notable exceptions are Garg and Morduch (1998) who show that children with relative more sisters than brothers benefit in the intra-household allocation of resources in resource-constrained households in Ghana; and Klasen (1996), who uses population and mortality data for Sub-Saharan Africa to construct an adjusted aggregate reference ratio and finds a small and increasing anti-female bias.

[4]Anderson and Ray (2010) find that the sources of excess female mortality among adult women in Sub-Saharan Africa are primarily HIV infection and maternal mortality. The World Development Report 2012 corroborates these findings observing that maternal mortality and morbidity related to childbirth and HIV/AIDS are the two mechanisms driving excess female mortality during women's reproductive years (World Bank, 2011).

2009).[5] In the African context, anthropological and demographic evidence emphasizes the dominant role of males in traditional patrilineal societies where descent and inheritance are 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 social pressure for bearing sons is associated with specific patterns in fertility behavior and family structure. It makes at least two novel 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 child fostering and polygyny by arguing that the preference for biological sons is among the factors affecting participation to these traditional institutions.

I use the Nigeria Demographic and Health Survey (NDHS) data 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

[5]Recently, Alesina *et al.* (2011) show that cross-country differences in fertility and female labor force participation can be explained by the form of agriculture traditionally practiced (in which the historical participation of women varied depending on the degree of physical strength needed). Similarly, Carranza (2012) shows that soil texture explains a large part of the variation in women's relative participation in agriculture and in infant sex ratios across districts in India.

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 the first two born children, those who have a third or fourth daughter are 3.6 (1.8) percentage points more likely to wait less than 24 (15) months. These findings have implications for child and maternal health that, according to the medical literature, are adversely affected by short spacing (i.e., birth intervals shorter than 24 and 15 months for child and maternal health, respectively).

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 (i.e., that their husband marries another woman). Importantly, this effect is specific to first-rank wives (who entered the current union with their husband as monogamous) while no effect is found for higher-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.

Altogether, these findings suggest that parental gender preferences *do* affect fertility behavior and play a role in shaping complex family structure in Nigeria. More specifically, this paper puts forward at least two channels through which adult women might be missing in some countries (the phenomenon unveiled by Anderson and Ray, 2010, and 2012). The first is son-preferring fertility behavior which can have negative effects on women's health by increasing the risk of maternal mortality and morbidity through short spacing and maternal depletion. Maternal mortality is in fact among the highest in the world in Nigeria (second only to India in absolute number of deaths), which alone accounts for 14 percent of all maternal deaths worldwide (WHO, UNICEF, UNFPA, and World Bank, 2010). The second channel refers to the association between parental gender preferences and specific marital outcomes and living arrangements in which women are often found to be particularly vulnerable and disadvantaged (i.e., female headship and polygyny). To my knowledge, this is the first paper arguing that these outcomes can arise as unintended consequences of son preference.

Theoretically, son-preferring fertility behavior would be generated by a simple dynamic model of mothers' decision to continue or stop having children, similar to the model developed in Jayachandran and Kuziemko (2011).[6] In their model, the utility depends on the quantity of children (with decreasing returns) and their sex composition. In particular, sons give additional utility which is scaled up depending on an exogenous parameter for the degree of son preference. Considering that breastfeeding can be similarly modeled as longer time interval between births, the predictions of such a model (under the mentioned assumptions and conditional on the number of children) would be that mothers are more likely to wait longer after the birth of boys rather than girls. Similarly, they will have fewer children the more boys they had among previously-born children.

Although these results cannot be generalized to the Sub-Saharan African continent or to other contexts, they might represent a starting point for further research studying the underlying motivations

[6]One of the key assumptions of their model is that breastfeeding (which can be thought as longer birth spacing in this paper) perfectly inhibits fecundity, meaning that the mother's decision of breastfeeding a child is equivalent to the decision of stopping having children.

for the existence of specific institutional features and understand whether they represent a constraint to economic development.

This paper is related to the literature on parental gender preferences in different areas of the world, mostly Asian countries and the United States (Ben-Porath, and Welch, 1976; Williamson, 1976; Sen, 1990; Dahl and Moretti, 2008; Abrevaya, 2009; Anderson and Ray, 2010, 2012; Jayachandran and Kuziemko, 2011). Compared to this literature, I investigate if fertility decisions are gender-biased in Nigeria, a context where little is known about parental gender preferences and complements findings in Anderson and Ray (2010, 2012).

A second related strand of literature analyzes the motivations and the economic implications 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, it relates to the sociological literature on the effects of a child's gender 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 1.2 provides some insights on the context and cultural background; section 1.3 describes the data; section 1.4 introduces the empirical methodology used to analyze fertility outcomes and birth spacing;

section 1.5 presents the results; Section 1.6 focuses on the interactions between son preference and family structure, by showing how the preference for own biological male children may affect child fostering decisions as well as marital outcomes, and section 1.7 concludes.

1.2 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.[7]. The anthropological and demographic literature emphasizes 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

[7] 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%.

the respondents 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).[8,9]

'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

[8]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.

[9]Female labor force participation (the proportion of population age 15 and older that are economically active) is 48% in 2010 (World Bank WDI, 2012). In the NDHS 2008, 59% of women 15–49 reported to be currently employed; half of these women worked in the sales and services sector, 24% in agriculture, and 14% as skilled manual workers.

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).[10]

'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 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).

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

Smith (2004) argues that the traditional high value associated with 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 the interaction between son preference, fertility behavior and complex family structure related to traditional institutions.

1.3 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.[11]

Table 1.1 reports summary statistics for this sample.[12] The mean number of children ever born is 4.3.[13] This does not reflect

[11] One of the findings presented in Section 1.6 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.

[12] Summary statistics for other variables are reported in appendix Table 1.14. Table 1.15 reports statistics for the full sample (which includes unmarried women and those not in first union).

[13] The number of children ever born includes children who have died. The WHO estimates that in Nigeria infant mortality rates (the probability of dying

completed fertility because part of the women aged 15–49 in the sample will bear more children. 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 a TFR of 5.9.[14] Knowledge of any contraceptive method is widespread in Nigeria, while 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).[15] As described in the previous section, high fertility is strongly desired in the Nigerian context. This is reflected both in the high number of children ever born

before reaching age 1) are 92 and 80 deaths per 1,000 live births for male and female children respectively in 2009 (WHO, 2011). For an international comparison, the male and female mortality rates in 2009 are: 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 (including China). 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).

[14]The WHO (2011) reports that the TFR is 5.2 in Nigeria in 2009. 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).

[15]68% of all married women and 90% of all married men know at least one method of contraception, and 29% (45%) of all married women (men) reported ever using a method of contraception at some time. Ever use peaks at 36% (49%) among married women and men age 35-39. 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). 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 knows about their use of contraception.

and desired fertility; 65 percent of women in fact answer that they want to have more children. About 30 percent of women have a polygynous husband, 9 percent are the head of the household, and for 90 percent of women the husband lives with them. Finally, husbands have completed on average 1.4 more years of education with respect to their wives.

Sex ratio at birth 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 reported births.[16,17] Figure 1.3 plots the sex ratio by the year of birth of all children reported in the NDHS using Kernel-weighted local polynomial smoothing. It shows a clear negative relationship between the sex ratio at birth and the birth year of the child for births occurred 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 an association between maternal malnutrition and the probability of a male birth.

[16] Based on several data sources, Anderson and Ray (2010) report that sex ratio at birth is about 0.514 in developed countries (including 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 NDHS 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 NDHS surveys, where the sex ratio is 0.512 (2008) and 0.515 (2003).

According to this literature, women exposed to adverse nutritional conditions are more likely to give birth to females. This is mainly because of the biological 'weakness' of male fetuses (relative to female's) that makes them more likely to die (Waldron, 1983). Therefore, it might be the case that the decreasing sex ratio reflects variations in nutritional conditions.[18] To check for the biological explanation, it is useful to understand whether the sex ratio is consistently high for births occurred during the period 1975–1994 by using data from surveys carried out in different periods. In absence of detailed fertility histories in Census data, I use all the survey rounds of the NDHS (1990, 1999, 2003, and 2008). Figure 1.4 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 for the sex ratio using data from the India National Family Health Survey (NFHS).[19] By comparing the age structure and

[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. If the biological explanation holds, these negative events would be consistent with lower sex ratio (more females born) for births occurred in the past, but this is not what is shown in figure 1.3, suggesting that there may be other explanations. In all the regressions, I also 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.

[19] Using a different dataset for India (with retrospective fertility histories as

health indicators of women by the sex of their first child, I argue that son preference might be associated with selective maternal mortality through son-preferring fertility behavior. Specifically, female births are associated with higher fertility and shorter spacing, which might put mothers at higher risk of dying for complications during or after childbirth and maternal depletion of nutritional resources. The analysis of health indicators indicates that women with first-born daughter are significantly more likely to be anemic when young (especially under age 30) while they show better anthropometric indicators and no different incidence of anemia as they get older. Interestingly, differences in anemia prevalence arise only some time after the birth of the first child. These findings are consistent with a selection effect in which healthier women of higher socioeconomic status are more likely to survive the birth of daughters.

Selective maternal mortality is also consistent with patterns found using the Nigerian data. I pool all the available NDHS surveys (carried out in 1990, 1999, 2003, and 2008) and plot the share of women with first-born girl (as opposed to first-born boy) by the age of the mother observed at the time of the survey. Figure 1.5 shows that the share of women with first-born girl is lower among older women, in particular those older than 30. Figure 1.6 shows that the declining trend is specific to first births, while the sex ratio for births of higher order is constant for women of all ages. This maybe due to the association between first-born girl, realized fertility, and spacing. Figure 1.7 plots the share of women with first-born daughter for women with no education ('No edu'), and those with

in the NDHS), 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.

at least one year of education ('Edu'). The sex ratio is declining only among uneducated women; this is consistent with the idea that women of 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 30 with first-born girl exhibit better nutritional conditions than those with first-born boy (see table 1.2 and 1.3). 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).

Therefore, son preference might partly contribute to the high number of deaths due to maternal reasons in Nigeria. Maternal mortality is in fact among the highest in the world in Nigeria with an estimated number of 50,000 maternal deaths in 2008, second only to India (with 63,000 maternal deaths). The WHO estimates that one in 23 women dies for complications related to childbirth in Nigeria (WHO, UNICEF, UNFPA, and World Bank, 2010).[20] Wall (1998) observes that a series of factors contribute to high maternal mortality and morbidity in Nigeria, including low female education, marriage and first pregnancy often occurring at an early age, inadequate health facilities, and cultural and religious practices that restrict women's access to adequate care. Improving maternal

[20]The adult lifetime risk of maternal death is defined as the probability of dying from a maternal cause during a woman reproductive lifespan. For comparison, the WHO estimates for broad world regions are: 1 in 32 in the African Region; 1 in 150 in South-East Asia (incl. India); 1 in 84 in the Eastern Mediterranean Region; 1 in 2900 in the European region; 1 in 670 in the Region of the Americas; and 1 in 1100 in Western Pacific Region (incl. China)

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.[21]

Finally, being based on self-reported retrospective information, birth histories contained in the NDHS might suffer from misreporting and recall bias, especially for births of children who died in infancy (Smith, 1994; Byass *et al.*, 2007). To check whether recall bias might be a gender-biased phenomenon (e.g., mothers more likely to under-report female births and deaths than males), I test if age at first birth predicts the sex at first birth.[22] The idea is that if mothers do not recall their first-born daughters who died and report a later born son as being the first-born instead, they will look like they had a boy as first-born, but later. If this is the case, then age at first birth should be negatively associated with the sex of the first-born being female. Table 1.16 shows that age at first birth is positively (but not significantly) associated with a female first birth: this evidence is against the possibility of selective recall bias. If present, selective recall bias would lead to an underestimation of the effect of first-born girl on fertility, because women of lower socio-economic status (thus, with a higher number of children on average) would be more likely to under-report female births.[23]

[21]The target of the fifth MDG consists in reducing the maternal mortality ratio (MMR) by three quarters between 1990 and 2015, and achieving universal access to reproductive health by 2015.

[22]I thank James Fenske for suggesting this check.

[23]Using data for India, a context with strong preference and where selective underreporting of female births should be a serious issue, I find evidence that is not consistent with recall bias. In particular, among all women with at least one child, those with first-born girl are younger on average than women with

Women's characteristics by the sex of the first born Table 1.2 compares some observable characteristics of women by the sex of their first-born child. On average, women with first-born girl are more educated, have a better educated husband, are more likely to report that they want to have more children, and to be the head of the household. 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 15–29, and 30–49. I expect to find smaller differences between women with first-born girl and first-born boy among younger women. This is because women's reproductive behavior differs depending on the sex of their earlier-born children, and women might have developed differences over time. If son-preferring fertility behavior leads to selective maternal mortality as explained above, better off and healthier women would be more likely to survive the birth of girls. As expected, table 1.3 shows that the differences are much smaller for younger women. Among older women (aged 30–49), those with first-born girl are 'better off' in terms of their own and husband's education, they married and had the first child later, are more likely to be head of the household, and

first-born daughter (with age observed at the time of the survey): this 'age disadvantage' of women with first-born daughter appears to be smaller in more recent waves. This is consistent with selective maternal mortality decreasing over time due to improved maternal health conditions, while it is in contrast with selective recall bias since son preference did not show any declining trend in the Indian context (if anything, evidence suggests it is worsening over time).

are also characterized by higher weight-for-height and body mass index. This evidence is consistent with selective maternal mortality as a plausible explanation (see Milazzo, 2012, for similar evidence 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.[24] 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 (NDHS, 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).[25] Also, there are effects on health of older children (e.g., breastfeeding duration, see Jayachandran, and Kuziemko, 2011). Palloni and Millman (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

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

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

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 1.4 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, 8 percent are less than 15 months apart, 13 percent less than 18 months, 33 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 and maternal health outcomes, it is important to understand if the desire for having sons affects the pace at which mothers try to conceive and give birth.

1.4 Empirical strategy

1.4.1 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[26]. 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, i.e. women with a first-born daughter tend to have more children, and their reproductive behavior may influence their health status and ultimately their survival. 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 the potential 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*

[26]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.

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.[27]

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

$$y_{i,t,r} = \beta_1(firstborngirl)_i + \gamma X_{i,t,r} + \gamma_t + \alpha_r + \epsilon_{i,t,r} \qquad (1.1)$$

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. $firstborngirl$ 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 husband) 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 husband), a wealth index, urban, ethnicity, and religion.[28,29] 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 re-

[27]The variable for contraceptive use is defined as equal to one if the woman reports she is currently using *any* contraceptive method, zero otherwise.

[28]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).

[29]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.

gressions 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 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 1.1 is estimated using OLS and a poisson model for count variables when the dependent variable is the # of children ever born, and using a probit model for the probability of desiring more children and 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 1.6) are likely to be taken simultaneously, the sex of children born after the first cannot be considered as random anymore. With this limitation in mind, it is worth noting that the effects should (and are found to) 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 the sex of 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 can be considered as random, the 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.[30]

1.4.2 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, not the woman as in the previous section).[31] 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.[32] Given that the NDHS includes information on the interval between each birth for

[30]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.

[31]I exclude mothers with twin births and those with more than 12 children.

[32]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.

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:

$$y_{i,j,t} = \beta_1 \ girl_{i,j,t-1} + \sum_k \gamma_k birthorder_{i,j,t-1,k} + \alpha_j + \epsilon_{i,j,t-1} \tag{1.2}$$

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 the succeeding birth interval is shorter than 24 or 15 months, the critical intervals below which child and maternal 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 regression 1.2 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.[33]. The prediction with son preference is that the succeeding birth interval at each birth should be shorter

[33]The percentages of women with at least three, four, five children ever born in this sample are the following: 70%, 55%, and 41.5%, respectively.

($\beta_1 < 0$) if a girl is born rather than a boy; given the 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.

1.5 Main results

1.5.1 Fertility results

Table 1.5 shows the OLS estimates of the effect of a first-born girl on the number of children ever born.[34] 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 larger in magnitude and significant at 5% 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.[35]

The rising coefficient of $firstborngirl$ across specifications in Table 1.5 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

[34]The estimated effect obtained using a poisson model for count variables is shown in appendix table 1.17 and it is consistent with the OLS estimates.

[35]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).

because of selective maternal mortality discussed in the descriptive section). Since education is negatively 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 she wants more children. Table 1.6 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 1.6. Women with first-born girl are 1.1 percentage points less likely to use any method of contraception. This is an important effect considering that a small fraction of women are currently using contraceptives.

Table 1.18 shows the estimated effects of the birth of two consecutive daughters on fertility and contraceptive use for the subsample of women with at least two children ever born. As expected, the percent effects are stronger than those found for first-born daughter.

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.

1.5.2 Birth spacing results

I turn now 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 1.7 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 birth of each successive girl 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 1.7 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 1.32 months less after the birth of any successive girl (this effect is significant at 5% 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 3.25 months, and this effect is significant at 1% level. Appendix table 1.19 shows additional results for the subsamples of women with first-born boy (column 2), first and second-born boy (column 4), and mixed gender composition (columns 5, 8, 9). 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 and 15 months which, as indicated in the medical literature, are the critical lengths below which child and maternal health are negatively affected, respectively. The results are shown in Table 1.8: column (1) and (3) show that, after the birth of any girl, mothers who had two girls as the first two born children are 2.1 percentage points more likely to wait less than 24 months, though not significantly so, and 1.5 percentage points more likely to wait less than 15 months (significant at 10 percent level).

In columns (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 only 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. Considering that about 32% of women wait less than 24 month in this subsample, the effect is important and equal to a 11.2% increase. Column (4) shows that the likelihood that the next birth is spaced less than 15 months apart increases by 1.8 percentage points (equivalent to a 22% increase given that on average 8% of births are spaced less than 15 months apart in this subsample) after the birth of a girl for mothers who already had two girls.

Figure 1.1 plots the distribution of the length of the interval for the subsample of women with two girls as the first two born children (and at least three children ever born) 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 and in 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, the distribution of the length of the interval after the birth of a girl lies to the left of the distribution of the interval after a boy is born. 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, with a p-value of 0.004). Appendix figure 1.8 shows that the distributions in the subsample of women who had two boys or mixed gender composition are not significantly different from each other. Similarly, figure 1.2 uses the subsample of women with three girls as the first three born children and shows that the whole distribution of the interval after each successive girl lies to the left of the distribution after each boy (and the two distributions are statistically different from each other as shown by the Kolmogorov-Smirnov test with p-value 0.047). 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.

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.[36]

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.[37]

Column 1 in Table 1.9 shows that a child born less than 24 months after the preceding child is 4.2 percentage points more likely to die (at age 2 or before) compared to a child born after a longer interval (the coefficient is significant at 1% level). It also shows that female children are less likely to die than males, which is consistent with descriptive evidence on mortality and with the medical literature documenting that males are biologically weaker than females at birth (Waldron, 1983). Column 2 includes the interaction term between 'short preceding interval' and the female dummy. Interestingly, a short birth interval has an additional positive and significant effect on the likelihood of dying for female children as opposed to males. This result, together with the results obtained in the previous section that birth spacing is shorter after the birth

[36] 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).

[37] The effects of short succeeding interval on the health status of the previous (older) child mainly acts through inadequate child care a interrupted breastfeeding.

of girls, suggests that son preference might be among the factors affecting child mortality.

1.6 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.

1.6.1 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 parents'* (Isiugo-Abanihe, 1985, p. 53). It is a widely accepted and practiced social institution in many parts of Sub-Saharan Africa.[38]

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 accord-

[38] 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).

ing 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 the sex composition of children. If non-biological girls (boys) are substitutes for biological daughters (sons), I should find symmetric responses for the imbalance with more girls or boys.[39] If instead there is preference for *own* biological son but not for *own* daughter, responses should not be symmetric.

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 mother 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).[40] The focus is on children (alive) age 6–14 (included), for which fostering is more common.[41]

[39] 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.

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

[41] 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

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)).[42] 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 includes households with no children in the relevant age group, which are the ones that typically receive more children in Nigeria.[43] 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

young) are likely to live elsewhere in a separate household.

[42]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 (2009).

[43]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.

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 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 1.20 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.[44] 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.[45] Moreover, female children are more involved in fostering than males. 24% percent of households have an excess of daughters or sons, while 51% are

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

[45]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).

balanced (including households with no children aged 6–14, which represent 40% of the sample). The bottom panel of Table 1.20 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.

Household fostering decisions, results

Table 1.10 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 (as the coefficients in columns 7 to 9 are all negative). 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 (see column 8 and 9), than the excess of sons[46]. 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 aged 6–14 I can focus on the effect of the sex 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

[46] A Wald test rejects the equality of the coefficients for the excess of sons and daughters separately in each of the regressions in columns 7 and 8, while accepts it in column 9

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 (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 1.11 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.[47] 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 (e.g., in fostering-in regressions, if there is excess of daughters, households should foster-in boys more in rural areas). Appendix table 1.21 shows that this is not the case, even though the asymmetries appear to be 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.

1.6.2 Marital outcomes

The anthropological and demographic evidence introduced in Section 1.2 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

[47] 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.

a polygamous relationship compared to women with sons.[48] 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.[49] 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 children).

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).[50] Female headship has often, but not always,

[48] 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.

[49] 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).

[50] 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

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 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 and living arrangements, which in turn affect children's and women's well-being.

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 rank one wives, 71% of rank two, 58% of rank three; while 92% for monogamous).[51]

I analyze the effect of first-born daughter on four outcomes: that the husband marries another woman after the first wife (the husband is polygynous), that he resides in the household with his wife, and that the woman is divorced or head of the household.

for Cote d'Ivoire.

[51] 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.

The empirical strategy is similar to that employed in the analysis of fertility outcomes illustrated in section 1.4.1.[52] 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. Table 1.12 shows the results. Column (1) shows that women with a first-born daughter are 1.3 percentage points (significant at 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

[52] 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.

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).[53] This evidence clearly suggests the preferences for the sex of children affects marriage outcomes.

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. Given that many women with a polygynous husband do not co-reside with him in this context, it is not possible to know the number and the sex composition of children born to every wife of the same husband for all the households in the sample. 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 in the dataset, and both dates of marriage of the first

[53]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.

and second wife are needed). Therefore, I end up with a sample of 1621 women who are first-rank wives of the head: these women were monogamous until the date of their husband's marriage with the second wife.[54] With this dataset, I compare monogamous women with women ranked first in a polygynous relationship. The fraction of polygynous women in this sample is 9.5% (all the other women in the sample are monogamous by construction), 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: first, the median number of children born before re-marriage is 2 (it's 3 after excluding women who did not have children); second, 19% of married women ranked first (305 out of 1621) did not have any child before re-marriage; and third, the median number of years between 1st and 2nd marriage is 7 years (only 3 years if the first wife did not bear children).

Turning to the results, column (1) in table 1.13 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 findings in the previous section.

[54] The sample restrictions also yield a very few women of higher order (rank three or higher) who are not considered in the empirical analysis).

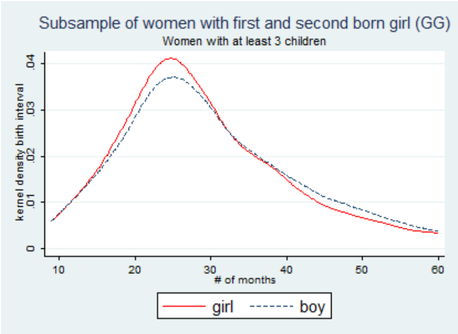
1.7 Conclusions

The economic literature has long studied the prevalence and the consequences of parental gender preference in developed and developing countries. In the latter context, most studies have focused on many Asian countries where the sex ratio at birth is abnormally high. As noted by Anderson and Ray (2010), 'normal' sex ratio (at birth) in Sub-Saharan countries might be one of the reasons 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 *not* at birth. Excess female mortality is in fact largest among women of adult and reproductive ages. This paper uses individual-level data for Nigeria and shows that mothers (or couples) make fertility decisions that seem to be motivated by the desire for giving birth to sons. Indeed, 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. 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.

The results of this paper suggest that son preference affects reproductive behavior and has potential implications for women's and children's well-being. High fertility and especially short birth spacing are in fact associated with poor maternal (and child) health outcomes.

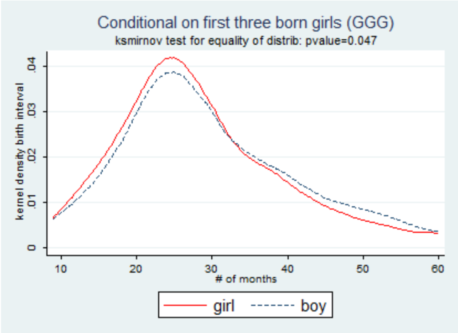
Figures

Figure 1.1: 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.

Figure 1.2: Distribution of length of birth interval after the birth of each girl or boy. First five births to each woman.



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

Tables

Table 1.1: Summary statistics

	mean	median	st. dev.	N
# children ever born	4.31	4	2.70	18426
wants another child	0.65	1	0.48	18314
contraceptive use	0.17	0	0.37	18426
polygynous husband	0.30	0	0.46	18346
female head	0.09	0	0.29	18426
husband living in	0.90	1	0.30	18307
age	31.62	30	8.39	18426
age first marriage	17.90	17	4.66	18426
age first birth	19.50	19	4.48	18426
woman eduyrs	5.03	5	5.36	18414
husband eduyrs	6.40	6	5.82	18055

2008 NDHS. Using sample weights. Sample of married women 15–49 in first union with at least one child ever born.

Table 1.2: Summary statistics, by the sex of the first born

	first-born girl	first-born boy	difference	st. error diff.	n
# children ever born	4.286	4.330	-0.044	(0.043)	18426
wants another child	0.657	0.642	0.015*	(0.008)	18314
contraceptive use	0.165	0.167	-0.002	(0.006)	18426
polygynous husband	0.297	0.297	0.000	(0.007)	18346
female head	0.095	0.086	0.009*	(0.005)	18426
husband living in	0.893	0.906	-0.013***	(0.005)	18307
age	31.587	31.648	-0.061	(0.133)	18426
age first marriage	17.992	17.820	0.172**	(0.077)	18426
age first birth	19.583	19.430	0.153**	(0.074)	18426
woman education yrs	5.177	4.913	0.264***	(0.088)	18414
partner education yrs	6.595	6.227	0.368***	(0.093)	18055
weight-for-height	-0.524	-0.554	0.030	(0.020)	17845
body mass index	23.137	23.032	0.105	(0.077)	17953

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.

Table 1.3: Summary statistics, by the sex of the first born and mother's age

	Women age 15-29					Women age 30-49				
	first girl	first boy	diff	se diff	n	first girl	first boy	diff	se diff	n
# children ever born	2.587	2.590	-0.003	(0.035)	8187	5.626	5.652	-0.026	(0.058)	10239
wants another child	0.874	0.856	0.018**	(0.008)	8142	0.485	0.479	0.006	(0.011)	10172
contracept. Use	0.124	0.127	-0.003	(0.009)	8187	0.197	0.197	0.000	(0.009)	10239
polygynous husband	0.245	0.256	-0.011	(0.010)	8156	0.339	0.328	0.011	(0.010)	10190
female head	0.072	0.082	-0.010	(0.006)	8187	0.113	0.089	0.024***	(0.007)	10239
husband living in	0.904	0.904	0.000	(0.007)	8132	0.885	0.908	-0.023***	(0.007)	10175
age	23.920	23.796	0.124	(0.082)	8187	37.635	37.616	0.019	(0.123)	10239
age first marriage	17.008	16.793	0.215**	(0.088)	8187	18.770	18.601	0.169	(0.115)	10239
age first birth	18.423	18.277	0.146*	(0.083)	8187	20.499	20.307	0.192*	(0.110)	10239
woman education yrs	4.733	4.694	0.039	(0.125)	8182	5.526	5.079	0.447***	(0.121)	10232
partner education yrs	6.425	6.189	0.236*	(0.135)	8005	6.729	6.256	0.473***	(0.129)	10050
weight-for-height	-0.581	-0.588	0.007	(0.026)	7920	-0.480	-0.528	0.048*	(0.029)	9925
body mass index	22.104	22.093	0.011	(0.098)	7977	23.950	23.744	0.206*	(0.110)	9976

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.

Table 1.4: 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.

Table 1.5: The effect of first-born girl on the number of children ever born.

	Age 15-49				30-49
dep var.	# children ever born				
	(1)	(2)	(3)	(4)	(5)
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)
woman eduyrs		-0.022***	-0.022***	-0.022***	-0.027***
		(0.004)	(0.004)	(0.004)	(0.006)
husband 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 husband'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.

Table 1.6: 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)
woman eduyrs	-0.003*	0.005***
	(0.001)	(0.001)
husband 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 husband'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.

Table 1.7: 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 women			first-born girl			girl, girl			girl, girl, girl		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
girl	0.300	0.139	0.015	-0.022	-0.171	-0.257	-1.021*	-1.316**	-1.682***	-2.724***	-3.252***	-3.232***
	(0.224)	(0.258)	(0.209)	(0.377)	(0.427)	(0.332)	(0.562)	(0.616)	(0.507)	(0.874)	(0.980)	(0.842)
girl*dead child		0.530	0.706*		0.550	0.276		1.198	0.619		2.120	0.936
		(0.491)	(0.396)		(0.689)	(0.564)		(1.034)	(0.883)		(1.568)	(1.317)
dead child		-4.928***	-7.459***		-4.993***	-7.450***		-5.650***	-7.514***		-6.694***	-7.204***
		(0.360)	(0.276)		(0.579)	(0.452)		(0.938)	(0.768)		(1.440)	(1.202)
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	51081	51081	51081	24082	24082	24082	11086	11086	11086	4808	4808	4808
R-squared	0.313	0.320	0.025	0.308	0.315	0.026	0.271	0.280	0.028	0.219	0.232	0.035

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.

Table 1.8: Probability of a short birth interval (shorter than 24 or 15 months), conditional on the first two born girls

	<24 months		<15 months	
Subsample	All children	first 4 children	All children	first 4 children
	(1)	(2)	(3)	(4)
girl	0.021	0.036*	0.015*	0.018*
	(0.015)	(0.022)	(0.008)	(0.011)
girl*dead child	-0.046	-0.069	-0.015	-0.015
	(0.033)	(0.048)	(0.023)	(0.032)
dead child	0.184***	0.193***	0.117***	0.109***
	(0.028)	(0.044)	(0.020)	(0.029)
Birth order dummies	Yes	Yes	Yes	Yes
Mother FEs	Yes	Yes	Yes	Yes
Observations	11086	8149	11086	8149
R-squared	0.244	0.319	0.276	0.373

Subsample of women with at least four children ever born. 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.

Table 1.9: The effect of short birth interval on child mortality

	(1)	(2)
short preceding interval	0.042***	0.034***
	(0.004)	(0.006)
female child	-0.007*	-0.012***
	0.004	0.003
short preceding interval*female		0.016*
		0.009
Birth order dummies	Yes	Yes
Mother FEs	Yes	Yes
Observations	51081	51081
R-squared	0.291	0.291

Subsample of women with at least three children ever born. 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.

Table 1.10: 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).

Table 1.11: 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).

Table 1.12: 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 marriage	-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)
woman 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)
husband eduyrs	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)
first child dead	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 husband (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 husband, 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 husband, 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 husband.

Table 1.13: 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 husband, 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.

Appendix

Figure 1.3: Male births (out of total births), by birth year of the child, NDHS 2008

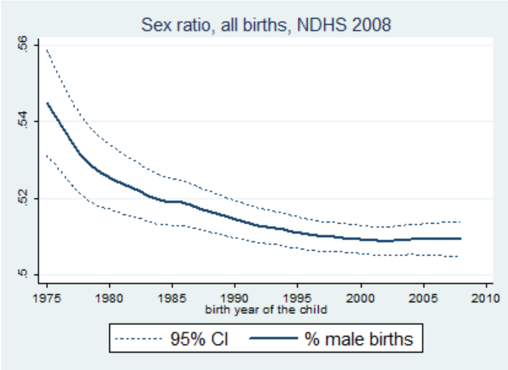


Figure 1.4: Male births (out of total births), by birth year of the child, by NDHS survey

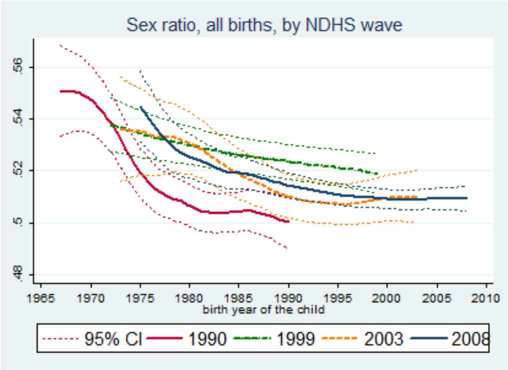


Figure 1.5: Share of women with first-born girl, by mother's age (at the time of the survey), Pooled NDHS 1990, 1999, 2003, and 2008

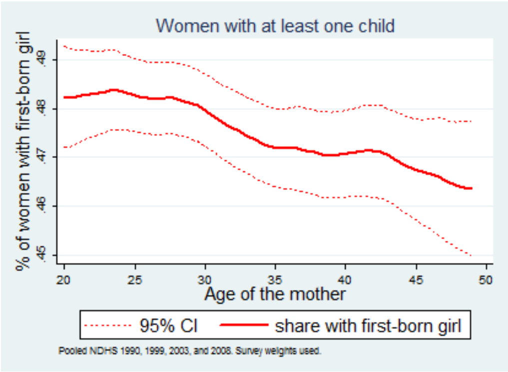


Figure 1.6: Share of women with first-born girl and higher-order female birth, by mother's age (at the time of the survey), Pooled NDHS 1990, 1999, 2003, and 2008

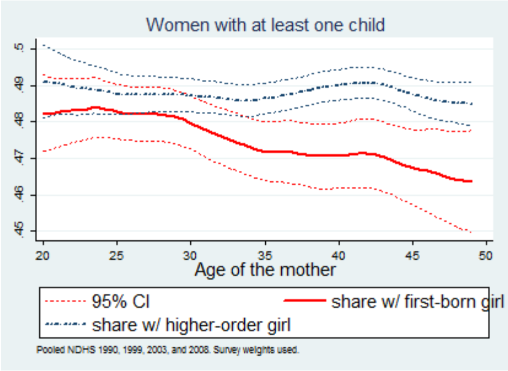


Figure 1.7: Share of women with first-born girl, by mother's age (at the time of the survey) and education level, Pooled NDHS 1990, 1999, 2003, and 2008

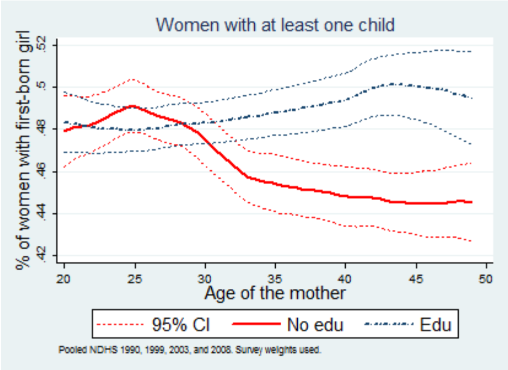
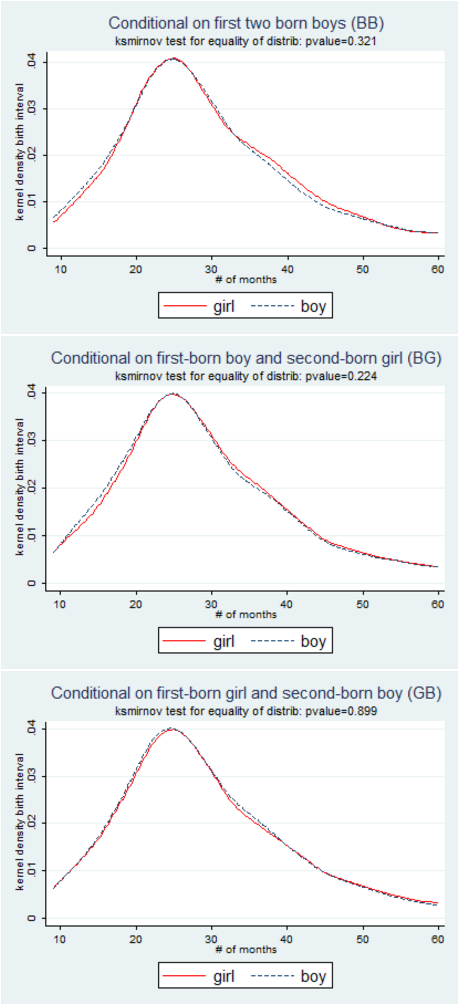


Figure 1.8: Distribution of length of birth interval after the birth of each girl or boy. First 4 children (or less) ever born to a woman



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

Table 1.14: 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
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	7	45	18426
age first birth	19.50	19	4.48	9	44	18426
husband 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
weight-for-height	-0.54	-0.67	1.21	-4	5.92	17845
body mass index	23.08	22.14	4.64	12.1	59.7	17953

2008 NDHS. Using sample weights.

Table 1.15: 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	7	45	23163
age first birth	19.36	19	4.44	9	44	23751
husband 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
weight-for-height	-0.55	-0.68	1.21	-4	5.92	22986
body mass index	23.07	22.14	4.66	12.1	59.81	23115

2008 NDHS. Using sample weights.

Table 1.16: Age at first birth and sex of the first-born child

dep. var:	first-born girl
age at first birth	0.001
	(0.001)
Observations	18189
Pseudo R-squared	0.003

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 six region dummies, urban dummy, birth year fixed effects, ethnicity, and religion. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 1.17: The effect of first-born girl on the number of children ever born, poisson estimates.

dep. var:	# children ever born
first-born girl	0.017***
	(0.006)
Observations	17587

Poisson 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, age at first birth, age at first marriage, interaction between first-born girl and a dummy for whether the first-born child is dead, mother and husband's age (and square), mother and husband's education, 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.

Table 1.18: Sex of the first two-born children, realized (and desired) fertility and the use of contraception

dep. var:	# children ever	=1 wants another child	=1 currently using contrac.
	(1)	(2)	(3)
girl,girl	0.127***	0.051***	-0.015*
	(0.039)	(0.015)	(0.008)
one boy one girl	0.038	-0.003	-0.002
	(0.032)	(0.012)	(0.007)
Observations	15065	14975	15027
R-squared (or Pseudo)	0.655	0.262	0.230
Percent effect	2.6	8.5	8.8
Mean boy,boy	4.88	0.60	0.17

OLS estimates in column (1). Probit estimates in columns (2)-(3), 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 husband's age (and square), six region dummies, birth year fixed effects, ethnic and region-specific time trends, urban dummy, ethnicity, religion, and a wealth index. The # children ever born is included in columns (2)-(3). ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 1.19: Length of the birth interval (# months), conditional on the gender of earlier-born children

	firstG	firstB	G,G	B,B	mixed	G,G,G	B,B,B	two boys one girl	two girls one boy
Women with:	3+ children		4+ children			5+ children			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
girl	-0.171	-0.039	-1.316**	0.437	0.167	-3.252***	1.015	0.267	-0.391
	(0.427)	(0.380)	(0.616)	(0.568)	(0.364)	(0.980)	(0.904)	(0.425)	(0.438)
girl*dead child	0.550	0.531	1.198	-1.612	1.036	2.120	-1.919	0.675	0.840
	(0.689)	(0.733)	(1.034)	(1.001)	(0.716)	(1.568)	(1.557)	(0.861)	(0.824)
dead child	-4.993***	-4.849***	-5.650***	-4.348***	-4.995***	-6.694***	-3.870***	-4.958***	-5.328***
	(0.579)	(0.458)	(0.938)	(0.574)	(0.520)	(1.440)	(0.762)	(0.551)	(0.715)
B.order dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24082	26999	11086	12920	21685	4808	5748	14750	13278
R-squared	0.315	0.324	0.280	0.277	0.268	0.232	0.250	0.220	0.237

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.

Table 1.20: 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 1.21: 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
Pseudo R-squared	0.04	0.04	0.05	0.07	0.06	0.06

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.

Bibliography

[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, pp. 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, Vol. 44, No. 4, pp. 976–997.

[4] Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2011. "Fertility and the Plough." *American Economic Review: Papers & Proceedings 2011*, Vol. 101, No. 3, pp- 499-503.

[5] Almond, Douglas, Lena Edlund, and Kevin Milligan. 2009. "Son Preference and the Persistence of Culture: Evidence from Asian Immigrants to Canada." NBER Working Paper No. 15391.

[6] 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.

[7] Anderson, Siwan and Debraj Ray. 2010. "Missing Women: Age and Disease." *Review of Economic Studies*, Vol. 77, No. 4, pp. 1262-1300.

[8] —. 2012. "The Age Distribution of Missing Women in India." Working Paper, March 2012.

[9] 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.

[10] 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.

[11] Barcellos, Silvia H., Leandro Carvalho, and Adriana Lleras-Muney. 2012. "Child Gender And Parental Investments In India: Are Boys And Girls Treated Differently?" NBER working paper No. 17781.

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

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

[14] Bergstrom, Theodore C. 1997. "A Survey of Theories of the Family." *Handbook of Population and Family Economics*, Vol. 1, No. 1, pp. 21–80.

[15] Bledsoe, Caroline. 1990. "Transformations in Sub-Saharan African Marriage and Fertility." *Annals of the American Academy of Political and Social Science*, Vol. 510, World Population: Approaching the Year 2000 (Jul., 1990), pp. 115–125.

[16] 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.

[17] Boserup, Ester. 1970. "Woman's Role in Economic Development." London: Allen and Unwin.

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

[19] 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.

[20] 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*, Vol. 5, No. 12.

[21] Cameron, Stephen V., and James J. Heckman. 1998. "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males." *Journal of Political Economy*, Vol. 106, No. 2, pp. 262-333.

[22] Carranza, Eliana. 2012. "Soil Endowments, Production Technologies and Missing Women in India." World Bank Policy Research Working Paper No. 5974.

[23] Chowdhury, Mridul K., and Radheshyam Bairagi. 1990. "Son Preference and Fertility in Bangladesh." *Population and Development Review*, Vol. 16, No. 4, pp. 749–757.

[24] Clark, Shelley. 2000. "Son Preference and Sex Composition of Children: Evidence from India." *Demography*, Vol. 37, No. 1, pp. 95–108.

[25] 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*, Vol. 196, No. 4, pp. 297–308.

[26] Conde-Agudelo, A., and Belizan, J.M. 2000. "Maternal Morbidity and Mortality Associated with Interpregnancy Interval: Cross Sectional Study" *British Medical Journal*, Vol. 321, No. 7271, pp. 1255–1259.

[27] Dahl, Gordon B., and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*, Vol. 75, No. 4, pp. 1085–1120.

[28] Das Gupta, Monica, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, and Bae Hwa-Ok. 2003. "Why is Son preference so persistent in East and South Asia? a cross-country study of China, India and the Republic of Korea." *Journal of Development Studies*, Vol. 40, No. 2, pp. 153–187.

[29] Dreze, Jean, and Mamta Murthi. 2001. "Fertility, Education, and Development: Evidence from India." *Population and Development Review*, Vol. 27, No. 1, pp. 33–63.

[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*, Vol. 14, No. 4, pp. 571–594.

[31] Fayehun O.A, Omololu O.O., and Isiugo-Abanihe Uche C. 2011. "Sex of Preceding Child and Birth Spacing Among Nigerian Eth-

nic Groups." *African Journal of Reproductive Health*, Vol. 15, No. 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] Garg, Ashish, and Jonathan Morduch. 1998. "Sibling Rivalry and the Gender Gap: Evidence from Child Health Outcomes in Ghana." *Journal of Population Economics*, Vol. 11, No. 4, pp. 471–493.

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

[35] Horrell, Sara, and Pramila Krishnan. 2007. "Poverty and Productivity in Female-headed Households in Zimbabwe." *The Journal of Development Studies*, Vol. 43, No. 88, pp. 1351–1380.

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

[37] —. 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, No. 6, pp. 913-933.

[43] Lloyd, Cynthia B., and Sonalde Desai. 1992. "Children's 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] Palloni, Alberto, and Sara Millman. 1986. "Effects of Inter-Birth Intervals and Breastfeeding on Infant and Early Childhood Mortality." *Population Studies*, Vol. 40, No. 2, pp. 215–236.

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

[51] 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.

[52] Smith, Herbert L. 1994. "Nonreporting of Births or Nonreporting of Pregnancies? Some Evidence from Four Rural Counties in North China." *Demography*, Vol. 31, No. 3, pp. 481–486.

[53] Smith, Daniel Jordan. 2004. "Contradictions in Nigeria's Fertility Transition: The Burdens and Benefits of Having People." *Population and Development Review*, Vol. 30, No. 2, pp. 221–238.

[54] Stansfield, W. D., and M. A. Carlton. 2007. "Human sex ratios and sex distribution in sibships of size 2." *Human Biology*, Vol. 79, No. 2, pp. 255-260.

[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, No. 6, 1983, pp. 321-333.

[57] Wall, Lewis L. 1988. "Dead Mothers and Injured Wives: The Social Context of Maternal Morbidity and Mortality Among the Hausa of Northern Nigeria" *Studies in Family Planning*, Vol. 29, No. 4, pp. 341–359.

[58] 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.

[59] World Bank. 2011. "World Development Report 2012. Gender Equality and Development." World Bank, Washington D.C.

[60] —. 2012. "World Development Indicators" World Bank, Washington D.C.

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

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

[63] 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.

Chapter 2

Why Are Adult Women Missing? Son Preference and Maternal Survival in India

This paper explores the phenomenon of missing women of reproductive and adult ages in India. By using individual-level data, I compare the age structure and health indicators of women by the sex of their first-born and uncover several new findings. First, among mothers with at least one child, the share of those with first-born daughter decreases with women's age (observed at the time of the survey). Second, women with first-born daughter are significantly more likely to be anemic when young while they show better anthropometric indicators and no different incidence of anemia as they get older. Interestingly, differences in anemia prevalence arise only some time after the birth of the first child. Third, women with first-born daughter have significantly higher (realized and desired) fertility and are less likely to use contraceptives. Moreover, a female birth is associated with shorter spacing which might put mothers at higher risk of dying for complications during or after childbirth and maternal depletion. I argue that all these findings are consistent with a selection effect in which healthier women of higher socioeconomic status are more likely to survive the birth of daughters. To my knowledge, this is the first paper arguing that women's excess mortality can (at least partly) be an unintended consequence of son preference.

2.1 Introduction

There is an extensive literature on parental preferences over a child's gender. Sen's 1990 article documented high sex ratios at birth (the ratio of males to females who are born) in many South and East Asian countries: he argued that women might be missing because of sex-selective abortion, infanticide and neglect of female children. Sex ratios for children aged 0–6 have been steadily increasing since 1961 in India, and evidence of female foeticide has been widely documented (Jha *et al*, 2006).[1]. The diffusion of prenatal sex determination technologies starting from the end of the 1980s has been found to be associated with the increasing sex ratios and preferential prenatal treatment for boys (Bhalotra and Cochrane, 2010; Bharadwaj and Nelson, 2012).[2] Other authors have studied fertility behavior and shown that subsequent fertility is higher for women who had girls among earlier-born children (son-preferring fertility stopping rules) (Chowdhury and Bairagi, 1990; Clark, 2000; Dreze and Murthi, 2001; Arnold *et al.*, 2002). The 'try until you have a son' fertility rule may also 'passively' lead to lower amount of resources being allocated to female children within the household. This may happen either as a direct consequence of the fact that the weaning time for girls will be shorter than that for boys as parents want to try again for a son, or because girls tend be born in larger households (e.g., see Jayachandran and Kuziemko, 2011, for gender disparities in breastfeeding duration, and Jensen, 2003, on the

[1] According to decennial Census data, the number of girls per 1000 boys aged 0–6 was 976 in 1961 and reached 914 (its lowest level) in 2011.

[2] Abortion was legalized in 1971 in India. In 1994, the Government of India passed the Prenatal Diagnostic Techniques Regulations and Misuse Act (PNDT Act) to make the use of ultrasound or amniocentesis for the purpose of sex determination illegal. However, there is evidence that this Law is often ignored and not enforced (Arnold *et al.*, 2002; George, 2002).

gender differential in education).[3]

Previous studies mostly focused on gender bias at birth or young ages in explaining the phenomenon of missing women. A notable exception is Anderson and Ray (2010), who provide a decomposition of the excess female deaths by age and causes of death using aggregate data for different parts of the world. While they still find evidence of severe bias against young (and unborn) female children, they show that the majority of missing women in India are of reproductive and adults ages.[4] Recently, Anderson and Ray (2012) replicate this exercise for India. They confirm that most missing women are found in adulthood and add that there is 'significant state-wise variation in the distribution of missing women across the age groups' (p. 3). Moreover, they observe that 'if we compare maternal mortality rates and the percentage of the female population that is missing at reproductive ages across the states, we do indeed see a positive correlation' (p. 15). The World Development Report (WDR) 2012 corroborates these findings and argues that the 'female disadvantage in mortality during the reproductive ages is in part driven by the risk of death in pregnancy and childbirth and associated long-term disabilities' (World Bank, 2011, p. 78).[5]

[3]Barcellos *et al.* (2012) find that, accounting for the effect of son-preferring stopping rules on family size, boys still receive more parental investment (including childcare time, breastfeeding, vaccinations, and vitamin supplementation) than girls.

[4]Anderson and Ray (2010) find that the sources of excess female mortality among women of reproductive ages (15–44) in India are mainly maternal mortality and injuries. Among among infants and adolescents (age 0–14) respiratory and infectious diseases are the main causes of excess mortality, while among older women (age 45 and older) the main cause is cardiovascular diseases.

[5]One of every 140 women die from causes related to childbirth in India (estimates for 2008 from the WHO/UNICEF/UNFPA/World Bank, 2010). The adult lifetime risk of maternal death is defined as the probability of dying from a maternal cause during a womans reproductive lifespan.

While previous recent studies have documented the excess of female mortality among adult women, this paper attempts to make a step forward in understanding the causes of excess female mortality among women of reproductive and adult ages. It makes at least two contributions to the existing literature. First, while Anderson and Ray (2010, 2012) use aggregate mortality data and find that, contrary to previously-held wisdom, many women are missing in adulthood, this paper uses individual-level data for India and explores whether and how fertility behavior and preferences can partly explain this pattern. Second, it compares health outcomes across different age groups to uncover a selection effect in which healthier women of higher socioeconomic status might be more likely to survive the birth of daughters. The main novelty of this paper is that it provides an original explanation for the phenomenon of missing adult women that essentially coincides with the one that has been brought forward to explain overt discrimination at younger ages, namely parental preferences for sons.

I pool three rounds of the India National Family Health Survey (NFHS) and find several interesting new patterns. First, I provide non-parametric evidence that, among women with at least one child ever born, the share of women with first-born daughter decreases with mother's age (reported at the time of the survey). To confirm this result, I also estimate an ordered logit model in which the dependent variable is a categorical variable for each age-group to show that the conditional age distribution shifts to the left for women with first-born girl (compared to first-born boys). Second, women with first-born daughter are 1.2 percentage points more likely to be anemic when young (especially under age 30) while they show significantly higher anthropometric indicators (weight-for-height and body mass index) and no different incidence of anemia as they get older. Interestingly, differences in anemia prevalence arise only some time after the birth of the first child. This suggests that women are similar at the time of first-birth, thus ruling out other interpretations based on biological differences among women with daughters or sons. This results are obtained after controlling for age at first birth, height (a proxy for health status before the first pregnancy), pregnancy and breastfeeding status, and other observable women's characteristics. Third, compared to women with first-born son, women with first-born daughter have 0.28 more children, are 7.3 percentage points less likely to use contraceptives (and in particular less likely to be sterilized), 1.3 more likely to have had a terminated pregnancy, and 5.1 percentage points more likely to desire more children. I focus on the sex of the first-born child to maintain a causal interpretation of the results. 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. Lastly, using mother fixed effects to control for unobserved mother's heterogeneity, I estimate the effect of the sex of the child on birth spacing. I find that the birth of a girl increases the probability that the mother waits less than 24 (15) months by 2 (0.8) percentage points (equivalent to about a 6 and 9 percent effect, respectively).

My hypothesis is that the patterns identified in the data are consistent with selective mortality through son-preferring fertility stopping behavior. In particular, as it is shown in the empirical part, the birth of a daughter is associated with higher fertility and closely spaced births. This reproductive behavior may lead to higher risk of death for complications during or after childbirth and maternal depletion. The link between son-preferring fertility behavior and maternal mortality is supported by medical evidence. Specifically, a high number of pregnancies is associated with higher lifetime risk of death due to pregnancy and short birth intervals are associated with poor child and maternal health outcomes. This is because close birth spacing does not allow a woman to regain her physical strength and nutrients required to have a successful following pregnancy. This fertility behavior is strongly associated with anemia, which in turn is an underlying cause of maternal mortality. I argue that a *selection effect* in which women of higher socioeconomic status are more likely to survive the birth of daughters can in part explain why adult women with first-born daughter are fewer and better-off in terms of nutritional indicators and other observable characteristics.

The remainder of the paper is organized as follows: Section 2.2 provides some background information on the prevalence of maternal mortality and anemia in India; Section 2.3 describes the data; Section 2.4 analyzes nonparametric evidence; Section 2.5 introduces the empirical methodology used to analyze fertility outcomes, birth spacing, health outcomes and describes the results; Section 2.6 discusses three alternative interpretations of the results; and Section 2.7 concludes.

2.2 Fertility behavior and maternal health outcomes

A high number of pregnancies is associated with higher lifetime risk of death due to pregnancy (WHO, 2008). Short intervals between births have been found to be related to poor child health and maternal outcomes.[6] Medical research has shown that short birth intervals (<24 months) are associated with poor child and mater-

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

nal health outcomes (Setty-Venugopal and Upadhyay, 2002).[7] For example, 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).[8]

The World Bank (2011) observes that maternal mortality and morbidity related to childbirth are one of the two mechanisms driving excess female mortality during the reproductive years.[9] The WHO estimates that 63000 maternal deaths occurred in India (out of 358000 worldwide) in 2008, the highest number in the world. India is making progress in the effort to improve maternal health (the fifth MDG) and the maternal mortality ratio went from 570 in 1990 to 230 in 2008.[10] Maternal mortality is difficult to measure and likely to go underestimated. It is defined by the WHO as the 'death of a woman while pregnant or within 42 days of termination

[7]With reference to child health, 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). Also, there are effects on health of older children (e.g., breastfeeding duration, see Jayachandran, and Kuziemko, 2011). Palloni and Millman (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.

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

[9]The other mechanism, HIV/AIDS, is more pertinent for the Africa region.

[10]The maternal mortality ratio is the number of maternal deaths during a given time period per 100 000 live births during the same time-period (WHO, 2008).

of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.' (WHO, 2008). This definition only refers to complications occurring at the time of birth and within 42 days of delivery. However, maternal mortality 'implies not only death during childbirth but also concurrent morbidities brought on by the experience of pregnancy and childbirth.' (WDR, p. 128). These morbidities include, for example, anemia, malaria, hepatitis, tuberculosis, cardiovascular disease, and obstetric fistula.

Anemia is pervasive in India, affecting more than 55% of women age 15–49 (NFHS-3). It usually results from a nutritional deficiency of iron, folate, vitamin B12, or some other nutrients (iron-deficiency anaemia). It is more prevalent during pregnancy and among breast-feeding mothers, when the nutritional requirements increase substantially. Iron-deficiency anemia is an important cause of maternal morbidity and, when severe, mortality in India (World Bank, 1996, p.31).

In the WDR, the World Bank (2011) observes that 'maternal mortality is fundamentally different from excess female mortality at other ages in that, to reduce it, societies must focus on an intrinsically female condition and specifically on improving the maternal health care system.' (p. 128). In India, many women do not receive adequate maternal health care.[11]

[11] The NFHS-3 reports that only 39% of births took place in health facilities and 58% of women did not receive any postnatal check-up after their most recent birth.

2.3 Data

The data used in this paper are from the India National Family Health Survey (NFHS), which contains individual-level information on birth histories (including children who have died), birth intervals, health and anthropometric indicators (for children and their mothers), and other variables. These are repeated cross-sectional survey data for a representative sample of households in India. I pool the three available survey rounds conducted in 1992/3, 1998/9, 2005/6. Pooling different survey waves allows to focus on age-effects rather that birth cohort-effects. The women covered in the NFHS are aged 15–49, and I consider the sample of all those who had at least one child ever born.

The total fertility rate in the NFHS-3 is 2.7, down to 2.9 in the NFHS-2. This means that, on average, an Indian woman will give birth to 2.7 children by the end of her childbearing years. Knowledge of contraception is nearly universal in India: 98% of women and 99% of men age 15-49 know one or more methods of contraception and contraceptive prevalence rate for currently married women in India is 56%, up from 48% in NFHS-2.[12]

Using the NFHS-3 birth history data, the sex ratio at birth (here defined as the fraction of male births) is 0.52 for all births.[13]

[12]Female sterilization, with a prevalence of 37%, accounts for 66% of all contraceptive use, down from 71% of all contraceptive use at the time of NFHS-2. The highest adoption rate of female sterilization, at 67%, is among women with three children who have two sons. The most common spacing methods are condoms and the rhythm method, each used by 5% of currently married women (India NFHS-3).

[13]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

Figure 2.6 and figure 2.7 show the sex ratios (at birth) by the birth year of the child in the pooled dataset and by survey year, respectively. Pooling different survey rounds 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 a particular period.[14] The most recent survey (NFHS-3) also shows an increasing sex ratio at birth from the end of the 1980s onward, which corresponds to the timing of the diffusion of ultrasound technologies and is consistent with evidence in Census data. Finally, it is important to note that the decreasing trend is specific to first births, while no similar pattern is found for higher birth-order births (figure 2.8). This is consistent with the association between sex of the first-born, reproductive behavior, and maternal survival. In the next sections, I'll explore why there are systematically higher sex ratios for births (especially at birth order one) that occurred back in the past from the time of the survey.

Table 2.1 reports summary statistics for the pooled sample of women age 15–49 with at least one child ever born, by the sex of the first-born. The upper panel of table 2.1 reports socio-economic and demographic variables, while the bottom panel shows health and anthropometric indicators.[15] The table shows that some charac-

are similar, both around 0.514. Several biological, environmental, and genetic factors can partly explain the variation across different areas.

[14]Using a different dataset for India (with retrospective fertility histories as in the NFHS), 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.

[15]The Body Mass Index (BMI) is a measure of women's nutritional status. It is the ratio of the weight in kilograms to the square of the height in meters. Undernutrition in women is defined as a BMI of less than 18.5 and is classified as chronic energy deficiency, and a BMI of 25.0 or above usually indicates obesity. Weight for Height is another indicator of women's nutritional status and it's

teristics significantly differ between the two subsamples of women. In particular, it shows that women with first-born girls are more educated, younger, married and had the first child later in life, had more children, are more likely to be the head of the household, and more likely to be anemic. While the higher number of children ever born is a direct consequence of the widespread use of son-preferring fertility rules in India, the differences in the other characteristics needs further investigation.

Table 2.2 reports the same summary statistics as in table 2.1 for the subsamples of women aged 15–30 and 31–49. The leftmost upper panel shows that younger women with first-born girl or boy appear to be quite similar in terms of observable characteristics. Women with first-born girl aged 15–30 still have significantly more children (because of son-preferring fertility rules), but their education level and age are more similar to women with first-born boy. The rightmost upper panel shows a very different picture: most of the differences reported in table 2.1 are to be found among women aged 31–49. Older women with first-born girl seem to be 'better off' in terms of their own and husband's education, they married and had first child later, and are more likely to be head of the household.

The bottom panel shows the differences for the health indicators. The leftmost panel shows significant differences between younger women in the two groups: younger women with first-born daughter are significantly more likely to be anemic and are also characterized by lower weight-for-height, while the rightmost panel shows that these differences disappear and become positive for older women.

---

defined in terms of standard deviations from the reference median based on the DHS reference standard.

This information is consistent with a selection effect through son-preferring fertility behavior. In particular, a female birth is associated with higher fertility and shorter spacing (as it will be shown later), thus putting mothers at higher risk of dying because of maternal depletion. Younger women in their reproductive years might physically suffer from having children shortly spaced, and this is reflected in the incidence of anemia and anthropometric indicators. Among women with first-born girl, the ones who survive to older ages are probably the ones who are healthier and from a higher socio-economic background and the descriptive evidence above described seems to reflect this fact. More evidence in support of this channel will be shown in the next sections.

Birth spacing In the India NFHS pooled dataset, of all births reported by each woman with at least two children ever born, 10 percent are spaced less than 15 months apart, 16 percent less than 18 months, 35 percent less than two years. The mean birth interval is 31.8 months, and the median is 27.[16] The birth interval increases with birth order (for given # of children as the preferences for family size affects the interval at each birth order). The sex of the first child seems to be associated with the interval between the first and second child ever born (as it will be investigated further in the following sections): it's 32.03 months after the birth of a girl, and 32.88 after a boy. The survival status of the child who opens the birth interval also matters: the median interval is 28 months if previous child is alive, 23 months if dead.

[16]To have at least one birth interval for each woman, it's necessary to consider the sample of women with at least two children ever born

2.4 Nonparametric evidence

Figure 2.1 reports non-parametric evidence (using Kernel-weighted local polynomial smoothing) of the share of women with first-born girl, by the age of the mother observed at the time of the survey. The confidence intervals at the 95% level of significance are also reported. The figure shows that the sex ratio for first births is in the biological range for women age 15–30, while the share of women with first-born girl (the inverse of the sex ratio, defined as the ratio of male to female children born) shows a steep decreasing trend for women older than 30 (below the biological range). This is equivalent to saying that, among women with at least one child ever born, the share of those who had a first-born girl (as opposed to a first-born boy) is lower among older women, especially older than 30. By pooling surveys for different years, it is possible to focus on age-effects rather than cohort of birth effects.

Appendix figure 2.9 shows that the decreasing sex ratio is found only among women with zero years of education, while the sex ratio is stable for women with at least one year of education. Also, figure 2.10 shows that the decreasing trend is found for women in rural areas, while no clear pattern is visible for those in urban areas. Since education can be considered a proxy for socioeconomic status and that urban households have better access to health facilities, this suggests an intuitive result: women who are worse-off and with poorer access to health services are the ones who are most likely to suffer from causes related to childbirth (aggravated by son-preferring fertility stopping rules) and thus most likely to die. Figure 2.11 shows the non-parametric graph for the subsample of women with at least two children ever born. As expected, the share of women with two boys (girls) increases (decreases) with the mother's age, while the shares of women with mixed sex composition (boy-girl or girl-boy) lies always in the middle.

Next, I turn to the health indicators. The upper panel of figure 2.2 shows non-parametric evidence of the incidence of anemia (the % of women who are anemic) by the sex of the first-born and mother's age. It is first worth noting that the overall pattern reflects the common prevalence of anemia; higher during adolescence and during the reproductive years, decreasing and stabilizing afterwards. The incidence of anemia is significantly higher for women with first-born girl up to age 30, while there is no difference for older women. The bottom panel of figure 2.2 shows the non-parametric graph for the incidence of anemia among women with at least two children ever born. As it would be expected in a context where son-preferring fertility stopping rules are commonly practiced, the incidence of anemia for young women with two daughters is higher than that for women who had only one daughter as first-born (shown in the upper panel), and it is significantly higher than the prevalence of anemia for women with two boys. Intuitively, this results might be the consequence of the fact that women with girls among earlier-born children tend to be pregnant or breastfeeding more often during their reproductive years than women who had boys. I replicate the same non-parametric graphs using the subsamples of women who are not currently pregnant or breastfeeding. Appendix figure 2.12 shows that the patterns are confirmed also when using these subsamples. This suggests that the differences found in the prevalence of anemia between women with first-born daughter and son might be driven mostly by birth spacing or other factors.

Figure 2.3 shows the non-parametric graph for weight-for-height.[17]

[17] Weight-for-Height standard deviations from the reference median based on

Both the upper and bottom panels (for women with at least one or two children, respectively) show a similar pattern. Overall, weight-for-height increases for women older than 40, reflecting the fact that women tend to gain weight as they get older. There are no differences in weight-for-height for women below age 40 (even though the bottom panel shows a slightly higher weight-for-height for women with two sons). Among older women, those with daughters among first-born children show better anthropometric outcomes (similar results obtained for the body mass index in appendix figure 2.13)

Lastly, figure 2.4 reports nonparametric evidence for height by the sex of the first-born and mother's age. Height reflects nutritional status in childhood and therefore should be affected by fertility behavior to a lower extent than other nutritional indicators.[18] As expected, the figure shows that women's height does not significantly differ by the sex of their first-born child.

2.5 Empirical strategy and results

To understand whether son preference may lead to selective maternal mortality through son-preferring fertility behavior, I test how the sex of the first-born affects reproductive behavior. In addition to confirming findings from the demographic literature documenting son-preferring reproductive behavior in India, I also provide estimates of the effects on child spacing that include mother fixed effects. This allows to control for unobserved mother's heterogeneity that may affect the estimates. To my knowledge, this approach has not been previously used. In this section I first describe the

the DHS reference standard.

[18] However, evidence suggests that weight and height can be affected if pregnancies occur before they have completed their adolescence growth spurt (World Bank, 1996).

empirical strategy used to estimate the effects of the sex of the first-born on fertility behavior, fertility preferences, birth spacing, and health outcomes, and then discuss the results.

2.5.1 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[19]. 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 (i.e., 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

[19]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.

desire to have more children (which may differ from *realized* fertility), I construct a dummy variable based on the 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 (and separately, of sterilization that is the most common contraceptive method in India), and the probability to have had a terminated pregnancy differs for women with a first-born daughter and first-born son.

I estimate the following regression (for women with one child or more):

$$y_{i,s,b,r} = \beta_1(firstborngirl)_i + \gamma X_{i,r,b,s} + \alpha_s + \mu_b + \delta_s + \epsilon_{i,s,b,r} \tag{2.1}$$

with mother i, resident in state s, born in year b, surveyed in round r. $y_{i,s,b,r}$ 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, of being sterilized, and of having had an abortion. $firstborngirl$ indicates whether the first child ever born is a girl (as opposed to a boy); $X_{i,s,b,r}$ is the set of covariates including: age of the mother (and age squared), age of the mother's partner, 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 dummy, caste, and religion.[20,21]

[20]The caste and religion variables are from the woman dataset. Caste include Scheduled Caste, and Scheduled Tribe (other backward classes (OBC) and other castes are the omitted category). Religion include Hindu, Sikh, Muslims, Christian (Buddhist, Jain, and other religions are the omitted category).

[21]The wealth index is a principal component index of services and durable goods owned by the household, including electricity, radio, television, refriger-

α_s, γ_b, δ_r are state, cohort of birth, and survey round 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 $firstborngirl$). 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 2.1 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, of using a contraceptive method, of being sterilized, and of having had an abortion.

If women practice son-preferring fertility stopping behavior, I expect to find $\beta_1 > 0$ for the regressions for realized (and desired) fertility, and abortion, while $\beta_1 < 0$ for the use of contraceptives and sterilization. The analysis mainly focuses on the sex of the first-born to maintain a causal interpretation (Dahl and Moretti, 2008).

Table 2.3 reports the results. Column (2) shows that women with first-born girl have 0.28 more children ever born (significant at 1% level), which implies a percent effect of 8.7 (with respect to women with first-born son). The probabilities of using contraceptives and of being sterilized decrease by 7.3 and 8.2 percentage points for women who had a first-born girl, respectively. Moreover, an abortion is 1.3 percent more likely for women with first-born girl. Finally, women with first-born girl are 5.1 percentage points more likely to report that they want more children. All these results suggest that the sex of the first-born strongly affects subsequent *realized* and *desired* fertility and are consistent with the idea that women intentionally continue bearing children until they have the

ator, bicycle, motorcycle, and car.

desired number of sons. In particular, the effect for the use of contraceptives and abortion points to a conscious decision.

2.5.2 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.[22] 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. Given that the NFHS 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:

$$y_{i,j,t} = \beta_1 \; girl_{i,j,t-1} + \sum_k \gamma_k birthorder_{i,j,k} + \alpha_j + \epsilon_{i,j,t} \qquad (2.2)$$

with child i, mother j, 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), or a dummy =1 if interval <24 or 15 months. α_j mother fixed effects. $birthorder_{i,j,k}$ set of k dummies for birth order. Identification with FEs requires to consider mothers with at least 3 children ever born. I also control for the survival

[22] I exclude births of order higher than 10.

status of the previous child (the one who opens the interval) and its interaction with her/his sex (*girl*). The succeeding birth interval at each birth should be shorter ($\beta_1 < 0$) if a girl is born rather than a boy.

A potential source of bias is represented by sex-selective abortion. There is evidence that female fetuses might be selectively aborted and that this is more common among second or higher order children (Bhalotra and Cochrane, 2010). If this is the case, abortion would be more likely after the birth of a first-born girl rather than a boy. Since the NFHS does not contain detailed information on abortion (which would anyway be highly subject to under-reporting) it is not possible to directly control for this. However, it is important to note that this would lead to an underestimation of the (negative) effect of first-born girl on the length of the birth interval as more time will be needed to have a following completed pregnancy after an abortion.

The first two columns of table 2.4 show the estimated effect of having a daughter on the length of the birth interval (in # of months) without including the mother fixed effects. The coefficients are -1.03 and -0.72 (both highly statistically significant at 1% level) meaning that, compared to mothers who had a boy, those who have a girl wait on average 0.7-1 months less to have the next child. Column (3) includes the mother fixed effects and yelds a similar coefficient (-0.83 months). Columns (4)-(9) show the results for the probability that the interval is shorter than 24 (columns 4-6) or 15 months (columns 7-9), which are the intervals below which child's and mother's health can be negatively affected. Column (9) shows that having a girl increases the likelihood of the interval being shorter that 15 months by 0.8 percentage points (significant at 1% level), which is equivalent to a large 8.9 percent effect. These findings are consistent with son-preferring fertility stopping rules.

2.5.3 Age structure

To confirm the non-parametric evidence shown in figure 2.1, I estimate an ordered logit model in which the dependent variable is a categorical variable with 7 categories for each of the 5-year age groups (from 15–19 to 45–49). The controls are similar to the ones included in regression 2.1. After estimating the ordered logit model, I compute changes in the outcomes predicted probabilities when the independent variable 'first-born girl' (or girl-girl when considering the subsample of women with at least two children) changes from 0 to 1, with all the other independent variables are set at their mean value. Changes are reported with 95% confidence intervals. Columns (1) and (3) in table 2.5 show the changes in the predicted probabilities for women with at least one or two children, respectively. The results confirm that the age distribution for women who had a first-born girl lies to the left of the distribution for women with first-born boy, conditional on the number of children and all other covariates (set at their mean).

2.5.4 Health outcomes

Next, I estimate the effect of the sex of the first-born on several health and anthropometric outcomes for women in different age groups. I estimate the following regression:

$$y_{i,s,b,r} = \beta_1(firstborngirl)_i + \beta_2(firstborngirl)_i(ageover30)_i + \\ + \gamma X_{i,r,b,s} + \alpha_s + \mu_b + \delta_s + \epsilon_{i,s,b,r} \tag{2.3}$$

with mother i, resident in state s, born in year b, surveyed in round r. $y_{i,s,b,r}$ is the the incidence of anemia (the % of women who are anemic), weight-for-height, and body mass index (BMI). $firstborngirl$ indicates whether the first child ever born is a girl (as opposed to a boy), which is also interacted with a dummy equal to one if the woman is older than 30); $X_{i,s,b,r}$ is the same set of covariates as in regression 2.1. α_s, γ_b, δ_r are state, cohort of birth, and survey round fixed effects, respectively. In addition, I control for potential confounding factors such as pregnancy and breastfeeding status of mothers which may be correlated with the error term. Moreover, as a proxy for women's health status before birth, I control for height which reflects nutritional status in childhood. 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 2.3 is estimated using OLS.

Columns (1) to (4) of table 2.6 shows the results for the incidence of anemia. Consistent with the non-parametric evidence shown in figure 2.2, women under the age of 30 with a first-born girl are 1.5 percentage points more likely to be anemic, while this difference disappears for women older than 30 (column 1). Columns (2) to (4) show that this result is robust to the inclusion of the number of children ever born, dummies for pregnancy and breastfeeding status and woman's height among the regressors. Consistent with the medical evidence, women with many children, pregnant or breastfeeding are more likely to be anemic. Columns (5) to (12) of table 2.6 show the results for weight-for-height and BMI. First, both these indicators of women's nutritional status are negatively correlated with the number of children ever born (controlling for all other covariates, including proxies for socioeconomic status). The estimates show that, among women age 30–49, those with first-born girl show significantly higher weight-for-height and BMI.

2.6 Alternative interpretations

I next turn to discussing and testing the plausibility of three alternative interpretations of the results. First, I check whether the results might be driven by biological differences. There is evidence that a female birth is more likely if the mother is malnourished (because male fetuses are more susceptible to death in utero than female) (Andersson and Bergstrom, 1998; Almond and Mazumder, 2011). If this is the case, this interpretation would not be compatible with a pattern in which differences in the nutritional indicators for women with first-born daughter or son arise only *after* the time of first birth. In order to check for this possibility, I plot the prevalence of anemia by the age and sex of the first-born (using Kernel-weighted local polynomial smoothing as in Section 4) to understand when these differences arise. The upper panel of figure 2.5 shows the % of women (with at least one child) with anemia by the age of the first child and the sex of the first-born. The figure shows that the incidence of anemia is the same for women with first-born daughter or son at the time of birth of the first child (i.e., when the first child is aged 0) and that difference develop over time. Specifically, women with first-born daughter start being more likely to be anemic approximately one year after the birth of their daughter, and differences increase over time. The bottom panel of figure 2.5 focuses on women with at least two children and shows the prevalence of anemia by the age of the second child for women who had at least two daughters or two sons among their earlier-born children. The figure shows that, at the time of birth of the second child, women who already had a first-born daughter are more anemic (but not significantly so) and, again, differences in the incidence of anemia increase over the years. This evidence is not consistent with the explanation based on biological differences. Instead, it probably reflects son-preferring fertility behavior in which women with daughters want to 'try again for a son' by having more children (and shortly spaced). This behavior is reflected in higher anemia prevalence, which is an indicator that responds quickly to reproductive behavior and it is associated with maternal mortality. Finally, it is important to note that the results for the anemia prevalence in table 2.6 are unchanged when mother's height is included among the controls, thus suggesting that the results are not driven by the nutritional status before birth.

The second check deals with selective recall bias. With son preference, women might under-report female births, especially those those occurred back in the past. In this regard, it is important to note that several studies find that son preference is not decreasing over time in India (if anything, Census data report increasing male-biased sex-ratios at birth). Consistent with nondecreasing son preference, if the results of this paper are driven by selective recall bias (motivated by son preference), then I should not find evidence of decreasing maternal mortality among women with first-born girl in most recent surveys. To check for this possibility, I regress the mother's age on the variable 'first-born girl' and its interaction with the survey round dummies. Table 2.7 shows that women with first-born girl are younger than those with first-born boy especially among women in the first wave of the NFHS. Interestingly, this effect is attenuated for women in the last two survey rounds. This pattern would not be consistent with sex-selective recall bias but instead with the reduction in maternal mortality that has been taking place in India over the last years.

Finally, one last concern is that the practice of sex-selective abortion may be driving the results. One could argue that older and uneducated women might be the ones who selectively aborted female children, and this might explain why they are fewer as shown in figure 2.9 (as they do not report aborted girls). However, the evidence on sex-selective abortion in India suggests that it is more prevalent among more educated women (who have access to the technologies for the detection of the sex of the fetus) and for higher-order children (*after* the first-born). Also, ultrasound technologies became increasingly available only after the end of the 1980s. Therefore, sex-selective abortion should be more frequent among younger women and not among the ones that in figure 2.9 appear to be 'missing' (older and uneducated women). As a further check, I estimate that the probability that the second-born child is a male on the first-born being a girl. The estimates suggest that this probability is positive and significant only for younger women (especially if educated and in urban areas): women below 30 are 1.2 percentage points more likely to have a second-born boy if the first is a girl, while no difference for women 30 or older are found. Hence, sex-selective abortion should not affect the result obtained for maternal mortality.

2.7 Conclusions

This paper has shown some evidence that there might be differential mortality for mothers of daughters (as opposed to mothers of sons) among earlier-born children in India. This finding might partly explain the phenomenon of missing women at reproductive and older ages recently uncovered by Anderson and Ray (2010, 2012). While more research in this area is needed, it proposes an original explanation in which maternal mortality can (at least partly) be an unintended consequence of son preference through son-preferring fertility behavior.

Figures

Figure 2.1: Share of women with first-born girl by mother's age (observed at the time of the survey)

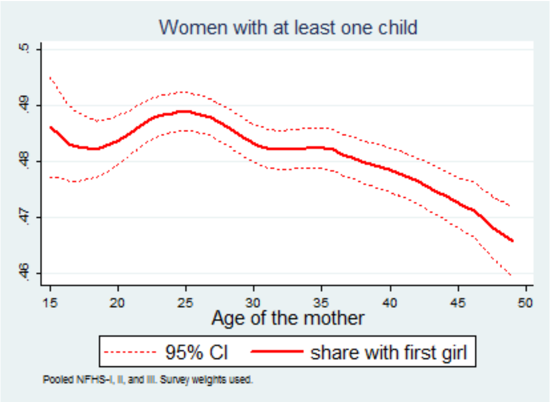


Figure 2.2: % anemic women, by age of the mother (observed at the time of the survey)

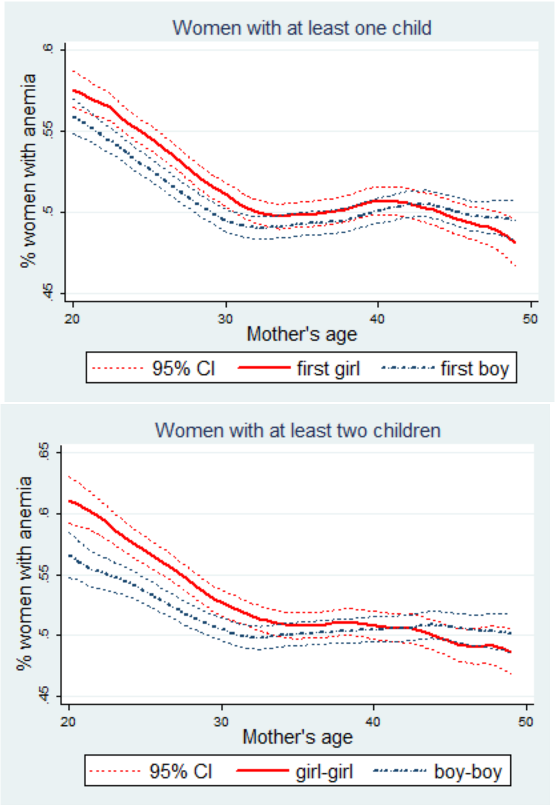


Figure 2.3: Weight-for-height, by age of the mother (observed at the time of the survey)

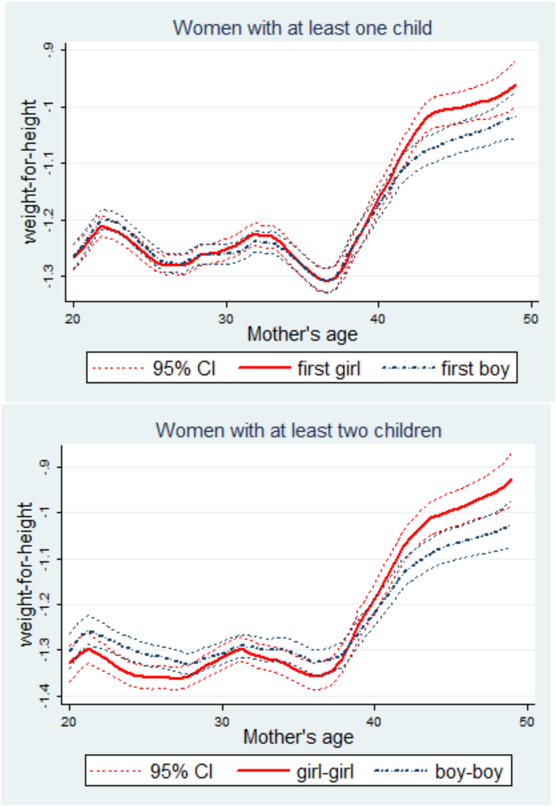
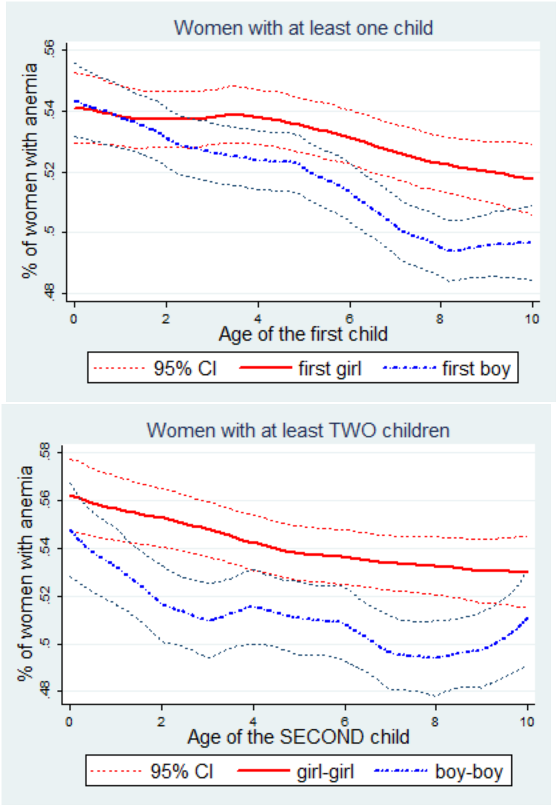


Figure 2.4: Height, by age of the mother (observed at the time of the survey)

Figure 2.5: Share of women with anemia, by age of the first (or first two) child (children) and the sex composition



Tables

Table 2.1: Summary statistics, by sex of the first-born child

	first-born girl	first-born boy	diff	se	Obs
education years	3.482	3.438	0.044**	(0.022)	244512
husband eduyrs	6.036	6.014	0.022	(0.026)	243334
age	32.062	32.239	-0.177***	(0.042)	244803
age first birth	18.913	18.878	0.035**	(0.017)	244803
age first marriage	16.773	16.743	0.030*	(0.016)	244690
# children ever born	3.441	3.246	0.195***	(0.010)	244803
husband living in	0.929	0.93	-0.001	(0.001)	230290
female head	0.099	0.095	0.004**	(0.002)	244800
muslim	0.125	0.127	-0.003	(0.002)	244729
hindu	0.817	0.814	0.003	(0.002)	244729
sikh	0.018	0.018	-0.000	(0.000)	244729
sched caste	0.167	0.167	0.000	(0.002)	241080
sched tribe	0.088	0.086	0.002	(0.001)	241080
Health indicators					
anemic	0.548	0.538	0.011***	(0.003)	149779
weight-for-height	-1.362	-1.358	-0.004	(0.007)	154633
height	151.55	151.519	0.031	(0.037)	156402
body mass index (BMI)	20.595	20.586	0.009	(0.024)	156146

All women aged 15–49 with at least one child ever born. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. Anemia, height, weight-for-height, BMI only in waves 1998 & 2005.

Table 2.2: Summary statistics, by age group and sex of the the first-born

	Women age 15-30					Women age 31-49				
	firstG	firstB	diff	se	obs	firstG	firstB	diff	se	obs
education years	3.676	3.72	-0.044	(0.033)	110396	3.309	3.193	0.116***	(0.030)	134116
husband eduyrs	6.152	6.185	-0.033	(0.037)	109858	5.931	5.865	0.066*	(0.035)	133476
age	24.621	24.628	-0.007	(0.027)	110533	38.753	38.873	-0.120***	(0.036)	134270
age first birth	18.377	18.37	0.007	(0.021)	110533	19.394	19.32	0.074***	(0.025)	134270
age first marriage	16.532	16.53	0.002	(0.020)	110471	16.988	16.928	0.060**	(0.023)	134219
# children ever born	2.494	2.352	0.142***	(0.010)	110533	4.292	4.025	0.267***	(0.014)	134270
husband living in	0.912	0.914	-0.002	(0.002)	107346	0.944	0.944	0.000	(0.002)	122944
female head	0.083	0.081	0.002	(0.002)	110532	0.114	0.108	0.006***	(0.002)	134268
muslim	0.136	0.137	-0.001	(0.003)	110491	0.115	0.119	-0.004*	(0.002)	134238
hindu	0.815	0.813	0.002	(0.003)	110491	0.819	0.815	0.004	(0.003)	134238
sikh	0.015	0.015	0.000	(0.001)	110491	0.019	0.02	-0.001	(0.001)	134238
sched caste	0.176	0.172	0.004	(0.003)	108996	0.16	0.163	-0.003	(0.003)	132084
sched tribe	0.098	0.094	0.004*	(0.002)	108996	0.078	0.079	-0.001	(0.002)	132084
Health indicators										
anemic	0.573	0.554	0.019***	(0.005)	65400	0.528	0.524	0.004	(0.004)	84379
weight-for-height	-1.389	-1.368	-0.021**	(0.010)	67850	-1.338	-1.349	0.011	(0.011)	86783
height	151.51	151.495	0.015	(0.054)	68432	151.585	151.54	0.045	(0.050)	87970
BMI	19.779	19.807	-0.028	(0.029)	68305	21.289	21.236	0.053	(0.036)	87841

All women aged 15–49 with at least one child ever born. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels. Anemia, height, weight-for-height, BMI only in waves 1998 & 2005.

Table 2.3: Fertility outcomes by the sex of the first-born

dep. var.	# children ever born		contraceptives		sterilized		abortion		wants more ch	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
first-born girl	0.196***	0.283***	-0.048***	-0.073***	-0.052***	-0.082***	0.011***	0.013***	0.047***	0.051***
	(0.011)	(0.008)	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
first child dead		0.783***		-0.160***		-0.151***		0.029***		0.165***
		(0.019)		(0.005)		(0.004)		(0.005)		(0.006)
first child dead*		-0.289***		0.073***		0.094***		-0.007		-0.044***
first-born girl		(0.026)		(0.008)		(0.009)		(0.007)		(0.003)
Survey dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Fes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Birth year Fes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Religion, Caste	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Age group	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Other controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	244803	224534	244803	224534	244803	224493	244783	224516	235206	224311
R2 (or Pseudo R2)	0.008	0.549	0.017	0.209	0.007	0.250	0.001	0.029	0.012	0.392
Percent effect	6.0	8.7	8.8	13.5	13.0	20.5	5.8	6.8	23.5	25.5
Mean first boy	3.25	3.25	0.54	0.54	0.40	0.40	0.19	0.19	0.20	0.20

OLS estimates in columns 1-2. Probit estimates in columns 3-10 (marginal effects reported). Robust standard errors adjusted for clustering at the household level in parentheses. All women aged 15–49. Using survey weights. Each regression also includes 7 five-year age groups dummies, age at first birth and marriage, mother and partner's age and education years, urban dummy, whether first child died (and its interaction with first born girl), and a durable index. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 2.4: Birth spacing and sex of the child

dep. var.	# months			interval < 24 months			interval < 15 months		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
girl	-1.031***	-0.723***	-0.834***	0.022***	0.017***	0.020***	0.009***	0.007***	0.008***
	(0.058)	(0.058)	(0.089)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
dead child	-6.540***	-4.735***	-4.692***	0.197***	0.165***	0.150***	0.141***	0.127***	0.111***
	(0.114)	(0.114)	(0.154)	(0.003)	(0.003)	(0.005)	(0.002)	(0.002)	(0.003)
girl*dead child	1.266***	0.891***	1.150***	-0.032***	-0.025***	-0.033***	-0.013***	-0.010***	-0.009*
	(0.146)	(0.146)	(0.189)	(0.005)	(0.005)	(0.006)	(0.004)	(0.004)	(0.005)
# children ever born		-1.776***			0.032***			0.014***	
		(0.018)			(0.000)			(0.000)	
Birth-order Fes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother Fes	No	No	Yes	No	No	Yes	No	No	Yes
Observations	552431	552431	552431	552431	552431	552431	552431	552431	552431
R-squared	0.017	0.044	0.505	0.021	0.033	0.401	0.030	0.036	0.402
% effect	3.2	2.2	2.6	6.5	5	5.9	10	7.8	8.9
Mean first-born boy	32.2	32.2	32.2	0.34	0.34	0.34	0.09	0.09	0.09

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. All women aged 15–49. Excludes births of birth order 10 or higher. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 2.5: Differences in the conditional age distribution between women with first-born girl or boy from ordered logit regression. Dep. variable: 5-year age groups.

women with:	at least one child		at least two children	
	first girl vs first boy		girl-girl vs boy-boy	
	change	95% CI	change	95% CI
	(1)	(2)	(3)	(4)
$Pr(y = 15 - 19\|x)$	0.0001	[0.0001, 0.0001]	0.0000	[0.0000, 0.0000]
$Pr(y = 20 - 24\|x)$	0.0022	[0.0019, 0.0026]	0.0018	[0.0016, 0.0021]
$Pr(y = 25 - 29\|x)$	0.0210	[0.0182, 0.0239]	0.0240	[0.0209, 0.0272]
$Pr(y = 30 - 34\|x)$	0.0022	[0.0015, 0.0030]	0.0346	[0.0305, 0.0388]
$Pr(y = 35 - 39\|x)$	-0.0218	[-0.0247, -0.0189]	-0.0468	[-0.0525, -0.0412]
$Pr(y = 40 - 44\|x)$	-0.0035	[-0.0039, -0.0030]	-0.0124	[-0.0140, -0.0109]
$Pr(y = 45 - 49\|x)$	-0.0003	[-0.0004, -0.0003]	-0.0013	[-0.0015, -0.0011]
Observations	224534		186992	
Pseudo R-sq	0.446		0.423	

Computed from the ordered logit estimates. The values reported are the changes in predicted probabilities when the 'first-born girl' changes from 0 to 1, with all the other independent variables set at their mean value. All regressions include individual controls as in table 2.3. Changes are reported with 95% confidence intervals by the delta method in parentheses. Robust standard errors adjusted for clustering at the household level in parentheses. All women aged 15–49. Excludes births of birth order 10 or higher. Using survey weights. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 2.6: Health outcomes and sex of first-born, by age group

dep. var.	% anemic				weight-for-height				BMI			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
first girl	0.015***	0.012***	0.012***	0.012***	-0.008	-0.000	0.005	0.005	0.011	0.003	0.018	0.018
	(0.004)	(0.004)	(0.004)	(0.004)	(0.011)	(0.011)	(0.011)	(0.011)	(0.032)	(0.032)	(0.032)	(0.032)
first girl*	-0.012**	-0.011*	-0.012**	-0.012**	0.029**	0.025*	0.033***	0.033***	0.084**	0.091**	0.119***	0.118***
age≥30	(0.006)	(0.006)	(0.006)	(0.006)	(0.013)	(0.013)	(0.013)	(0.013)	(0.041)	(0.041)	(0.041)	(0.041)
currently		0.042***	0.039***	0.039***		-0.071***	-0.038***	-0.037***		-0.278***	-0.171***	-0.178***
breastfeeding		(0.005)	(0.005)	(0.005)		(0.008)	(0.008)	(0.008)		(0.027)	(0.027)	(0.027)
pregnant		0.017**	0.016**	0.017**		-0.267***	-0.262***	-0.262***		1.096***	1.112***	1.112***
		(0.008)	(0.008)	(0.008)		(0.013)	(0.013)	(0.013)		(0.045)	(0.045)	(0.044)
# children			0.004***	0.004***			-0.037***	-0.037***			-0.118***	-0.117***
			(0.001)	(0.001)			(0.003)	(0.003)			(0.008)	(0.008)
height				-0.001***				0.005***				-0.036***
				(0.000)				(0.001)				(0.002)
Observations	136419	136419	136419	136216	140895	140895	140895	140895	142201	142201	142201	142200
R-squared	0.047	0.048	0.048	0.048	0.232	0.235	0.236	0.237	0.248	0.253	0.255	0.257

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. All women aged 15–49. Using survey weights. Includes age at first birth and marriage, mother and partner's age and education years, urban dummy, whether first child died, and a durable index, birth year dummies, wave dummies, state FEs, religion and caste. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Table 2.7: Dep. variable: Mother's age

	(1)
first-born girl	-0.287***
	(0.031)
first-born girl*wave2	0.026
	(0.044)
first-born girl*wave3	0.080*
	(0.044)
wave2	0.273***
	(0.036)
wave3	0.926***
	(0.039)
Obs	224534
R-squared	0.830

OLS estimates. Robust standard errors adjusted for clustering at the household level in parentheses. All women aged 15–49. Using survey weights. Includes age at first birth and marriage, mother and partner's age and education years, urban dummy, whether first child died, state dummies, # children ever born, and a durable index. ***, **, and * indicate significance at 1%, 5% and 10% levels.

Appendix

Figure 2.6: Sex ratio (% male births), by birth year of the child

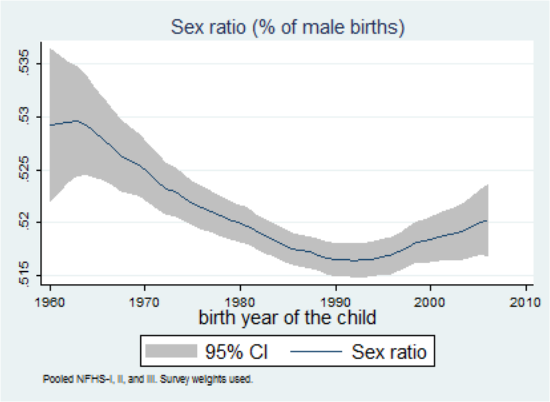


Figure 2.7: Sex ratio (% male births), by birth year of the child and DHS survey round

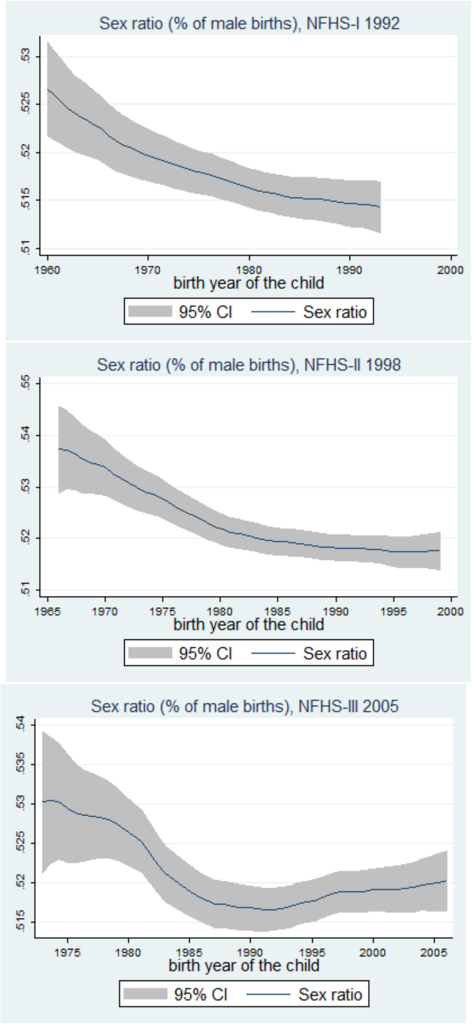


Figure 2.8: Sex ratio (% male births), by birth year of the child and birth order

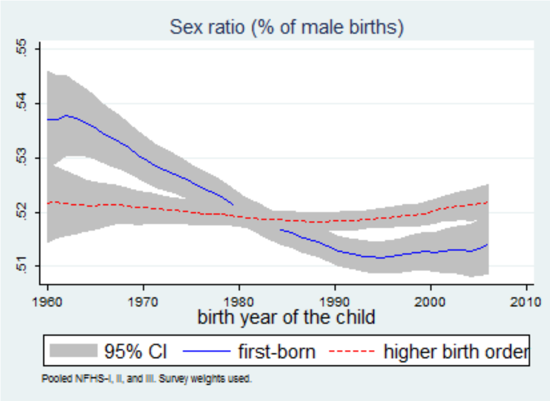


Figure 2.9: Share of women with first-born girl by mother's age and education

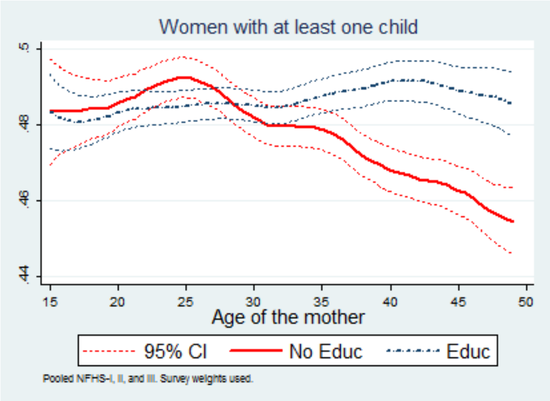


Figure 2.10: Share of women with first-born girl by mother's age and urban/rural location

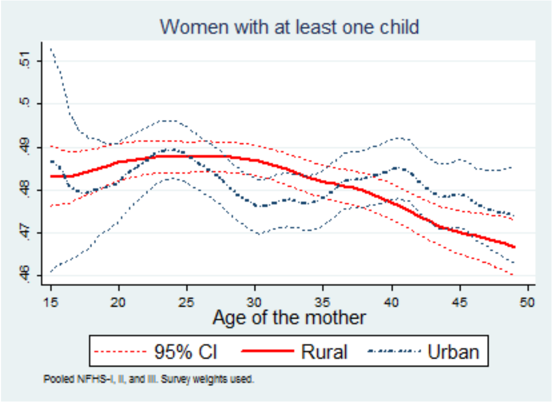


Figure 2.11: Share of women by the sex of the first two children mother's age

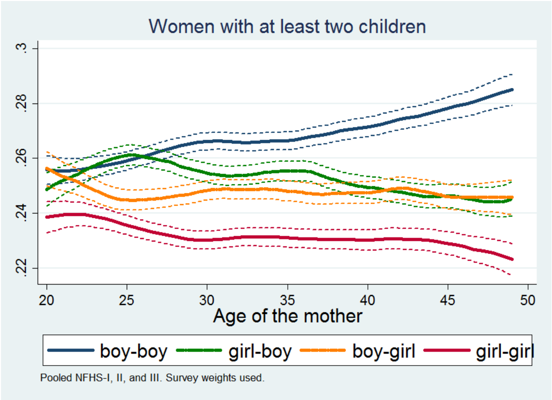


Figure 2.12: % anemic women, by age of the mother at the time of the survey. Subsamples of women not currently breastfeeding or pregnant

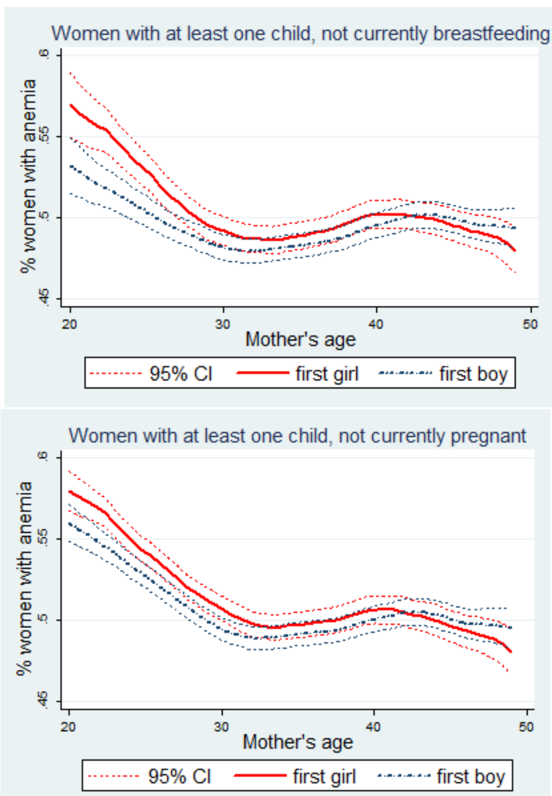
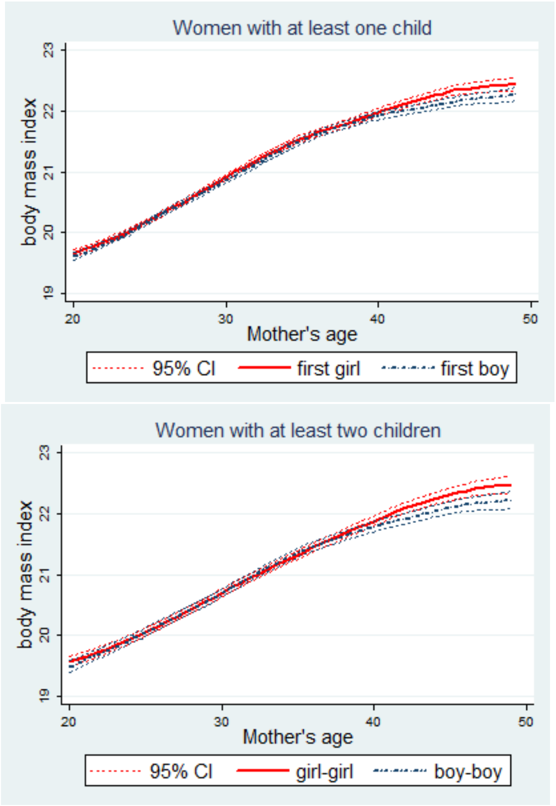


Figure 2.13: Body mass index, by sex of the first-born and age of the mother at the time of the survey



Bibliography

[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] AbouZahr, Carla. 2003. "Global burden of maternal death and disability." *British Medical Bulletin*, 67: pp. 111.

[3] Akbulut-Yuksel, Mevlude, and Daniel Rosenblum. 2012. "The Indian Ultrasound Paradox." IZA Discussion Paper No. 6273.

[4] Alesina, Alberto, Paola Giuliano, and Nathan Nunn. 2011. "Fertility and the Plough." *American Economic Review: Papers & Proceedings 2011*, Vol. 101, No. 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, No.4, pp. 1262-1300.

[7] —. 2012. "The Age Distribution of Missing Women in India." Working Paper, July 2012.

[8] 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.

[9] 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.

[10] Arnold, Fred, Minja Kim Choe, and T.K. Roy. 2002. "Sex-selective Abortions in India." *Population and Development Review*, Vol. 28, No. 4, pp. 759–785.

[11] Barcellos, Silvia H., Leandro Carvalho, and Adriana Lleras-Muney. 2012. "Child Gender And Parental Investments In India: Are Boys And Girls Treated Differently?" NBER working paper No. 17781.

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

[13] Becker, Gary and Nigel Tomes. 1976. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy*, 84(4), pp. S143–S162.

[14] —. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3), pp. S1-S39.

[15] Behrman, Jere R. 1988. "Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?." *Oxford Economic Papers*, Vol. 40, No. 1, pp. 32–54.

[16] —. 1997. "Intrahousehold Distribution and the Family." *Handbook of Population and Family Economics*, Vol. 1, No.1, pp. 125–187.

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

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

[19] Bergstrom, Theodore C. 1997. "A Survey of Theories of the Family." *Handbook of Population and Family Economics*, Vol. 1, No. 1, pp. 21–80.

[20] Bhalotra, Sonia, and Tom Cochrane. 2010. "Where Have All the Young Girls Gone? Identification of Sex Selection in India." IZA Discussion Paper No. 5381.

[21] Bharadwaj, Prashant and Leah Nelson. 2012. "Discrimination Begins in the Womb: Evidence of Sex Selective Prenatal Investments." *Journal of Human Resources*, forthcoming.

[22] Boserup, Ester. 1970. "Woman's Role in Economic Development." London: Allen and Unwin.

[23] 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.

[24] Carranza, Eliana. 2012. "Soil Endowments, Production Technologies and Missing Women in India." World Bank Policy Research Working Paper 5974.

[25] Chowdhury, Mridul K. and Radheshyam Bairagi. 1990. "Son Preference and Fertility in Bangladesh." *Population and Development Review*, Vol. 16, No. 4, pp. 749–757.

[26] Clark, Shelley. 2000. "Son Preference and Sex Composition of Children: Evidence from India." *Demography*, Vol. 37, No. 1, pp. 95–108.

[27] 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*, Vol. 196, No. 4, pp. 297–308.

[28] Conde-Agudelo, A., and Belizan, J.M. 2000. "Maternal morbidity and mortality associated with interpregnancy interval: Cross sectional study" *British Medical Journal*, Vol. 321, No. 7271, pp. 1255–1259.

[29] Dahl, Gordon B. and Enrico Moretti. 2008. "The Demand for Sons." *Review of Economic Studies*, Vol. 75, No. 4, pp. 1085–1120.

[30] Das Gupta, Monica. 2005. "Explaining Asia's 'Missing Women': A New Look at the Data." *Population and Development Review*, Vol. 31, No. 3, pp. 529–535.

[31] Das Gupta, Monica, Jiang Zhenghua, Li Bohua, Xie Zhenming, Woojin Chung, Bae Hwa-Ok. 2003. "Why is Son preference so persistent in East and South Asia? a cross-country study of China, India and the Republic of Korea." *Journal of Development Studies*, Vol. 40, No. 2, pp. 153–187.

[32] Dreze, Jean and Mamta Murthi. 2001. "Fertility, Education, and Development: Evidence from India." *Population and Development Review*, Vol. 27, No. 1, pp. 33–63.

[33] George, Sabu M. 2002. "Sex Selection/Determination in India: Contemporary Developments." *Reproductive Health Matters*, Vol. 10, No. 19, pp. 190–192.

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

[35] Goody, Jack, and Stanley J. Tambiah. 1973. "Bridewealth and Dowry" Cambridge: Cambridge University Press.

[36] Hu, Luojia, and Anala Schlosser. 2011. "Prenatal Sex Selection and Girls Well-Being: Evidence from India." IZA Discussion Paper No. 5562.

[37] 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.

[38] 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.

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

[40] Jha, Prabhat, Rajesh Kumar, Priya Vasa, Neeraj Dhingra, Deva Thiruchelvam, and Rahim Moineddin. 2006. "Low Male-to-Female Sex Ratio of Children Born in India: National Survey of 1 1 Million Households." *The Lancet*, Vol. 367, No. 9506: 211-218.

[41] Kalaivani K. 2009. "Prevalence & consequences of anaemia in pregnancy." *Indian Journal of Medical Research*, Vol. 130, No. 5, pp. 627–633.

[42] King, Janet C. 2003. "The Risk of Maternal Nutritional Depletion and Poor Outcomes Increases in Early or Closely Spaced Pregnancies." *Journal of Nutrition*, Vol. 133, No. 5, pp. 1732–1736.

[43] Klasen, Stephan. 1996. "Nutrition, Health, and Mortality in Sub Saharan Africa: Is There a Gender Bias?" *Journal of Development Studies*, Vol. 32, No. 6, pp. 913-933.

[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. "Son Preference, Fertility, and Family Structure. Evidence from Reproductive Behavior among Nigerian Women" Bocconi University, working paper.

[46] Oster, Emily. 2005. "Hepatitis B and the Case of the Missing Women." *Journal of Political Economy*, 2005, vol. 113, No. 6, pp. 1163–1216.

[47] Palloni, Alberto, and Sara Millman. 1986. "Effects of Inter-Birth Intervals and Breastfeeding on Infant and Early Childhood Mortality." *Population Studies*, Vol. 40, No. 2, pp. 215–236.

[48] 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.

[49] 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.

[50] Rosenblum, Daniel. 2012. "The Effect of Fertility Decisions on Excess Female Mortality in India." Working Paper, Dalhousie University.

[51] 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.

[52] Sen, Amartya. 1990. "More Than 100 Million Women Are Missing." *The New York Review of Books*, Vol. 37, No.20.

[53] 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.

[54] Smith, Herbert L. 1994. "Nonreporting of Births or Nonreporting of Pregnancies? Some Evidence from Four Rural Counties in North China." *Demography*, Vol. 31, No. 3, pp. 481–486.

[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, No. 6, 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 Bank. 1996. "Improving women's health in India." Development in Practice, World Bank.

[59] —. 2011. "World Development Report 2012. Gender Equality and Development." World Bank.

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

[61] —. 2008. "Worldwide prevalence of anaemia 1993-2005. WHO Global Database on Anaemia." World Health Organization.

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

[63] Thomas, Duncan. 1992. "Like Father, like Son; Like Mother, like Daughter: Parental Resources and Child Height." *The Journal of Human Resources*, Special Issue: The Family and Intergenerational Relations, Vol. 29, No. 4, pp. 950-988.

Chapter 3

Customary Norms, Inheritance, and Human Capital: Evidence from a Reform of the Matrilineal System in Ghana
(joint with Eliana La Ferrara)

This paper explores the effects of different descent rules on human capital accumulation. In a context where parents are constrained in the possibility of passing land on to their children (e.g., because it is considered property of the extended family, or clan), investment in their children's human capital might not be optimal. We focus on matrilineal inheritance rules and exploit a policy experiment in Ghana, the introduction of the 1985 Intestate Succession Law. The Law introduced minimum quotas for the land that parents can devolve on their children through intestacy, substantially reducing the share going to the matriclan. In our setting, the Law allowed parents to move closer to the unconstrained optimum. We find evidence that, compared to the other patrilineal ethnic groups in Ghana (unaffected by the Law), children in matrilineal groups exposed to the reform received significantly less education. This effect is specific to males for whom the matrilineal constraint was binding, while there is no effect for females. This evidence suggests that before the reform matrilineal groups in Ghana invested more in their children's education to substitute for land inheritance and more generally that traditional norms are important in determining the intergenerational accumulation of property and human capital investments.

3.1 Introduction

Parental investment in human capital has been shown to have important consequences for children's well being and is a crucial input into their social mobility prospects. Available evidence mostly comes from industrialized countries, where parents are generally free to pass their wealth on to their offspring according to their will. In many developing countries, however, bequests respond not only to parental decisions, but to a series of claims by extended family and lineages that are enforced through customary norms. The presence of these norms is a potential source of distortions in parents' allocation decisions, the extent of which is not yet fully understood (e.g., Platteau, 2000; Pande and Udry, 2006).

This paper exploits a policy change introduced by the Government of Ghana, the Intestate Succession Law, 1985 (PNDCL 111), which radically changed the inheritance transmission rules followed by matrilineal groups in the country (notably, the Akan). According to customary matrilineal practices, Akan fathers could not devolve their land to their sons, but only to their sisters' sons or other eligible males of their own maternal kin. Akan mothers were instead free to pass their property on to their daughters (even though women have very weak rights over land in this context). This traditional inheritance mode can be viewed as a constraint, which we refer to as the "matrilineal constraint", on the amount of land that can be passed on to a man's offspring. This constraint generates a distortion in the amount of human capital investment chosen by the parent, leading to an allocation where the mix of human and physical capital is different from the unconstrained optimum.

We formalize this intuition presenting a simple model where a parent has to allocate his income between own consumption and investment in education of the child, and where education and land enter the child's income-generating function. The matrilineal constraint is represented by an exogenous cap on the amount of land that the child can inherit. We show that a binding matrilineal constraint leads to a higher than optimal investment in education of the child, to compensate for the lower amount of land. The Intestate Succession Law introduced minimum quotas for the land that fathers should devolve on their children, virtually reducing the claims of the matriclan to a small fraction of the land. This Law can thus be interpreted as a relaxation of the matrilineal constraint, leading to a higher amount of land devolved on children. In our model this leads to a decrease in education investment for children. We also provide an extension of the model and consider a case in which the parent has two children, a son and a daughter, and allocates resources between his consumption and investment in education of the son and of the daughter. We assume that land inheritance is gender-linked (i.e., land is controlled by the male parent who can devolve it on his son upon his death). In this case, the prediction of the model is that the Law leads to a decrease in education investment for sons, and an ambiguous effect on daughters' education.

In the context of Ghana, where not all ethnic groups are matrilineal, this simple theoretical framework implies that after the passage of the Law we should expect relatively lower levels of investment in education for groups that were previously subject to the matrilineal constraint (i.e., the Akan), but not for other groups. Furthermore, because the customary norm implied restrictions on inheritance that were most binding for boys, we expect the reform to affect boys' human capital more than girls'.

We test the above predictions using five rounds of data from the Ghana Living Standards Survey (GLSS). To analyze the impact of the reform on investment in education, we restrict the sample to individuals aged 20 to 50. Then we compare educational outcomes of those who have been exposed to the reform to those whose education should not have been affected by the Law according to their age the year in which the reform was passed. Our empirical strategy is thus a difference in difference strategy akin to the ones used by Duflo (2001) and Lavy (2011): we exploit differences across cohorts and ethnic groups separately for males and females.

Our results can be summarized as follows. Akan males exposed to the reform experienced approximately a one-year reduction in the number of years of education completed compared to non-Akan males in the same cohort. Considering that the average years of education of Akan males pre-reform were 8.6, this represents a 12 percent reduction in education for this group compared to non-Akan males. We show that there was a parallel trend in education for Akan and non-Akan men before the passage of the Law, but the trend for Akan men slowed down for those Akan cohorts exposed to the reform. The same does not hold for Akan women: the (positive) difference between their education levels and those of non-Akan women is constant before and after the reform. We also find that Akan men affected by the reform have a 9.7 (11.4) percentage points lower probability of completing at least primary (secondary) school, while the change in the probability of completion for women is null.

We also find a negative effect on attendance for children in later survey rounds, i.e. when the reform has been fully internalized by the parents. Interestingly, the negative effect on attendance is attenuated for boys who have more siblings. Although we cannot make any causal claim, this is consistent with the fact that their paternal land should be shared among more heirs, leading to lower land per child, hence lower disinvestment in education.

We show that our results are robust to different definitions of the threshold for treatment, to excluding from the sample Northern regions where Akan are in very low numbers, to considering the effects for the subsample of cocoa-growing villages among which Akans are highly represented, and for the subsample of households in which at least one member had a government job. This allows us to make sure that our results are not driven by omitted time-varying factors that are specific to the regions where the Akan are spatially concentrated, or to changes in returns to education due to fluctuations in cocoa production or to cuts in public employment following the structural adjustment program started in 1986. We also conduct a test in which our control group is composed by the Ewe, a sub-group of the non-Akan characterized by a higher level of education, to rule out the possibility that our results might be affected by convergence. In our last check, we exploit the fact that Akan are also present in the neighboring Côte d'Ivoire (but were not affected by the Intestate Law introduced in Ghana in 1985) to understand if our results are driven by shocks specific to the Akan ethnicity.

Finally, we also find evidence that the Law affected occupational choice. In particular, compared to non-Akan males, Akan males in later rounds show a higher probability of being farmers, while this is not the case for females. This is consistent with the fact that the possibility of inheriting the father's land may have affected the choice of becoming farmers (an occupation in which the returns to education are lower than in other sectors).

To sum up, we therefore uncover significant and sizable effects of customary norms that constrain bequest allocation on the accumulation of human capital, effects that differ across genders and that may account for part of the education gender gap.

This paper is related to several strands of the literature. A first body of literature looks at the economic consequences of social norms, kinship systems and inheritance rules, modeling household behavior as a rational response to traditional customs (e.g., La Ferrara, 2003; La Ferrara, 2007; Goetghebuer and Platteau, 2010; Mobarak *et al.*, 2009; Quisumbing *et al.*, 2001; Quisumbing and Otsuka, 2001; Lambert, Ravallion, and van de Walle, 2011; Carranza, 2012). Gneezy, Leonard and List (2009) show that differences between societies with distinct kinship systems and customary rules (the matrilineal Khasi and patriarchal Maasai) correlate with economically relevant behaviors such as the inclination toward competitiveness. Compared to this literature, our paper is the first to exploit a policy change that constitutes an exogenous shock to the strength of the customary norm. This allows us to make some progress toward establishing a causal link going from social constraints to economic choices.

A second strand of the literature explains the presence of different social norms regarding the transmission of property from parents to children as a rational response to different economic environments (e.g., Botticini and Siow, 2003; Platteau and Baland, 2001). In this paper we take the existence of matrilineal customs as given and study the effects of an exogenous shift in these customs operated through legislation.

Thirdly, our paper relates to the literature on the constraints generated by extended families ties. In particular, this literature suggests that family networks, mainly in the African context, limit the possibility of undertaking profitable economic opportunities and create incentives for hiding income due to informal risk-sharing arrangements (Baland, Guirkinger and Mali, 2011; Jakiela and Ozier, 2011).

Finally, our paper also speaks to the literature on land rights security and investment in agriculture (e.g., Besley, 1995; Goldstein and Udry, 2008; Hornbeck, 2010). By showing that control over the inter-generational transmission of land affects parental investment in human capital, we show that policies for the individualization of land rights have far reaching consequences that go well beyond changes in agricultural investment and productivity.

The remainder of the paper is organized as follows. Section 3.2 briefly reviews some basic notions on matrilineal and patrilineal descent principles taken from the anthropological literature and gives some background on the Intestate Succession Law. In section 3.3 we present a simple model that serves as a guide for our empirical analysis. Section 3.4 introduces our empirical strategy and Section 3.5 provides a descriptive analysis of the data. Sections 3.6, 3.7, 3.8, and 3.9 contain our main econometric results and robustness checks, and Section 3.10 concludes.

3.2 Matrilineal inheritance and the Intestate Succession Law

Kinship systems form the basis for social organization in many developing countries. Kinship is usually built around a unilineal descent group in which kin membership is transmitted from one generation to the next only through ancestors of one gender. In patrilineal societies only males can pass kin membership on to their offspring and children are considered to be part of their father's kin group. In matrilineal societies instead, only females can pass kin membership on to their offspring and children are part of their mother's kin group.

The principle of matrilineal descent is illustrated in figure 3.1. Following the notation in social anthropology, triangles indicate males, circles females, vertical links indicate a descent bond, horizontal ones a codescent bond, and the sign "=" stands for a marriage relationship. The colored symbols indicate members of the same matrikin. The top part of the diagram indicates a couple, and the first of the three generations represented in the diagram. Following the descent bonds (vertical lines), it is easy to see how descent is traced only through females from a founding female ancestor. In particular, if we consider the eldest man in the diagram, indicated as "$head_0$", we see that his children belong to his wife's kin group, not his. When we look at figure 3.1 from the point of view of the male symbol referred to as "$head_1$", we see again that his children do not belong to his kin group, and that the members of his maternal kin are the children of his sisters.

The relationship between father and child in matrilineal societies is thus somewhat weaker than in patrilineal ones, and some of the responsibilities generally assigned to fathers are instead taken on by the mother's brother. Importantly for our study, though, among the matrilineal Akan of Ghana the father remains responsible for food and education expenditures of his children.[1]

The matrilineal and patrilineal systems involve important differences not only in terms of social organization, but also for the intergenerational transmission of property. A general principle common to both is that rights to inheritance are usually gender-linked: males-to-males and females-to-females. In patrilineal systems a

[1] For extensive studies on matrilineal traditions in Ghana, see among others Fortes (1950) and Okali (1983).

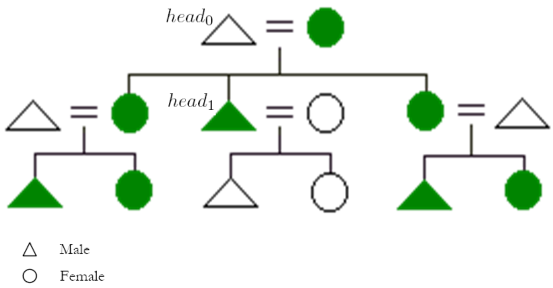


Figure 3.1: Matrilineal descent

man's property is transmitted to his children (typically the sons), while in matrilineal systems the man's children do not belong to his kin group and are not entitled to inherit his property. The man's property is instead transferred to males members of his matrikin, the preferred order of inheritance being: the man's uterine brother, the son of a uterine sister, and the son of the deceased's mother's sister.[2] The woman's property instead is typically passed on to the daughters in both matrilineal and patrilineal societies.

Among the matrilineal Akan of Ghana, children have the customary obligation to work on the father's land before moving out of the parents' home to go work on the uncle's land. This can generate tensions between members of the nuclear family and the matriclan over the rights to inherit the father's land. On the other hand, "fathers are expected to set up their male children in life (...) Today, setting up a child in life includes providing a western type of education and/or apprenticeship" (Awusabo-Asare, 1990,

[2]The matrilineal inheritance principle applies to inherited property, which belongs to the matrikin, and to self-acquired property in case of a man's death intestate. According to Akan customary norms, a man could dispose of his self-acquired property through inter-vivos gifts, sales or by writing a will before death, but this practice could only entail a limited portion of land and required formal approval by the matrikin.

p. 7). Duncan (2010) reports qualitative evidence of the growing importance of a mutual understanding between husband and wife in favor of educating children as an acceptable substitute for land. This evidence corroborates our theoretical prediction that Akan fathers traditionally considered education as a substitute for land inheritance for their sons.

On June 14 1985 the Intestate Succession Law (PNDCL 111) was promulgated by the Government of Ghana.[3] The main innovation brought by the Law was the specific protection granted to members of the nuclear family (as opposed to the extended family) in the distribution of the man's self-acquired property on his death intestate. The Law states that, after the house and household chattels are devolved entirely to the spouse and children, the residue of a man's intestate property has to be distributed as follows: nine-sixteenth to children, three-sixteenth to surviving spouse, one-eighth to surviving parents and the remaining one-eighth, in the case of the matrilineal Akan, to the matrikin.[4,5] The Law applies

[3] The Intestate Succession Law was introduced two years after Ghana launched the Economic Recovery Program in 1983.

[4] The interpretation section of the PNDCL 111 provides a definition for household chattels, which includes jewelery, clothes, furniture and furnishings, refrigerator, television, radiogram, other electrical and electronic appliances, kitchen and laundry equipment, simple agricultural equipment, hunting equipment, books, motor vehicles other than vehicles used wholly for commercial purposes, and household livestock.

[5] At the time of the passage of the Law, there was some debate regarding the application of the Law to any type of marriage, including marriages celebrated under customary law (which are the most common form of marriage in Ghana). The Customary Marriage and Divorce (Registration) Law, 1985 (PNDCL 112) states that the 1985 Law on intestate succession applies also to all customary marriages, as long as they have been registered. Woodman (1985) states that following the PNDCL 112 Law "all customary marriages are required to be registered, but it seems likely that unregistered marriages will continue to be valid for the purpose of the Intestate Succession Law"(p. 123). In 1991, the

to all property which a deceased could have but did not dispose of through will (Woodman, 1985).[6] Prior to the passage of the Law, this property was automatically devolved to the kin group and allocated to individual members following customary rules. As mentioned above, the need for a Law regulating intestate succession especially for matrilineal groups became more pressing given the conflicting interests between the nuclear family and the matriclan over the rights to inheritance. Duncan (2010) documents that the growing role of cocoa farming in the rural economy together with the customary obligation placed on women to assist their husbands in their economic pursuits intensified the use of conjugal labor giving rise to the "conjugal unit" as the major unit of production and consumption. The Law was introduced by the government to reflect changes in society and give more importance to the nuclear family.[7] It represented a radical change in the ways property was transmitted across generations in matrilineal groups, giving fathers greater control over the amount of physical capital their children would inherit. We argue below that this allowed fathers to choose an allocation of human capital closer to the one they would have

PNDCL 112 was amended by the Customary Marriage and Divorce (Registration) (Amendment) Law, 1991 (PNDCL 263) to solve this applicability issue. The amended Law states that the Intestate Succession Law applies also in cases "where a court or tribunal is satisfied by oral or documentary evidence before it that a customary law marriage had been validly contracted between a deceased and surviving spouse."(Fenrich and Higgings, 2001, p. 293).

[6] The practice of disposing of individually owned property by will seems to be rare among Ghanaians as "many view drafting a will as inviting death "(Fenrich, and Higgins, 2001, p. 293).

[7] The PNDCL 111 was amended by the Intestate Succession (Amendment) Law, 1991 (PNDCL 264) to further protect spouse and children from ejection from the matrimonial home before the distribution of the estate and prescribe punishment for those who deprive a person entitled to inherit from his portion of the estate. The Children's Act, 1998 (Act 560) amended PNDCL 111 mainly providing that "where there is a child who is a minor undergoing educational training, reasonable provision shall be made for the child before distribution.".

chosen in the absence of the (customary) matrilineal constraint.

One final consideration relates to women's vs. men's land rights. According to customary principles, women could already pass their land on to daughters before the reform. Obviously, in a context where women do not own land or have weak rights over it, this may translate into few bequests. Qualitative evidence from Ghana suggests that often "male kin play the land-allocating role in both matrilineal and patrilineal societies: secure access rights for women therefore depend on the nature of their relationships with male relatives"(Duncan, 2010, p. 302). The Law on intestate succession in principle allowed fathers' property to be transmitted both to sons and to daughters, but in practice inheritance rights have remained gender-linked, with male property being (the bulk of land property) being passed on to male heirs. This means that empirically we expect to see most of the effects of the reform to be on male children, and to a lesser extent on female children.

3.3 Theoretical framework

In this section we propose a simple theoretical framework to highlight the effects of the reform of matrilineal inheritance on human capital investment. The model is very stylized and serves as a motivation for our empirical strategy.

3.3.1 Allocation with and without matrilineal norms

Consider an environment where an altruistic parent has one child, and allocates resources between own consumption and investment in education of the child (E). We assume that there is no saving technology so all income is either consumed or invested in children's human capital.

The parent is endowed with an exogenous amount of land L. The amount of land that the son inherits is denoted by $L^c \leq L$. Importantly, the share of parental land that can be transferred to the son depends on the prevailing social norm.

First, we derive the optimal allocation in the absence of claims by the matrikin, i.e. the unconstrained optimum. In this case, land is entirely devolved on to the child, eg. $L^c = L$. Second, we model the matrilineal constraint as an upper bound on the amount of land the parent can transfer to the child, eg. $L^c = \overline{L^c} < L$. Third, we model the Intestate Succession Law as an increase in the upper bound of the land the parent can transfer to the child and show that the equilibrium allocation is closer to the one he would have chosen in absence of the matrilineal constraint. We then conduct some comparative statics to determine how the optimal allocation changes by varying the exogenous amount of land for the child (consistently with the change in the matrilineal rule).

The child's preferences are represented by the utility function $U^c = u(C^c)$, with $u'(\cdot) > 0$, $u''(\cdot) < 0$. The variable C^c denotes the child's consumption, which is equal to the child's income due to the absence of savings. Income, in turn, depends on the child's endowment of physical and human capital, i.e. $C^c = Y^c(E, L^c)$. As previously noted, L^c is the fraction of father's land going to his child. We assume that the child's income depends positively on each of its arguments, $Y^c_E(\cdot) > 0$, $Y^c_L(\cdot) > 0$, and that there are decreasing marginal returns.

The parent's utility depends on own and child's consumption (or income): $V^p = v\left(C^p; Y^c(E, L^c)\right)$, with partial derivatives $v_{C^p}(\cdot) > 0$, $v_{Y^c}(\cdot) > 0$, $v_{C^p,C^p}(\cdot) < 0$, and $v_{Y^c,Y^c} < 0$. Moreover, we assume that $v_{C^p,Y^c}(\cdot) = 0$, which is a common separability assumption between parental consumption and child income in models on intrahousehold allocations with parental altruism (Behrman, 1997). The parent's income, Y^p, is taken as exogenous. This income has to be allocated between the parent's own consumption C^p, and expenditure on child's education, pE.

Let us start by considering an equilibrium in which no matrilineal constraint exists, where the parent's exogenous land endowment $L > 0$ is automatically passed to the child upon the parent's death. The amount of land inherited by the child is thus $L^c = L$, which can be considered also the endowment of the child.

The parent's optimization problem can thus be written as:

$$\max_{C^p,E} v\left(C^p; C^c\right)$$

$$\text{s.t.} \quad C^c = Y^c(E, L^c)$$
$$Y^p = C^p + pE$$

After substituting the budget constraint into the parent's utility, we derive the following first order condition:

$$E: \quad v_{C^p}\left(C^p; Y^c(\cdot)\right)(-p) + v_{Y^c}\left(C^p; Y^c(\cdot)\right)\frac{\partial Y^c(\cdot)}{\partial E} = 0 \qquad (3.1)$$

Let us assume that the conditions for the existence of an interior solution are satisfied and focus on this case. We refer to this as the 'unconstrained' solution, and we denote the equilibrium values of the above problem as C^{p*}, $E^* > 0$, (with $L^c = L$). As discussed above, according to traditional matrilineal principles (i.e. pre-reform), there is no systematic transfer of land rights from a man to his child upon the man's death. Inter-vivos gifts from parent to child are allowed provided that the matriclan formally approves, and in any case can entail only a limited portion of the man's land, while the rest goes to the matriclan. In terms of the model, the matrilineal constraint can therefore be represented as an upper bound on the amount of land that the parent can pass on to his child:[8]

$$L^c = \overline{L_1^c} < L$$

The Intestate Succession Law of 1985 included a provision that a given fraction of the property should go to the man's children and substantially decreased the share going to the matriclan. In our framework this reform can be modeled as an increase in the upper bound of the land that can be allocated to the child, e.g. to $\overline{L_2^c} \in (\overline{L_1^c}, L)$.

We now derive the sign of the change in the optimal amount of E^* implied by an exogenous change of L^c using the implicit function theorem. The sign of $\frac{\partial E}{\partial L^c}$ is unambiguously negative if the following sufficient (not necessary) condition hold: $\frac{\partial^2 Y^c}{\partial E \partial L^c} \leq 0$, which indicates that the returns to education are non-increasing in the amount of land.[9]

Under the assumptions made on the child's production function and the parent's utility function, the equilibrium allocation with the matrilineal constraint implies a higher amount of education for the child compared to the unconstrained case: $E_1^* > E^*$, (with $L^c = \overline{L_1^c}$). This equilibrium represents a situation in which parents 'overinvest' in their children's education because they are restrained in the amount of physical capital they can devolve on them. It is now easy to see that the reform of matrilineal inheri-

[8]We assume that the parent derives no utility from land going to other kin members upon his death.

[9]We are not aware of any evidence explicitly testing the former assumption, although Kingdon and Soderbom (2008) show that returns to education in the agricultural sector in Ghana are very low.

tance would lead to a decrease in schooling investment for the child (compared to the pre-reform situation). The corresponding equilibrium allocation will involve a lower level of education, $E_2^* < E_1^*$ (with $L^c = \overline{L_2^c} > \overline{L_1^c}$). In other words, the effect of the reform is to bring the parent closer to his unconstrained optimal allocation.

Testable predictions

Despite its simplicity, our basic theoretical framework has several interesting implications. To frame these implications in the context of the empirical analysis that we conduct for Ghana, we refer to a group that follows matrilineal descent principles (Akan) and a group that does not (non-Akan).

Prediction 1: Ceteris paribus, after the reform Akan parents should invest relatively less in their children's education compared to non-Akan parents (for whom L^c in unchanged).
Prediction 1 is the key one for our paper, and will be tested through a difference-in-difference estimation strategy. To assess the plausibility of the identification assumption, we shall look at pre-reform trends for Akan and non-Akan.

Prediction 2: Prediction 1 should apply to landed households, for which the reform actually changes the budget set, and not to landless ones.

3.3.2 Extension: one son and one daughter

Now we extend the model and consider a case in which the parent has two children, a son and a daughter, and allocates resources between own consumption and investment in education of the son (E^s) and of the daughter (E^d). Throughout the analysis, we use superscripts p, s and d to denote parent, son and daughter, respectively.

As in the general case, the parent is endowed with an exogenous amount of land L. Consistent with the evidence described in the previous section, we assume that land inheritance is gender-linked, i.e. land is controlled by the male parent who can pass it on to his son upon his death (or to other male kin members depending on the specific inheritance rule). The amount of land that the son inherits is denoted by $L^s \leq L$, that depends on the prevailing social norm.

The daughter's rights are instead very weak and limited to the use of the land controlled by the husband or other male relatives. Accordingly, we denote the amount of land that can be used by the daughter as L^d.

The child's preferences are defined as before, $U^i = u(C^i)$, where $C^i = Y^i(E^i, L^i)$, and $u'(\cdot) > 0$, $u''(\cdot) < 0$, $Y^i_E(\cdot) > 0$, $Y^i_L(\cdot) > 0$ (and decreasing marginal returns), with i=(son, daughter). L^s is the fraction of father's land going to his son, and L^d is the land that can be used by the daughter.

The parent's utility now depends on own, son and daughter's consumption (or income): $V^p = v\left(C^p; Y^s(E^s, L^s); Y^d(E^d, L^d)\right)$, with partial derivatives $v_{C^p}(\cdot) > 0$, $v_{Y^s}(\cdot) > 0$, $v_{Y^d}(\cdot) > 0$, $v_{C^p,C^p}(\cdot) < 0$, $v_{Y^s,Y^s}(\cdot) < 0$, and $v_{Y^d,Y^d}(\cdot) < 0$. Moreover, we assume that $v_{C^p,Y^s}(\cdot) = v_{C^p,Y^d}(\cdot) = 0$.

The exogenous parent's income Y^p has to be allocated between the parent's own consumption C^p, and expenditure on son's and daughter's education, $p_s E^s$ and $p_d E^d$, respectively. The price of education is assumed to differ across genders.

The parent's optimization problem in the case that $L^s = L$ (ie, no matrilineal constraint exists) can be written as:

$$\max_{C^p,E^s,E^d} v\left(C^p;Y^s(E^s,L^s);Y^d(E^d,L^d)\right)$$
$$\text{s.t.}\quad Y^p = C^p + p_d E^d + p_s E^s$$

After substituting the budget constraint into the parent's utility, we derive the following first order conditions:

$$E^s: \quad v_{Y^s}\left(C^p;Y^s(\cdot);Y^d(\cdot)\right)\frac{\partial Y^s(\cdot)}{\partial E^s} - p_s v_{C^p}\left(C^p;Y^s(\cdot);Y^d(\cdot)\right) = 0 \tag{3.2}$$

$$E^d: \quad v_{Y^d}\left(C^p;Y^s(\cdot);Y^d(\cdot)\right)\frac{\partial Y^d(\cdot)}{\partial E^d} - p_d v_{C^p}\left(C^p;Y^s(\cdot);Y^d(\cdot)\right) = 0 \tag{3.3}$$

Taking the ratio of (3.2) and (3.3), we obtain the following equilibrium condition:

$$\frac{v_{Y^s}(\cdot)}{v_{Y^d}(\cdot)} = \frac{p_s}{p_d}\frac{\partial Y^d(E^d,L^d)/\partial E^d}{\partial Y^s(E^s,L^s)/\partial E^s} \tag{3.4}$$

This condition states that the ratio of resources invested in son and daughter's education reflects differences in returns to education and relative prices of education across genders. We assume that the conditions for the existence of an interior solution are satisfied and derive the 'unconstrained' solution. The equilibrium values of the above problem can be denoted as C^{p*}, $E^{s*} > 0$, $E^{d*} > 0$ (with $L^s = L$).

The matrilineal constraint is defined as an upper bound to the transferable land, $L^s = \overline{L_1^s} < L$, and the Law is represented as a relaxation of this constraint, ie $\overline{L_2^s} \in (\overline{L_1^s}, L)$.

We now differentiate the first order conditions and the budget constraint with respect to L^s to derive the sign of the change in the optimal amount of E^{s*} implied by an exogenous change of L^s. The sign of $\frac{\partial E^s}{\partial L^s}$ is unambiguously negative if the following sufficient (not necessary) conditions hold: $\frac{\partial^2 Y^s}{\partial E^s \partial L^s} \leq 0$, which indicates that the returns to education are non-increasing in the amount of land, and $v_{Y^s,Y^d}(\cdot) = 0$, meaning that there are no complementarities between son and daughter's consumption.

Therefore, the equilibrium allocation with the matrilineal constraint implies a higher amount of education for the son compared to the unconstrained case: $E_1^{s*} > E^{s*}$, (with $L^s = \overline{L_1^s}$), while the allocation after the Law implies a lower level of education, $E_2^{s*} < E_1^{s*}$ (with $L^s = \overline{L_2^s} > \overline{L_1^s}$).

The effect of the reform on daughters' education ($\frac{\partial E^d}{\partial L^s}$) is ambiguous: it may be positive due to a relaxation in the household's budget constraint, but this depends on the shape of the utility function.[10] Whether the effect of the reform on girls is zero or not is ultimately an empirical question, but under realistic assumptions the main effect of the reform should be seen on boys, as before and after the reform males were the main recipients of land bequests.

Additional testable predictions

Compared to the predictions outlined in section 3.3.1 here we further distinguish, within the Akan group, between a population that according to customary norms could not inherit paternal property

[10]Depending on the parent's preferences for own consumption and investments in children's education, the additional resources made available by the reduction in E^s can be reallocated differently. Therefore, the effect on daughter's education is ambiguous.

but that acquires this right after the passage of the reform (male children) and a population that typically does not have control over land, which is unchanged after the reform (female children). Within Prediction 1 in section 3.3.1, we can thus distinguish among the children's gender.

Prediction 1a: Ceteris paribus, after the reform Akan parents should invest relatively less in their sons' education compared to non-Akan parents.

Prediction 1b: The effect should be ambiguous for daughters. Under plausible assumptions we expect either an increase in education or no change when we look at the difference between Akan and non-Akan females after the reform compared to before.

Prediction 2 applies only to Prediction 1a in that we expect a reduction in education for Akan males in landed households, for which the reform actually changes the budget set, and not to landless ones.

Our stylized model also has implications related to the levels of human capital investments in the pre-reform equilibrium allocations. In particular, there should be differences in the pre-reform education levels across ethnic groups and genders, but we only view this as a consistency check since unobserved characteristics other than the matrilineal inheritance norm may be driving these initial gaps. Our key identification strategy relies on changes in allocations before and after the reform across different groups, as highlighted in the above predictions.

3.3.3 Discussion

The above framework is extremely stylized and involves a set of simplifying assumptions. However, it has the advantage of capturing the essential workings of the reform in the most parsimonious way. In this section we briefly discuss alternative modeling assumptions and the effects they may have on our results. While our simple model is static, one could extend it to a dynamic setting using an overlapping generations framework in which each individual lives for two periods. In such a setting, an individual's income (and consumption) would depend on his/her stock of human and physical capital, and each generation of parents would choose optimal amounts of investment in education.[11] This problem would have a recursive structure, and the equilibrium without matrilineal constraint would entail a lower level of education compared to the 'constrained' equilibrium, similar to the static model.

A simplifying assumption of our setting is that children can only inherit land from their parents. We know that in real matrilineal systems males inherit lineage land from their maternal uncles, and this may be cause of concern because the reform of matrilineal inheritance should affect uncles' behavior as well as parents'. We can explore the implications of a richer model, one in which each generation can inherit both from the father and from the uncle, in a qualitative way. Consider first a setting where each young individual has exactly one parent and one uncle, and where the uncle has one child of his own. Assume also that the amount of land owned by the parent and the uncle is the same. In this setting the matrilineal reform would have no effect on equilibrium allocations, because the increase in the land that a parent can pass on to his child after the reform would perfectly compensate the decrease in the land received by the uncle (who, in turn, is devolving more land

[11] In the static model, we have assumed that the parent's income does not depend on human capital, as this is equivalent to assuming that it depends on an exogenous (pre-determined) level of human capital.

to his own child).

In a realistic model of the matrilineal system, however, it may be preferable for children to inherit their own father's land rather than their uncle's for several reasons. First, this may be due to land-specific investments made by the children while working for their father, quality and quantity of the father's and uncle's land, or to the security of property rights. Bruce and Migot-Adholla (1994), for example, argue that rights are better enforceable when the land is inherited from the father than when it is allocated by the matrikin. Second, one should capture the fact that, even if the child has a maternal uncle (which obviously depends on the realized sibling composition of the child's mother), fathers do not have full control over the land the child will inherit. This may lead the father to invest more in his son's education to compensate for the uncertain amount of physical capital the child will be endowed with in the future. Indeed, the fact that the customary matrilineal norm would lead to impoverishment of those children who had no maternal uncles was one of the motivations for the Ghanaian legal reform that we examine.

A way to capture this in the model would be to assume that there is uncertainty about the probability to inherit the uncle's land. In this case the reform would still have a negative impact on education investment, which is what we test in the data. Ideally, we would like to have data on the composition of the entire extended family (notably, the existence of a maternal uncle) and on land owned by each member. This information is not available in the GLSS, nor in the Demographic and Health Survey (DHS). Hence we cannot control for inheritance received by the uncle, nor assess the extent of substitutability between father's and uncle's land.

3.4 Empirical strategy

3.4.1 Years of schooling and completion

To study the impact of the reform on investment in education, we compare educational outcomes for two age groups: those aged 0–17 (the most exposed to the reform) and those older than 18 (whose education should not have been affected by the Law) when the Law on intestate succession was passed in 1985. We restrict the sample to individuals 20 to 50 in each survey round. The group of children aged 0–17 includes children that are in primary school (or will enroll in primary), children in junior secondary/middle school, and those who are completing senior secondary school. In Ghana children normally enroll in primary school when they are 6 and complete it when they are 11.[12,13]

Basic specification

Our basic estimation strategy exploits differences across cohorts and ethnic groups. We analyze the impact of the Law separately for males and females. The basic difference-in-difference regressions

[12] Even though the minimum age required to start primary is six, late enrollment is not uncommon as there are many children starting primary at 7 or 8 (White, 2005). This obviously also affects all successive education levels. In terms of our empirical strategy, this would lead to an underestimation of the effect on education because also individuals in the 18–30 age bracket would be partly treated. Because of this choice, our estimates can be considered conservative.

[13] As far as secondary education is concerned, Ghana embarked on a reform of the education system in 1987 that essentially reduced the duration of pre-university education. Before the 1987 education reform, the school system was structured as a 6+4+7 system: 6-year primary, 4-year middle, 5-year secondary plus 2 years in preparation for entering tertiary education. The reform replaced the 6+4+7 with a 6+3+3 system with 6-year primary, 3-year junior secondary and 3-year senior secondary.

is as follows:

$$y_{itkrv} = \beta_1(A_i \cdot P_{it}) + \beta_2 A_i + \beta_3 X_{itkrv} + \beta_4 C_v + \beta_5(X_{itkrv} \cdot P_{it}) + \\ + \beta_6(C_v \cdot P_{it}) + \alpha_r + \mu_k + \nu_t + \gamma_r \cdot t + \epsilon_{itkrv} \quad (3.5)$$

where $y_{i,t,k,r,v}$ is number of years of schooling or the highest level completed of an individual i, born in year t, and observed in survey year k, in region r, and in village v. A_i denotes individuals belonging to the Akan ethnic group. α_r, μ_k, and ν_t are region, wave, and birth year fixed effects, respectively. Since Akan are spatially concentrated in the Southern and Western part of Ghana and grow different types of crops than the non-Akan, we also include is a set of nine dummies, denoted with C_v, for the major crop grown at the village level (among which: cocoa, cassava, maize, yam, tomato, plantains, beans/peas, groundnut/peanut, and pepper).[14]

We also include a region specific linear time trend, $\gamma_r \cdot t$, to capture region and cohort-specific effects that may be correlated with the error term, e.g. variation across regions and over time in the supply of education. $X_{i,t,k,r,v}$ is a set of individual and household-level covariates observed at the time of each survey including: age and age squared, age of the household head, household size, a principal component index of durable goods owned by the household, a dummy for female headed households, mother's and father's ed-

[14]This information is retrieved from the rural community questionnaire. In particular, the survey asks the respondents for the village the following question: "what are the major crops grown by the people of this community?". Up to 9 crops can be named. Unfortunately, in survey rounds 3 and 4 it's not possible to know the ranking of the crops. Thus, each dummy is equal to one if the specific crop is among the ones mentioned in the community questionnaire as being among the major grown in the village. Finally, among all crops, the first 9 most mentioned were considered.

ucation, and religion of the head.[15]

The variable $P_{i,t}$ indicates the post-reform period, which is constructed based on the individual's age when the Law was passed in 1985. We use two alternative definitions of $P_{i,t}$ which take into account the structure of the education system. The first definition is a dummy equal to one if the individual was born in 1968 or after (aged 17 or younger in 1985). Since those born in 1968 belong the first cohort being affected for the completion of secondary school, we can define this cutoff as "treated by the end of secondary". The second cutoff birth year is 1974. Strictly speaking, children born in 1974 were completing the last year of primary school in 1985 (aged 11 or younger in 1985). Accordingly, we define this cutoff as "treated by the end of primary". For these children, the Law did not only affect the decision of completing a given level of education, but also the decision of enrolling in secondary school. We expect that educational outcomes should be most affected for children who were at an earlier stage of education when the Law was passed.

We also include the interactions between the individual characteristics and $P_{i,t}$ to take into account the possibility that the effect of these variables has changed over time and that this might correlate with the error term.[16]

Moreover, we add the interactions between major community crops and the post-reform dummy to control for the possibility that our results may be confounded by movements in returns to

[15]The durable goods that enter our principal component index are: radio, tape-player, television, sewing machine, refrigerator, air conditioner, bicycle, motor cycle, and car. The religion of the household head include Catholic, Protestant, other Christians, Animist, and Muslim.

[16]The characteristics interacted with the post-reform dummy are the following: age of the household head, household size, a principal component index of durable goods owned by the household, a dummy for female headed households, parental education, and religion of the head.

education for Akan compared to non-Akan males due to changes in agricultural production (ie, by changes related to the production of a specific crop) over the period considered.

In regression (3.5), our coefficient of interest is β_1, the coefficient of the interaction term between A_i and $P_{i,t}$. In line with our theoretical predictions we expect $\beta_1 < 0$, i.e., compared to non-Akan males of the same age, Akan males exposed to the reform should have significantly less years of schooling. The same should not hold (or should hold to a lesser extent) for females.

The regressions for years of education are estimated with simple OLS. To estimate the impact on highest grade completed we instead use a non-linear probit model for the probability of completing at least primary or secondary school. We also estimate an ordered logit model for the highest grade completed, which is a categorical dependent variable for the different educational levels completed.

Our identification strategy assumes that, conditional on the controls we include, changes in education outcomes for Akan and non-Akan males would have been the same in the absence of the reform. Below we discuss the plausibility of this assumption examining pre-trends and conducting some falsification tests.

Cohort-specific effects

The specification (3.5) allows us to compare educational outcomes across ethnicities for two broad age groups: those aged 0-17 and those older than 18 when the Law was passed in 1985. We can go further in understanding when the reduction of investment in education started and evolved by estimating year-by-year differences as follows.[17]

[17] We do not include the interactions between the observables and the post-reform period because of the small sample size at each bin.

$$y_{itkrv} = \sum_{l=1960}^{1985} \beta_{1,l}(I_l \cdot A_i) + \beta_2 A_i + \beta_3 X_{itkrv} + \beta_4 C_v + \alpha_r + \\ + \mu_k + \nu_t + \gamma_r \cdot t + \epsilon_{itkrv} \quad (3.6)$$

where I_l is a birth-year dummy, and all the other variables are defined as in (3.5). The omitted group is formed by individuals born from 1937 and 1959 (which represent about 34 percent of all individuals in our sample).[18] Our coefficients of interest are the $\beta_{1,l}$ that estimate the year by year difference in education across ethnicities compared to the control group of those born before 1960.[19] We expect the $\beta_{1,l}$ to be equal to zero for $l < 1968$ and become negative after 1968. If the reform mainly affected the decision to enroll in secondary school (while those who were already in secondary school in 1985 completed that level regardless of the Law), then we expect the effect to become negative and significant for the individuals born in 1974 or after (who were in primary school in 1985).

3.4.2 Current attendance rates

After having estimated the impact of the reform on our measures of "accumulated human capital", we turn to the analysis of the effects on current attendance rates.[20] In order to do this, we select indi-

[18] Given our age restriction on the sample of individuals 20 to 50, the oldest individuals are in fact those aged 50 in the first wave of the GLSS in 1987.

[19] Mean bin size in the period 1960-85 is 550 individuals.

[20] The definition of attendance rates used in this paper, the same used by White (2005), is the fraction of children who are supposed to be enrolled in a particular school level for their age who are currently in school. For example, for primary (secondary) school age children, it is the fraction of the number of children aged 6–11 (12–17) currently in school. Therefore, this is an age-based

viduals aged 6 to 17 from each of the five rounds of the GLSS, and compare attendance rates for those observed if the first two waves (GLSS 1987/88 and 1988/89) to those observed in later rounds of the survey (GLSS 1991/92, 1998/99 and 2005/2006).[21] Since the first two waves were carried out already after the passage of the Law in June 1985, the treatment is not very accurate. Ideally, we would like to have information on attendance rates before 1985, but this is data is not available in GLSS. However, it is likely that it took some time before the reform was fully operative, so we can consider that parents internalized the effects of the reform more in rounds 3 to 5 of the GLSS. If the reform led to an sudden reduction in attendance levels in 1985, which is somewhat unlikely, our results would *underestimate* the true effect.

We use a probit model to estimate the probability of being currently in school in the post reform period for individuals belonging to the Akan versus non-Akan ethnic group. We run a regression similar to (3.5), where edu_{itkrv} is a dummy variable equal to one if the individual is currently in school and zero otherwise. P_{it} is defined as a dummy equal to one if the individual was observed in GLSS wave 3, 4 or 5, and zero otherwise. We also estimate interaction terms with each individual survey round to show that the effect is gradual and not dependent on the particular cutoff. We include wave fixed effects but not birth year fixed effects. All the other control variables are the same as in (3.5). We also run a regression with a triple interaction between $akan * post * age1217$ to understand if the effect on attendance is different for primary and secondary age school children (eg, aged 6–11 and 12-17, respec-

measure not specifically related to the actual level attended.

[21]For the analysis of attendance rates, we restrict the sample to children of the household head.

tively). We expect to find a smaller effect on attendance of Akan boys aged 6–11, while a significant reduction for those in the 12–17 cohort ($\beta_1 < 0$). This is mainly because Akan males had very high primary school completion rates already in the pre-reform period. If primary school is considered by the Akan as the basic or necessary education level a male should complete, the effect should be mainly at higher levels than primary.

3.5 The data: a descriptive analysis

The theoretical predictions illustrated in the previous sections will be tested using individual-level data from all five rounds of the Ghana Living Standard Survey: GLSS1 (1987/88), GLSS2 (1988/89), GLSS3 (1991/92), GLSS4 (1998/99), and GLSS5 (2005/06). Because matrilineal inheritance norms essentially apply to the allocation of land, we restrict our attention to the rural subsample of the GLSS.

We analyze the effects of the Intestate Succession Law on two main educational outcomes: the "stock of schooling" (eg., number of years of education and highest grade completed), and current attendance rates. For the part of the analysis of the impact on accumulated schooling we focus on individuals aged 20 to 50, and test the theoretical predictions first using the full sample and then using the sample of individuals whose father is/was a farmer. Having a father farmer is a proxy for whether the father owned land at the time when his children (who are the individuals included in our sample) were of school age. This variable is preferred to the one that can be derived from the question "Does any member of the household own any land?" for two reasons. First, since this information refers to land owned by the household at the time of the survey, we do not know whether this land was already owned at the time when the individual was in school. Second, if the father of the individual is not a household member, there is no question asked about land ownership of the father. However, while having a father who was a farmer is a 'lagged' proxy for land ownership (which is what we need for our analysis), it is a very imperfect proxy: being a farmer does not necessarily mean owing land.

In the second part of the empirical analysis, when we consider current attendance rates of younger cohorts (aged 6 to 17), we present evidence for the full sample and for the subsample of landed households, defined as households who answer "yes" to the above question. In this case the contemporaneous land ownership status of the household is the appropriate measure because the population under study is that of individuals younger than 18, most of whom are still living with their father in the interviewed household. We also run the attendance regressions on the subsamples of individuals with and without a father farmer to provide support to the validity of our proxy for land ownership.

For the first part of the analysis on accumulated human capital, we use all five rounds of the GLSS and restrict the sample to individuals aged 20–50 in rural areas. The restricted sample includes $22,156$ individuals. Summary statistics for the main variables of interest are shown in panel A of Table 3.1. Summary statistics for the other variables used in the regressions as well as a breakdown by ethnicity and gender are shown in appendix table 3.19.

The percentage of Akans (mostly Asante and Fanti) in our rural sample is 42 percent.[22] The other (patrilineal, and classified as

[22]In GLSS1-3 the criterion according to which an individual is classified as Akan is the language of the household head. In GLSS 4 and GLSS 5 Akan is instead defined according to the ethnicity of the household head, due to changes in the questionnaire. The percentages are similar across rounds.

non-Akan) groups are the Ewe, Mole Dagbani, Ga-Dangme and others. Inter-ethnic marriages among Akans and non-Akans tend to be rare, especially in rural areas. In the 2003 Ghana Demographic and Health Survey (DHS) the percentage of unions in which the household head and his/her spouse belong to different ethnic groups is 7.2 percent of all unions, while it is 5.7 percent in rural areas.[23] The age at first marriage is similar across ethnic groups: in the rural sample of the 2003 Ghana DHS the median age at which females got first married is 18, while it is 23 for males, both for Akans and not.[24] Average years of education in the sample is equal to 4.82, which varies widely across ethnicity and gender. Akans males in our sample have on average 8.61 years of education, compared to 4.93 for non-Akan males. For females, the corresponding figures for Akan and non-Akan are 5.02 and 2.36, respectively. Consistently, a higher fraction of Akan have attained primary or higher levels of education.[25]

As far as our proxy for land ownership is concerned, we see that 79 percent of the individuals in our sample have a father who is/was a farmer: this percentage is higher among the non-Akan (84 and 85 percent for males and females, respectively) than among the Akan (72 percent).

The remaining part of Table 3.1 shows the summary statistics for the sample of younger cohorts used in the second part of the

[23] The percentages obtained when using the other rounds of the Ghana DHS surveys (1993, 1998, and 2008) and the fourth and fifth round of GLSS (it is not possible to use the earlier waves for this purpose because only the language/ethnicity of the head is known, and not that of all household members) are similar.

[24] The median age at first marriage is somewhat constant across the waves of the DHS surveys (1993, 1998, 2003 and 2008) for both ethnic groups.

[25] We constructed the variables years of education and attainment rates using data on highest grade and level completed.

analysis. Panels B.1 and B.2 refer to the sample of individuals aged 6–11 and 12–17, respectively. More descriptive statistics for the whole sample of those aged 6 to 17 are shown in appendix table 3.20. Attendance rates for Akan and non-Akan boys of primary school age are 88 and 65 percent, respectively. The percentage of girls going to school is lower than that of boys by 2 to 4 percentage points in both ethnic groups. In panel B.1 we see that the gap in attendance rates between Akans and non-Akans boys aged 12 to 17 is slightly lower than that among younger cohorts. Finally, we see that about 59 to 69 percent of individuals live in a household in which at least one member owns land. The fraction of those owning land is similar in all groups, somewhat higher for Akans.

3.6 Main results: years of education

Table 3.2 shows the simple difference-in-difference estimates of the effect of the Intestate Law on years of schooling. We consider mean years of education for four different age groups: 0–11, 12–17, 18–25 and 26–30. We expect that the 1985 Law affected parental educational decisions for children of primary or secondary school-age, but not for those older than 18 that had already completed secondary school in 1985. Each panel in the table reports education estimates separately for Akan and non-Akan individuals (columns), in different age brackets when the reform was implemented (rows), and by gender (leftmost versus rightmost part of the table). For example, in panel (a) we see that Akan males who were aged 0 to 11 in 1985 on average completed 8.57 years of education, while non-Akan males in the same cohort completed 5.74. On the other hand, for the older cohort who was 18 to 25 years old when the reform was passed, completed years of education average 8.86 for Akan males and 5.11 for non Akan males. Our diff-in-diff estimate of the effect of the reform is thus -0.92 with a standard error of 0.31, indicating that the reform induced Akan parents to give on average 0.92 less years of education to their male children compared non-Akan children in the same age group. When we repeat a similar exercise for females (rightmost part of Panel a), we find no statistically significant effect of the reform (if anything, the effect is in the opposite direction). In panel (b) of table 3.2 we estimate the effect of the reform on a group of children who were in secondary school age when the reform was implemented (i.e., who were 12 to 17 years old in 1985), keeping as control group cohorts who were 18 to 25 in the same year. We find no effect of the reform on Akan males and females. This might suggest that individuals who were already enrolled in secondary school when the reform was introduced continued going to school and completed that level. Thus, the effect of the reform seems to be mainly concentrated on those who were at an earlier stage of their education and had to decide whether to enroll in secondary.

Finally, in panel (c) of table 3.2 we conduct a falsification test, comparing individuals aged 18 to 25 in 1985 to individuals aged 26 to 30. Neither group should be affected according to our theory, hence if we found a significant difference in the same direction when comparing these two groups this would suggest that our effects may be spurious. As can be seen in panel (c), no significant difference emerges when we conduct this placebo test, which increases our confidence in our identification strategy.

As a robustness test, we repeated the same exercise on the subsample of individuals whose father is (or was) a farmer. According to our theory, we should expect a greater impact on education for individuals whose father owned land at the time when they made educational choices for their children. The results are reported in Appendix table 3.14 and table 3.15. The estimated effects in table 3.14 are consistent with those in table 3.2, and higher in magnitude in panel a and b. We also find a positive statistical effect for females aged 0-11. As expected, in table 3.15 we do not find any significant effect for the sample of individuals whose father is/was not a farmer.

We expect that the 1985 Law should have affected parental educational decisions for children of primary or secondary school-age, but not for those older than 18 who had already completed secondary school in 1985. Moreover, the negative effect on education should be stronger the younger individuals were when the Law was introduced. The results of the estimation of regression (3.5) are shown in table 3.3. We run separate regressions for males (columns 1 and 2) and females (columns 3 and 4) using the full sample and the two birth year cutoffs: $post68$ and $post74$. We expect to find a negative and significant coefficient for the interaction term $akan^{*}post$ for males, but no effect (or a smaller effect) for females. The negative effect should be stronger for individuals born in 1974 and after because their decision of enrolling (not just completing, as for those born in 1968 or after) in secondary school was affected. Thus, $akan^{*}post74$ should be bigger in absolute terms than $akan^{*}post68$.

Table 3.3 confirms the descriptive evidence that individuals belonging to the Akan group have an initial education advantage. This advantage is about 1.6 additional years of schooling for males and approximately one year for females. In column 1 of table 3.3 the coefficient of $akan^{*}post68$ is equal to -0.658 and is significant at the five percent level. This means that, compared to non-Akan males, ceteris paribus Akan males exposed to the reform (born in 1968 or after, thus aged 0-17 in 1985) experienced a 0.658-year reduction in the number of years of completed education. As expected, the coefficient of $akan^{*}post74$ is bigger and equal to -0.918. Columns 3 and 4 of table 3.3 show the estimates for females. As expected, after the reform, there is no significant change in years of education for Akan females with respect to non-Akan females in the same cohorts. If anything, the change is positive (the coefficients on $akan^{*}post$ are 0.075 and 0.128, none of them significant). Among the other regressors, we see that both mother and father's education are strong predictors of educational outcomes of their children.

In table 3.4 we estimate the effect of the Law on the sample of individuals whose father is/was a farmer. The negative effect on education for Akan males should be stronger for those with a father farmer, which is our proxy for land ownership of fathers at the time when individuals were in school, and no effect should be found for the subsample of individuals whose father is/was not a farmer. In column (1) and (3) we see that the coefficient of both the interaction terms $akan^{*}post68$ and $akan^{*}post74$ are bigger in magnitude when restricting the sample to those with father farmer: the corresponding reductions are 0.672 and 1.115 years of schooling, significant at the five and one percent level, respectively. Columns (2) and (4) in table 3.4 show that for the subsample of individuals whose father is/was not a farmer the coefficients of the interaction term are not significant, which is consistent with our theoretical prediction that the reform affected educational decisions only for fathers who owned land. Again, we do not find any significant effect for females.

We next turn to the results of the interaction term analysis. Figure 3.2 shows the results for males using the full sample, and the subsamples of males whose father is/was a farmer and not.

The graphs plot the coefficients of the interaction term (regression equation 3.6) with 95 percent confidence bands. Average years of schooling are displayed along the vertical axis and year of birth along the horizontal axis. According to our theory, we expect to observe a similar trend in education for the cohorts born before 1968 and a reduction of the difference in years of education between Akan and non-Akan males born in 1968 or 1974 onwards. Cohorts born in 1968 or after should have been increasingly affected by the Intestate Law as there was more and more scope for parents to adjust their schooling investment choices.

Starting from the top, the first panel of figure 3.2 shows that before 1968 Akan and non-Akan males share a parallel trend: this is an important piece of evidence to support our identification strategy. Furthermore, the gap in years of education between Akan and non-Akan males starts closing for cohorts born after 1968, it remains somewhat constant until birth year 1973, and then it significantly decreases for the cohorts born in 1974 and after. Consistently with our predictions, the middle and bottom panels of figure 3.2 show that Akan individuals whose father is/was a farmer experienced a sharp decrease in education, while there is no effect for those whose father was not a farmer (even though estimates are much less precise due to small sample size for this subsample). Figure 3.3 in the appendix shows the results for females. Contrary to what happens for males, the gap between Akan and non-Akan remains fairly constant between cohorts affected and not affected by the reform. This is consistent with the fact that the Intestate Law effectively relaxed the transmission of property from father to son, while girls kept inheriting their mothers' land which was either nonexistent or "unconstrained" to start with.

3.6.1 Alternative interpretations and robustness checks

We next run several robustness checks. Our first robustness check deals with the fact that, compared to non-Akan, Akan have a higher initial level of education. This generates a concern that our results may be driven by mean reversion. It is first worth noting that, as shown in figure 3.2, there is no differential trend in education between Akan and non-Akan males in the pre-reform period. This supports the idea that there was no convergence in education for cohorts born before 1968 who were too old to be affected by the reform. Moreover, we have shown that there is no decrease in years of education for Akan females (rather, a slight increase) even though there is about one year of education gap between Akan and non-Akan females.

To further check whether our results might be affected by convergence, we conduct an additional test in which our control group is composed by the Ewe, a sub-group of the non-Akan characterized by a higher level of education. In our sample of individuals age 20–50, the Ewe constitute about 25% of all the non-Akan in the sample and are mainly concentrated in the Eastern and Volta regions of Ghana. Their average number of years of education is much higher than the rest of the non-Akan (Ewe males have on average 7.61 education years compared to 4.03 for the rest of the non-Akan) which is closer to the Akan level (8.61). Table 3.5 shows the results. We can first note that the initial education advantage of the Akan with respect to the Ewe is smaller than the gap between Akan and non-Akan, and ranges between 0.8–1 years of education, both for males and females. Compared to the Ewe, Akan males experience about a 1.1 and 0.8 years of education reduction when using the *post68* and *post74* cutoffs, respectively. These coefficients are statistically significant and not too different from the results in our benchmark regression in table 3.3. This provides further evidence that our results are not driven by a convergence effect.

Second, we consider the fact that the Akan are concentrated in the Southern and Western regions of Ghana. Although all our regressions include region fixed effects and region-specific linear trends, in panel a of Table 3.6 we estimate our years of education regression excluding from the sample the individuals in the Northern regions. We find that for males the coefficients of the interaction term of our interest are very similar in magnitude to the results in table 3.3 (-0.682 compared to -0.658 for the 1968 cutoff and -0.886 compared to -0.918 for the 1974 cutoff), and significant at 5 and 1 percent level. As before, we find no significant effect on the subsample of females. This evidence confirms that our results are not driven by omitted time-varying factors that are specific to the regions where the Akan are spatially concentrated.

Third, we check whether our results may be confounded by changes in returns to education for Akan compared to non-Akan males due to changes in agricultural production over the period considered. Akan and non-Akan are spatially concentrated in different regions where also different types of crops are grown (see summary statistics for the 9 major crops grown at the village level in table 3.19). In all our regressions, in addition to control for the major crops grown at the community level, in order to check whether our results are affected by changes related to the production of a specific crop, we include the interactions between major community crops and the post-reform dummy to account for this possibility. We also consider the fact that one characteristic of Akans is that they are highly represented among cocoa producers. In fact, cocoa is mainly grown in the Southern and Western parts of Ghana, the regions where Akans are concentrated. If returns to education in cocoa farming changed over this period, or if Akans switched into or out of cocoa farming differentially from other groups (and this changed their returns to education), this may generate a pattern of results like the one we find. We thus use the dummy '*cocoa major crop*' to distinguish between villages where cocoa is listed as being one of the major crops (*cocoa major crop*=1), and villages where cocoa is not grown (*cocoa major crop*=0). In panel b of table 3.6 we re-estimate our main regressions for years of education introducing a triple interaction term *akan*post* cocoa major crop*. If our results were entirely driven by changes in the returns to education associated with cocoa farming, we would expect a negative and significant coefficient on this triple interaction, and an insignificant coefficient on *akan*post*. We find that for the male sample (column 1 and 3) the coefficient on the interaction term of our interest is negative and significant (and close to our benchmark estimates). On the other hand the coefficient on *akan*post* cocoa major crop*, is insignificant. No significant effect is found for females (column 2 and 4). Overall, we interpret these findings as evidence that our results are not driven by changes in returns to education for cocoa farmers or by reallocation of Akans in or out of the cocoa sector correlated with the post-reform period.

A fourth check relates to changes in returns to education due to shocks to specific sectors in which Akan and non-Akan are differentially employed. Over the period 1987 and 1990 about 47500 civil servants were laid off as part of the Redeployment program (Alderman, Canagarajah, and Younger, 1993). Since Akans are relatively more employed in public sector jobs with respect to the non-Akan, the Redeployment program might have lowered investment in secondary education due to reduced access to the public sector jobs. This may constitute a shock specific to Akan in the post-reform period. To check whether this affected our results, we constructed a variable '*any member in government*' which is a dummy equal to one if at least one member of the household works for the government. The idea is to see whether the expectation of lower returns to education, proxied by the presence of a household member with a job in the public sector, is driving the reduction in investments in children's education. Since we have information on the type of employer only for resident members, we can do this check only for the regressions for attendance rates in which we use the sample of younger cohorts (aged 6 to 17) and we know the occupations and the employer for all resident members. Similarly to our robustness check for cocoa villages, in Appendix table 3.18, panel d, we re-estimate the attendance attendance regression introducing a triple interaction term *akan*post*any member in government*. In column (3) we see that the coefficient of *akan*post* is negative and statistically significant at 1 percent level (equivalent to a reduction in attendance of 12.4 percentage points for individuals living in households in which none of the members works for the government). The coefficient of the triple interaction *akan*post*any member in government* is instead equal to zero. For females, in column (4) we see that only the coefficient of *akan*post* is negative and significant while there is no additional effect for households with a government employee. These results support our idea that the reduction of investment in education has not been driven by households in which the father or any other member of the household had a job in the public sector.

Panel c of table 3.6 provides a further check that our results are not driven by factors related to changes in occupation in specific sectors among different ethnicities. We construct dummies for the type of mother's and father's occupation including: farmer, sales, clerical job, and professional (eight dummies). We also interact these eight dummies with the post-reform dummy and include these terms in our benchmark regression. The results in panel c of table 3.6 show that the inclusion of these variables does not affect our estimates, confirming our results.

3.6.2 Placebo test: Akan in Côte d'Ivoire

We next check whether shocks specific to the Akan ethnicity might be driving our results. We exploit the fact that Akan are also present in the neighboring Côte d'Ivoire, but were not affected by the Intestate Law introduced in Ghana in 1985.[26] They represent about the 34% of the population and are geographically concentrated in the Southern and Eastern regions, the regions closest to Ghana. As in the context of Ghana, all the ethnic groups other than the Akan are patrilineal.

We use the Côte d'Ivoire Demographic and Health Survey (DHS) for 1994 and 2005 and, as in the analysis conducted in the case of Ghana, we consider individuals age 20-50 in rural areas.[27] Average number of years of education for Akan males is 4.71, and 3.07 for non-Akan males (for females is: Akan 1.93, and non-Akan 1.31). Therefore, Akan are more educated than non-Akan but to a lesser degree than in Ghana. The gap in education between the two ethnicities is greater in the three neighboring regions to Ghana (North-East, Center-East, and South), where Akan males have on average 5.54 years of education, compared to 2.72 for non-Akan males. We focus on the full rural sample and subsample of those

[26] In Côte d'Ivoire, the 1964 Family Code regulated inheritance and designated the nuclear family (spouse and children) as the sole inheritors.

[27] We cannot use the 1998 DHS survey because the information on the region of residence is not available. The Akan ethnicity is from the woman and man individual datasets and merged with the household recode dataset.

living in North-East, Center-East, and South regions.

We regress the number of years of education for Akan and not in the Ghana post-reform period (born after 1968 or 1974), similar to (3.5).[28] The results are shown in table 3.7. When considering the sample with all regions, we see that Akan males actually have lower initial education than non-Akan males (columns 1 and 2), while if we restrict the sample to the three neighboring regions to Ghana in columns (3) and (4) we see that Akan males have about one-year of education initial advantage (similar to Ghana). In neither samples we can detect any effect on education for individuals born after 1968 or 1974, both for males and females. This evidence supports our setting in that the reduction in education found for Akan males in Ghana is not due to shocks specific to the Akan ethnicity, but are due to the 1985 reform of the inheritance system that affected only the Akan in Ghana.

3.7 Completion rates

Table 3.8 shows the probit estimates (the marginal effects) for the probability of completing at least primary or secondary school, by gender. The upper panel of table 3.8 shows that, when using the 1968 cutoff, the probability of completing at least primary is negative but not statistically different from zero (column 1). This is probably due to the timing of the reform: individuals born in 1968 or after were younger than 18 in 1985. In this age group, the older individuals had already completed primary school when the Intestate Law was introduced, so we expect a smaller effect on completion of primary in this group. Not surprisingly, Column (2)

[28]It is not possible to run this analysis for attendance because the first survey was carried out in 1994 and we would not have observations from the 'pre-reform' period.

shows that, compared to the non-Akan, Akan males born after 1974 are 9.7 percentage points less likely to complete at least primary school. Using the 1974 cutoff, we are considering individuals who were younger than 12 in 1985 and most likely to be affected for the completion of primary school.

The bottom panel of table 3.8 shows the results for the probability of completing at least secondary school.[29] Column (1) and (2) show that the Akan males are 8.3 and 11.4 percentage points less likely to complete secondary school or higher level, respectively. As expected, individuals born in 1968 or after (most of whom were treated for the completion of secondary) are less likely to attain secondary school and the effect is stronger for individuals born after 1974 who were the most exposed to the reform. Finally, columns (3) and (4) show the results for females. Consistent with the results obtained for years of education, we do not find any significant effect on female's education.

Table 3.9 shows the results for completion for the subsample of individuals with a father farmer and not. By comparing columns (1) with (2) and (3) with (4) for the 1968 and 1974 cutoffs respectively, we see that only those with a father farmer experienced a reduction in the probability of completing at least primary or secondary school, while there is no effect on those whose father is not a farmer. This is consistent with our theoretical prediction. The rightmost part of table 3.9 (columns 5 to 8) shows the results for females: again, we cannot detect any significant change.

As a further check, we also estimate an ordered logit model where the dependent variable is the highest level of education com-

[29]Completion of at least secondary school is defined as a dummy equal to one if the individual has completed either middle school or junior secondary (in the old and new education system, respectively) or higher level, and zero otherwise.

pleted, which has four categories: no education, incomplete primary, completed primary, and secondary or higher. The estimates are reported in the appendix table 3.16 and confirm our probit results in table 3.8. Appendix figures 3.4 to 3.7 show the graphs of the cohort analysis for the probability of completing primary or higher level for males and females, respectively (estimated with OLS and probit). The results are similar to those found for the cohort analysis done for the number of years of education. Finally, we conduct all the robustness checks done for the regressions for years of education also for the completion regressions. Results are shown in appendix table 3.17 and confirm our main results.

3.8 Attendance and sibling composition

We now analyze the effects of changes in matrilineal inheritance rules on current attendance rates for individuals of primary and secondary school-age. In the previous sections, we used the sample of individuals who had completed their education (aged 20 to 50) and studied the effects on their accumulated human capital. Here we want to understand how the reform affected parental choices regarding their children's current enrollment in school. To do this, we use the sample of all individuals 6 to 17-year-old in each survey round. Then we compare attendance rates of individual observed in wave 1 and 2, our control group, to those surveyed in wave 3, 4, and 5, our treatment group. We expect to find a reduction in attendance rates for Akan males surveyed in the last three rounds of GLSS (those most affected by the reform), compared to non-Akan males. Moreover, the effect should be stronger for individuals in landed households. We also expect that secondary school age individuals (eg., aged 12 to 17) should be the most negatively affected, since the vast majority of Akan males complete at least primary school.

Table 3.10 shows the results for attendance rates of individuals aged 6 to 17, by gender. In column (1) we see that the coefficient of our variable of interest, $akan^{*}post$, is negative and highly significant at the 1 percent level. The estimated coefficient is equal to -0.119, which means that the probability of being in school for those between 6 and 17 decreased for Akan males in the post-reform period by 11.9 percentage points. To see whether this reduction was stronger for Akan males of primary or secondary school-age (age 6–11 or 12–17, respectively), in column (2) we estimate a regression with a triple interaction term between $akan * post * agegroup1217$. The results suggest that, while there is weak evidence of lower attendance of primary school age, the reduction in attendance occurred mostly for Akan males aged 12–17. Column (3) shows that there was a significant reduction in attendance also for Akan females, which is weaker than that for males, but significant at the 5 percent level. Results in column (4) are similar to those obtained for males.

Panel a of table 3.11 shows the results for the subsample of landed households and not. We focus on the subsample of individuals aged 12–17, where the effect on attendance for Akan males is stronger. First of all, it should be noted that Akan males have a strong initial advantage in attendance rates compared to non-Akan males that is limited to landed households (by looking at the coefficient of *akan* in column (1) and column (3) for landed households and not, respectively). Moreover, again as expected, the reform had an effect on the probability of being in school only for individuals in households with land. This reduction is equal to 15.4 percentage points. This evidence strongly supports our theoretical setting. In order to see if this reduction was gradual and understand when it occurred, we break down the variable *akan*post* into the interactions between *akan* and dummies for the three most recent survey waves (wave 1 and 2 are taken as the omitted category).

Column (2) shows that the reduction was gradual, starting from wave 3 and becoming stronger and significant later waves. The negative effect of the reform on attendance rates for Akan males is very large for those in households with land. Column (4) shows that there is no effect whatsoever for landless households. Finally, column (5) to (8) report the estimates for females. Column (5) shows that there is a strong negative effect on attendance of females as well, even though the breakdown in column (6) suggests that this evidence is weak. Column (5) and (6) show that there is no effect for females in landless households. In panel b of table 3.11, we report the estimates for the subsample of individuals whose father is a farmer and not. The results are similar to those obtained in panel a. This strongly support the validity of our proxy for land ownership used for the analysis of the number of years of education and completion.

The next exercise is to look at whether there are heterogeneous effects of the reform on attendance rates depending on the number of siblings within the household.[30,31] If bequests are to be divided among all children, the higher the number of children, the smaller the fraction going to each of them. We expect that, after the reform, disinvestment in education should be lower for boys who have more siblings. Table 3.12 shows the results. In column (1) we see that the reduction in attendance is equal to 34.6 percentage points for Akan

[30]Unfortunately, we only have information about siblings currently *residing* in the household, and not about those who are living elsewhere. Therefore, we are underestimating the total number of siblings.

[31]The number of siblings is calculated taking into account all biological children of the household head.

males with no siblings. The coefficient on the triple interaction term ($akan^{*}post^{*}\#siblings$) is positive and significant at 1 percent level: it suggests that the negative effect on attendance is attenuated by 5.7 percentage points for each additional sibling.

In columns (2) and (3) of table 3.12 we explore whether this "attenuation effect" on attendance rates varies according to the age and gender composition of siblings. We expect that the attenuation effect is mainly driven by those who have older siblings. In fact, in relation to the order of inheritance according to traditional matrilineal rules we know that 'if more than one person qualifies to inherit the property of the deceased, age and achievement become other important criteria' (Awusabo-Asare, 1990, p. 7). Moreover, since land is typically controlled (and consequently, inherited) by males, we expect that the number of older male siblings should have a positive mitigating effect (in the post-reform period) on attendance for Akan boys. Column (2) of Table 3.12 shows that the coefficient on the $akan^{*}post^{*}(\#oldersiblings)$ is positive and significant. This is consistent with our expected result. We also find evidence that it is the number of older brothers that is important in mitigating the fall in attendance for Akan boys (column (3)), again confirming our prediction. Finally, no similar effects are found on female's attendance (columns 5 to 6). Although we cannot make any causal claim, this evidence is consistent with the fact that their paternal land should be shared among more heirs, leading to lower land per child, hence lower disinvestment in education.

Finally, we conduct all the robustness checks also for the attendance regressions. Results are shown in in appendix table 3.18 and support our main findings.

3.9 Occupational choice

We next analyze whether the reform also affected occupational choice. With the introduction of the 1985 Intestate Law, fathers acquired the possibility of passing land on to their children. In the previous sections, we have shown substantial evidence that this reform led to a reduction of investment in education. This is consistent with the fact that, for fathers, the newly acquired opportunity for their children of working their land as farmers (an activity in which the return to education are lower than in other sectors) outweighed the reduced opportunities of working in sectors requiring a higher level of education.

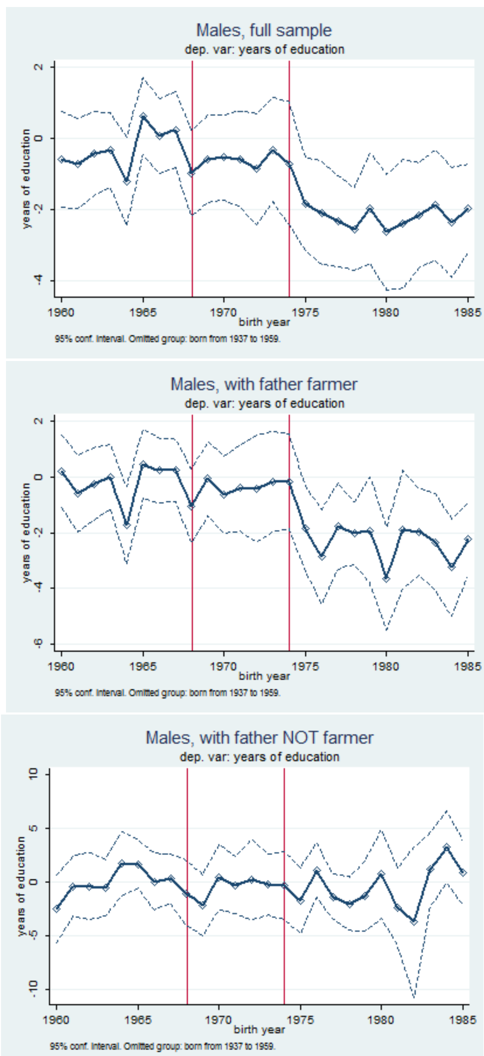
We analyze whether the probability of being a farmer differs between Akan and non-Akan males in the post-reform period. Using the sample of individuals age 20–50, we run a regression similar to the specification (3.5) used in the regressions for years of education and completion. The post-reform dummy is defined as equal to one in the last three rounds of the GLSS (similar to the analysis for attendance). We expect that, compared to non-Akan, Akan males in the most recent survey rounds have switched more into the agriculture sector because of the possibility of inheriting their father's land. Table 3.13 shows the results. As expected, Akan males are 6.9 more likely to be farmers in the post-reform period compared to the non-Akan. Consistent with our education results, there is no significant change for females.

3.10 Concluding remarks

This paper has provided some evidence that descent rules have important implications for human capital investment. We exploit the introduction of the 1985 Intestate Succession Law that radically changed traditional inheritance practices among matrilineal groups in Ghana. The Law allowed fathers to transfer a substantial fraction of their property to their own children, reducing the share going to the matrikin. We interpret this evidence as consistent with the idea that the reform allowed parents to choose an allocation where the mix of human and physical capital was closer to the one they would have chosen in the absence of customary bequest constraints. Our results speak to a growing literature on the economic effects of traditional norms in developing countries and have relevant policy implications. In particular, our paper suggests that the individualization of land rights may have far reaching implications that go beyond agricultural investment and productivity, and affect human capital accumulation in the long run.

Figures

Figure 3.2: Cohort-specific effects. Coefficients of the interaction term Akan*birth year dummies, males aged 20-50 in all waves, full sample and sub-samples of those with and without a father farmer.



Tables

Table 3.1: Summary statistics

Panel A. Sample of individuals aged 20 to 50

	Full sample		Males				Females			
			Akan		non-Akan		Akan		non-Akan	
	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.
akan	0.42	(0.49)	-	-	-	-	-	-	-	-
education years	4.82	(4.98)	8.61	(4.16)	4.93	(5.22)	5.02	(4.51)	2.36	(4.00)
primary or higher	0.47	(0.50)	0.83	(0.38)	0.47	(0.50)	0.52	(0.50)	0.23	(0.42)
middle/jun sec. or higher	0.34	(0.47)	0.67	(0.47)	0.35	(0.48)	0.32	(0.47)	0.14	(0.35)
sec./senior sec. or higher	0.05	(0.22)	0.11	(0.31)	0.08	(0.27)	0.03	(0.16)	0.02	(0.14)
father farmer	0.79	(0.41)	0.72	(0.45)	0.84	(0.37)	0.72	(0.45)	0.85	(0.36)
land (at the hh level)	0.58	(0.49)	0.60	(0.49)	0.55	(0.5)0	0.61	(0.49)	0.57	(0.49)

Panel B.1. Sample of individuals aged 6 to 11

	Full sample		Males				Females			
			Akan		non-Akan		Akan		non-Akan	
	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.
akan	0.41	(0.49)	-	-	-	-	-	-	-	-
currently in school	0.73	(0.44)	0.88	(0.33)	0.65	(0.48)	0.86	(0.35)	0.61	(0.49)
education years	0.78	(1.34)	0.95	(1.39)	0.63	(1.16)	0.95	(1.44)	0.58	(1.15)
land (at hh level)	0.59	(0.49)	0.63	(0.48)	0.57	(0.50)	0.64	(0.48)	0.56	(0.49)

Panel B.2. Sample of individuals aged 12 to 17

	Full sample		Males				Females			
			Akan		non-Akan		Akan		non-Akan	
	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.	mean	s.dev.
akan	0.44	(0.50)	-	-	-	-	-	-	-	-
currently in school	0.68	(0.47)	0.83	(0.38)	0.62	(0.49)	0.75	(0.43)	0.56	(0.50)
education years	3.91	(3.23)	5.02	(3.00)	3.24	(3.20)	4.84	(3.06)	2.94	(3.12)
land (at hh level)	0.63	(0.48)	0.69	(0.46)	0.61	(0.49)	0.67	(0.47)	0.57	(0.50)

Rural sample. Pooled GLSS1-5. Using survey weights.

Table 3.2: Simple difference-in-difference estimates of the effect of the Intestate Law on years of education, by gender. Dependent variable is number of years of education. Full sample of individuals aged 0–30 in 1985 (N=17848)

	Males			Females		
	Akan	non-Akan	diff	Akan	non-Akan	diff
Panel a. Pre-school and primary school age versus control						
0 to 11 years in 1985	8.57	5.74	2.82	6.39	3.33	3.06
se	(0.146)	(0.146)	(0.260)	(0.167)	(0.130)	(0.266)
18 to 25 years in 1985	8.86	5.11	3.74	5.17	2.38	2.80
se	(0.157)	(0.225)	(0.271)	(0.155)	(0.148)	(0.216)
difference	-0.29	0.63	-0.92	1.21	0.95	0.26
se	(0.216)	(0.259)	(0.314)	(0.20)	(0.228)	(0.303)
Panel b. Secondary school age versus control						
12 to 17 years in 1985	8.82	5.16	3.66	5.89	2.74	3.15
se	(0.191)	(0.198)	(0.313)	(0.220)	(0.139)	(0.287)
18 to 25 years in 1985	8.86	5.11	3.74	5.17	2.38	2.80
se	(0.157)	(0.225)	(0.271)	(0.155)	(0.148)	(0.216)
difference	-0.04	0.04	-0.08	0.71	0.36	0.35
se	(0.220)	(0.278)	(0.356)	(0.242)	(0.197)	(0.303)
Panel c. Placebo						
18 to 25 years in 1985	8.86	5.11	3.74	5.17	2.38	2.80
se	(0.157)	(0.225)	(0.271)	(0.155)	(0.148)	(0.216)
26 to 30 years in 1985	8.79	5.14	3.64	4.89	1.92	2.97
se	(0.244)	(0.263)	(0.329)	(0.190)	(0.151)	(0.241)
difference	0.07	-0.03	0.10	0.29	0.46	-0.17
se	(0.265)	(0.276)	(0.391)	(0.209)	(0.157)	(0.267)

OLS estimates. Pooled GLSS1-5. Robust standard errors are adjusted for clustering at the village level. Sample of all individuals aged 0–30 in 1985 (from the sample of all individuals aged 20-50 in each survey round). Using survey weights.

Table 3.3: Effect of the Intestate Law on years of education, by gender. Dependent variable is number of years of education.

	Males		Females	
	(1)	(2)	(3)	(4)
akan*post68	-0.658**		0.075	
	(0.290)		(0.250)	
akan*post74		-0.918***		0.128
		(0.320)		(0.295)
akan	1.665***	1.606***	0.961***	0.964***
	(0.194)	(0.183)	(0.168)	(0.157)
durables	0.908***	0.936***	0.734***	0.745***
	(0.085)	(0.081)	(0.073)	(0.066)
hh size	-0.055**	-0.078***	-0.039***	-0.050***
	(0.022)	(0.020)	(0.014)	(0.014)
female head	0.192	0.209	1.120***	0.987***
	(0.353)	(0.313)	(0.163)	(0.138)
mother eduyrs	0.103**	0.109***	0.221***	0.189***
	(0.045)	(0.033)	(0.030)	(0.024)
father eduyrs	0.176***	0.165***	0.197***	0.184***
	(0.024)	(0.020)	(0.020)	(0.015)
age head	0.023***	0.019***	-0.010**	-0.008*
	(0.008)	(0.006)	(0.004)	(0.004)
age	-0.033	-0.029	-0.070	-0.084*
	(0.067)	(0.067)	(0.049)	(0.049)
age sq	-0.001	-0.001	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Observations	8337	8337	10285	10285
R-squared	0.390	0.393	0.369	0.372

OLS estimates. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, interactions between the individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head) and community crops with post68 (or post74).

Table 3.4: Effect of the Intestate Law on years of education on the subsamples of individuals with a father farmer and not, by gender. Dependent variable is number of years of education.

	Males				Females			
	father farmer	father not farmer	father farmer	father not farmer	father farmer	father not farmer	father farmer	father not farmer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
akan*post68	-0.672**	-0.394			0.076	-0.093		
	(0.315)	(0.650)			(0.278)	(0.653)		
akan*post74			-1.115***	0.036			0.211	-0.325
			(0.343)	(0.664)			(0.335)	(0.697)
akan	1.799***	0.928**	1.788***	0.733*	1.055***	0.556	1.043***	0.580*
	(0.211)	(0.443)	(0.198)	(0.393)	(0.181)	(0.356)	(0.171)	(0.297)
Observations	6700	1460	6700	1460	8481	1726	8481	1726
R-squared	0.382	0.333	0.387	0.326	0.335	0.332	0.337	0.337

OLS estimates. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.5: Effect of the Intestate Law on years of education, by gender. Subsample of Akan and Ewe (non-Akan). Dependent variable is number of years of education.

	Males		Females	
	(1)	(2)	(3)	(4)
akan*post68	-1.092***		-0.360	
	(0.397)		(0.370)	
akan*post74		-0.771*		-0.356
		(0.399)		(0.432)
akan	1.008***	0.791***	1.019***	0.959***
	(0.258)	(0.236)	(0.257)	(0.232)
Observations	4410	4410	5433	5433
R-squared	0.201	0.199	0.273	0.274

OLS estimates. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.6: Robustness checks. Years of education regressions.

	Males		Females	
	(1)	(2)	(3)	(4)
Panel a. Excluding Northern regions				
akan*post68	-0.682**		0.105	
	(0.288)		(0.257)	
akan*post74		-0.886***		0.202
		(0.317)		(0.298)
akan	1.716***	1.656***	0.943***	0.942***
	(0.195)	(0.184)	(0.173)	(0.159)
Observations	5914	5914	7151	7151
R-squared	0.238	0.240	0.282	0.284
Panel b. Subsample of cocoa-growing villages				
akan*post68	-0.798**		0.261	
	-0.399		-0.347	
akan*post74		-0.774*		0.342
		-0.412		-0.429
akan*cocoa major crop*post68	0.274		-0.402	
	-0.529		-0.435	
akan*cocoa major crop*post74		-0.294		-0.398
		-0.572		-0.517
cocoa major crop*post68	-0.194		-0.069	
	-0.486		-0.385	
cocoa major crop*post74		0.373		0.065
		-0.488		-0.458
akan*cocoa major crop	-0.896**	-0.728**	-0.352	-0.432
	-0.354	-0.332	-0.318	-0.296
cocoa major crop	0.295	0.141	0.29	0.269
	-0.324	-0.299	-0.29	-0.265
akan	2.149***	2.005***	1.159***	1.198***
	-0.269	-0.259	-0.237	-0.212
Observations	8337	8337	10285	10285
R-squared	0.391	0.394	0.37	0.373
Panel c. Including parental occupations				
akan*post68	-0.680**		0.109	
	(0.288)		(0.254)	
akan*post74		-0.926***		0.209
		(0.313)		(0.296)
akan	1.749***	1.676***	0.983***	0.990***
	(0.198)	(0.186)	(0.169)	(0.158)
Observations	7893	7893	10061	10061
R-squared	0.392	0.394	0.376	0.379

OLS estimates. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.7: Placebo regression, Côte d'Ivoire

	Males				Females			
	All regions		Eastern regions		All regions		Eastern regions	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
akpost68	0.142		0.232		-0.112		0.153	
	(0.339)		(0.557)		(0.226)		(0.314)	
akpost74		0.232		-0.155		-0.368		-0.205
		(0.407)		(0.616)		(0.308)		(0.337)
akan	-0.623*	-0.591*	0.273	0.437	-0.293	-0.243	0.025	0.174
	(0.330)	(0.306)	(0.354)	(0.302)	(0.180)	(0.188)	(0.187)	(0.182)
Observations	3530	3530	1443	1443	4246	4246	1923	1923
R-squared	0.255	0.247	0.301	0.292	0.199	0.197	0.213	0.208

OLS estimates. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. All regressions include birth year and region fixed effects, a survey round dummy, regional trends, age of the individual (and square), other individual controls (female headed hh, age of the head, household size, a durables index, and religion (merged from woman or man dataset)), and interactions between the other individual controls and the post68 (or post74) dummy.

Table 3.8: Effect of the Intestate Law on the probability of completing at least primary or secondary schooling. Full sample.

	Males		Females	
	(1)	(2)	(3)	(4)
dep. var:	**completion of primary or higher level**			
akan*post68	-0.045		0.010	
	(0.040)		(0.030)	
akan*post74		-0.097*		0.042
		(0.051)		(0.036)
akan	0.213***	0.214***	0.113***	0.110***
	(0.023)	(0.022)	(0.021)	(0.019)
Observations	8323	8323	10283	10283
Pseudo R2	0.324	0.326	0.296	0.296
dep. var:	**completion of secondary or higher level**			
akan*post68	-0.083**		-0.017	
	(0.039)		(0.019)	
akan*post74		-0.114***		0.011
		(0.043)		(0.025)
akan	0.176***	0.168***	0.066***	0.057***
	(0.024)	(0.022)	(0.015)	(0.013)
Observations	8323	8323	10192	10192
Pseudo R2	0.242	0.244	0.248	0.250

Probit estimates, marginal effecte reported. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.9: Effect of the Intestate Law on the probability of completing at least primary or secondary schooling. Subsamples of individuals with a father farmer and not, by gender.

	Males				Females			
	father farmer	father not farmer	father farmer	father not farmer	father farmer	father not farmer	father farmer	father not farmer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
dep. var: completion of primary or higher level								
akan*post68	-0.047	-0.004			0.002	0.011		
	(0.045)	(0.051)			(0.029)	(0.087)		
akan*post74			-0.115**	-0.014			0.035	0.032
			(0.054)	(0.061)			(0.038)	(0.089)
akan	0.229***	0.089***	0.235***	0.087***	0.114***	0.047	0.109***	0.050
	(0.026)	(0.034)	(0.025)	(0.030)	(0.021)	(0.050)	(0.019)	(0.039)
Observations	6690	1455	6690	1455	8411	1714	8411	1714
Pseudo R2	0.326	0.275	0.328	0.274	0.288	0.232	0.288	0.235
dep. var: completion of secondary or higher level								
akan*post68	-0.075*	-0.040			-0.006	-0.098		
	(0.041)	(0.081)			(0.017)	(0.080)		
akan*post74			-0.104**	-0.066			0.016	-0.040
			(0.044)	(0.096)			(0.024)	(0.084)
akan	0.183***	0.064	0.178***	0.056	0.051***	0.137***	0.044***	0.105**
	(0.027)	(0.048)	(0.025)	(0.043)	(0.013)	(0.051)	(0.012)	(0.044)
Observations	6690	1443	6690	1443	8330	1709	8330	1709
Pseudo R2	0.244	0.206	0.245	0.208	0.232	0.223	0.231	0.217

Probit estimates, marginal effects reported. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.10: Effect of the Intestate Succession Law on attendance rates, full sample, by gender. Dependent variable is a dummy equal to one if the individual is currently enrolled in school.

	Male		Female	
	(1)	(2)	(3)	(4)
akan*post	-0.119***	-0.062	-0.110**	-0.065
	(0.046)	(0.050)	(0.050)	(0.058)
akan*post*age1217		-0.107*		-0.097
		(0.062)		(0.074)
akan*age1217		0.051		0.038
		(0.043)		(0.057)
post*age1217		-0.025		0.073*
		(0.034)		(0.043)
age1217		0.031		-0.049
		(0.034)		(0.046)
akan	0.111***	0.083**	0.150***	0.134***
	(0.036)	(0.040)	(0.041)	(0.047)
Observations	8001	8001	6762	6762
Pseudo R2	0.244	0.245	0.264	0.265

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. Sample of individuals 6–17. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes region, and survey round fixed effects, regional trends, community major crops, log of distances to primary, junior and senior secondary, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post.

Table 3.11: Effect of the Intestate Succession Law on attendance rates, secondary school age children, by gender. Dependent variable is a dummy equal to one if the individual is currently enrolled in school. 12–17. Subsamples.

Panel a. Subsample of households with and w/o land

	Males				Females			
	w/ land		w/o land		w/ land		w/o land	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
akan*post	-0.154*		0.147		-0.194*		0.027	
	(0.088)		(0.118)		(0.106)		(0.252)	
akan*wave2		-0.063		-0.508		0.104		-0.015
		(0.105)		(0.400)		(0.103)		(0.301)
akan*wave3		-0.049		0.001		-0.104		0.083
		(0.109)		(0.219)		(0.129)		(0.244)
akan*wave4		-0.288**		-0.033		-0.203		0.046
		(0.131)		(0.241)		(0.146)		(0.269)
akan*wave5		-0.357***		-0.068		-0.217		-0.129
		(0.133)		(0.257)		(0.152)		(0.330)
akan	0.213***	0.264***	-0.213	0.000	0.166*	0.140	0.037	0.063
	(0.069)	(0.078)	(0.168)	(0.212)	(0.089)	(0.100)	(0.246)	(0.270)
Observations	2295	2295	1140	1140	1751	1751	984	984
Pseudo R2	0.277	0.280	0.221	0.223	0.304	0.305	0.278	0.282

Panel b. Subsample of individuals with father farmer and not

	father farmer		father not farmer		father farmer		father not farmer	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
akan*post	-0.143*		0.035		-0.132		0.109	
	(0.083)		(0.090)		(0.104)		(0.120)	
akan*wave2		-0.062		0.052		-0.000		0.165***
		(0.109)		(0.079)		(0.121)		(0.035)
akan*wave3		-0.064		0.065		-0.102		0.194***
		(0.102)		(0.078)		(0.123)		(0.043)
akan*wave4		-0.295**		0.033		-0.111		0.160*
		(0.118)		(0.107)		(0.129)		(0.096)
akan*wave5		-0.245**		0.079		-0.263*		0.174**
		(0.124)		(0.083)		(0.135)		(0.071)
akan	0.207***	0.239***	-0.058	-0.088	0.144	0.168*	0.017	-0.143
	(0.062)	(0.071)	(0.080)	(0.099)	(0.088)	(0.096)	(0.114)	(0.140)
Observations	2630	2630	767	767	2019	2019	675	675
Pseudo R2	0.251	0.253	0.261	0.262	0.280	0.281	0.300	0.306

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. Sample of individuals 12–17. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes region, and survey round fixed effects, regional trends, community major crops, log of distances to primary, junior and senior secondary, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post.

Table 3.12: Effect of the Intestate Succession Law on attendance rates and number of siblings, secondary school age children, by gender. Dependent variable is a dummy equal to one if the individual is currently enrolled in school.

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
akan*post	-0.346***	-0.355***	-0.343***	0.112	0.122	0.122
	(0.096)	(0.095)	(0.094)	(0.115)	(0.115)	(0.115)
akan*post*# siblings	0.057***			-0.039*		
	(0.018)			(0.023)		
post*# siblings	-0.014			0.053**		
	(0.023)			(0.025)		
akan*# siblings	-0.038**			0.052***		
	(0.015)			(0.020)		
# siblings	-0.004			-0.046**		
	(0.020)			(0.023)		
akan*post*# younger siblings		0.027			-0.038	
		(0.020)			(0.028)	
akan*post*# older siblings		0.136***			-0.053	
		(0.034)			(0.042)	
post*# younger siblings		-0.009			0.067**	
		(0.024)			(0.028)	
post*# older siblings		-0.011			0.025	
		(0.028)			(0.033)	
akan*# younger siblings		-0.022			0.055**	
		(0.017)			(0.025)	
akan*# older siblings		-0.070**			0.057	
		(0.028)			(0.037)	
# younger siblings		-0.010			-0.063**	
		(0.021)			(0.026)	
# older siblings		0.000			-0.011	
		(0.025)			(0.030)	
akan*post*# younger sisters			0.049			-0.058
			(0.036)			(0.048)
akan*post*# younger brothers			0.001			-0.019
			(0.032)			(0.044)
akan*post*# older siblings			0.076			-0.040
			(0.067)			(0.067)
akan*post*# older brothers			0.172***			-0.060
			(0.045)			(0.063)
akan	0.271***	0.271***	0.261***	-0.100	-0.109	-0.109
	(0.067)	(0.067)	(0.066)	(0.111)	(0.112)	(0.113)
Observations	3559	3559	3559	2825	2825	2825
Pseudo R2	0.231	0.235	0.238	0.263	0.265	0.266

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. Sample of individuals 12–17. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes region, and survey round fixed effects, regional trends, community major crops, log of distances to primary, junior and senior secondary, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post.

Table 3.13: Effect of the Intestate Law on occupational choice, by gender. Dependent variable is a dummy equal to one if the individual is a farmer. Full sample.

	Males	Females
	(1)	(2)
akan*post	0.069*	0.039
	(0.041)	(0.039)
akan	-0.053	-0.032
	(0.037)	(0.036)
Observations	7588	9360
Pseudo R2	0.146	0.157

Probit estimates, marginal effects reported. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head, and community crops), and interactions between the individual controls and community crops with post.

Appendix

Figure 3.3: Coefficients of the interaction term Akan*birth year dummies, females aged 20-50 in all waves, full sample and subsamples. Dep. variable: number of years of education.

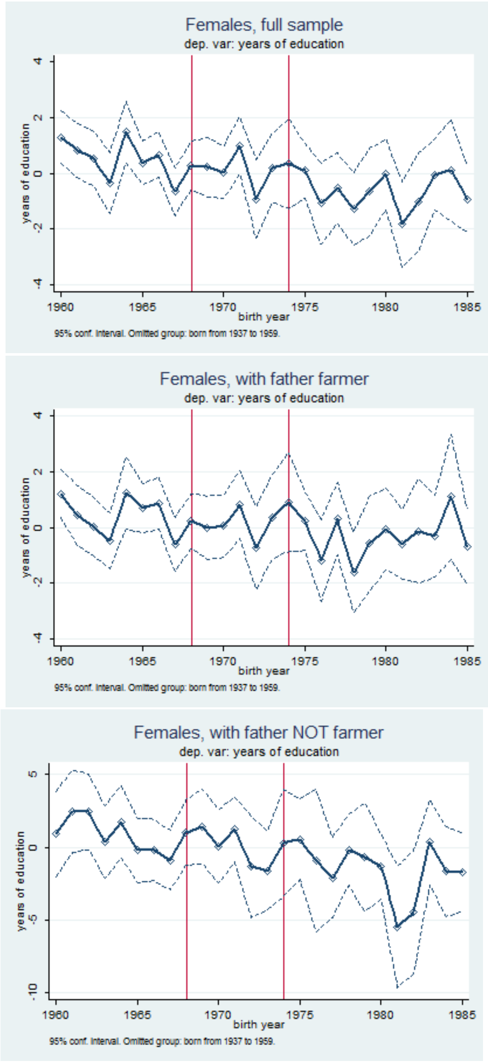


Figure 3.4: Coefficients of the interaction term Akan*birth year dummies, males aged 20-50 in all waves, full sample. Dep. variable: completion rate of primary school or higher level. OLS linear probability.

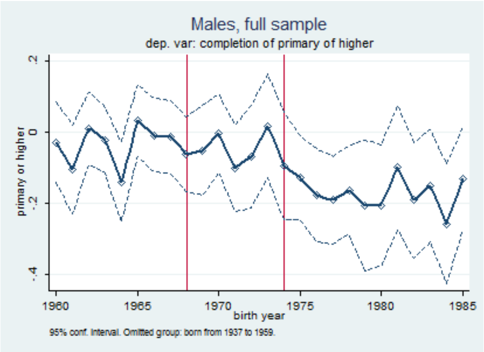


Figure 3.5: Coefficients of the interaction term Akan*birth year dummies, females aged 20-50 in all waves, full sample. Dep. variable: completion rate of primary school or higher level. OLS linear probability.

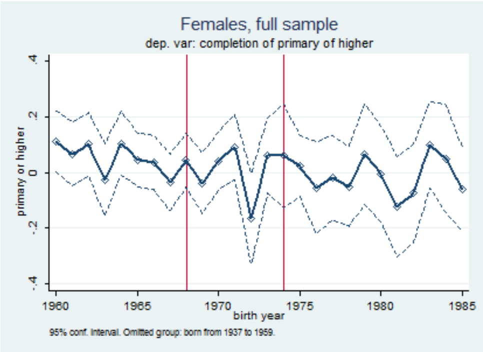


Figure 3.6: Coefficients of the interaction term Akan*birth year dummies, males aged 20-50 in all waves, full sample. Dep. variable: completion rate of primary school or higher level. Non-linear probit.

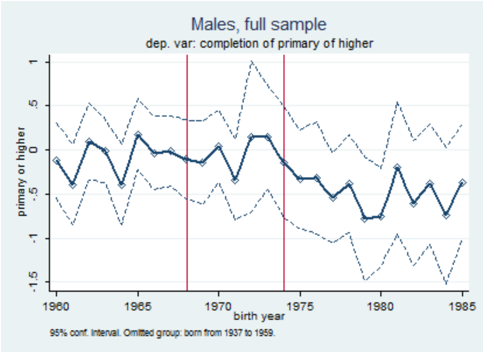


Figure 3.7: Coefficients of the interaction term Akan*birth year dummies, females aged 20-50 in all waves, full sample. Dep. variable: completion rate of primary school or higher level. Non-linear probit.

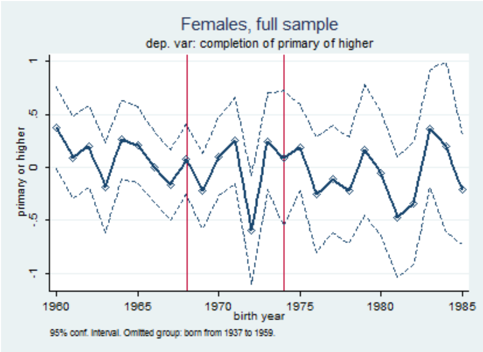


Table 3.14: Simple difference-in-difference estimates of the effect of the Intestate Law on years of education, by gender. Dependent variable is number of years of education. Sample of individuals aged 0–30 in 1985 with father farmer (N=13935).

	Males			Females		
	Akan	non-Akan	diff	Akan	non-Akan	diff
Panel a. Pre-school and primary school age versus control						
0 to 11 years in 1985	8.26	5.22	3.04	6.17	3.58	3.59
se	(0.149)	(0.163)	(0.282)	(0.202)	(0.210)	(0.273)
18 to 25 years in 1985	8.55	4.48	4.07	4.60	2.69	2.69
se	(0.172)	(0.219)	(0.280)	(0.157)	(0.209)	(0.209)
difference	-0.28	0.74	-1.03	1.57	0.68	0.89
se	(0.216)	(0.269)	(0.345)	(0.236)	(0.204)	(0.31)
Panel b. Secondary school age versus control						
12 to 17 years in 1985	8.37	4.52	3.85	5.08	2.14	2.94
se	(0.196)	(0.212)	(0.324)	(0.231)	(0.143)	(0.285)
18 to 25 years in 1985	8.55	4.48	4.07	4.60	2.69	2.69
se	(0.172)	(0.219)	(0.280)	(0.157)	(0.209)	(0.209)
difference	-0.18	0.04	-0.22	0.49	0.24	0.24
se	(0.252)	(0.274)	(0.381)	(0.266)	(0.195)	(0.318)
Panel c. Placebo						
18 to 25 years in 1985	8.55	4.48	4.07	4.60	1.90	2.69
se	(0.172)	(0.219)	(0.280)	(0.157)	(0.099)	(0.209)
26 to 30 years in 1985	8.18	4.55	3.62	4.52	1.56	2.97
se	(0.205)	(0.269)	(0.328)	(0.217)	(0.137)	(0.254)
difference	0.37	-0.07	0.44	0.06	0.34	-0.28
se	(0.250)	(0.285)	(0.394)	(0.243)	(0.154)	(0.286)

OLS estimates. Pooled GLSS1, GLSS2, GLSS3, GLSS4, and GLSS5. Robust standard errors are adjusted for clustering. Sample of all individuals aged 0–30 in 1985 (from the sample of all individuals aged 20-50 in each survey round). Using survey weights.

Table 3.15: Simple difference-in-difference estimates of the effect of the Intestate Law on years of education, by gender. Dependent variable is number of years of education. Sample of individuals aged 0–30 in 1985 with father NOT farmer (N=3622).

	Males			Females		
	Akan	non-Akan	diff	Akan	non-Akan	diff
Panel a. Pre-school and primary school age versus control						
0 to 11 years in 1985	9.16	8.29	0.87	6.74	6.28	0.46
se	(0.271)	(0.331)	(0.428)	(0.273)	(0.398)	(0.522)
18 to 25 years in 1985	9.64	8.13	1.50	6.65	5.05	1.60
se	(0.293)	(0.453)	(0.520)	(0.328)	(0.420)	(0.533)
difference	-0.47	0.16	-0.63	0.09	1.27	-1.13
se	(0.371)	(0.597)	(0.675)	(0.422)	(0.556)	(0.712)
Panel b. Secondary school age versus control						
12 to 17 years in 1985	9.66	7.87	1.79	7.31	5.17	2.14
se	(0.354)	(0.513)	(0.570)	(0.360)	(0.412)	(0.662)
18 to 25 years in 1985	9.64	8.13	1.50	6.65	5.05	1.60
se	(0.293)	(0.453)	(0.520)	(0.328)	(0.420)	(0.533)
difference	-0.02	-0.26	0.28	0.66	0.12	0.54
se	(0.400)	(0.539)	(0.679)	(0.452)	(0.564)	(0.749)
Panel c. Placebo						
18 to 25 years in 1985	9.64	8.13	1.50	6.65	5.05	1.60
se	(0.293)	(0.453)	(0.520)	(0.328)	(0.420)	(0.533)
26 to 30 years in 1985	10.7	9.71	0.99	6.09	4.64	1.45
se	(0.675)	(0.711)	(0.973)	(0.350)	(0.542)	(0.652)
difference	-1.07	-1.58	0.51	0.55	0.41	0.14
se	(0.732)	(0.812)	(1.122)	(0.409)	(0.67)	(0.819)

OLS estimates. Pooled GLSS1, GLSS2, GLSS3, GLSS4, and GLSS5. Robust standard errors are adjusted for clustering. Sample of all individuals aged 0–30 in 1985 (from the sample of all individuals aged 20-50 in each survey round). Using survey weights.

Table 3.16: Ordered logit estimates of the effect of the Intestate Law on completion rates, by gender. Dependent variable is the highest level completed.

	Males		Females	
	(1)	(2)	(3)	(4)
akan*post68	-0.373**		-0.046	
	(0.150)		(0.124)	
akan*post74		-0.585***		0.044
		(0.172)		(0.149)
akan	0.858***	0.837***	0.516***	0.489***
	(0.097)	(0.091)	(0.089)	(0.081)
Observations	8323	8323	10283	10283
Pseudo R2	0.210	0.211	0.207	0.207

Ordered logit estimates. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head, and community crops), and interactions between the individual controls and community crops with post68 (or post74).

Table 3.17: Robustness checks. Completion regressions.

	Akan vs Ewe				No Northern regions				Cocoa villages				Parental occupations			
	Male		Female		Male		Female		Male		Female		Male		Female	
Panel a. Primary or higher																
akan*post68	-0.077*		-0.025		-0.039		0.010		-0.020		0.028		-0.056		0.011	
	(0.040)		(0.051)		(0.031)		(0.035)		(0.059)		(0.044)		(0.040)		(0.030)	
akan*post74		-0.083*		0.012		-0.071*		0.048		-0.048		0.075		-0.118**		0.046
		(0.050)		(0.059)		(0.041)		(0.041)		(0.077)		(0.057)		(0.052)		(0.037)
akan	0.092***	0.083***	0.134***	0.122***	0.176***	0.176***	0.135***	0.130***	0.273***	0.272***	0.124***	0.123***	0.219***	0.219***	0.117***	0.114***
	(0.023)	(0.024)	(0.036)	(0.031)	(0.020)	(0.019)	(0.025)	(0.021)	(0.032)	(0.030)	(0.032)	(0.027)	(0.024)	(0.021)	(0.023)	(0.019)
akan*cocoa majcr *post68									-0.040		-0.035					
									(0.076)		(0.050)					
akan*cocoa majcr *post74										-0.087		-0.053				
										(0.099)		(0.058)				
cocoa majcr* *post68									0.021		0.015					
									(0.057)		(0.047)					
cocoa majcr* *post74										0.026		0.017				
										(0.065)		(0.056)				
akan*cocoa majcr									-0.118**	-0.113**	-0.018	-0.023				
									(0.047)	(0.045)	(0.039)	(0.034)				
cocoa majcr									0.039	0.040	0.009	0.014				
									(0.038)	(0.036)	(0.035)	(0.031)				
Observations	4394	4394	5432	5432	5901	5901	7149	7149	8323	8323	10283	10283	7872	7872	10059	10059
R-squared	0.139	0.137	0.185	0.182	0.180	0.180	0.200	0.199	0.325	0.327	0.296	0.297	0.324	0.326	0.303	0.304

continues on next page...

Table 3.17: Robustness checks. Completion regressions. *continued from previous page*

	Akan vs Ewe				No Northern regions				Cocoa villages				Parental occupations			
	Male		Female		Male		Female		Male		Female		Male		Female	
Panel b. Secondary or higher																
akan*post68	-0.117**		-0.037		-0.088**		-0.029		-0.070		0.004		-0.093**		-0.013	
	(0.056)		(0.041)		(0.040)		(0.027)		(0.056)		(0.029)		(0.040)		(0.019)	
akan*post74		-0.105*		0.005		-0.112**		0.013		-0.114*		0.040		-0.132***		0.014
		(0.057)		(0.049)		(0.047)		(0.034)		(0.059)		(0.039)		(0.044)		(0.026)
akan	0.119***	0.100***	0.101***	0.085***	0.177***	0.168***	0.092***	0.077***	0.238***	0.236***	0.083***	0.077***	0.183***	0.174***	0.066***	0.058***
	(0.032)	(0.029)	(0.027)	(0.025)	(0.023)	(0.022)	(0.019)	(0.017)	(0.033)	(0.030)	(0.020)	(0.018)	(0.025)	(0.023)	(0.014)	(0.013)
akan*cocoa majcr *post68									-0.025		-0.041					
									(0.072)		(0.030)					
akan*cocoa majcr *post74										-0.001		-0.045				
										(0.083)		(0.034)				
cocoa majcr* *post68									0.007		0.015					
									(0.061)		(0.034)					
cocoa majcr* *post74										0.054		0.014				
										(0.066)		(0.042)				
akan*cocoa majcr									-0.112***	-0.123***	-0.027	-0.034*				
									(0.042)	(0.039)	(0.024)	(0.020)				
cocoa majcr									0.053	0.045	0.033	0.037*				
									(0.040)	(0.035)	(0.025)	(0.021)				
Observations	4397	4397	5380	5380	5901	5901	7086	7086	8323	8323	10192	10192	7872	7872	9971	9971
R-squared	0.112	0.112	0.171	0.170	0.132	0.132	0.180	0.179	0.244	0.245	0.249	0.250	0.239	0.240	0.255	0.256

Probit estimates, marginal effects reported. Pooled GLSS1-5. Robust standard errors adjusted for clustering at the village level in parentheses. Sample of individuals 20-50. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes birth year, region, and survey round fixed effects, regional trends, community major crops, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post68 (or post74).

Table 3.18: Robustness checks. Attendance regressions. Dependent variable is a dummy equal to one if the individual is currently enrolled in school.

	Males	Females	Males	Females
	(1)	(2)	(3)	(4)
	a) Akan vs Ewe		**b) No Northern regions**	
akan*post	-0.068*	-0.033	-0.082***	-0.079**
	(0.041)	(0.050)	(0.032)	(0.038)
akan	0.072	0.080	0.091***	0.123***
	(0.052)	(0.057)	(0.030)	(0.036)
Observations	4167	3652	5480	4810
Pseudo R2	0.167	0.219	0.169	0.224
	c) Parental occupation		**d) any member in the govmt**	
akan*post	-0.124***	-0.116**	-0.124***	-0.105**
	(0.046)	(0.050)	(0.048)	(0.052)
akan	0.122***	0.160***	0.124***	0.149***
	(0.036)	(0.041)	(0.037)	(0.043)
akan*post*any member in Govmt			0.013	-0.036
			(0.100)	(0.139)
akan*any member in Govmt			-0.136	0.015
			(0.098)	(0.104)
post*any member in Govmt			-0.010	0.080
			(0.077)	(0.088)
any member in Govmt			0.135***	0.005
			(0.033)	(0.088)
Observations	7420	6260	7863	6649
Pseudo R2	0.247	0.265	0.248	0.265

	e) Cocoa villages	
	Male	Female
akan*post	-0.168**	-0.048
	(0.073)	(0.071)
akan	0.138**	0.106*
	(0.058)	(0.062)
akan*cocoa majcr*post	0.078	-0.113
	(0.071)	(0.100)
akan*cocoa majcr	-0.054	0.078
	(0.078)	(0.074)
cocoa majcr*post	-0.061	0.124*
	(0.070)	(0.071)
cocoa majcr	0.073	-0.094
	(0.055)	(0.072)
Observations	8001	6762
Pseudo R2	0.244	0.264

Probit estimates, marginal effects reported. Robust standard errors adjusted for clustering at the household level in parentheses. Sample of individuals 6–17. ***, **, and * indicate significance at 1%, 5% and 10% levels. Using survey weights. Each regression also includes region, and survey round fixed effects, regional trends, community major crops, log of distances to primary, junior and senior secondary, dummies for religion of the head of the household, age (and its square), other individual controls (durables, female head, household size, mother and father's education, age of the head, religion of the head), and interactions between other individual controls and community crops with post.

Table 3.19: Summary statistics, sample of individuals aged 20–50.

	All			Male						Female					
				Akan			non-Akan			Akan			non-Akan		
	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n
akan	0.42	0.49	22156	1.00	0.00	3897	0.00	0.00	5958	1.00	0.00	5086	0.00	0.00	7215
eduyrs	4.82	4.98	22156	8.61	4.16	3897	4.93	5.22	5958	5.02	4.51	5086	2.36	4.00	7215
in school	0.03	0.16	22156	0.04	0.19	3897	0.05	0.22	5958	0.01	0.11	5086	0.01	0.11	7215
no education	0.43	0.49	22136	0.10	0.29	3888	0.44	0.50	5949	0.34	0.47	5086	0.68	0.47	7213
incompl primary	0.10	0.30	22136	0.07	0.26	3888	0.09	0.29	5949	0.14	0.35	5086	0.09	0.29	7213
primary or higher	0.47	0.50	22136	0.83	0.38	3888	0.47	0.50	5949	0.52	0.50	5086	0.23	0.42	7213
jun sec/middle or higher	0.34	0.47	22136	0.67	0.47	3888	0.35	0.48	5949	0.32	0.47	5086	0.14	0.35	7213
senior sec/sec or higher	0.05	0.22	22156	0.11	0.31	3897	0.08	0.27	5958	0.03	0.16	5086	0.02	0.14	7215
female	0.55	0.50	22156	0.00	0.00	3897	0.00	0.00	5958	1.00	0.00	5086	1.00	0.00	7215
age	33.17	8.91	22156	33.05	8.97	3897	33.09	9.00	5958	33.34	8.95	5086	33.16	8.77	7215
own land (hh)	0.58	0.49	20943	0.60	0.49	3658	0.55	0.50	5700	0.61	0.49	4753	0.57	0.49	6832
acres of land	23.05	91.20	20885	27.42	115.46	3649	20.35	70.71	5681	25.22	108.11	4746	21.16	75.18	6809
hh size	5.88	3.42	22156	5.05	3.01	3897	5.88	3.58	5958	5.42	2.74	5086	6.71	3.77	7215
female head	0.17	0.38	22156	0.10	0.30	3897	0.05	0.23	5958	0.36	0.48	5086	0.16	0.37	7215
age head	42.52	13.03	22156	40.31	12.17	3897	41.53	13.13	5958	42.61	12.80	5086	44.57	13.31	7215
durables	-0.34	1.02	20933	-0.09	1.19	3712	-0.49	0.88	5674	-0.17	1.11	4736	-0.50	0.89	6811
catholic (head)	0.16	0.37	22130	0.16	0.37	3893	0.17	0.37	5951	0.16	0.37	5083	0.15	0.36	7203
protestant (head)	0.16	0.37	22130	0.21	0.40	3893	0.12	0.33	5951	0.22	0.42	5083	0.12	0.32	7203
other christian (head)	0.28	0.45	22130	0.42	0.49	3893	0.19	0.39	5951	0.42	0.49	5083	0.18	0.38	7203
muslim (head)	0.15	0.35	22130	0.06	0.24	3893	0.21	0.41	5951	0.06	0.23	5083	0.21	0.41	7203
animist (head)	0.18	0.38	22130	0.06	0.23	3893	0.25	0.43	5951	0.06	0.24	5083	0.28	0.45	7203
married head	0.75	0.43	22148	0.68	0.47	3896	0.79	0.41	5957	0.66	0.47	5084	0.84	0.37	7211
polygynous head	0.12	0.33	22156	0.03	0.17	3897	0.13	0.34	5958	0.05	0.21	5086	0.23	0.42	7215
farmer (strict def)	0.70	0.46	20182	0.66	0.47	3561	0.76	0.43	5406	0.67	0.47	4703	0.72	0.45	6512
cocoa (hh)	0.23	0.42	19378	0.43	0.50	3268	0.11	0.31	5413	0.39	0.49	4220	0.11	0.31	6477
agric-fish (industry)	0.74	0.44	20075	0.71	0.45	3545	0.82	0.38	5358	0.68	0.47	4683	0.73	0.44	6489
mining (industry)	0.01	0.09	20075	0.02	0.15	3545	0.01	0.08	5358	0.00	0.04	4683	0.00	0.06	6489
manufacturing (industry)	0.08	0.27	20075	0.06	0.24	3545	0.05	0.21	5358	0.09	0.29	4683	0.10	0.30	6489
utilities (industry)	0.00	0.03	20075	0.00	0.04	3545	0.00	0.04	5358	0.00	0.00	4683	0.00	0.00	6489
construction (industry)	0.01	0.10	20075	0.03	0.17	3545	0.01	0.12	5358	0.00	0.01	4683	0.00	0.05	6489
sales (industry)	0.10	0.30	20075	0.05	0.22	3545	0.03	0.16	5358	0.18	0.38	4683	0.14	0.34	6489
transport (industry)	0.01	0.10	20075	0.03	0.18	3545	0.02	0.12	5358	0.00	0.03	4683	0.00	0.02	6489
financing (industry)	0.00	0.05	20075	0.01	0.08	3545	0.00	0.04	5358	0.00	0.02	4683	0.00	0.02	6489
civil (industry)	0.05	0.21	20075	0.08	0.27	3545	0.06	0.24	5358	0.04	0.19	4683	0.02	0.14	6489
other (industry)	0.02	0.15	20075	0.03	0.16	3545	0.02	0.15	5358	0.03	0.17	4683	0.01	0.12	6489

Rural sample of individuals aged 20–50. Pooled GLSS1-5. Using survey weights.

continues on next page...

Table 3.19: Summary statistics, sample of individuals aged 20–50. *continued from previous page*

	All			Male						Female					
				Akan			non-Akan			Akan			non-Akan		
	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n
mother eduyrs	0.89	2.78	21974	1.40	3.30	3858	0.56	2.27	5917	1.34	3.35	5021	0.52	2.20	7178
father eduyrs	2.43	4.64	21667	3.94	5.34	3786	1.54	3.84	5900	3.57	5.28	4870	1.47	3.84	7111
father farmer	0.79	0.41	21837	0.72	0.45	3874	0.84	0.37	5766	0.72	0.45	5053	0.85	0.36	7144
mother farmer	0.80	0.40	21538	0.81	0.40	3834	0.78	0.41	5634	0.82	0.39	4961	0.80	0.40	7109
father sales sector	0.02	0.14	21839	0.03	0.16	3874	0.01	0.12	5766	0.03	0.16	5054	0.02	0.14	7145
mother sales sector	0.15	0.36	21540	0.15	0.35	3834	0.16	0.37	5635	0.14	0.35	4962	0.14	0.35	7109
father clerk	0.02	0.14	21839	0.03	0.18	3874	0.01	0.11	5766	0.03	0.18	5054	0.01	0.11	7145
mother clerk	0.00	0.04	21540	0.00	0.05	3834	0.00	0.03	5635	0.00	0.04	4962	0.00	0.03	7109
father professional	0.04	0.21	21839	0.07	0.26	3874	0.03	0.18	5766	0.06	0.24	5054	0.03	0.17	7145
mother professional	0.01	0.08	21540	0.01	0.09	3834	0.01	0.08	5635	0.01	0.09	4962	0.01	0.07	7109
cocoa major crop	0.37	0.48	20290	0.95	0.21	3506	0.68	0.47	5462	0.96	0.20	4667	0.67	0.47	6655
cassava major crop	0.79	0.41	20290	0.89	0.31	3506	0.88	0.33	5462	0.90	0.30	4667	0.88	0.33	6655
maize major crop	0.88	0.32	20290	0.33	0.47	3506	0.50	0.50	5462	0.33	0.47	4667	0.49	0.50	6655
yam major crop	0.43	0.49	20290	0.43	0.50	3506	0.31	0.46	5462	0.43	0.50	4667	0.32	0.47	6655
tomato major crop	0.36	0.48	20290	0.73	0.45	3506	0.25	0.43	5462	0.72	0.45	4667	0.23	0.42	6655
plantains major crop	0.44	0.50	20290	0.12	0.32	3506	0.60	0.49	5462	0.12	0.32	4667	0.61	0.49	6655
nuts major crop	0.40	0.49	20290	0.43	0.49	3506	0.37	0.48	5462	0.43	0.49	4667	0.37	0.48	6655
pepper major crop	0.40	0.49	20290	0.13	0.34	3506	0.56	0.50	5462	0.12	0.33	4667	0.58	0.49	6655
beans/peas major crop	0.38	0.49	20290												

Rural sample of individuals aged 20–50. Pooled GLSS1-5. Using survey weights.

Table 3.20: Summary statistics, sample of individuals aged 6–17.

	All			Male						Female					
				Akan			non-Akan			Akan			non-Akan		
	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n
akan	0.42	0.49	17413	1.00	0.00	3819	0.00	0.00	5653	1.00	0.00	3269	0.00	0.00	4672
eduyrs	2.11	2.81	17371	2.80	3.04	3813	1.75	2.61	5642	2.64	2.99	3258	1.52	2.45	4658
inschool	0.71	0.45	17375	0.86	0.35	3813	0.64	0.48	5644	0.81	0.39	3258	0.59	0.49	4660
no education	0.39	0.49	17364	0.25	0.43	3808	0.47	0.50	5640	0.27	0.45	3258	0.52	0.50	4658
incompl. primary	0.44	0.50	17364	0.52	0.50	3808	0.40	0.49	5640	0.50	0.50	3258	0.36	0.48	4658
primary or higher	0.17	0.38	17364	0.23	0.42	3808	0.13	0.34	5640	0.23	0.42	3258	0.12	0.32	4658
jun sec/middle or higher	0.03	0.18	17364	0.05	0.22	3808	0.03	0.16	5640	0.05	0.21	3258	0.02	0.13	4658
female	0.46	0.50	17413	0.00	0.00	3819	0.00	0.00	5653	1.00	0.00	3269	1.00	0.00	4672
age	10.83	3.37	17413	11.01	3.40	3819	10.88	3.40	5653	10.93	3.37	3269	10.56	3.31	4672
own land (hh)	0.61	0.49	16672	0.66	0.48	3663	0.59	0.49	5410	0.65	0.48	3131	0.56	0.50	4468
acres of land	26.24	103.87	16628	29.84	125.62	3655	23.13	79.95	5394	33.70	134.19	3121	21.23	78.97	4458
hh size	7.32	3.30	17413	6.54	2.49	3819	7.86	3.68	5653	6.66	2.57	3269	7.83	3.68	4672
female head	0.20	0.40	17413	0.28	0.45	3819	0.12	0.33	5653	0.29	0.45	3269	0.14	0.35	4672
age head	46.85	11.42	17413	45.46	10.69	3819	48.24	11.86	5653	45.31	10.69	3269	47.55	11.73	4672
durables	-0.39	1.00	16452	-0.21	1.09	3576	-0.55	0.89	5347	-0.16	1.13	3089	-0.51	0.90	4440
cocoa (hh)	0.25	0.43	15624	0.44	0.50	3327	0.11	0.32	5194	0.45	0.50	2849	0.11	0.31	4254
mother eduyrs	2.61	4.06	17357	3.98	4.48	3810	1.60	3.38	5630	3.96	4.46	3261	1.62	3.43	4656
father eduyrs	4.99	5.43	17318	7.35	5.13	3780	3.18	4.88	5629	7.37	5.25	3253	3.35	5.00	4656
catholic	0.15	0.36	17398	0.16	0.36	3818	0.15	0.35	5647	0.16	0.37	3267	0.15	0.36	4666
protestant	0.15	0.36	17398	0.21	0.41	3818	0.11	0.31	5647	0.21	0.41	3267	0.11	0.31	4666
other christian	0.27	0.45	17398	0.42	0.49	3818	0.15	0.36	5647	0.42	0.49	3267	0.18	0.38	4666
muslim	0.14	0.35	17398	0.06	0.24	3818	0.20	0.40	5647	0.07	0.25	3267	0.20	0.40	4666
animist	0.21	0.40	17398	0.06	0.25	3818	0.32	0.47	5647	0.06	0.24	3267	0.29	0.46	4666
married head	0.82	0.38	17403	0.73	0.44	3816	0.88	0.32	5652	0.75	0.44	3267	0.88	0.33	4668
# siblings	3.77	2.51	17413	3.28	2.02	3819	4.12	2.76	5653	3.36	2.06	3269	4.10	2.77	4672
# sisters	1.69	1.47	17413	1.49	1.31	3819	1.79	1.54	5653	1.63	1.37	3269	1.80	1.56	4672
# brothers	2.08	1.74	17413	1.79	1.48	3819	2.33	1.91	5653	1.73	1.42	3269	2.29	1.86	4672
# older siblings	1.63	1.72	17413	1.42	1.43	3819	1.74	1.87	5653	1.45	1.46	3269	1.80	1.90	4672
# younger siblings	2.15	1.83	17413	1.86	1.58	3819	2.38	1.98	5653	1.91	1.58	3269	2.30	1.97	4672
# older sisters	0.67	0.95	17413	0.61	0.86	3819	0.68	1.00	5653	0.67	0.89	3269	0.72	1.01	4672
# younger sisters	1.02	1.14	17413	0.88	1.02	3819	1.11	1.20	5653	0.96	1.08	3269	1.08	1.20	4672
# older brothers	0.95	1.21	17413	0.81	1.03	3819	1.06	1.32	5653	0.79	1.01	3269	1.08	1.33	4672
# younger brothers	1.13	1.23	17413	0.99	1.10	3819	1.27	1.33	5653	0.95	1.04	3269	1.22	1.30	4672

Rural sample of individuals aged 6–17. Pooled GLSS1-5. Using survey weights.

continues on next page...

Table 3.20: Summary statistics, sample of individuals aged 6–17. *continued from previous page*

	All			Male						Female					
				Akan			non-Akan			Akan			non-Akan		
	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n	mean	s.dev.	n
farmer	0.82	0.39	5266	0.81	0.39	1003	0.84	0.37	2077	0.75	0.43	798	0.83	0.38	1388
any hh member in govmt	0.08	0.27	17134	0.09	0.29	3799	0.07	0.25	5518	0.10	0.30	3248	0.07	0.25	4569
father farmer	0.73	0.44	16740	0.67	0.47	3756	0.79	0.41	5345	0.66	0.47	3227	0.77	0.42	4412
mother farmer	0.74	0.44	16570	0.72	0.45	3728	0.75	0.43	5286	0.73	0.45	3177	0.74	0.44	4379
father sales	0.02	0.15	16754	0.03	0.17	3757	0.02	0.13	5349	0.04	0.19	3228	0.02	0.14	4420
mother sales	0.14	0.34	16586	0.15	0.36	3728	0.13	0.33	5289	0.15	0.36	3181	0.13	0.34	4388
father clerk	0.02	0.13	16754	0.03	0.16	3757	0.01	0.10	5349	0.03	0.18	3228	0.01	0.11	4420
mother clerk	0.00	0.04	16586	0.00	0.04	3728	0.00	0.04	5289	0.00	0.06	3181	0.00	0.02	4388
father professional	0.06	0.24	16754	0.09	0.28	3757	0.04	0.20	5349	0.08	0.27	3228	0.04	0.20	4420
mother professional	0.02	0.13	16586	0.02	0.14	3728	0.01	0.11	5289	0.02	0.15	3181	0.01	0.12	4388
ln (1+dist primary)	0.21	0.54	16514	0.16	0.42	3643	0.24	0.60	5356	0.15	0.41	3103	0.25	0.63	4412
ln (1+dist jsec)	0.59	0.86	16472	0.38	0.65	3641	0.76	0.96	5331	0.39	0.66	3102	0.73	0.97	4398
ln (1+dist ssec)	1.86	1.06	16300	1.74	1.02	3574	1.94	1.08	5312	1.73	1.02	3040	1.97	1.07	4374
cocoa major crop	0.37	0.48	16054	0.65	0.48	3518	0.17	0.38	5211	0.64	0.48	3008	0.18	0.39	4317
cassava major crop	0.80	0.40	16054	0.97	0.18	3518	0.66	0.47	5211	0.96	0.19	3008	0.69	0.46	4317
maize major crop	0.89	0.32	16054	0.89	0.31	3518	0.88	0.33	5211	0.90	0.30	3008	0.88	0.33	4317
yam major crop	0.42	0.49	16054	0.33	0.47	3518	0.49	0.50	5211	0.34	0.47	3008	0.49	0.50	4317
tomato major crop	0.37	0.48	16054	0.43	0.50	3518	0.32	0.47	5211	0.44	0.50	3008	0.33	0.47	4317
plantains major crop	0.44	0.50	16054	0.73	0.44	3518	0.22	0.41	5211	0.74	0.44	3008	0.23	0.42	4317
nuts major crop	0.41	0.49	16054	0.13	0.33	3518	0.63	0.48	5211	0.12	0.33	3008	0.60	0.49	4317
pepper major crop	0.40	0.49	16054	0.42	0.49	3518	0.37	0.48	5211	0.42	0.49	3008	0.38	0.49	4317
beans/peas major crop	0.39	0.49	16054	0.12	0.32	3518	0.59	0.49	5211	0.12	0.33	3008	0.56	0.50	4317

Rural sample of individuals aged 6–17. Pooled GLSS1-5. Using survey weights.

Bibliography

[1] Alderman Harold, Sudharshan Canagarajah, and Stephen D. Younger. 1993. “Consequences of Permanent Lay-Off from the Civil Service: Results from a Survey of Retrenched Workers in Ghana.” Working paper, Cornell Food and Nutrition Policy Program.

[2] Awusabo–Asare, Kofi. 1990. “Matrilinity and the New Intestate Succession Law of Ghana.” *The Canadian Journal of African Studies*, Vol. 24, No. 1, pp. 1–16.

[3] Becker, Gary and Nigel Tomes. 1976. “Child Endowments and the Quantity and Quality of Children.” *Journal of Political Economy*, Vol. 84, No. 4, pp. S143–S162.

[4] Baland, Jean-Marie, Catherine Guirkinger and Charlotte Mali. 2011. “Pretending to Be Poor: Borrowing to Escape Forced Solidarity in Cameroon.” *Economic Development and Cultural Change*, Vol. 60, No. 1, pp. 1-16.

[5] Behrman, Jere R. 1988. “Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?.” *Oxford Economic Papers*, Vol. 40, No. 1, pp. 32–54.

[6] Behrman, Jere R. 1997. “Intrahousehold Distribution and the Family.” *Handbook of Population and Family Economics*, Vol. 1, No. 1, pp. 125–187.

[7] Besley, Tim. 1995. “Property Rights and Investment Incentives: Theory and Evidence from Ghana.” *Journal of Political Economy*, Vol. 103, No. 5, 1995: 903–937.

[8] Botticini, Maristella and Siow, Aloysius. 2003. “Why Dowries?” *American Economic Review*, American Economic Association, vol. 93, No. 4, po. 1385–1398.

[9] Bruce, J., Migot-Adholla. 1994. "Searching for Land Tenure Security in Africa." Kendal/Hunt Publishing Co., Dubuque.

[10] Carranza, Eliana. 2012. "Islamic Inheritance Law, Son Preference and Fertility Behavior of Muslim Couples in Indonesia." World Bank Policy Research Working Paper No. 5972.

[11] Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, Vol. 91, No. 4, pp. 795–813.

[12] Duncan, Beatrice Akua. 2010. "Cocoa, Marriage, Labour and Land in Ghana: some matrilineal and patrilineal perspectives." *The Journal of the International African Institute*, Vol. 80, No. 2, 2010, pp. 301–321.

[13] Fenrich, Jeanmarie and Tracy E. Higgins. 2001. "Promise Unfulfilled: Law, Culture, and Womens Inheritance Rights in Ghana." *Fordham International Law Journal*, Vol. 25, No. 2, pp. 259–341.

[14] Fortes M., 1950. "Kinship and Marriage Among the Ashanti." in Radcliffe-Brown A.R. and D. Forde (eds) "African Systems of Kinship and Marriage." London: Oxford University Press.

[15] Gneezy Uri, Kenneth L. Leonard, and John A. List. 2009. "Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society." *Econometrica*, Vol. 77, No. 5, pp. 1637–1664.

[16] Goetghebuer, Tatiana and Platteau, Jean-Philippe. 2010. "Inheritance patterns in migration-prone communities of the Peruvian Highlands." *Journal of Development Economics*, Vol. 93, No. 1, pp. 71–87.

[17] Goldstein, Markus and Christopher Udry. 2008. "The Profits of Power: Land Rights and Agricultural Investment in Ghana." *Journal of Political Economy*, Vol. 116, No. 6, pp. 981–1022.

[18] Hornbeck, Richard. 2010. "Barbed Wire: Property Rights and Agricultural Development." *Quarterly Journal of Economics*, Vol. 125, No. 2, pp. 767–810.

[19] Jakiela, Pamela and Owen Ozier. 2011. "Does Africa Need a Rotten Kin Theorem? Experimental Evidence from Village Economies." mimeo.

[20] Kingdon, Geeta and Mans Söderbom. 2008. "Education, Skills, and Labor Market Outcomes: Evidence from Ghana" World Bank Education Working Paper Series, No. 12.

[21] Kutsoati, Edward and Randall Morck. 2012. "Family Ties, Inheritance Rights, and Successful Poverty Alleviation: Evidence from Ghana." NBER Working Paper No. 18080.

[22] La Ferrara, Eliana. 2003. "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana." *The American Economic Review*, Vol. 93, No. 5, pp. 1730–1751.

[23] La Ferrara, Eliana. 2007. "Descent Rules and Strategic Transfers. Evidence from Matrilineal Groups in Ghana" *The Journal of Development Economics*, Vol. 83, No. 2, pp. 280–301.

[24] Lambert, Sylvie, Martin Ravallion and Dominique van de Walle. 2011. "Is It What You Inherited or What You Learnt? Intergenerational Linkage and Interpersonal Inequality in Senegal " World Bank Policy Research Working Paper No. 5658.

[25] Lavy, Victor, and Alexander Zablotsky. 2011. "Mothers Schooling and Fertility under Low Female Labor Force Participation: Evidence from a Natural Experiment." NBER Working Paper No. 16856.

[26] Mobarak, A. Mushfiq, Randall Kuhn, and Christina Peters. 2009. "Marriage Market Effects of a Wealth Shock in Bangladesh" mimeo, Yale University.

[27] Okali, C. 1983. "Cocoa and Kinship in Ghana; The Matrilineal Akan of Ghana." Boston: Kegan Paul International.

[28] Pande, Rohini and Christopher Udry. 2006. "Institutions and Development: A View from Below." in Blundell, Newey and Persson (eds), *Advances in Economics and Econometrics.*

[29] Platteau, Jean-Philippe. 2000. "Institutions, Social Norms, and Economic Development." Harwood Academic Publishers.

[30] Platteau, Jean-Philippe and Jean-Marie Baland. 2001. "Impartible Inheritance versus Equal Division: A Comparative Perspective Centered on Europe and Sub-Saharan Africa", in De Janvry, A., G. Gordillo, E. Sadoulet and J.-P. Platteau (eds.), *Access to Land, Rural Poverty, and Public Action*, Oxford: Oxford University Press.

[31] Quisumbing, A., E. Panyongayong, J.B. Aidoo and K. Otsuka. 2001. "Women's Land Rights in the Transition to Individualized Ownership: Implications for the Management of Tree Resources in Western

Ghana." *Economic Development and Cultural Change*, Vol. 50, No. 1, pp. 157–81.

[32] Quisumbing, A. and K. Otsuka. 2001. "Land Inheritance and Schooling in Matrilineal Societies: Evidence from Sumatra." *World Development*, Vol. 29, No. 12, pp. 2093–2110.

[33] White, Howard. 2005. "Using Household Survey Data to Measure Educational Performance: the Case of Ghana" *Social Indicators Research*, Vol. 74, No. 2, pp. 395-422.

[34] Woodman, Gordon R. 1985. "Ghana Reforms the Law of Intestate Succession." *Journal of African Law*, Vol. 29, No. 2, pp. 118–128.

[35] World Bank. 1999. "Ghana Living Standards Survey (GLSS), 1987-88 and 1988-89 Basic Information." The World Bank: Washington D.C.