Université de Montréal

Culture, Gender and Development in Sub-Saharan Africa

Par

Lucienne TALBA PAPITE

Département de sciences économiques

Faculté des arts et sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en sciences économiques

Aout, 2020

Université de Montréal
Faculté des Études Supérieures et Postdoctorales

Cette thèse intitulée :

Culture, Gender and Development in Sub-Saharan Africa

présentée par :

Lucienne TALBA PAPITE

sera évaluée par un jury composé des personnes suivantes :

Baris Kaymak,	president du jury
Joshua Lewis,	directeur de recherche
Raphael Godefroy,	co-directeur de recherche
Andriana Bellou,	membre du jury
Arthur Blouin,	examinateur externe
Michael Hubermann,	représentant du doyen de la FAS

à mon amour Idriss, mes parents Talba Radjab et Mbialeu Julienne

Remerciements

Ce travail de thèse de doctorat est pour moi un dur mais bel exercice qui me fait réaliser que je n'aurais jamais pu finir ce projet sans le soutien de plusieurs personnes. C'est l'occasion pour moi aujourd'hui de leur témoigner ma gratitude.

Je tiens à remercier grandement mon directeur de recherche Joshua Lewis, pour son soutien, sa patience et sa disponibilité sans mesure. Ses encouragements et la confiance qu'il a mis en moi, m'ont été très utiles pour l'aboutissement de ma thèse. Je tiens également à lui dire merci pour son appui financier tout au long de ma période de thèse.

Mes profonds remerciements vont également à l'endroit de Raphael Godefroy, Andriana Bellou et Emanuela Cardia qui ont bien voulu accepter de faire partir de mon comité de thèse et m'ont soutenu durant mon année de marché de travail.

Je suis reconnaissante envers Baris Kaymak, Michael Hubermann et Christopher Rauh, ainsi que tous les participants du seminaire midi de l'Université de Montréal. Merci à tous les enseignants chercheurs et aux personnels administratifs du département des sciences économiques et du CIREQ. Cette thèse a été ralisée avec l'appui financier du département et du CIREQ.

Je ne pouvais pas terminer ces remerciements sans témoigner toute ma reconnaissance aux personnes les plus chères pour moi. J'envoie un merci spécial à mes parents Talba Radjab et Mbialeu Julienne pour leur soutien et encouragement perpetuels. Je suis profondément reconnaissante envers mon conjoint Idriss, pour sa patience et son soutien durant ces six dernières années. Enfin je dit merci à ma meilleure amie Claire ainsi qu'à mes frères et soeurs pour leur soutien inconditionnel tout au long de mes études.

Résumé

Cette thèse analyse les causes profondes du sous-développement et des inégalités de genre en Afrique Sub-Saharienne. Le premier chapitre teste empiriquement l'hypothèse de Engels (1884) selon laquelle l'origine des inégalités de genre en Afrique serait la domestication historique des bovins. Pour résoudre les problèmes de biais d'endogenéités dans le choix de la domestication des bovins, j'adopte la stratégie des variables instrumentales en exploitant les facteurs géo-climatiques qui prédisent les terres les plus adaptées pour l'élevage des bovins. Les résultats des analyses montrent que les sociétés qui ont historiquement domestiquées les bovins ont eu plus d'inégalités de genre qui persistent jusqu'aujourd'hui.

Le deuxième chapitre est co-écrit avec Raphael Godefroy et Joshua Lewis. Dans ce chapitre, nous analysons les effets de court et long termes de la grande peste bovine survenue en Afrique dans les années 1890. Nous utilisons la méthode des doubles differences combinant les differences entre les terres favorables pour l'élevage des bovins et les conditions climatiques contemporaines. Notre étude montre que les sociétés les plus touchées par la peste bovine sont moins peuplés et ont moins de bovins, de plus les descendants de ces sociétés sont aujourd'hui les plus pauvres.

Le dernier chapitre analyse comment le genre des frères et sœurs influence l'âge au mariage des femmes. Les analyses basées sur le sexe du deuxième enfant revèlent que les femmes qui ont une petite sœur se marient plus tôt, ont plus d'enfants et sont moins éduquées. Les effets sont plus accentués dans les groupes ethniques qui utilisent la dot comme norme culturelle du mariage.

Mots-clés :

Valeurs patriarcales, bovins, pouvoir de négociation, transmission culturelle, peste des bovins, groupe ethniques, développement, sécheresse, genre des frr̀es et sœurs, institution du mariage, dot, contraintes économiques, Afrique.

Abstract

This dissertation investigates the deep roots of differences in gender roles and development across societies in Sub-Saharan Africa. Chapter 1 examines the deep origins of differences in gender roles in Africa. I test empirically Engels (1884) hypothesis, that the origin of differences in gender roles in Africa was the historical domestication of cattle. To address potential endogeneity in historical cattle adoption, I adopt an instrumental variables approach that leverages geo-climatic factors affecting the suitability of lands for cattle-raising. I find empirical support for Engels (1884) hypothesis. Further, the results show that these differences in gender roles have persisted to present day.

In chapter 2, which is co-authored with Raphael Godefroy and Joshua Lewis, we study the short- and long-run consequences of the 1890s African Rinderpest Epizoodic. We adopt a difference-in-differences strategy that combines differences across ethnic homelands in cattle-suitability with contemporaneous local drought conditions to identify ethnic groups that were more or less exposed to the outbreak. We find that the societies exposed to rinderpest experienced relative decreases in cattle-ownership in the decades after the outbreak. We uncover large relative long-run decreases in wealth among descendants of affected ethnic groups. These persistent economic losses appear to be partially driven by distressed migration.

In chapter 3, I investigate how sibling gender composition affects women's transition to first marriage in sub-Saharan Africa. To address potential endogeneity in the final sibling gender composition, I exploit the random assignment of the second child's gender in household with at least two children. I find that female with a younger sister get married younger, with negative consequences for her education and literacy. The effects are stronger within countries that traditionally pay bride price at marriage.

Keywords :

Patriarchal values, cattle, bargaining power, cultural transmission, rinderpest, ethnicity, development, drought, sibling gender, marriage patterns, bride price, resource constraints, Africa.

Table des matières

Liste des tableaux

Table des figures

Chapitre 1

The Historical Determinants of Gender Roles in Sub-Saharan Africa

"The domestication of cattle constituted the world historical defeat of the female sex."

–Engels (1884)

1.1 Introduction

There are vast disparities in women's labour force participation across Sub-Saharan Africa. In 2000, female participation rates ranged from 33% in Senegal to 88% in Mozambique (World Bank, 2001). These differences are mirrored in reported attitudes about the role of women in society.[1] An understanding the deep roots of these differences is crucial for the effective design of policies aimed at reducing gender inequality and promoting the empowerment of women (World Bank 2015 ; Ashraf et al. 2018).

This paper examines a prominent anthropological hypothesis first proposed by Engels (1884), that differences in gender roles in Africa have their origin in the domestication of cattle. Cattle were highly valued in pre-colonial Africa, and were the dominant source of wealth and medium of exchange.[2] Because cattle-herding required

1. The proportion of women in the *Demographic and Health Survey* (DHS) who respond "husband alone" to the statement : "person who should have greater say : visits to family or relatives" ranged from 72% in Senegal to 24% in Mozambique.

2. Cattle provided food, clothing, cooking fuel, and were used as a form of insurance against climatic

travelling long distances and working in close proximity to large animals, activities that were incompatible with child-rearing, this wealth was typically held by men. Thus, anthropologists argue that the introduction of cattle around 5,000 BC led to shift towards patriarchal norms that favoured men and placed women in subservient position (Holden and Mace 2003; Aberle 1961).

I investigate the relationship between historical cattle adoption and women's outcomes in both pre-colonial and contemporary Africa. To begin, I study the link between historical cattle presence and women's outcomes in the pre-colonial period. This analysis combines information on pre-colonial characteristics (cattle presence, female participation in agriculture, and other norms over women's role in society) across 533 African ethnic groups from Murdock's (1967) *Ethnographic Atlas* with geographical data on historical ethnic homelands from Murdock's Map (1959). I find a significant negative relationship between historical cattle presence and female participation in agriculture. The results are stable across a range of specifications, and are robust to controls for a number of historical ethnic group characteristics including political development, economic complexity, as well as controls for the geographic conditions of the ethnic homeland (tropical climate, longitude, absolute latitude, a measure of agricultural suitability and malaria ecology index). I also find that cattle-based societies were more likely to adopt other norms that disadvantaged women, including patrilineal inheritance rules – in which family wealth is transferred to sons, and patrilocality marriage customs – in which the wife co-resides with the husband's family after marriage.

I explore whether the historical presence of cattle continues to influence women's outcomes today. To investigate this question, I use data from the Demographic and Health Surveys (DHS) on more than 400,000 women from 24 Sub-Saharan African countries for the period 1992 to 2016. I link individuals to historical ethnic groups in the *Ethnographic Atlas*, building on a procedure developed by Michalopoulos, Putterman and Weil (2016) and Alesina, Brioschi and Ferrara (2016), from which I am able to match 85 percent of the sample to an historical ethnic group.

shocks (Schneider 1964; Marshall and Hildebrand 2002). In Africa, cattle were less valued for tractive power, given the unsuitable land for plough cultivation (Pryor 1985; Green 2013).

The results show consistent evidence that women whose ancestors herded cattle are significantly less likely to work today. The coefficient estimates are large in magnitude, and stable across specifications, implying that female descendants from cattle-based societies have participation rates that are 4 percentage points lower than those from non cattle-based societies. Consistent with the labour market evidence, I also find that historical cattle presence is associated with higher rates of fertility and polygynous marriage, and a decreased in women's role in household decision-making.

The baseline analysis show a strong link between historical cattle presence and women's role in society in both pre-colonial and contemporary Africa, consistent with Engel's hypothesis. Nevertheless, these patterns may not reflect the causal impact of cattle on women's outcomes, due to issues of omitted variables bias, reverse causality, and measurement error. Consider, for example, the Maasai tribe, a society with strong patriarchal norms that herded cattle. Since there are no records on the Maasai prior to the introduction cattle, it is impossible to assess whether the adoption of cattle led to the creation of these patriarchal norms, or whether the adoption of cattle was simply a consequence of pre-existing male-oriented culture.

To establish the causal impact of historical cattle presence on gender norms, I adopt an instrumental variable strategy that exploits differences across ethnic homelands in suitability of geo-climatic conditions for cattle-raising. Agronomists have identified three factors as critical determinants for cattle prosperity : abundant pasture, available water sources, and sufficiently flat terrain for migration (Marshall and Hilderbrand 2002 ; Mattioli et al. 2000 ; and Murray, Morrison and Whitelaw 1982). Motivated by the agronomic literature, I construct an index for cattle suitability (CSI) as the interaction of these three factors, using grid-cell level data from FAO's *Global Agro-Ecological Zones* (GAEZ v3.0) 2011 database, and the ESRI's Natural Earth dataset. [3] I aggregate my grid-cell level CSI measure to an average cattle suitability across historical ethnic homeland, and use this variable as an instrument for historical cattle presence. [4]

3. This functional form reflects the fact that inputs are not substitutable. For example, an increase in available pasture for grazing is of limited value if there are no available sources of water.

4. There is a strong first-stage relationship, with F-statistics for the excluded instrument typically exceeding 16.

The IV estimates support the baseline findings. I find a negative and statistically significant relationship between historical presence and women's participation in both the pre-colonial and modern periods. I also find similar effects for various measures of women's empowerment and gender norms.

Next, I explore the mechanisms underlying the relationship between historical cattle presence and women's outcomes. The results cannot be attributed to plough cultivation. The estimates for historical cattle presence are unchanged when I control for the traditional practice of plough agriculture, as reported in the *Ethnolographic Atlas*, or when I instrument for traditional plough presence using local geo-climatic suitability for "plough-positive" crops (Alesina, Giuliano and Nunn, 2013). These results suggest that, in Africa at least, cattle presence did not influence women's outcomes through historical agricultural practices. These findings contrast with the views of Boserup (1970), who argues that men's physical comparative advantage in controlling large animals during plough cultivation ultimately led to a shift in gender roles. Instead, these patterns likely reflect the fact that soil characteristics in Africa were largely unsuitable for plough agriculture (Pryor 1985; Green 2013).[5] Whereas traditional plough cultivation played an important role in gender norms in other regions (Alesina, Giuliano and Nunn, 2013), its effect was limited in Africa, and the technology cannot account for the large differences in participation across cattle-based and non cattle-based societies.

I assess a number of alternative explanations for the observed link between historical cattle presence and women's outcomes. First, I examine whether the patterns can be attributed to pastoralism rather than cattle-ownership per se. Anthropologists have argued that pastoralist societies often adopt patriarchal norms (i.e. patrilocality and low female participation in agriculture) as a means of reducing paternal uncertainty due to extended male absences associated with herding (Becker 2018; Xia 1992). To assess this hypothesis, I include controls for other herded animals – sheep and goats – in the main specifications. The effects for historical cattle presence are unaffected by these covariates. In addition, I show that the results cannot be attributed to the tsetse fly, historical exposure to the slave trade, or a number of historical ethnic group characteristics. Taken

5. Just 7 percent of pre-colonial ethnicities used plough cultivation (Murdock, 1969).

together, these results provide strong support for the view that historical cattle presence, by creating large gender-imbalance in wealth-holdings in pre-industrial Africa, led to a historical shift in gender norms that continues to influence women's outcomes today.

Why are modern gender roles still shaped by historical cattle presence ? One possibility is that these animals continue to play a crucial role in economic and social life, so the forces that operated historically are still at work today. Alternatively, the traditional roles that emerged in pre-industrial Africa may continue among the descendants of cattle-herders in the modern era, even if cattle no longer play a dominant role in society. This persistence could arise from the continuity of cultural norms or endogenous changes in formal institutions that inhibited women's participation over the long-run. To shed light on these questions, I first estimate the effect of historical cattle presence among descendants who reside in urban versus rural areas. Intuitively, since cattle are rarely held in urban areas, this comparison distinguishes the direct effect of contemporary cattle ownership from the legacy of the animal's presence among one's ancestors. The coefficient estimates are virtually identical across the two sub-samples, suggesting an important role for the transmission of cultural beliefs. Furthermore, I exploit the fact that roughly half of the contemporary sample no longer resides within their historical ethnic homeland to separately identify the effects of historical ethnic-based cattle presence from historical location-based cattle presence. I find that the ancestral cattle-ownership has a significant negative impact on contemporary women's participation decisions, whereas historical location-based ownership has no significant impact on participation. Taken together, these findings suggest that historical cattle ownership primarily shaped long-run outcomes through a persistent change in cultural attitudes regarding the role of women in society.

This paper contributes to two strands of the literature. First, it is directly related to the literature on determinants of gender roles. Previous studies have documented the importance of economic development (Goldin 1994 and 2006 ; Ross 2008 ; Duflo 2012), medical progress (Goldin and Katz (2002) ; Albanesi and Olivetti 2007 ; Miller 2009), improvements in labor-saving consumer durables in the household (Greenwood et al. 2005 ; Graham, Hirai and Kim 2016 ; Lewis 2018). While others emphasized social and cultural factors including patrilocality (Dyson and Moore 1983; Chakraborty and Kim 2010; Ebenstein 2014; Bau 2016), patrilineality (Deininger et al. 2013; Lowes 2017 and 2018), brideprice and dowry (Arnold et al. 1998; Das Gupta 2003; Lowes and Nunn 2017; Ashraf et al. 2016). I contribute to this literature by showing how the historical cattle presence created large gender-imbalance in wealth, which led to a shift in gender norms that continues to affect women's outcome today.

Second, the paper generally contributes to the literature on the persistence of historical and cultural causes of gender roles. Previous work has demonstrated the persistent influence of historical factors on women's outcomes, including the role of traditional plough cultivation on women's participation, and fertility (Alesina, Giuliano, and Nunn, 2013, and 2011), the long-run effect of the Neolithic revolution on current gender norms (Hansen, Jensen, and Skovsgaard 2015), the effect of the Africa's slave trade on polygyny and women's participation in the labor force (Edlund and Ku 2011; Fenske 2015; Dalton and Cheuk Leung 2014; and E. Teso 2018), the role of pastoralism on women's outcomes (Voigtlander and Voth 2013; Becker 2018). I contribute to the literature by showing how the introduction of cattle in Africa around 5,000 BC led to shift towards patriarchal norms, and persistently affects women's outcomes even after. By studying these relationships in both the historical and modern contexts, I am able to track the evolution of these norms over an extended time horizon.

The paper proceeds as follows : Section 1.2 discusses the history of cattle and the role of women in Africa; Section 1.3 describes the data construction; Section 1.4 presents the empirical strategy; Section 1.5 presents the results for pre-colonial analysis; Section 1.6 presents the results for contemporary data; and Section 1.7 concludes.

1.2 Cattle and the Role of Women in Sub-Saharan Africa

Anthropologists argue that early African societies were matriarchal, with lineage and inheritance that were passed through mother's line, but that over time many adopted patriarchal norms (Morgan 1871 ; Engels 1884). Since Engels (1884), a number of scholars have argued that the source of this transition was the introduction of cattle (Aberle 1961 ; Holden and Mace 2003).

Cattle were first introduced to Africa around 5,000 BC, through the migration of farmers from the Middle East (Epstein and Mason 1984 ; Mukasa 1989). The animals were crossbred with wild African aurochs to constitute indigeneous African cattle population which can be classified into four major categories : humpless Bos taurus, humped Bos indicus (zebu), sanga (crossbreding of Bos taurus and humped Bos indicus) and Zenga (sanga and zebu backcross).[6]

Vegetation type, percent slope of soils and availability of water were the main factors that drove geographic spread of cattle throughout Africa (Senft, Rittenhouse and Woodmansee 1985 ; Ganskopp et al. 2007 ; Owens, Launchbaugh and Holloway 1991). Specifically, Ganskopp et al. (2007) argues that locations suitable for keeping cattle must have high quality and quantity of pasture, with sufficient water, and flat slope. More precisely, the physiological features of cattle do not allow them to travel more than 3.2 kilometers per day for water, or to graze on slope above ten percent, they must consume at least 36 lbs of pasture per day. These requirements are consistent with the geographic pattern of cattle throughout the continent. For instance, the zebu cattle, one of the major type of cattle in Africa, were widespread in the Sahel and Savannah, but were nearly absent in the Sahara and Kalahari desert (Green, 2013). The tsetse fly had been identified as another factor that slowed down the spread of cattle in the continent. As noted Diamond (1997, p. 186) "the spread southward of Fertile Crescent domestic animals through Africa was stopped or slowed by climate and disease, especially by trypanosome diseases carried by tsetse flies. [...] The advance of cattle, sheep, and goats halted for 2,000 years at the northern edge of the Serengeti Plains, while new types of human economies and livestock breeds were being developed." Over time, cattle that were resistant to trypanosome disease were bred (Hannote et al., 2000). For example, the Bos taurus subspecies, which are resistant, are commonly found in the tsetse fly belt around Central and West Africa. However, their small size and their low productivity

6. See Hanotte et al., (2000) for further details on African cattle characteristics.

made them useless in the fields, and they were primarily valued for cultural purposes (Murray, Morrison and Whitelaw 1982).

Anthropological records highlight the essential economic role of cattle in pre-colonial Africa. For agricultural societies, cattle provided an important source of insurance against volatile climatic conditions (Marshall and Hilderbrand 2002). They provided the main source of animal protein, they were used as draft power in the fields, and their dung was used as fertilizer and fuel (Schneider 1957). Cattle were also used as means of exchange to acquire grain from another man who has surplus. Given these economic benefits, cattle were highly valued in pre-colonial societies, and considered the central source societal wealth (Schneider 1957 ; Comaroff 1990). As Burchell (1822, p. 347) states : “From the possession of cattle, the distinction of men into richer and poorer classes has followed as natural consequence".

According to custom, men held ownership over cattle. Deshler (1963) identifies two major reasons for this unequal distribution of societal resources. First, the cattle indigenous to Africa were large, with sharp horns that made them dangerous to corral, especially for women with young children (see Figure 1.1). For instance, the Ankole longhorn cattle from central and eastern Africa usually weigh 900lb for males and 700lb for females, and their horn can measure six to eight feet in length.[7] Second, because cattle consume large amounts of pasture and drink an average of 10 gallons of water per day, in the dry season herders had to travel long distances and spend many days away from camp, an activity that was incompatible with child-rearing. This view is outlined by Dupire (1963, p.75), who argues, “to look after the cattle, [...] demands activities of which a woman is physically incapable." She continues "It would be beyond a woman’s strength to draw water for the herd in the dry season, to go on long marches to reconnoitre for grazing lands, to protect the herd against wild animals and thieves, to hold her own with a buyer at the market, to castrate bulls, or to train the pack oxen. This hard, dangerous life, full of uncertainty and of prolonged absences from the camp, would be incompatible with the duties of motherhood, which require a more sedentary and more regular life". In line with the narrative, cattle herding was male dominated activity in almost 80% of

7. See Deshler (1963) for more details.

the societies in Africa according to the Murdock's (1967) Ethnographic Atlas.

The presence of cattle in African societies coincided with a number of patriarchal social norms. One such norm is patrilineality, in which lineage and inheritance are passed through the father's line. Engels (1884) argues that early hunter gatherer societies practiced matrilineality custom, but that the domestication of cattle improved men's position in the family, which they used to overthrow matrilineal descent and established a new norm in favor of their own children. Under patrilineal custom, lineage and inheritance were transmited from father to sons, so women became dependent on husbands for their wealth. Empirical support for this argument is provided by Aberle (1961), who finds that matilineality is more likely to evolve in horticultural societies, while patrilineality is practiced in pastoralist societies with a large presence of cattle. He concluded that, "the cow is the enemy of matriliny, and the friend of patriliny". Further support for this hypothesis was provided by Holden and Mace (2003).

Cattle presence has also been linked to patrilocality norms, in which the wife takes residence with the husband's kinship group after marriage (Vroklage, 1952). The practice of patrilocality is often associated with a bride price, for which cattle provide the typical form of payment (Goldschmidt, 1974). This payment provides the husband rights to women's labor and reproductive capabilities (Anderson, 2007). Further, patrilocal residence is associated with lower investment in female health and education, this might reflect the fact that parents, who anticipate these future post-marriage living arrangements, will direct resources to their sons, who will continue to work on family farms (Jayachandran, 2015).

Finally, the presence of cattle might increase women's fertility. Cattle-based societies may have developed gender preferences that favored sons (Gitungwa, 2018). These preferences could reflect the value of sons as helpers in management, and also their insurance value for livestock herding, in the case of the father's death. These preferences could promote fertility, in an effort to acquire a son, at the cost for females of participating in work outside the home.

1.3 Data Construction

This paper explores whether the historical presence of cattle affected women's outcomes in the pre-colonial and contemporary periods. The pre-colonial analysis relies on the *Ethnographic Atlas* data, and the contemporary investigation on the combination of the *Ethnographic Atlas* with the Demographic and Health Surveys (DHS). The next sections present these datasets, and describe the matching procedure I use to combine the *Ethnographic Atlas* with the DHS data.

1.3.1 Pre-colonial Data

The information on ethnic groups pre-colonial traits comes from Murdock's (1967) *Ethnographic Atlas*. The Atlas is a compilation of ethnographic, anthropological, and archives of missionary studies that George Peter Murdock assembled and published in 29 installments in the journal *Ethnology* between 1962 and 1980. It contains information on historical characteristics of 1,267 societies from around the world in which 533 are from Africa.

Although the Atlas is the best available cross-cultural data, its greatest shortcoming is that ethnic groups are observed in different periods of time. The approximate date of observation ranges from 1830 to 1960 [8]. Thus, ethnic groups observed in 1830 may sytematically differ from societies observed in 1960 due to the passage of time (Henderson and Whatley, 2014). Fortunately, this is not an issue that prohibits the use of the data set since most of the characteristics of African societies are thought to remain stable over time.

Murdock's Ethnographic atlas has been widely used. Alesina, Giuliano and Nunn (2013) aggregated the *Ethnographic Atlas* at country level to show that the historical use of plough agriculture predicts contemporary gender inequalities. Michalopoulos, Putterman and Weil (2016) have combined this data with Demographic and Health Surveys dataset in order to show that individuals whose ancestors relied on agriculture

8. The Ethnographic Atlas is intended to describe as much as possible the characteristics of the ethnicities before their contacts with Europeans.

for subsistence are today more educated and wealthier. Alesina, Brioschi and Ferrara (2016) have investigated the determinants of violence against women in Africa, they showed that ancestral socioeconomic conditions explain intrafamily violence today.

In order to determine the spatial locations of the ethnic groups, I follow previous work and combine the Atlas with the Murdock's (1959) Map [9]. This map provides the geographic boundaries of approximatively 830 African ethnic groups. In total, 522 African ethnic groups from the Ethnographic Atlas are matched to the Murdock's Map. Figure 1.2 (a) portrays boundaries of African ethnic groups in the late 19th and early 20th century.

My main explanatory variable is an indicator for the pre-colonial presence of large domesticated cattle. I also construct a continuous measure of the society's dependence on cattle, which is calculated as the interaction of cattle presence and the degree to which the society relied on husbandry for subsistence (ranges 0 to 100%). Figure 1.2 (b) shows the spatial variation in pre-colonial presence of cattle for African ethnic groups in the Ethnographic Atlas. Cattle were historically present among 58% of ethnic groups. Their presence was concentrated in the Sahel and along the East and West coasts, regions that had ample pasture and flat soil. Meanwhile, few ethnic groups raised cattle in the tropical forest, due to the inadequate pasture and the endemic tsetse fly.

I use the following outcomes for the precolonial analysis. *Female Participation in Agriculture* is a dummy variable that takes the value of 1 if in the ethnic group women performed the majority of agricultural tasks. This outcome is relevant to capture females empowerment since societies characterized by fewer participation of females in agriculture tend to unequal gender norms that subardinate women over men (Alesina, Giuliano and Nunn (2013)). *Patrilineal inheritance* is an indicator variable for whether ethnic group practiced patrilineal inheritance systems relative to matrilineal inheritance. In patrilineal inheritance systems, inheritance is traced through male members. Studies showed that women are less competitive relative to men in patrilineal kinship system (Gneezy, Leonard and List 2009; Andersen et al. 2013). *Patrilocality* takes the value of 1 if patrilocal

9. Since the ethnicity names from the Murdock's Map do not always coincided with the names from the Ethnographic Atlas, Fenske (2014) and Michalopoulos and Papaioannou (2013) constructed a concordance rule that allows to link the groups from the Atlas to the groups in the Map.

residence was the rule within ethnic group as compared to matrilocality. The patrilocality tradition is the fact that married couple lives with the parents of the husband. In this case, women are more likely to experienced violence (Alesina, Brioschi and Ferrara, 2016). *Brideprice* is a dummy equals to 1 if ethnic group practiced brideprice marriage, and zero otherwise. *Polygyny* equals to 1 if polygyny marriage was the rule, and zero otherwise.

In the precolonial analysis, I include a set of controls variables measured at the ethnicity-level. Longitude controls for differences between the eastern and the western part of the continent. Absolute latitude and proportion of land in tropical and sub-tropical capture differences in the agricultural and ecological areas. Altitude is a I include a measure of agricultural suitability from the Food and Agricultural Organization (FAO). Malaria ecology index from Kiszewski et al. (2004) control for different forms of malaria. Date of observation controls for differences in the years of observation of ethnic groups in the Ethnographic Atlas. Finally, I control for political complexity and economic development from the *Ethnographic Atlas*. Political complexity is measured by the levels of juridictional hierarchies in the ethnic group. Economic development, measured using density of ethnic groups settlements.

1.3.2 Contemporary data

To study the effects of historical cattle adoption on contemporary women's outcomes, I use data from the *Demographic and Health Surveys* (DHS). The DHS are nationally representative households surveys which cover over 90 developing countries. They consist in individual-level datasets which provide information on a range topics including female labor participation, education, female decision making within household, and marriage market outcomes (age at first marriage, fertility, polygyny and sons preferences). The DHS also provides individual characteristics such as age, religion, and literacy. Importantly, the surveys provide information on ethnicity, allowing me to identify descendants of the ethnic groups reported in the Murdock datasets. I use 84 surveys from DHS, covering 24 Sub-Saharan African countries over the period 1992-2016. My final dataset is a sample of 1,065,768 individuals, in which about 72 percent are women aged between 15 to 49.

The main outcomes of interest is an indicator for female labor force participation, FLFP, that equals to one if the respondent is currently working or had worked in the twelve months before the survey. Since the DHS does not provide information whether individuals are looking for jobs, I consider that an individual does not work if she did not work for more than a year. I also construct indicators for employment by occupation for the follow occupations : domestic service, agriculture, manual, clerical, and sales. Other outcomes include *fertility*, measured by the woman's number of children ever born ; *education* ranges from 0 to 5 for categories "no education", "incomplete primary", "complete primary", "incomplete secondary", "complete secondary", and "higher" ; *son preferences* is an indicator variable that takes the value of one if respondent prefers to have a son child ; *age at first marriage* measuring woman's age at first marriage ; and *polygyny*, which is equal to one if the respondent has co-wives. *female decision-making*, captures whether females played the dominant role in household decisions over specific issues such as health care ; large purchases ; visits to relatives ; how many children to have ; food to be cooked ; and what to do with women's earnings. This variable ranges from 0 to 4 and is increasing with women's ability to make household decisions. I also construct the variable *attitude towards violence*, which measures the share of specific cases for which female agrees that violence against women is justified. [10]

I include a set of controls including respondent's age and age squared, a quadratic in national income per capita, survey year fixed effects. I also add covariates used for precolonial analysis. Table 1.2 reports the summary statistics from the DHS data.

1.3.3 Linking Contemporary Outcomes to Historical Ethnic Groups

To investigate the long-run impact of cattle on contemporary outcomes, I match individual-level data from the DHS to Murdock's ethnic group data based on ethnic names reported in the modern surveys. Because ethnic groups in DHS often do not

10. The respondents are asked whether a husband is justified in hitting or beating his wife under different events : the wife goes out without telling him ; the ife neglects their children ; the wife argues with him ; the wife refuses to have sex with him ; the wife burns the food.

coincide with historical ethnic names, I employ a matching procedure that builds on extends those used by Michalopoulos, Putterman and Weil (2016) and Alesina, Brioschi and Ferrara (2016).

The matching techniques is as following. First, I match DHS ethnic groups that coincide with ethnic names in the Murdock Atlas. If no match is found, I use the dataset from Nunn and Wantchekon (2011), who constructed a concordance between the ethnic groups names in the Afrobarometer and the names in the Atlas, I then match Afrobarometer ethnicities that match uniquely to ethnicities in the DHS. For example, the ethnicity "Pokot" which is included in DHS and Afrobarometer, which the Nunn (2011) file links to the "SUK" ethnic group in Murdock's *Atlas*. For the remaining unmatched data, I use the 22nd edition of the Ethnologue, a pamphlet containing information of more than 7,111 spoken languages in the world, to identify alternate variants of DHS ethnic names and use these to link to the Atlas. For example, the ethnic group in the DHS "Pular" can be listed as alternative names by Ethnologue as "Foula Fouta" ; "Fouta Djallon" ; "Fulbe" ; "Fulfulde Jalon" ; "Fullo Fuuta" ; "Futa Fula" ; "Futa Jallon" ; "Fuuta Jalon" ; "Jalon" ; and "Poular". I match these groups with the ethnic names "Futajalon" in the *Ethnographic Atlas*. For remaining unmatched groups, I use information from the Joshua Project, a database of more than 9,803 ethnic groups, to identify possible links between the DHS and the *Ethnographic Atlas*. Following Alesina, Brioschi and Ferrara (2016), I identify three potential links cases in which there are links between these two sources : 1) DHS and Murdock ethnic groups names are listed as alternative names by the Joshua Project ; 2) the ethnic name in the *Atlas* is identified as a superset of the ethnic group in the DHS ; 3) the ethnic group in the *Atlas*, is a subset of the DHS ethnicities. For remaining unmatched ethnicities in DHS, I use Wikipedia to identify alternate spelling of ethnic names. For example, For example, the DHS ethnic group "Djerma" is listed as an alternative name for "Zerma", the name reported in the *Atlas*. Finally, I draw on other sources as *Peoplegroups.org*, and *Zyama.com* to match the remaining ethnic groups name.

Overall, I am able to match 84.5 percent of observations from the DHS to the *Ethnorahic Atlas*. Totalling 1,065,768 respondents matched to their ancestral characteristics from the *Atlas*, equivalent to 169 ethnic groups.

1.4 Empirical Strategy

1.4.1 OLS Specification

To examine the effects of cattle presence on pre-colonial outcomes, I estimate the following equation :

$$Y_e = \alpha + \beta Cattle_e + Z_e^{'}\Gamma + \epsilon_e \tag{1.1}$$

Where e indexes ethnic group. $Cattle_e$ is an indicator for the historical presence of cattle among ethnic group e.[11] The coefficient of interest is β, which captures the relationship between cattle and the various pre-colonial outcomes.

The term Z_e denotes a vector of socioeconomic and geographic variables that may be correlated with both the adoption of cattle and pre-colonial gender roles. I use information reported in the Murdock Atlas to construct a control for the economic complexity, measured in integer values from 1 to 8 that are increasing in the density of ethnic group's settlement.[12] Because historical settlement density may have been endogenous determined by the presence of cattle, I report results with and without this covariate. I also construct a control for political hierarchy, which takes values from 0 to 4, based on the number of political jurisdictions.[13] I also control for the year the ethnic group was reported observed in the *Ethnographic Atlas*. In addition, I control for a number of geographic conditions of ethnic homelands that might simultaneously influence gender norms and the presence of cattle. These controls include an index for agricultural suitability and the proportion of the ethnic homeland that falls in tropical or subtropical

11. In some specifications, $Cattle_e$ is measured as a continuous variable for historical reliance on cattle, measured as the interaction of cattle presence and the degree to which the society relied on husbandry for subsistence.

12. The *Ethnographic Atlas* grouped ethnicities into the following categories : (1) nomadic or fully migratory, (2) semi-nomadic, (3) semi-sedentary, (4) compact but not permanent settlements, neighboohoods of dispersed family homesteads, (5) separate hamlets, (6) forming a single community, compact and relatively permanent settlements and (8) complex settlements.

13. A score of 0 designates stateless societies, a score of 1 is using to indicate paramount chiefdom, a score of 3 and 4 indicate groups that were part of large states.

zones from the FAO. I also control for the Malaria Stability Index from Kiszewski et al., (2004), given the link between malaria and women's labour force participation (Quisumbing et al., 2014). Finally, I control for ethnic homeland longitude and absolute latitude, that capture differences for the eastern and western part of the continent.

For inference, I cluster the standard errors at the level of ethnic provinces to account for potential correlation in outcomes across ethnicities that belong to the same ethnic hierarchy.[14] These provinces are described by Murdock (1959) capture both spatial and genealogical correlations. This approach allows for potential correlation across ethnicities that shared common ancestors, cultural histories, or geography. For example, the migration patterns during the Bantu expansion may have led to correlation in outcomes across geographically diffuse ethnicities that nevertheless shared a common lineage. I also report Conley (1999) standard errors which correct for spatial correlation between the observations as robustness checks.

To examine the effect of pre-colonial cattle presence on contemporary women's outcomes, I estimate the following equation :

$$Y_{i,e,c} = \alpha + \beta Cattle_e + Z_e^{'}\Gamma + X_{i,c}^{'}\Theta + \epsilon_i \tag{1.2}$$

where i indexes individuals, e ethnic groups, and c countries. My coefficient of interest β captures the average treatment effect of the historical adoption of cattle. The term $Z_e^{'}$ includes the historical ethnic group-level control variables, while $X_{i,c}^{'}$ represents a vector of individual-country covariates for respondent's age, age squared, survey years fixed effects, and a quadratic in national per capita income.[15]. Finally, I cluster the standard errors at the ethnic group level, in order to account for within-group correlation of the residuals. $Z_e^{'}$ includes previous historical ethnic group-level control variables.

1.4.2 IV Specification

The OLS specifications may not identify the causal effects of cattle presence on women's role in society. First, it is possible that there is an omitted variable that both determines

14. This method has been used by Alsan (2015).
15. Given endogeneity concerns, I report the results with and without the national income covariates.

cattle presence and women's outcomes. Consider, for example, the Maasai tribe, a society with strong patriarchal norms that herded cattle. Since there are no records on the Maasai prior to the introduction of cattle, it is impossible to assess whether the adoption of cattle led to the creation of these patriarchal norms, or whether the adoption of cattle was simply a consequence of pre-existing male-oriented culture. Second, the recording of cattle presence in pre-colonial societies is likely to be measured with error. Ethnographers observed ethnic groups each at a single point in time. As a result, some societies may have only recently acquired the animals, whereas others may have been formerly reliant on cattle only to lose them to disease, war, slave raids, or theft. The Tonga tribe for example, used to keep cattle, but only lost them on the eve of colonization due to raids and the rinderpest epidemic of 1896 (Dixon-Fyle, 1983). However, the Tonga has been observed in the *Ethnographic Atlas* in 1940, and recorded as no cattle keepers. This form of measurement error will tend to bias the OLS estimates towards zero.

To address these issues, I use an instrumental variable strategy. Agronomists have identified three important determinants of cattle presence : abundant pasture, available water sources, and sufficiently flat terrain for migration (Murray, Morrison and Whitela 1982 ; Bailey et al. 1996 ; Mattioli et al. 2000 ; Ganskopp et al. 2007). They argue that the physiology of cattle requires high quality and quantity of forage sources ; and makes them difficult to travel up to 3.2 kilometers a day for water, specially on land slope above ten percent. Motivated by these arguements, I construct an index for cattle suitability (CSI) that is the interaction between : (1) The soil suitability for pasture ; (2) distance to nearest water source ; and (3) the measure of land slope gradient.[16]

Two assumptions must hold for *CSI* to be a valid instrument. First, Cattle Suitability Index must be a strong predictor of the presence of cattle. This assumption is supported by the estimates of the first-stage regression in Section 1.5. The second assumption requires that changes in ethnic group's *CSI* was uncorrelated with the error term from the estimation equation. There is no reason the combination of three factors should independently have effects on women's outcomes. To verify this arguement, I control for

16. Tsetse fly has been identified as another predictive factor for cattle presence (Hanotte at al. 2000). While I do not include tsetse fly in the index, I control for the tsetse suitability index from Alsan (2015) in the robustness.

each component of the *CSI* in the robustness.

To map the spatial distribution for cattle suitability, I develop a Habitat Suitability Index (HSI) model for cattle population. HSI models was originally developed by the U.S. Fish and Wildlife Services in 1981. These models are constructed in three steps : (1) the first step identifies factors influencing animals distribution ; (2) the second step develops suitability index functions for each individual factors ; and (3) the third step combines these functions into a functional form equation for the HSI. [17] HSI models are commonly used to determine habitat suitability of species, generally when the actual distribution of these species is not observed. Further, they are simple to use and understand, and has been identified as overcoming the limitations of other models such as multiple regression models Pandey (2007).

Suitability for pasture data are from the FAO's *Global Agro-Ecological Zones* (GAEZ v3.0) 2011 database. The GAEZ data set provides information on land suitability for pasture and other vegetables such as cassava, millet, and shrubs, for 3 arc-seconds in Global Land Cover (GLC2000) (Fischer et al., 2012). These data are constructed by combining various land cover maps, including vegetated land, forest resources assessment, irrigated land areas, protected areas inventory, and population density inventory. Data on suitability for pasture is ranged 0 to 100. Data on terrain slope are also from the FAO's *Gaez* which provides terrain slope gradients data for each 3 arc second grid. The terrain slope have been compiled using elevation data from the Shuttle Radar Topography Mission (SRTM), which is available as 3 arc-second Digital Elevation Models. I use the median terrain slopes of 3 arc-second grid-cells, which are grouped into eight categories [18]. Data on location of rivers and lakes are from the ESRI's Natural Earth dataset available at 1 :10m scales. The data provides rivers and lakes centerlines of the world. I calculate the cost distance to water for each grid-cells (rivers or lakes), and I reclassify the resulting distance as followed : (1) distance below 1.6 km taking the value 1 ; (2) a value of 0.5 is given to distance between 1.6 to 3.2 kilometers ; and (3) distance more than 3.2 km is assigned a value of zero.

17. For more details see Pandey (2007).
18. Details on data are in appendix.

I restrict the geographic extent of my analysis to Africa, using a spatial resolution of 5×5 km grid cells. Rescaling the previous factors to lye between zero and one, I calculate the cattle suitability index (CSI) according to the following equation :

$$CSI = Pasturesuitability \times (1 - distancetowater) \times (1 - Terrainslope)$$

This functional form reflects the non-substitutability across geographic inputs. For example, a lack of local sources of water cannot be offset by more pastoral land for grazing. [19] The Cattle Suitability Index is the standardized value of this functional form.

Figure 1.3 reports values of the CSI. The Sahel and the temperate zones are zones with high index for cattle suitability. Further, Locations with cattle presence from the Ethnographic Atlas coincide with places suitable for cattle.

1.5 Pre-colonial Results

1.5.1 Cattle Presence and Female Participation in Agriculture

This subsection reports the relationship between historical cattle presence and pre-colonial outcomes. Table 1.3 reports the results from different OLS specifications. Column 1 reports the baseline results without controls. In column 3, I include geographic covariates for agricultural suitability, longitude and absolute latitude, malaria ecology index and date of observation. In column 5 and 7, I add covariates for the degree of economic complexity and political hierarchy in pre-colonial societies.

The results show that the historical presence of cattle is associated with lower relative participation of women in agriculture. [20] The point estimates are stable across the various specifications ranging from -0.198 to -0.205. These effects are large and statistically significant, implying that female participation was 42 percent lower in cattle-based societies.

Table 1.4 further explores the relationship between cattle presence and women's role in

19. In the empirical analysis, I explore the robustness of the results to an additive model of the three inputs.

20. Female participation is measured relative to male participation.

agricultural production. The analysis is based on specific agricultural tasks, information that is available for a subset of historical ethnicities from Murdock and White's (1969) Standard Cross-Cultural Sample (SCCS).[21] The estimates show that cattle presence is negatively associated with women participation in all agricultural tasks including land clearance, soil preparation, planting, crop tending, and harvesting.

One possible explanation for the negative relationship between cattle presence and women's participation in agriculture found in Tables 1.3 and 1.4 is that women were the primary caretakers for cattle. While this explanation contradicts the historical narrative, it could explain the withdrawal of women from agricultural work. I explore this possibility, estimating the relationship between historical cattle presence and women's relative participation in animal husbandry. Table 1.5, column 1 reports the results. The coefficient estimates are large, negative, and statistically significant. Another explanation for the negative relationship between cattle presence and women's participation in agriculture and husbandry is that women were responsible for fishing. Table 1.5, column 2 documents a negative, and statistically insignificant relationship between historical cattle presence and women's relative participation in fishing. Together, these findings suggest that historical cattle presence did not cause a shift of women's labour from agriculture to husbandry or fishing, consistent with the historiography. Instead, the domestication of cattle appears to have led to a broad decrease in pre-colonial female participation. In fact, to the extent that cattle domestication caused a decline in male agricultural labour, as they increased time to animal husbandry, the estimates for relative female participation in agriculture may understate the overall in their agricultural work.

The OLS estimates presented so far, while strong and statistically significant, might not be causal due to reverse causality, omitted variable bias, or measurement error. To address these issues, I estimate IV regression, using the cattle suitability index (CSI) as an instrument for historical cattle presence. I begin by verify the first-stage relationship between the CSI and historical cattle presence. Figure 1.7 supports the relationship showing that the presence of cattle is positively correlated with land suitability for keeping cattle. The bottom panel of Table 1.3 shows the first stages estimates. The

21. I use this data because the *Ethnographic Atlas* does not provide information on agricultural tasks.

Kleibergen-Paap F−statistics on the excluded instruments document that the *CSI* is a strong predictor of the cattle presence, with F-statistics that are all greater than 10.

Odd columns in Table 1.3 reports the IV results for female participation in agriculture. Overall, the IV estimates confirm the results from the OLS regressions in terms of sign and statistical significance. Historical cattle presence is negatively associated with historical participation of women in agriculture. In addition, Table 1.5 columns 2 and 4 show the IV estimates for female participation in husbandry and fishing. The results confirm the OLS estimates and document a negative relationship between historical cattle presence and historical participation of women in husbandry and fishing. In general, the IV estimates are larger than the OLS coefficients in term of magnitude, which could reflect measurement error in the OLS specification.

1.5.2 Cattle Versus Plough Cultivation

So far, I have presented evidence that the adoption of cattle was negatively associated with women's participation in agriculture. However, the presence of large domestic cattle may have allowed societies to adopt plough agriculture, which has been shown to independently contribute to reduced female labour force participation (Alesina, Giuliano and Nunn, 2013). Because the soil characteristics in Africa are generally unsuitable for plough agriculture (Pryor, 1985), [22] however, there is limited scope for this mechanism. In fact, just 7 percent of pre-colonial ethnicities used plough cultivation. Nevertheless, in this section I investigate the extent to which the effect of cattle on women's participation can be attributable to greater reliance on plough cultivation.

Table 1.6, col. 3 reports the OLS estimates for plough cultivation. I find that historical use of plough cultivation is negatively associated with female participation in agriculture, consistent with the patterns found by Alesina, Giuliano and Nunn (2013). Column 4 reports the results from IV regressions that use the suitability of locations for the cultivation of plough-positive versus plough-negative crops. [23] The IV estimates for plough

22. The thin topsoil of Africa combined with the concentration of the nutrients on the surfaces makes the use of plough inefficient Green (2013).

23. The suitability for plough-positive crops measures the average suitability for wheat, barley, and rye, normalized by the overall suitability for cultivation in general. Whereas the suitability for plough-negative

cultivation are negative and remain statistically significantly, despite the weak first stage relationship. [24]

Next, I explore whether the negative relationship between historical cattle presence and female participation in agriculture can be attributable to plough cultivation. Table 1.6, col. 5 and 6 report the results from OLS and IV regressions that include separate indicators for cattle presence and plough cultivation. [25] The coefficients on cattle remain large, negative, and statistically significant in these specifications. Together, these results suggest that the effects of cattle on women's historical participation in agriculture cannot be attributed to plough cultivation. The coefficient estimates for plough cultivation are also negative and statistically significant. Since all plough-based ethnic groups in the sample also raised cattle, these estimates capture the differential effect of cattle in societies that also used plough cultivation. The negative coefficients imply that some fraction of the overall effects of cattle presence on women's participation were mediated through the use of plough cultivation. Together, these findings suggest that the use of plough cultivation cannot fully account for the negative relationship between cattle presence and women's participation in agriculture in pre-colonial African societies.

1.5.3 Robustness checks

A first concern with the results presented so far is that pre-colonial cattle presence was systematically related to the tsetse fly, which could directly affect female participation in agriculture through the disease environment (Alsan, 2015). The IV specification, which relies on variation in local suitability for cattle-rearing based on the CSI rather actual cattle presence, should be less subject to these concerns. In fact, Figure 1.4 shows no correlation between the suitability of climate for tsetse flies and the CSI. To further investigate this question, I re-run the main specifications, controlling for

crops is the average suitability for sorghum, foxtail millet, and pearl millet, normalized by the overall suitability for cultivation in general.

24. The F-statistic of 3.3 on the excluded instruments is consistent with the limited adoption of this technology throughout Africa.

25. In the IV specification, I instrument for cattle with the CSI and plough cultivation with the suitability for plough-positive and plough-negative crops. The F-statistic from the excluded instruments in 17.

the tsetse fly suitability index (see Alsan (2015)). Table 1.7 column 1 shows that the OLS estimates for cattle presence remain large, negative, albeit imprecise. Note that controlling for TSI increases the standard errors for cattle since tsetse flies and cattle are strongly (negatively) correlated, reducing precision. However, the IV estimates presented in column 2 remain large, negative and statistically significant. This is consistent with the absence of correlation between the tsetse suitability and the CSI presented in Figure 1.4.

A second concern is that the results might be attributed to pastoralism rather than cattle-ownership per se. Anthropologists have argued that pastoralist societies often adopt patriarchal norms (i.e. patrilocality and low female participation in agriculture) as a means of reducing paternal uncertainty due to extended male absences associated with herding (Becker 2018 ; Xia 1992). To assess this hypothesis, I include controls for other herded animals – sheep and goats – in the main specifications. The effects for historical cattle presence are unaffected by these covariates.

A third shortcoming is that elements of CSI may have influenced women's participation in agriculture directly, independently of the presence of cattle. For example, proximity to rivers might cause women to devote more time to water collection, reducing the time available for agricultural production. Importantly, because the CSI is constructed as a combination of the three factors interacted, I am able to test the individual influence of each factor. Table 1.8 reports the results from IV estimates based on the CSI that also control for each factor individually. The results do not change.

Table 1.7, cols. 5 to 12 report the results from alternate specifications that control for a range of additional historical controls including export slave trade, ethnic group land area, indigenous slavery, initial population density, and the level of reliance in fishing. The estimates for cattle presence are generally unaffected by these various covariates. In addition to these various controls, I have also explored the robustness of the results to alternative method, including probit and biprobit specification. The results remain the same in term of sign, significance and magnitude.

Finally, I assess the robustness of inference to alternative forms of cross-group correlation in outcomes. The baseline specification corrects for within-ethnic language correlation. This form of clustering allows for correlation in outcomes, even across geographically disparate groups, for example, due to the Bantu expansion. Table 1.9 also reports results based on Conley (1999) standard errors with cutoff of 10 degrees longitude and latitude. In columns 2 and 3, I cluster the standard errors by country, and in columns 4 and 5, I use the multiway clustering by country and by family languages (Cameron, Gelbach, and Miller, 2011). These alternative forms of clustering have little impact on the standard errors.

1.5.4 Cattle Presence and Women's Role in Society

The historical presence of cattle may have had broad effects on women's role in society that extended beyond their participation in agricultural production. In this subsection, I explore the relationship between historical cattle presence and a broad range of pre-colonial women's outcomes including polygyny, brideprice marriage, the use of patrilocal residence, and patrilineal inheritance and lineage system.

Table 1.10 , cols. 7 and 8 report the results for patrilineal inheritance, as measured as a dummy variable for whether ethnic group used patrilineal inheritance as opposed to matrilineal system. Across both the OLS and IV specifications, I find positive and statistically significant effects, indicating that historical cattle presence is associated with a shift towards inheritance rules that favour men. These patterns are consistent with anthropological research that has linked the emerge of patrilineal norms to the control of private property by men (Engels, 1884), and have found that these norms are common among pastoralist societies (Holden and Mace, 2003).

Cols 5 and 6 report the results for patrilocality, measured by an indicator variable for whether patrilocal was the mode of residence as opposed to matrilocality. Historical cattle presence is associated with statistically significant differences in patrilocality, with cattle-based societies more likely to adopt norms in which the wife resides with the husband's family. This finding is in line with anthropological studies that argue that patrilocality norms emerge together with the practice of patrilineality (Trivers 1972, Xia 1992, and Becker 2018). [26]

26. This effect does not appear to be driven by paternal uncertainty, I will later explore the source of

In Cols 3 and 4, I investigate the relationship between cattle presence and the practice of brideprice marriage. The dependent variable *Brideprice* is an indicator for whether the prevalent mode of marriage of the ethnic group was characterized by a brideprice payment from the husband to the bride's family.[27] The results show no relationship between historical cattle presence and the traditional practice of bride price marriage.[28] This might result from two contracting effects. On the one hand, anthropologists have linked bride price marriage to patrilocality, as bride price compensates the family of the bride for taking the daughter from their group (Vroklage, 1952). If cattle presence is positively correlated with patrilocality, I would expect to see a positive relationship between cattle presence and bride price marriage. On the other hand, Boserup (1970) has hypothesized that bride price is linked to women participation in agriculture, paying a bride price is a right to the bride's labor. To the extent that cattle presence is negatively correlated with female particiation in agriculture, a negative relationship between cattle presence and brideprice marriage is expected.

Finally, cols 1 and 2 report the results for polygny. The dummy *Polygyny* is equal to 1 if the prevalent mode of marriage of the ethnic group was polygyny, the custom that allows men to have multiple wifes, as opposed to monogamous marriage. The coefficient estimates are small and statistically insignificant, which might reflect two offsetting effects.[29] On the one hand, polygyny is linked to wealth inequality among men (Becker 1974 ; Grossbard 1976). To the extent that the presence of cattle increased wealth inequality *among* men, I would expect higher polygyny in these societies. On the other hand, Boserup (1970) argues that having multiple wives is more profitable for men when women contribute to agriculture production. In this case, given that cattle presence is negatively correlated with female participation in agriculture, it may also be negatively associated with polygyny.

this effect.

27. The variable *Brideprice* includes brideprice or wealth to the bride's family, bride service, and token bride price to the the bride's family.

28. One should keep in mind that more than 95 percent of ethnic groups in the sample practiced brideprice marriage, the no correlation might only result of the no variation to identify in the dependent variable.

29. The results might also reflect the fact that there is no variation in the practice of polygyny, the mean for polygyny is 0.95.

Together, the patterns in Table 1.10 show that the historical presence of cattle had broad effects on the role of women in pre-colonial African societies that extended beyond their labour market participation. In the next section, I explore the extent to which the historical cattle presence continues to shape contemporary women's outcomes.

1.6 Historical Cattle Presence and Contemporary Women's Outcomes

I explore whether the historical adoption of cattle continues to influence contemporary female labour force participation. The analysis is based on DHS data for years 1992-2016, where I match ancestral cattle presence to individuals based on their ethnic names (see section 1.3.3). In Section 1.6.1 I report the OLS and IV results for female participation ; in Section 1.6.2 to 1.6.3, I explore the mechanisms driving the relationship between cattle and female LFP, including the role of plough cultivation, various historical norms related to women's role in society, and the link between historical cattle presence and other contemporaneous women's outcomes ; finally, in Section 1.6.4, I examine the extent to which the long-run effects women's participation were driven by persistence in cultural norms.

1.6.1 Historical Cattle Presence and Contemporary Female Labour Force Participation

Table 1.11 presents the OLS and IV estimates of the effect of the historical cattle presence on contemporary females participation in the labor force. The even columns document the OLS estimates, and the IV coefficients are reported in odd columns. In columns 1 and 2, I report the estimates including the same historical ethnic groups level controls as in precolonial analysis [30]. In columns 3 and 4, I add a set of survey-year fixed effects, in columns 5 and 6 I include individual-level controls (a quadratic measure of

30. The covariates include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absolute latitude coordinates, malaria ecology index, a measure of historical economic complexity, and a measure of historical political hierarchies.

the respondents age). Finally, in columns 7 and 8, I add country-level controls including country income per capita and income per capita squared.

I find that women whose ancestors raised cattle are less likely to participate in the labor force today. The OLS estimates are negative and statistically significant across the various specifications. The point estimates are large in magnitude : historical cattle presence is associated with a 3.1 to 4.9 percentage point decrease in modern female participation, which is equivalent to 5 to 7.7 percent decrease relative to mean participation among descendants of ethnic groups without cattle.

To address potential bias in the OLS results, I use the CSI as an instrument for historical cattle presence. Figure 1.5 shows a strong positive first-stage relationship between historical cattle presence and the CSI, indicating that cattle were more likely to be present in places suitable for herding activities. Figure 1.6 depicts a negative reduced-form relationship between the CSI and contemporary female LFP, consistent with OLS findings.

Odd columns of Table 1.11 reports the IV estimates. The bottom panel reports the first-stage results. The Kleibergen-Paap F-statistics on the excluded instruments are greater than 10, consistent with the strong first-stage relationship. The top panel reports the two-stage least squares results. The IV estimates support the results from the OLS regressions in terms of both sign and statistical significance. Across the various specification the IV estimates range from -0.11 to -0.13 and are all statistically significant. These IV coefficients are larger than the OLS estimates, which could reflect measurement error in the identification of pre-colonial cattle presence in the Murdock (1967) data (see section 1.5.1 for a discussion). The fact that the IV estimates exceed the OLS results, provides further evidence against reverse causality, which should tend to bias the OLS estimates away from zero.

To further explore the link between historical cattle presence and contemporaneous female participation, I estimate the OLS and IV relationship for different occupational outcomes, measured as a series of dummy variables for female employment in different occupation (domestic, agriculture, manual, clerical, and sales). Table 1.12 reports these results. The OLS results reveal heterogeneous effects across occupational categories :

Historical cattle presence is negatively associated with the females likelihood of working in sales occupation, but positively related to employment in domestic services. One potential explanation of these results is that ethnic groups that adopted cattle adopted patriarchal norms that kept women within the home. The IV estimates confirm the OLS regressions in term of sign, but the estimates are no longer statistical significant.

1.6.2 Alternative Mechanisms : Plough Cultivation, Pastoralism, and Other Historical Gender Norms

The historical adoption of cattle led to a large gender imbalance in wealth holdings, given the physiological requirements involved in raising and herding these large animals. Standard non-cooperative theories of marriage matching predict that this imbalance should lead to a shift in household bargaining power that favoured men (Manser and Brown 1980 ; McElroy and Horney 1981 ; Lundberg and Pollak 1993). In this subsection, I examine whether there were alternative channels through which cattle presence influence long-run female participation.

In Table 1.13, I explore whether the observed relationship between historical cattle presence and contemporaneous female LFP can be attributed to difference in plough cultivation. For reference, columns 1 and 2 report the baseline results from the OLS and IV specifications. Columns 3 and 4 report the direct estimate for historical plough cultivation on modern female participation. There is a strong negative relationship in the OLS specification. Nevertheless, the IV point estimates are very noisy and statistically insignificant, due to the weak first-stage relationship.[31] Columns 5 and 6 report the results from models for female LFP that include both historical cattle presence and plough cultivation. One should keep in mind that all ethnicities that historical relied on plough agriculture also owned cattle, so the plough coefficients capture the additional effects of plough cultivation on female LFP among societies that herded cattle.

The results for historical cattle presence are not driven by plough cultivation. Column

31. The F-statistic for the excluded instruments is 0.47. Unlike the pre-colonial analysis, which was based on variation across 297 ethnic groups, the contemporaneous effects rely on variation across 161 ethnicities. Given the low rates of plough cultivation, there is simply not enough to identify this first-stage relationship in the African context.

5 of Table 1.13 shows that the coefficient on cattle remains negative and statistically significant after controlling for plough. I also find a significantly negative correlation between plough and female labor force participation (column 5 of Table 1.12), consistent with Alesina, Giuliano and Nunn (2013). In column 6 of Table 1.12, I report IV results using the CSI and land suitability for plough positive versus plough negative crops as instruments for the two endogenous variables. The coefficient on cattle remains negative, while the coefficient on plough is positive. Both estimates are imprecise, due to the fact that the instruments for plough cultivation have little explanatory power in Africa. Given this weak instrument issue, in column 7 of Table 1.12, I report the results from regressions that use only the CSI as an instrument for historical cattle presence and control directly for historical plough cultivation. The coefficients on both explanatory variables are negative and statistically significant. In fact, the coefficient estimates for historical cattle presence are virtually identical to the baseline specification, suggesting that the influence of historical cattle presence on contemporaneous female labour force participation was not mediated through plough-based agricultural technology.

Cattle could affect female mobility through male absence. Men from pastoralist societies are more likely to spend days out of villages to herd animals. Recent research has shown that male absence generated incentives to regulate female sexuality and mobility (Becker, 2018). If male absence alone is sufficient to generate the observed results, I would expect to see similar association between female labor force participation and the presence of other herding animals such as sheep and goat. To test for this, I regress my main outcome on an indicator for whether ancestors had sheep or goat. The coefficient for sheep and goat herding in column 2 of Table 1.15 has the opposite sign. The effect of the presence of sheep and goat is positively associated with female participating in the workplace. This is consistent with the argument put by anthropologists that sheep and goats tended to be women property, and consequently the presence of these animals tends to increase female empowerment (Deere and Doss, 2006). In column 3, I control for sheep and goat herding, historical adoption of cattle continues to have sizeable negative impact on FLFP, albeit statistically insignificance. These findings suggest that the effect of cattle on today's FLFP is not driven by male absence culture.[32]

A third possibility is that historical cattle presence led to a broad shift towards patriarchal norms that were unconducive to women's employment. I explore this question in Table 1.15. These models report the OLS and IV results for historical cattle presence, controlling for other historical societal norms (patrineal descent, patrilocal marriage, polygyny marriage, and brideprice marriage).[33] Controlling for ancestral patrilineal norms reduces the coefficient on cattle from -0.044 to -0.029, although it remains statistically significant, consistent with patrilineal regimes acting as a potential transmission mechanisms for the long-run decrease in women's participation. Columns 5 to 12 show that none of the other historical norms (patrilocality, polygyny, or brideprice) have any impact on the main estimates.

The results in this section suggest that the negative relationship between historical cattle presence on contemporaneous women's participation was not driven by plough agricultural technology, male absence associated with pastoralism, or changes in the norms governing marriage. Instead, the results appear to be consistent with the historical adoption of cattle having led to shift in bargaining power that ultimately caused societies to adopt norms and cultures that favoured men.

1.6.3 Historical Cattle Presence, Fertility, Marriage, and Household Decision-Making

In this section I examine the effect of cattle on fertility, and on marriage market outcomes such as polygyny, early marriage, and women empowerment.

To explore the link between historical cattle presence and fertility, I use information from the DHS on the number of children woman has ever born, and re-estimate my preferred specification with fertility as dependent variable. Columns 7 of Table 1.16 documents a positive and statistically significant relationship between cattle and fertility. Descendants of cattle-based ethnicities had 0.35 to 0.63 additional children, which

32. Controlling for patrilocality does not change results, providing further evidence against pastoralism and male absence as mechanism driving he results.

33. These results provide suggestive evidence for the mediating influence of these various factors, given the endogeneity of the covariates.

represents a 12 to 21 percent increase in fertility. The results are consistent with the historical narrative, the historical cattle presence developed son preference for herding, and therefore is morelikely to increase women fertility. It is also in line with the general shift in bargaining power, since men prefer larger family.

Did cattle also affect marriage market outcomes such as age of marriage, polygyny ? The decline in women's labour force participation may have led to decrease in the age of marriage if women were dependent on husband's income for subsistence. Cattle may also have increased the practice of polygyny by increasing inequality in wealth holdings *among* males (Becker, 1974). I examine the effects on age at first marriage in columns 1 and 2 in Table 1.16. The results show that the historical presence of cattle is associated with higher rates of early marriage, with estimates ranging from -0.78 to -2.0.

Columns 3 and 4 report the results for polygynous marriage, measured as an indicator for whether the woman had co-wives. [34] In both the OLS and IV specifications, I estimate positive and statistically significant effects on historical cattle presence, suggesting that ancestors from cattle-based societies were more likely to enter into polygynous marriage, with effect sizes that range from 15 to 30 percent. It is notable that historical cattle presence influence contemporaneous polygynous marriage rates, despite having no influence on the practice in pre-colonial Africa (Table 1.10, cols. 1 and 2). One potential explanation is that the unequal distribution of assets across males arose only in the post-colonial period (Van de Walle, 2009). Alternatively, the divergence findings could simply reflect the near universality of this practice in pre-colonial Africa. [35]

Lastly, I investigate whether cattle were related to women's bargaining power within household. The analysis is based on self-reported questions from DHS that capture women's ability to make decisions within the home. These questions include who makes decisions on specific issues,such as health care ; large purchases ; visits to relatives ; how many children to have ; food to be cooked ; and what to do with money women earn. I construct an ordered variable, *Female decision within household*, that ranges from 0 to

34. Roughly 33 percent of respondents in the sample reported that they belonged in a polygynous marriage.

35. 483 in 508 societies in the Murdock database are reported to have used the practice of polygyny.

6, increasing with women empowerment within household.[36] Columns 5 and 6 of Table 1.16 document that women from ethnic groups that had cattle are today less likely to make decision within household. Column 5 induces that having ancestors that had cattle is associated with a reduction of 0.5 women decision within the home. These results provide further evidence that the historical adoption of cattle led to shift in household bargaining power that favoured men.

1.6.4 Persistence

Why does historical cattle presence still influence contemporaneous female participation ? One explanation is that animals still play a central role in these societies, and the forces that influenced women's outcomes in pre-colonial Africa continue to operate today. An alternative explanation is that the patterns may reflect cultural persistence. The historical presence of cattle may have fundamentally altered norms about women's role in society and these norms may have persisted even after the direct influence of the animals waned.[37] In this section, I shed light on this question, and explore the extent to which the long-run effects can be attributable to a persistent shift in cultural norms.

I begin by examining the extent to which the main results can be attributed to cultural persistence. Specifically, I investigate whether cattle have similar effects on women's employment in urban versus rural area. Intuitively, the estimates for female participation in urban areas should capture the independent effect of cultural norms since cattle are largely absent from these areas[38], whereas the effects among the rural population represent the combined influence of the direct effect of cattle and persistent cultural norms. Although this comparison may be biased due to endogeneity in individuals' locational choices, to the extent that migrants to urban areas are generally less attached to historical cultural values, the estimates for the urban population with tend to *understate*

36. Women typically report making 2 to 4 decisions within household.

37. Nunn (2012) proposed to view culture as a set of decision-making heuristics or rules of thumb that are used in the case of uncertain environments. According to Boyd and Richardson 1988 models, when information is costly or imperfect, it is optimal to develop rules of thumb in decisions making. These set of decision-making heuristics are considered as values and social norms transmitted across generations (E. Teso, 2018).

38. Only 5 percent of urban household from DHS own cattle.

this cultural persistence channel.

I split the sample into urban and rural residence, and re-estimate the baseline specifications. The results are presented in Table 1.17. The estimates for historical cattle presence are very similar across the two sub-samples. The point estimates are all large, negative, and statistically significant. The findings among the urban sample, in particular, provides strong evidence for a cultural persistence channel. Although cattle are largely absent from urban areas, the cultural norms they influenced over generations are portable and continue to shape women's employment outcomes.

To further explore the role of cultural persistence, I exploit information on current location of residence in the DHS data, to construct a location-based indicator for historical cattle presence. This variable is equal to one if the individual currently resides in an ethnic homeland territory in which cattle was historically present. I re-run my preferred specifications, including both ethnicity-based and location-based measures of historical cattle presence. Identification in these models is based on the fact that roughly half of the DHS respondents no longer reside within their former ethnic homeland. [39] Columns 1 and 2 from Table 1.20 report the effects of cattle presence on female labour participation, my main result presented so far. The estimates are all large, negative and statistically significant. Columns 3 and 4 report the coefficients for the OLS and IV estimates when the explanatory variable is the location-based cattle presence. The OLS estimates are large, negative, but insignificant. The IV estimates are negative and statistically significant, suggesting that cattle presence affects female participation through location fixed factors, such as local institutions. Columns 5 and 6 report the estimates when I include both ethnicity-based cattle presence and location-based cattle presence. The results show that controlling for location-based cattle presence does not alter the coefficients for ethnicity-based cattle. In addition, the estimates for the location-based cattle presence are no longer significant. Together, these results provide further evidence for the cultural legacy of cattle, independently from their contemporaneous influence on

39. This outmigration may have occurred either by the current generation or by previous generations, in the century since much of the Murdock data was assembled. In column 7 and 8 of Table 1.20 shows that historical ancestral cattle presence is unrelated to the probability that individuals have left their ancestral homeland.

women's participation.

Lastly, I investigate the role of colonization in the persistent effect of historical cattle presence. I exploit the fact that French and English colonizers developed different policies of assimilation. In contrast to the English indirect rule, French colonizers adopted a policy of assimilation which promoted French over African culture. I use data from Laporta et al. (2008), and I examine whether cattle affect differently women's employment in French versus English colonies. I re-estimate the baseline specifications including a dummy for French colonization and an interaction term between cattle presence and French colonization. The results are presented in Table 1.21. Columns 1 and 2 report the main result presented so far. Column 3 presents the OLS estimates of the effect of French colonization on female labour participation. The coefficients are large, negative and statistically significant. Women from French colonization are more likely to participate in the workplace today compared to those from English colonization. Columns 6 and 7 report the coefficients for the OLS and IV estimates when I include the interaction term, a dummy for French colonization and historical cattle presence. The OLS estimates on historical cattle presence remain stable, although statistically insignificant. The IV estimates are negative and statistically significant at 10 percent level. However, the estimates for the interaction term between French colonization and historical cattle presence are not statistically significant for both OLS and IV specifications. Taken together, these results provide consistent evidence that cattle have not influenced women's employment in French versus English colonies differently.

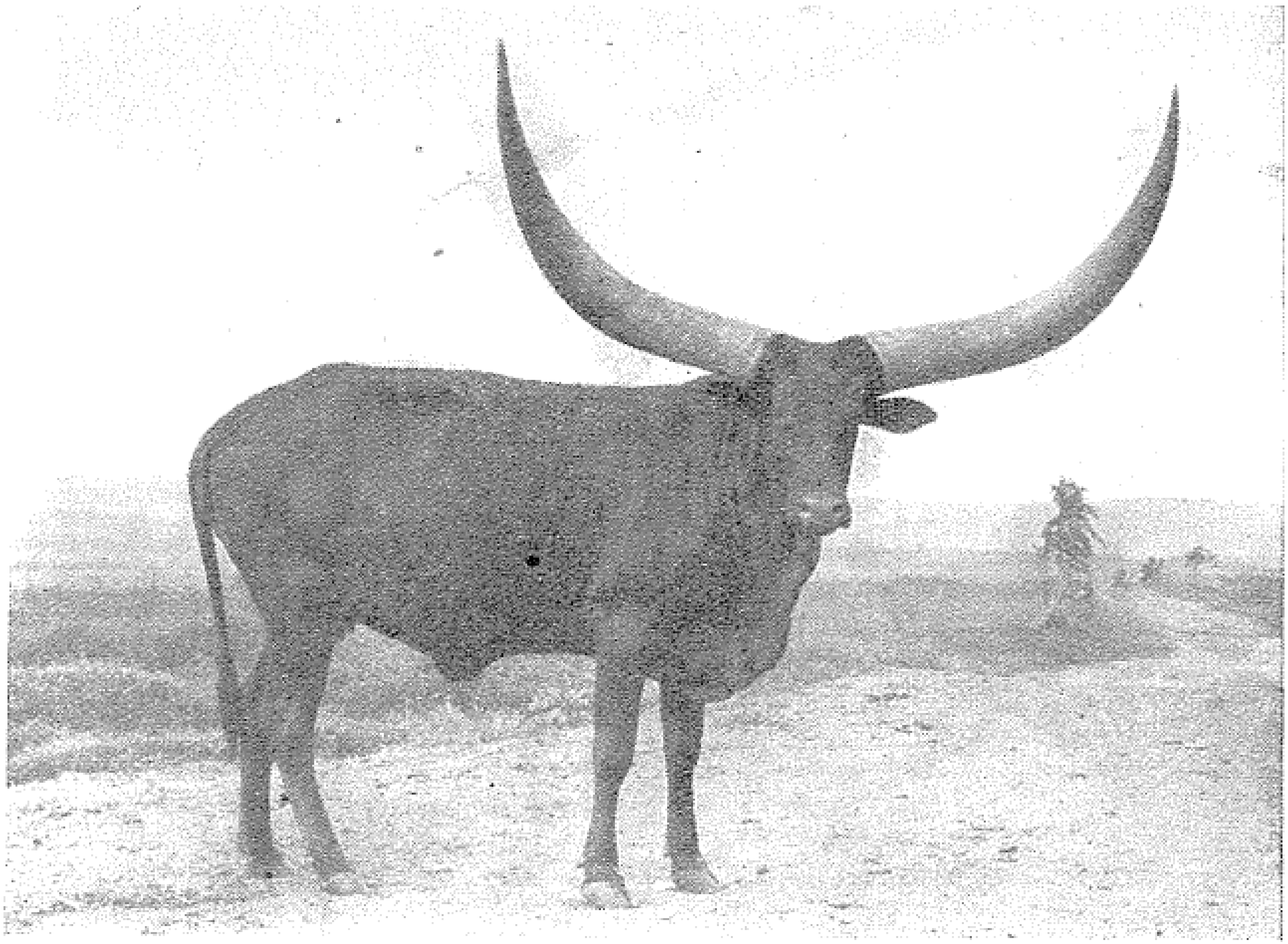
1.7 Conclusion

Anthropologists have long considered the presence of cattle as an important determinant of patriarchal norms in Africa. I empirically test this hypothesis in both the pre-colonial and modern era. My analysis relies on both information on historical ethnic norms from the *Ethnographic Atlas* with data from *Demographic and Health Surveys* on modern women's outcomes. To establish the causal impact of historical cattle presence on gender roles, I exploit geo-climatic factors as a source of plausibly exogenous variation in the likelihood of historical cattle adoption.

The paper's findings show that cattle-based societies had more historical gender inequality, as measured by female participation in agriculture, inheritance rules, and other marriage customs. Moreover, I find that gender inequality has persisted into the modern era, with descendants of cattle-based societies having lower labour force participation rates, marrying at younger ages, having higher fertility rates, and participating less in household decision-making. The results cannot be attributed to either plough agriculture or pastoralism. Instead, the findings are consistent with a setting in which the historical adoption of cattle led to a shift of resources to men, which ultimately caused societies to adopt patriarchal norms that have persisted to present day, even in areas where cattle no longer play an important role in society.

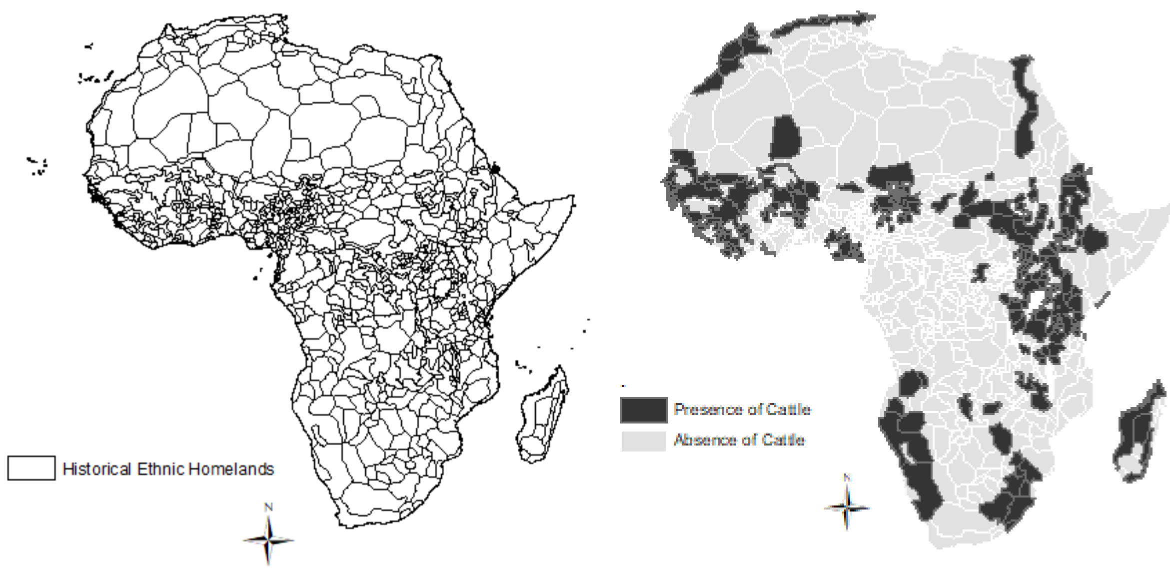
This paper adds to our understanding of the origins of the large differences in gender roles throughout Sub-Saharan Africa. In addition to establishing the role of cattle domestication in driving these patterns, the results demonstrate the persistence of traditional norms regarding the role of women in society, and how these norms can outlast the initial economic forces that created them. Understanding the sources of this persistence, and how policy can best be designed to break these patterns, is a potentially fruitful area for future research.

Figure 1.1 : An Ankole cow in Uganda



Notes : This figure from Joshi, McLaughlin and Phillips, (1957), shows an Ankole cow in Uganda. The horns measure 52 inches between the tips.

Figure 1.2 : Ethnic groups boundaries and pre-colonial presence of cattle



(a) Ethnics Boundaries (b) pre-colonial Presence of Cattle

Figure 1.3 : Cattle Suitability Index

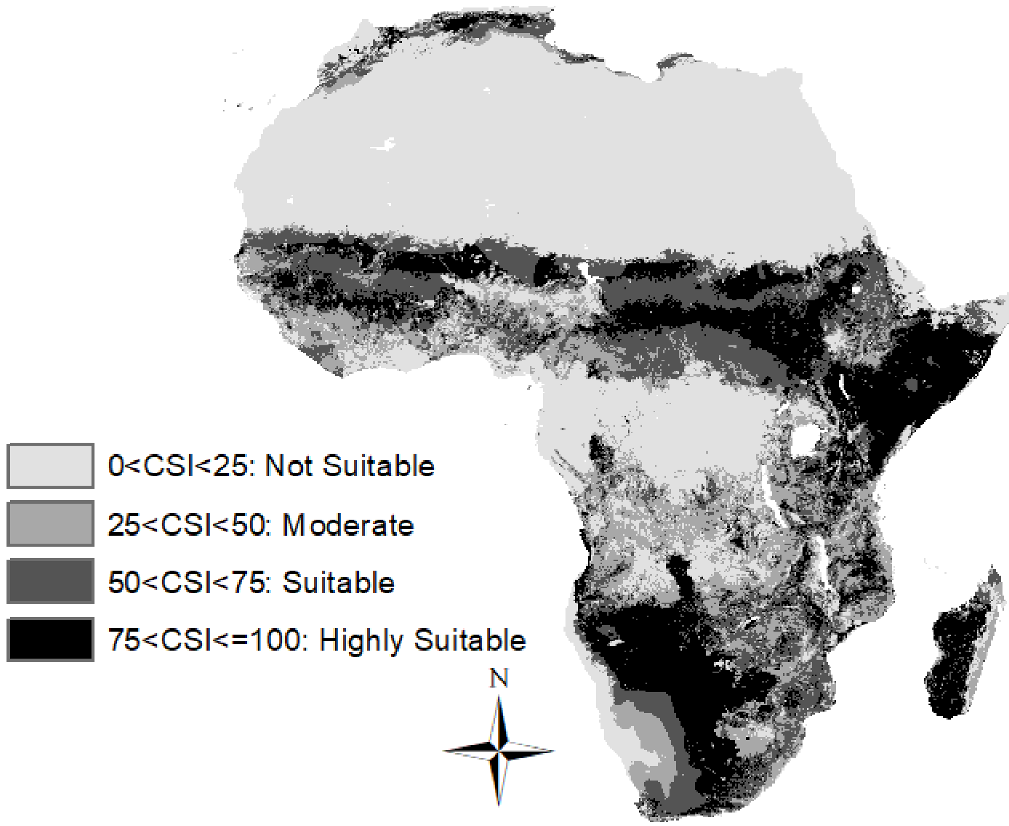
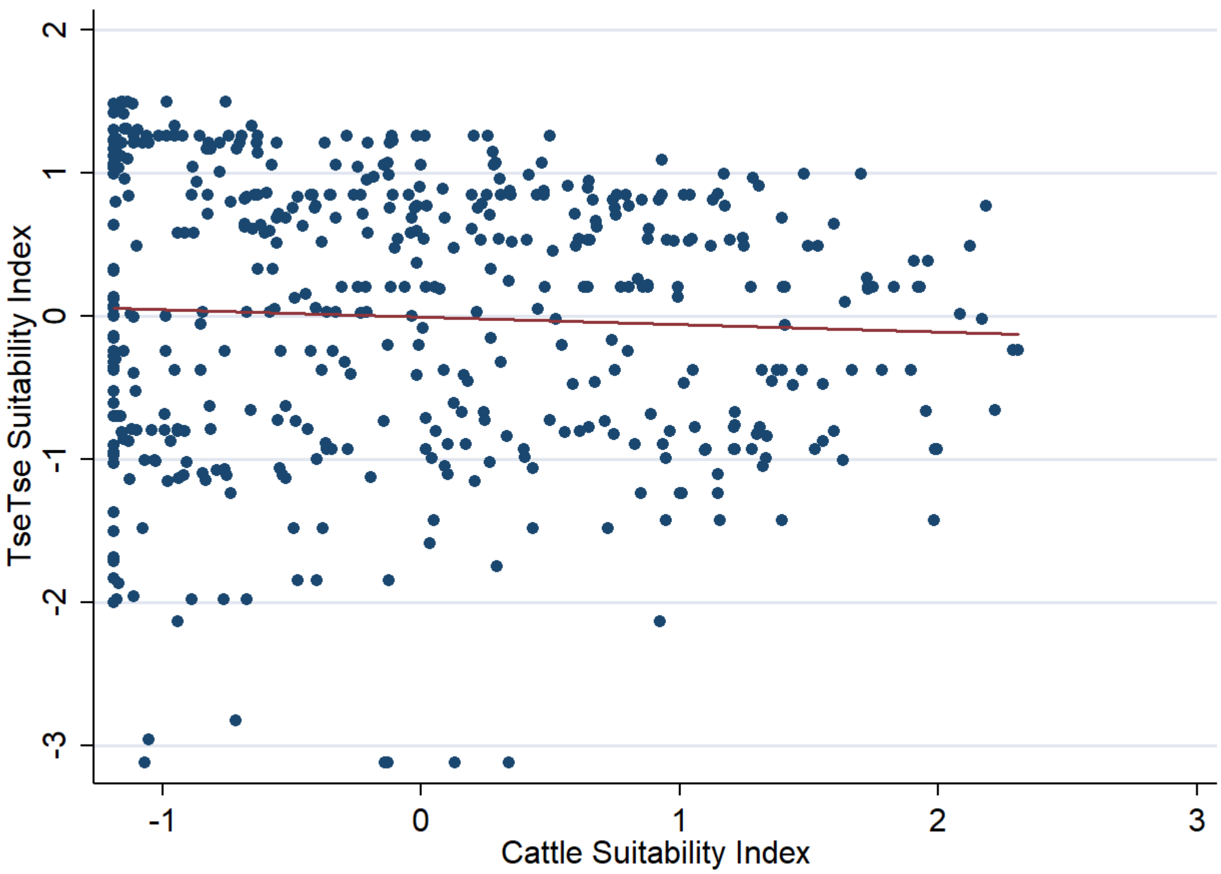
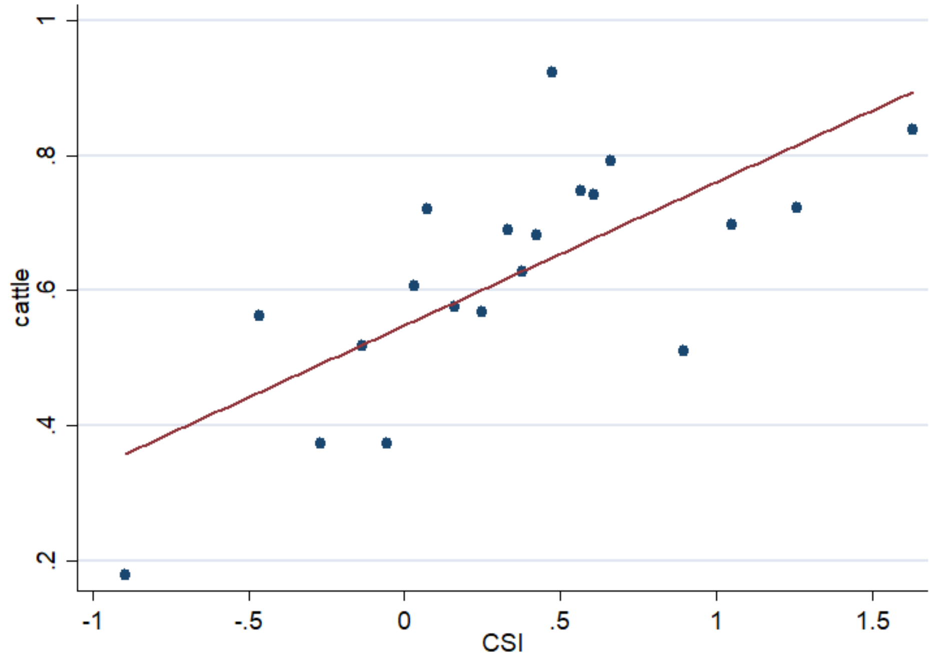


Figure 1.4 : TseTse Suitability Index versus Cattle Suitability Index



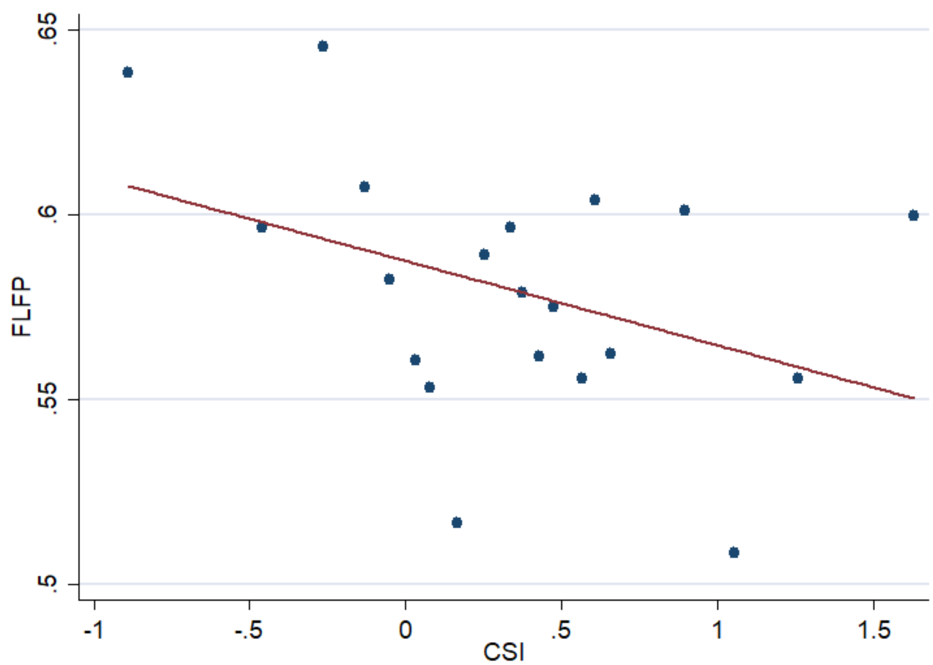
Notes : This figure shows that the tsetse suitability and cattle suitability index are not correlated.

Figure 1.5 : Historical adoption of cattle and Cattle Suitability Index



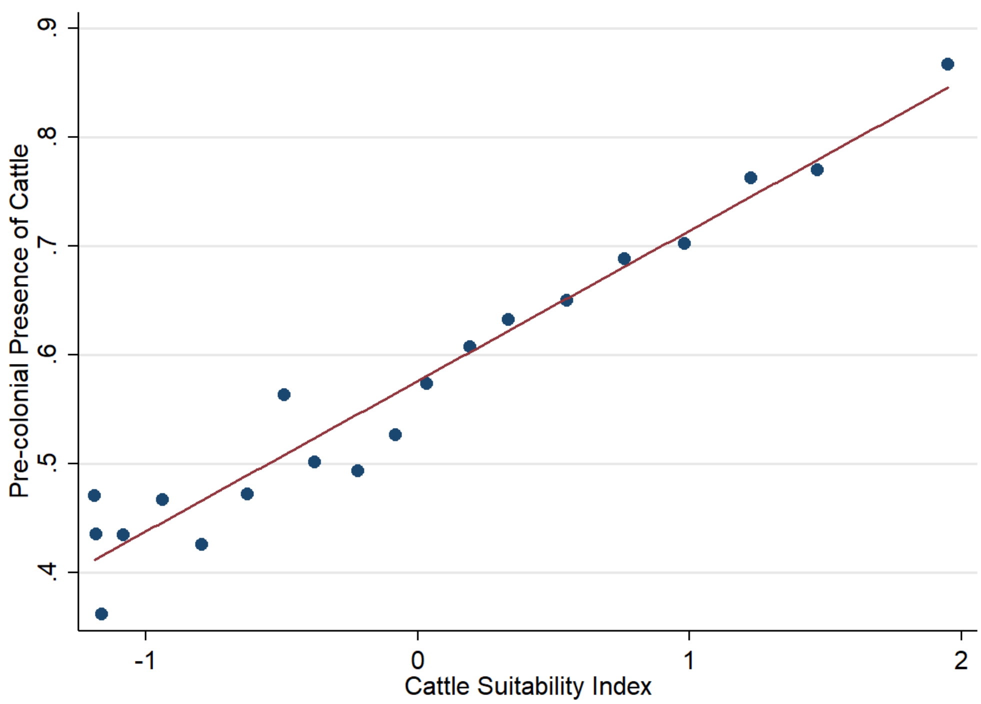
Notes : Binscatter plot. precolonial presence of cattle and cattle suitability index for African ethnic groups conditional on controls.

Figure 1.6 : FLFP and Cattle Suitability Index



Notes : Binscatter plot. FLFP and cattle suitability index conditional on controls.

Figure 1.7 : Pre-colonial presence of cattle and Cattle Suitability Index



Notes : Binscatter plot. pre-colonial presence of cattle and cattle suitability index for African ethnic groups conditional on geography controls.

Table 1.1 : Summary statistics

Variable	Mean	Std.Dev.	Min	Max	Obs	Mean	Std.Dev.	Min	Max	Obs
	Presence of cattle					Absence of cattle				
Ethnographic Atlas Variables										
Female participation agriculture	0.406	0.492	0	1	192	0.608	0.49	0	1	120
Female participation husbandry	0.183	0.388	0	1	164	0.457	0.504	0	1	46
Female participation fishing	0.25	0.435	0	1	92	0.342	0.477	0	1	79
Polygyny	0.956	0.205	0	1	274	0.95	0.218	0	1	201
Brideprice	0.96	0.196	0	1	277	0.932	0.252	0	1	206
Patrilocality	0.966	0.182	0	1	263	0.797	0.403	0	1	192
Patrilineal inheritance	0.921	0.271	0	1	189	0.68	0.468	0	1	128
Patrilineal lineage	0.909	0.289	0	1	230	0.736	0.442	0	1	182
Plough use	0.112	0.315	0	1	278	0.029	0.169	0	1	206
Economic development	5.661	1.657	1	8	277	5.79	1.98	1	8	205
Political complexities	1.883	0.666	1	3	266	1.965	0.681	1	3	201
Geographical variables										
CSI	0.193	0.958	-1.189	2.307	276	-0.304	0.827	-1.189	2.183	202
Slope	0.508	0.341	0	1	276	0.67	0.305	0	1	202
Distance to water	0.872	0.233	0	1	276	0.88	0.278	0	1	202
Pasture suitability	0.556	0.353	0	1	276	0.275	0.312	0	1	202
SI	0.515	0.182	0	0.911	278	0.547	0.215	0.091	0.885	206
Proportion in tropics	0.922	0.265	0	1	278	0.949	0.21	0	1	206
Malaria index	13.496	10.633	0	34.491	278	13.882	8.431	0	34.083	206
TSI	-0.362	0.987	-3.119	1.495	278	0.445	0.789	-2.136	1.495	206
Plough Negative	0.635	0.068	0.518	0.852	265	0.667	0.061	0.518	0.852	192
Plough Positive	0.114	0.123	0	0.423	265	0.099	0.099	0	0.365	192

Table 1.2 : Summary statistics (continued)

Variable	Mean	Std.Dev.	Min	Max	Obs	Mean	Std.Dev.	Min	Max	Obs
	Presence of cattle					Absence of cattle				
SCCS variables										
Land clearance	0.038	0.196	0	1	26	0	0	0	0	11
Soil preparation	0.154	0.368	0	1	26	0.364	0.505	0	1	11
Planting	0.308	0.471	0	1	26	0.636	0.505	0	1	11
Crop Tending	0.346	0.485	0	1	26	0.636	0.505	0	1	11
Harvesting	0.269	0.452	0	1	26	0.818	0.405	0	1	11
DHS Variables (women sample)										
FLFP	0.54	0.498	0	1	245000	0.639	0.48	0	1	153000
Agriculture	0.273	0.445	0	1	225000	0.316	0.465	0	1	150000
Clerical	0.007	0.082	0	1	225000	0.008	0.088	0	1	150000
Manual	0.077	0.267	0	1	225000	0.085	0.279	0	1	150000
Domestic	0.019	0.136	0	1	225000	0.008	0.088	0	1	150000
Sales	0.197	0.398	0	1	225000	0.196	0.397	0	1	150000
Number of Children	3.099	2.944	0	22	248000	2.87	2.741	0	18	155000
Age First Marriage	17.435	3.897	2	48	189000	18.176	4.103	5	49	114000
Household Decisions	1.67	1.512	0	4	58263	2.286	1.436	0	4	49162
Polygyny	0.383	0.486	0	1	165000	0.254	0.435	0	1	97415
Age	28.34	9.455	15	49	248000	28.444	9.422	15	49	155000
Rural	0.678	0.467	0	1	248000	0.608	0.488	0	1	155000
Education Attainment	1.045	1.389	0	5	248000	1.584	1.436	0	5	155000

Table 1.3 : OLS and IV estimates, the effect of the pre-colonial presence of cattle on female participation in agriculture

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A. OLS and second stage							
	Dependent variable : Female participation in agriculture relative to male							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.199**	-0.802***	-0.203***	-0.328*	-0.205***	-0.371**	-0.204***	-0.379**
	(0.084)	(0.292)	(0.073)	(0.187)	(0.073)	(0.185)	(0.072)	(0.192)
Geography Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	No	No	No	No	Yes	Yes	Yes	Yes
Political Complexity	No	No	No	No	No	No	Yes	Yes
Observations	308	308	308	304	308	308	308	304
No. Clusters	109	109	109	109	109	109	107	107
Dep. var. mean unaffected	0.605	0.606	0.605	0.606	0.605	0.607	0.602	0.603
	Panel B. First Stage							
	Dependent Variable is pre-colonial presence of cattle							
Cattle Suitability Index		0.133***		0.204***		0.207***		0.209***
		(0.031)		(0.035)		(0.036)		(0.038)
F-stat		15.4		29.07		27.72		25.59

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The variable *female participation in agriculture* takes the value 1 if female participated more than male and 0 otherwise. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors in parenthesis clustered at the level of cultural province. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.4 : OLS and IV estimates, the effect of the pre-colonial presence of cattle on female participation in agricultural tasks

	(1)		(2)		(3)		(4)		(5)	
	Land Clearance		Soil Preparation		Planting		Crop Tending		Harvesting	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.013	0.470	-0.113	-1.629	-0.199	-1.738	-0.154	1.917	-0.458***	-0.935
	(0.029)	(0.873)	(0.169)	(2.500)	(0.217)	(2.757)	(0.165)	(3.565)	(0.148)	(1.780)
Geography Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	37	37	37	37	37	37	37	37	37	37
Dep. var. mean unaffected	0	0	0.363	0.363	0.636	0.636	0.636	0.636	0.818	0.818
F-stat		0.3		0.3		0.3		0.3		0.3

Notes. The unit of observation is an ethnic group from the *the Standard Cross-Cultural Sample*. All dependent variables are dummies for female participation in agricultural task. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.5 : OLS and IV estimates, the effect of pre-colonial presence of cattle on female participation in husbandry and fishing

	(1)		(2)	
	Husbandry		Fishing	
	OLS	IV	OLS	IV
Historical Cattle Presence	-0.250***	-0.499**	-0.018	-0.047
	(0.082)	(0.246)	(0.073)	(0.258)
F-stat		11.90		9.78
Geography Controls	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes
Observations	206	204	169	167
R-squared	0.117	0.076	0.100	0.109
No. Clusters	91	91	82	82
Dep. var. mean unaffected	0.466	0.477	0.428	0.450

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The variables *Husbandry* and *Fishing* are dummies for female particiation in husbandry and fishing respectively. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors in parenthesis clustered at the level of cultural province. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.6 : OLS and IV estimates, the effect of cattle versus plough agriculture on female participation in agriculture

	(1)	(2)	(3)	(4)	(5)	(6)
	Female participation in agriculture relative to male					
	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.204***	-0.379**			-0.179**	-0.554***
	(0.0725)	(0.192)			(0.0695)	(0.167)
Historical plough use			-0.471***	-1.957***	-0.423***	-1.141**
			(0.132)	(0.756)	(0.128)	(0.552)
F-stat		25.59		3.28		17.46
Geography Controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes	Yes
Observations	307	304	307	300	307	297
R-squared	0.192	0.165	0.11	0.14	0.226	0.18
No. Clusters	107	107	107	103	107	103
Dep. var. mean unaffected	0.602	0.603	0.602	0.607	0.602	0.609

Notes. OLS estimates in odd numbered columns, and IV estimates in even numbered columns. I instrument for *Historical Cattle Presence* with the *CSI*, and for *Historical plough use* with the suitability location for cultivating plough-positive versus plough-negative crops from Alesina, Giuliano and Nunn (2013). The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable takes the value 1 if females participated more than males in agriculture and 0 otherwise. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors in parenthesis clustered at the level of cultural province. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.7 : Robustness tests of the effect of cattle on female participation in agriculture : controlling for observables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable : Female participation in agriculture relative to male									
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.112	-0.399**	-0.183**	-0.365*	-0.183**	-0.364*	-0.101	-0.282	-0.246**	-0.926
	(0.080)	(0.198)	(0.074)	(0.201)	(0.071)	(0.203)	(0.081)	(0.220)	(0.118)	(0.714)
TseTse Suitability Index	0.310***	0.211*								
	(0.074)	(0.105)								
Presence of sheep and goats			-0.053	-0.702						
			(0.136)	(0.685)						
Ethnic group land area					0.081	0.046				
					(0.051)	(0.058)				
Indigenous Slavery							0.002	-0.000		
							(0.097)	(0.095)		
Log of Population Density									-0.004	0.017
									(0.027)	(0.034)
F-stat		24.68		23.31		23.83		19.58		3.91
Observations	304	308	304	304	274	274	238	238	304	304
R-squared	0.285	0.277	0.200	0.172	0.226	0.197	0.207	0.182	0.192	0.094
Geography Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *female participation in agriculture* takes the value 1 if female participated more than male and 0 otherwise. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors in parenthesis clustered at the level of cultural province. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.8 : IV estimates. Robustness tests of the effect of cattle on female participation in agriculture : controlling for CSI factors

	(1)	(2)	(3)	(4)	(5)
	Female participation in agriculture relative to male				
	Second Stage estimates				
Historical Cattle Presence	-0.379**	-0.430**	-0.469**	-0.518***	-0.537***
	(0.192)	(0.168)	(0.188)	(0.162)	(0.162)
Slope index		✓			✓
Distance to water index			✓		✓
Pasture suitability index				✓	✓
F-stat	25.59	29.97	24.13	29.56	28.38
Observations	304	304	304	304	304
Geography Controls	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *female participation in agriculture* takes the value 1 if female participated more than male and 0 otherwise. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Standard errors are clustered at the level of cultural province dimensions. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.9 : Robustness tests of the effect of cattle on female participation in agriculture : alternative clustering

	(1)	(2)	(3)	(4)	(5)	(6)
	Female participation in agriculture relative to male					
	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.202**	-0.379**	-0.204**	-0.379	-0.204*	-0.379
	(0.096)	(0.146)	(0.101)	(0.245)	(0.105)	(0.251)
Conley SE	✓	✓				
SE clustered by country			✓	✓		
Multiway clustering					✓	✓
F-stat		25.59		9.89		27.5
Geography Controls	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes	Yes
Observations	304	304	304	304	304	304

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *female participation in agriculture* takes the value 1 if female participated more than male and 0 otherwise. The mean for the variable *historical presence of cattle* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. The T. G. Conley (1999) standard errors account for spatial correlation with cutoffs of 50 degrees latitude and 50 degrees longitude. The multiway clustering have been computed using the method developed by Cameron, Gelbach, and Miller (2011) and clusters standard errors along the country and cultural province dimensions. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.10 : Pre-colonial adoption of cattle and other gender norms : OLS, second stage estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Polygyny		Brideprice		Patrilocality		Patrilineal land inheritance		Patrilineal lineage	
Mean of dep. Var.	0.956		0.95		0.891		0.821		0.825	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	0.009	-0.006	0.026	0.014	0.173***	0.222**	0.257***	0.472*	0.174**	0.193
	(0.029)	(0.047)	(0.023)	(0.063)	(0.056)	(0.104)	(0.071)	(0.243)	(0.067)	(0.146)
F-statistic		39.76		40.37		41.26		14.03		45.27
Geography Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Economic Development	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Political Complexity	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	459	453	465	459	438	432	308	305	398	395
R-squared	0.128	0.124	0.011	0.010	0.126	0.125	0.193	0.178	0.121	0.120
No. Clusters	111	111	107	107	106	106	98	98	108	108

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variables are all dummies for the practice of each cultural norm. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Geography Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation. "Economic development" is an index of settlement density, and "Political complexity" is an index for the levels of jurisdictional hierarchies. Coefficients are reported with robust standard errors in parenthesis clustered at the level of cultural province. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

TABLE 1.11 : OLS AND IV ESTIMATES, THE EFFECT OF CATTLE ON FEMALE LABOR FORCE PARTICIPATION

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable : Female Labor Force Participation							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	OLS and Second Stage estimates							
Historical Cattle Presence	-0.049**	-0.110*	-0.046***	-0.134**	-0.044***	-0.130**	-0.031**	-0.106*
	(0.024)	(0.065)	(0.015)	(0.054)	(0.015)	(0.052)	(0.014)	(0.056)
Observations	393,407	393,407	393,407	393,407	393,407	393,407	391,931	391,931
R-squared	0.049	0.046	0.171	0.166	0.241	0.236	0.244	0.240
Dep. var. mean unaffected	0.638	0.638	0.638	0.638	0.638	0.638	0.645	0.645
	First Stage : Dependent variable is Historical Cattle Presence							
Cattle Suitability Index		0.254***		0.236***		0.236***		0.213***
		(0.057)		(0.059)		(0.059)		(0.064)
Observations		393,407		393,407		393,407		391,931
R-squared		0.329		0.406		0.406		0.419
F-statistic		19.79		16		16		10.94
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	No	No	Yes	Yes	Yes	Yes
Country-level Controls	No	No	No	No	No	No	Yes	Yes
No. Clusters	161	161	161	161	161	161	158	158

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The top panel reports the OLS and the second stage estimates of *Historical Cattle Presence*, the bottom panel shows the first stage etimates of *Cattle Suitability Index*. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *Country-level Controls* include : income per capita for year 2000, and income per capita squared for year 2000. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.12 : OLS and IV estimates, the effect of historical cattle adoption on Occupation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Domestic		Agriculture		Manual		Clerical		Sales	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	0.016***	0.014	0.025	0.074	0.001	-0.077**	-0.001	-0.006	-0.037**	-0.023
	(0.004)	(0.015)	(0.022)	(0.073)	(0.013)	(0.037)	(0.001)	(0.004)	(0.017)	(0.066)
F-statistic		11.34		11.34		11.34		11.34		11.34
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	369,072	369,072	369,072	369,072	369,072	369,072	369,072	369,072	369,072	369,072
R-squared	0.034	0.034	0.081	0.080	0.025	0.012	0.004	0.004	0.060	0.060
No. Clusters	155	155	155	155	155	155	155	155	155	155
Dep. var. mean unaffected	0.007	0.007	0.315	0.315	0.086	0.086	0.007	0.007	0.198	0.198

Notes. The unit of observation is a female respondent from *DHS*. *Domestic* is an indicator that takes the value of one if the respondent work within household as domestic servant. *Agriculture* is a dummy for the respondent being employed in agriculture. *Manual* is an indicator that takes the value one if the respondent is employed in manual activities. *Clerical* is an indicator that takes the value one if the respondent is employed in clerical jobs. *Sales* is an indicator that takes the value one if the respondent is employed in sales activities. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group kept cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.13 : OLS and IV estimates, the effect of cattle versus plow agriculture on female labor force participation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Female labor force participation						
	OLS	IV	OLS	IV	OLS	IV	IV†
Historical Cattle Presence	-0.044***	-0.130**			-0.044***	-0.073	-0.130**
	(0.015)	(0.052)			(0.015)	(0.061)	(0.052)
Historical plough use			-0.150***	46.51	-0.134***	10.24	-0.102***
			(0.032)	(56.74)	(0.029)	(48.67)	(0.035)
F−Statistic for cattle		16				8.77	16
F−Statistic for plow				0.47		0.31	
F−Statistic all		16		0.47		0.21	16
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	393,407	393,407	393,407	393,198	393,407	393,198	393,407
R-squared	0.241	0.13	0.239	0.00	0.241	0.21	0.13
No. Clusters	161	161	161	161	160	160	161
Dep. var. mean unaffected	0.645	0.645	0.490	2.161	0.490	2.871	2.871

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.14 : OLS and IV estimates : investigating the mechanisms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable : Female labor force participation							
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.044***	-0.130**	-0.049	-0.005	-0.042**	-0.122**	-0.044***	-0.130**
	(0.015)	(0.052)	(0.038)	(0.042)	(0.016)	(0.052)	(0.015)	(0.052)
Historical Sheep and Goats Presence			0.005	-0.279				
			(0.037)	(0.215)				
Historical Polygyny					-0.099	-0.009		
					(0.061)	(0.079)		
Historical Brideprice							-0.003	0.000
							(0.026)	(0.034)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	392,005	392,005	392,005	392,005	388,419	388,419	393,481	393,481
R-squared	0.243	0.239	0.243	0.229	0.238	0.234	0.239	0.235
F-stat		10.92		5.47		19.41		16.67

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.15 : OLS and IV estimates : investigating the mechanisms (continued)

	(9)	(10)	(11)	(12)	(13)	(14)
	Dependent variable : Female labor force participation					
	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.040**	-0.145**	-0.046***	-0.241**	-0.029*	-0.136**
	(0.018)	(0.073)	(0.017)	(0.111)	(0.016)	(0.054)
Historical Patrilocality	0.017	0.047				
	(0.028)	(0.040)				
Historical Patrilineal Inheritance			0.017	0.080		
			(0.023)	(0.065)		
Historical Patrilineal Lineage					-0.009	0.034
					(0.025)	(0.032)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	375,944	375,944	296,437	296,437	325,061	325,061
R-squared	0.232	0.226	0.239	0.219	0.223	0.217
F-stat		15.35		5.60		24.33

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.16 : OLS and second stage estimates : the effect of historical cattle presence on fertility, polygyny, and women empowerment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Age First Marriage		Polygyny		Female decision Household		Number of Children	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.780**	-2.011***	0.0469*	0.102*	-0.474***	-0.783***	0.354***	0.634***
	(0.367)	(0.681)	(0.0258)	(0.0560)	(0.119)	(0.258)	(0.105)	(0.226)
F-statistic		17.18		21.74		18.46		19.79
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	241,304	241,304	259,874	259,874	105,617	105,617	398,885	398,885
R-squared	0.047	0.19	0.134	0.21	0.202	0.21	0.601	0.20
No. Clusters	158	158	157	157	114	114	161	161
Dep. var. mean unaffected	18.17	18.17	0.254	0.254	2.286	2.286	2.87	2.87

Notes. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.17 : OLS and second stage estimates : Rural versus Urban location

	(1)	(2)	(3)	(4)
	Female Labor Force Participation			
	OLS	IV	OLS	IV
	Rural		Urban	
Historical Cattle Presence	-0.049***	-0.151**	-0.047**	-0.119**
	(0.016)	(0.063)	(0.016)	(0.051)
F-statistic		15.11		16.77
Observations	255,837	255,837	137,570	137,570
R-squared	0.253	0.246	0.238	0.235
Historical Controls	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes
No. Clusters	159	159	158	158
Dep. var. mean unaffected	0.610	0.610	0.484	0.484

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.18 : OLS and IV estimates, the effect of historical presence of cattle on FLFP, robustness to other measure of the presence of cattle

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable : Female labor force participation									
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.043***	0.0003	-0.025***	-0.035*	-0.022***	-0.041***	-0.022***	-0.040***	-0.019***	-0.036**
	(0.013)	(0.032)	(0.009)	(0.019)	(0.005)	(0.015)	(0.005)	(0.015)	(0.005)	(0.017)
Observations	397,727	397,727	393,481	393,481	393,481	393,481	393,481	393,481	392,005	392,005
F-stat		9.64		19.77		19.45		19.45		11.48
No. Clusters	161	161	161	161	161	161	161	161	160	160
Historical Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	No	No	No	No	Yes	Yes	Yes	Yes	Yes	
Individual Controls	No	No	No	No	No	No	Yes	Yes	Yes	Yes
Country-level Controls	No	No	No	No	No	No	No	No	Yes	Yes
Dep. var. mean unaffected	0.639	0.639	0.642	0.642	0.642	0.642	0.642	0.642	0.645	0.645

Notes. The unit of observation is a female respondent from *DHS*. Standard errors in parenthesis are clustered at the ethnicity level. The dependent variable *FLFP* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Dependence on Cattle* is range 0-9 measuring the degree to which society depended on cattle for subsistence. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *Country-level Controls* include : income per capita for year 2000, and income per capita squared for year 2000. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.19 : OLS estimates : Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Dependent variable : Female labor force participation									
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Historical Cattle Presence	-0.044***	-0.130**	-0.046	-0.156	-0.048*	-0.124	-0.044*	-0.111	-0.043*	-0.131
	(0.015)	(0.052)	(0.027)	(0.173)	(0.025)	(0.117)	(0.025)	(0.109)	(0.025)	(0.102)
Tansatlantic slave trade			-0.000	-0.001						
			(0.001)	(0.002)						
Distance to Sea					0.000	0.000				
					(0.000)	(0.000)				
Railway contact							-0.012	-0.011		
							(0.026)	(0.027)		
Total missions Area									-26.99	-17.87
									(20.99)	(24.39)
Observations	393,481	393,481	234,276	234,276	234,276	234,276	234,276	234,276	234,276	234,276
R-squared	0.239	0.235	0.227	0.221	0.228	0.225	0.227	0.225	0.228	0.223
Partial R-squared		0.13		0.02		0.04		0.04		0.06
F-stat		15.98		1.29		2.62		3.01		4.01

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Baseline Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity, age and age squared, and survey-year fixed effect. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.20 : OLS and second stage estimates : Ancestral versus Location effect

	(1)	(2)	(3)	(4)	(5)	(6)
	Female Labor Force Participation					
Mean of dep. Var.	0.578	0.578	0.578	0.578	0.578	0.578
	OLS	IV	OLS	IV	OLS	IV
Ethnicity-based cattle presence	-0.044**	-0.130**			-0.049***	-0.152*
	(0.015)	(0.052)			(0.014)	(0.085)
Location-based cattle presence			-0.023	-0.109*	0.007	0.032
			(0.015)	(0.059)	(0.013)	(0.095)
F-statistic		15.66		16.97		8.09
Observations	393,407	393,407	393,407	393,407	393,407	393,407
R-squared	0.240	0.235	0.239	0.234	0.240	0.235
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
No. Clusters	161	161	161	161	161	161

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The top panel reports the OLS and the second stage estimates of *Historical Cattle Presence*, the bottom panel shows the first stage etimates of *Cattle Suitability Index*. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Table 1.21 : French versus English colonization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent variable : Female Labor Force Participation						
Mean of dep. Var.	0.578	0.578	0.484	0.484	0.484	0.484	0.484
	OLS	IV	OLS	OLS	IV	OLS	IV
Historical Cattle Presence	-0.044***	-0.130**		-0.045***	-0.138***	-0.041	-0.206*
	(0.015)	(0.052)		(0.017)	(0.052)	(0.025)	(0.106)
French Colonization			0.046*	0.050**	0.057**	0.054*	0.001
			(0.026)	(0.024)	(0.027)	(0.031)	(0.077)
Cattle X French Colonization						-0.009	0.126
						(0.039)	(0.152)
F-statistic		16			16.39		10.39
Observations	393,407	393,407	393,198	393,198	393,198	393,198	393,198
R-squared	0.241	0.236	0.240	0.241	0.236	0.241	0.232
No. Clusters	161	161	160	160	160	160	160
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Survey-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes. The unit of observation is a female respondent from *DHS*. The dependent variable, *Female labor force participation* is an indicator variable which takes the value one if the respondent has worked in last 12 months. *Historical Cattle Presence* is an indicator variable taking the value of one if ethnic group had cattle in pre-colonial times. The top panel reports the OLS and the second stage estimates of *Historical Cattle Presence*, the bottom panel shows the first stage etimates of *Cattle Suitability Index*. The mean for the variable *Historical Cattle Presence* is 0.574 and the standard deviation is 0.495. The "Historical Controls" includes : the agricultural suitability, the proportion of area in tropical zone, longitude and absloute latitude coordinates, malaria ecology index, date of observation, economic development, and a measure of political complexity. The *Individual Controls* are : age and age squared. The *F-statistic* is the value of the Kleibergen-Paap F statistic on the excluded instrument. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. *, **, *** indicate the level of significance at the 10, 5 and 1 percent level.

Chapitre 2

Disease, Drought, and Development : Effects of the 1890s Cattle Plague in Sub-Saharan Africa

"But a far more general plague than the small-pox, and a much more terrible scourge than locusts, suddenly made its appearance and dogged our steps. This was rinderpest. No one who has not lived in Africa can form the least idea of this awful calamity."

– Francois Coillard (French missionary, 1897)

2.1 Introduction

In the 1890s, a severe outbreak of rinderpest – also known as 'cattle plague – spread throughout Sub-Saharan Africa. Although it did not directly affect human health, the disease devastated the cattle population, with mortality rates exceeding 90 percent among affected herds (Spinage 2003, Caceres 2011). For traditional cattle-based societies, this event represented not just the destruction of the main source of wealth, but the elimination of an asset that formed the basis of many social and economic arrangements. The 1890s Rinderpest Epizootic is widely considered to be the worst natural disaster in the continents history (Sunseri, 2018).

This paper studies the short- and long-run impacts of the 1890s rinderpest epizootic on cattle-based societies in Sub-Saharan Africa. We study the effects on ethnic groups in the immediate aftermath of the shock. Our analysis draws on detailed data on a range of ethnic group outcomes in the early 20th century (Murdock, 1967). We then track the long-run effects among descendants of tribes that were more exposed to the shock using contemporary individual-level data on wealth, education, fertility, and reported levels of trust.

Our empirical analysis is based on a difference-in-differences approach. We compare outcomes across cattle-based and non-cattle-based ethnic groups that were more or less exposed to the 1890s outbreak due to contemporaneous environmental conditions. Our approach combines two distinct sources of variation to identify rinderpest exposure. First, we use differences across ethnic homelands in the suitability of land for cattle-rearing, to identify which ethnic groups were more likely to own cattle and thus suffer from rinderpest episode. To measure ethnic homeland suitability for cattle-raising, we use information on the invariant geoclimatic conditions within each historical ethnic homeland. [1] Motivated by the agronomic literature, our 'cattle suitability index (CSI) is constructed as the interaction of three factors critical for cattle prosperity : abundance of pasture, availability of water sources, and a sufficiently flat terrain for migration. [2] We combine this variation with information on local climatic conditions from 1890-1896, to identify ethnic homelands that experienced a drought during the rinderpest period. The impact of drought on the local severity of rinderpest is well-documented (Spinage, 2003). This occurred both through increased disease transmission as animals crowded around scarce watering holes, and through increased fatality rates due to malnourishment. We measure rinderpest exposure to be those ethnic groups that *both* were likely to hold cattle *and* experienced a drought that preceded with the arrival of rinderpest to the continent.

We have three main sets of results. The first set focuses on the historical impacts of rinderpest. We find that cattle-based societies that *also* experienced a contemporaneous

1. We identify 533 historical ethnic homelands from Murdocks Map (1959).
2. The interaction of these three factors is motivated by the inability non-substitutability across the three different inputs (see Talba (2019), for a detailed discussion of the construction of this index).

drought were substantially less likely to own cattle in the early 20th century. Our estimates imply that rinderpest exposure reduced cattle ownership by 52 percent. The coefficient estimates are stable across specifications and unaffected by controls for invariant climatic or geographic conditions. Notably, the estimates from placebo regressions that interact cattle suitability with droughts in non-rinderpest years (1870-1876, 1880-1886) are small and statistically insignificant. Taken together, these findings support our identifying assumption that outcomes among cattle-based ethnic groups would have been similar, absent the rinderpest shock.

We find that cattle-based societies that were more exposed to rinderpest had lower population densities and higher level of indigenous slavery in the early 20th century. Given the close link between population density and development in pre-industrial settings, these findings suggest that cattle-based societies did not immediately recover from the episode. To explore the sources of these economic losses, we study how rinderpest influenced tribal economic activities. We find that the shock led a shift from cattle herding to shepherding. In comparison to sheep and goats, cattle offered a number of distinctive economic advantages : a source of draft power and fertilizer (Schneider, 1957), a better source of nutrition and more reliable food supply during droughts (Marshall and Hildebrand, 2002). Thus, the transition from cattle-based to shepherding societies may have contributed to the decline in population density. We also find some evidence that these societies shifted towards agriculture, although the ability to meet food surplus requirements may have hampered by the speed of the transition and lack of specific human capital. Despite these economic changes, we find no effects on the level of political centralization. These findings support the historiography which argues that the rinderpest epizootic did not cause institutional breakdown.

Our second set of results focus on the long-run economic impacts of rinderpest among the descendants of cattle-based societies. The analysis is based on socioeconomic data from the Demographic and Health Surveys (DHS). These surveys provide individual information for a large number of Sub-Saharan countries. We link individuals to historical ethnic groups, allowing us to identify descendants from societies that were more or less affected by the 1890s outbreak.

We find that the 1890s Rinderpest Epizootic had persistent effects on socioeconomic outcomes, more than a century after its end. We estimate that descendants of exposed ethnic groups have significantly lower levels of non-cattle wealth. These effects cannot be attributed to fertility or educational choices. Instead, we find that descendants from rinderpest exposed societies are significantly less likely to reside within the boundaries of their historical ethnic homeland, and they are more likely live in more rural areas with lower relative incomes. These results suggest that distressed migration may partly account for the persistent economic effects of rinderpest.

To conclude the empirical analysis, we explore the long-run social impact of the Rinderpest Epizootic. We use data from the 2005 Afrobarometer survey to study whether Rinderpest had persistent effects on self-reported levels of trust among descendants of cattle-based societies. We find that across several different measure of both inter- and intra-group trust, the outbreak had a persistent negative impact on self-assessed attitudes.

Our findings are consistent with a number of studies documenting the persistent impact of historical on contemporary levels of trust in Sub-Saharan Africa (Nunn and Wantchekon, 2011 ; Lowes, 2018). The results support a number of historical accounts that document the impact of the Rinderpest Epizootic on both inter- and intra-group hostilities (Rouget 1906, Van Onselen 1972). As cattle were destroyed, the social arrangements which the undergirded – marriage, tenant farming labor, etc. – broke down. In addition, widespread migration and cattle raiding may have contributed to inter-ethnic group violence. The timing of outbreak shortly after the arrival of European colonization may also have reinforced anti-colonial views. These attitudes were often reinforced by the actions of European colonizers. A number mine owners responded to the influx of labor by sharply lowering salaries. The following statement by J.H Johns (Manager of Fereira Gold Mining Co.), captures this view towards the outbreak : "At the present moment, the natives have to work because they cannot obtain food otherwise, and therefore I think we have a splendid opportunity."

This paper contributes to the literature on how societies adapt to environmental catastrophe in the short-run and long-run. Evidence from the U.S. suggests that migration is an important margin of adjustment that can significantly mitigate the immediate impacts (Blanchard and Katz, 1992 ; Hornbeck, 2012 ; Boustan, Kahn, and Rhode, 2012). In contrast, Feir, Gillezeau, and Jones (2019) find that the slaughter of the bison had persistent negative effects on Native Americans. They argue that constraints imposed on out-migration from reservations substantially hampered long-run adjustments. Our study provides an example of adaption in a context that lies between these two cases of high worker mobility and heavily restricted migration. Our context, in which individuals faced significant informal barriers to mobility, is similar to settings in many developing country.

Our paper also contributes to literature on the determinants of long-run development in Africa. Previous research has emphasized the role of both institutional factors (Acemoglu et al., 2001 ; 2012 ; Michalopoulos and Papaionnou, 2014) and cultural factors (Nunn, 2008 ; Alesina, Giuliano and Nunn (2013 ; Besley and Persson, 2010). We add to the small literature that identifies the role of environmental factors for long-run development (Alsan, 2015 ; Fenske and Kala, 2015).

The rest of the paper is organized as follows : Section 2.2 discusses the historical background ; Section 2.3 describes the data ; Section 2.4 presents the empirical strategy ; Section 2.5 presents the results ; and Section 2.6 concludes.

2.2 Historical Background

Rinderpest or cattle plague is an infectious cattle disease that used to ravage Europe, Asia and India (Spinage 2003). The virus is highly contagious affecting mostly cattle and buffalo with fatality rates that can exceed 80 percent ; it does not affect humans, however (Caceres 2011). The virus is present in the breath, secretions, and excretions of animals, and typically spread by direct contact between livestock, often at watering holes. Airborne spread is a secondary source of transmission (Refik-Bey and Refik-Bey 1899). Control techniques – quarantines and slaughter of infected animals – were first adopted in England in the 17th century, and widely used throughout Europe to limit its spread. An effective vaccine was developed in the twentieth century, and the disease was declared eradicated in 2011 (Morens et al. 2011).

The 1890s Rinderpest Epizootic originated when infected Indian cattle were introduced to Africa, imported by Italian traders through the Massawa port of Eritrea in. There is some uncertainty regarding the exact timing of introduction, with estimates between 1887-9 (Lugard 1893, Edington 1896, Zonchello 1917). As Crouzet (1889, p. 62-63) on January 8, 1889 at Massawa "For more than a year [...] a terrible epizootic has carried off the cattle, now in many places". Rinderpest spread rapidly westwards along the main trade route used by Fulani pastoralists. It reached the area of todays Burkina Faso in 1891 (Monteil 1895, Spinage 2003). It also spread southwards up to the Zambezi river in 1893. The natural barrier of the river halted its spread for three years, but the disease eventually crossed the Zambezi in 1896 and infected southern Rhodesia thereafter (Spinage 2003). The disease had unusually high fatality rates, and killed more than 90 percent of infected cattle (Spinage 2003).

In the 1880s and 1890s, the continent experienced a number of drought episodes (Pribyl et al. 2019, Nicholson et al. 2012). Drought conditions substantially increased the spread of the virus, as crop failures and water shortages increased the likelihood of closer contact of animals contact at fewer watering points (Marquardt 2005, Spinage 2003). Conditional on infection, cattle that were already weakened by drought conditions were also significantly more likely to die from disease (Pribyl et al. 2019).

The widespread death of cattle impeded the use of draught oxen, the development of transport system and plough oxen. This contributed to lower agricultural productivity and increase of food price (Schapera and Comaroff 1990). Since cattle were the dominant source of wealth in pastoralist societies, their loss greatly raised poverty and reduced the potential for Africans to invest in agriculture and to purchase additional land (Pribyl et al. 2019). The outbreak also led to a large displacement to non-husbandry activities such as hunters-gatherers, wage labour and mining sector (Spinage 2003).

2.3 Data

The next subsection describes the creation of the index of exposure to Rinderpest, which is used for our identification strategy. This is followed by a description of the historical and the contemporary data.

2.3.1 Index of exposure to Rinderpest

We construct the index of exposure to Rinderpest because a reliable map on the severity of the 1890s Rinderpest Epizotic is not available. The index can be constructed by combining the index of the suitability of land for cattle-rearing (CSI) with the index of drought exposure during the Rinderpest Epizotic period. Using the CSI as opposed to observed cattle presence purges the estimates of bias arising from ethnic groups with stronger institutions being able to better control the plague.

The CSI is constructed using a habitat suitability index model for cattle, which is a species habitat model used in research and management to predict an animal species occurrence over time (USFWS 1981). [3] The model combines three important factors for cattle presence : (1) The soil suitability for pasture ; (2) distance to nearest water source ; and (3) the measure of land slope gradient. Using the FAO's GAEZ v3.0 data along with the Esri's Natural Earth dataset, the CSI is computed according to the following functional form :

$$CSI = Pasture\ suitability \times distance\ to\ water \times terrain\ slope$$

This functional form reflects the non-substitutability across geographic inputs. For example, a lack of local sources of water cannot be offset by more pastoral land for grazing. The resulting data is ranged from zero to one and measure the likelihood of cattle being present in a grid-cell. The CSI is next computed for each ethnic group using the median of grid-cells within 10 kilometers of an ethnic group centroid. The CSI joined with a map of African ethnic groups (Murdock's map (1959)) is shown in Figure 2.1 panel (A).

3. For more details on the construction of the CSI see Talba (2019).

The main source of drought data is the 19th Century African Instrumental and Documentary Precipitation Data.[4] The database are compiled by Nicholson et al. (2012), and report annual rainfall conditions in terms of rainfall "anomalies" from the years 1801-1900. They reconstruct annual rainfall anomalies combining data from several historical archive of proxy and actual precipitation data in Africa. The proxy includes selected information, including historical verbal information on famine, drought, agriculture or the nature of the rainy season. They compute annual rainfall anomalies for 90 aggregated geographical regions in Africa using a principal component method of climate field reconstruction similar to Mann and Jones (2003). The anomaly classes are ranged from -3 to +3 representing severe drought, drought, dry, normal, good rains, wet and very wet. A rainfall anomaly of -1 in a given region at year t means that the region experienced an unusual dry year.

From the 90 geographical regions, we interpolate the baseline data using the Inverse Distance Weighted (IDW) method. This allows us to impute data wherever they are missing at a 0.005 degree by 0.005 degree (approximately 0.55km by 0.55km) grid cell level. We next convert the raw data into an annual drought anomaly for each grid cell from 1801-1900, and we re-scale the data to lie between 0 to 3 increasing with the severity of drought. A value of zero representing a no drought climate condition ; a value of one representing dry climate, two is assigned for drought, and three is for severe drought years. Given that drought data are measured at the grid-cell level, and our outcomes of interest involve historical location of ethnic groups, we construct a measure of ethnic groups drought anomaly as a median measure of grid-cells that lie within 10 kilometers of an ethnic group centroid. Figure 2.1 panel (B) depicts the drought anomaly joined with Murdock's map of African ethnic groups.

The index of exposure to Rinderpest is constructed as an interaction of the CSI and the index of drought anomaly preceding the arrival of the plague.[5] Two main reasons motivate our use of measure of drought preceding the arrival of rinderpest. First, drought

4. The dataset is available from the National Oceanic and Atmospheric Administrations (NOAAs) World Data Center for Paleoclimatology ftp ://ftp.ncdc.noaa.gov/pub/data/paleo/historical/africa/.

5. Data on the timing of the spread of Rinderpest in Africa are from Spinage (2003) and Botte (1985). They reconstruct the period of the outbreaks for each African country using numerous source of historical records from European explorers and missionaries.

favoured the spread of cattle plague as it reduced watering points and feeding grounds, increasing contacts between cattle and the transmission of the disease (Prebyl et al., 2019). Second, drought conditions during the year preceding the arrival of the plague made cattle weak and more susceptible to the disease. For instance, when Eastern Africa experienced severe cattle plague in 1889, the region was already facing a drought in 1888, and this contributed to kill more than 90 percent of cattle (Spinage 2003) . The resulting data consist of spatial data on Rinderpest exposure for years 1890-1896 at the ethnic group level.

2.3.2 Historical Data

The first set of analysis focuses on the historical impacts of rinderpest. We use data from Murdock's (1967) *Ethnographic Atlas*. The Atlas is a compilation of ethnographic, anthropological, and archives of missionary studies that George Peter Murdock assembled and published in 29 installments in the journal *Ethnology* between 1962 and 1980. It is a cross-sectional data which contains categorical variables describing cultural characteristics and subsistence strategy of 533 African ethnic groups.

Our main outcome for the historical analysis is an indicator for the presence of large domesticated cattle within ethnic group. For a robustness test, we use a continuous measure of the society's dependence on cattle, which is computed as the interaction of the cattle presence index and the degree to which the society relied on husbandry for subsistence (ranges 0 to 100%). To analyze the impacts on food production system, we construct a series of dummy variables for whether the main economic activity of ethnic groups was "animal husbandry", "agriculture", "gathering", "hunting" or "fishing". We also examine how rinderpest affected historical institutions. We measure institutions by a dummy variable for whether ethnic group practised indigenous slavery, and a measure for whether ethnic group was politically centralized. [6] Lastly, we examine the effect of rinderpest on historical population density using data from Murdock's book Africa : Its Peoples and Their Culture History.

6. from the literature, an ethnic group is considered politically centralized if it has a value above one for the jurisdictional hierarchy.

The *Ethnographics Atlas* is the best available cross-cultural data. However, its greatest shortcoming is that they are anachronistic. Ethnic groups are observed in different periods of time. Fortunately, this is not an issue that prohibits the use of the data set. We keep only ethnic groups observed after the arrival of rinderpest.

Murdock's Ethnographic atlas has been widely used. Alesina, Giuliano and Nunn (2013) aggregated the *Ethnographic Atlas* at country level to show that the historical use of plough agriculture predicts contemporary gender inequalities. Alsan (2015) used this data to analyze the effect of tsetse flies on African Development. Michalopoulos, Putterman and Weil (2016) have combined this data with Demographic and Health Surveys dataset in order to show that individuals whose ancestors relied on agriculture for subsistence are today more educated and wealthier.

We combine the Atlas with the Murdock's (1959) Map [7], which provides the geographic boundaries of approximatively 830 African ethnic groups. In total, 522 African ethnic groups from the Ethnographic Atlas are matched to the Murdock's Map. We can then determine the spatial locations of each ethnic group in the Atlas.

For the pre-colonial analysis, we include a set of controls variables measured at the ethnicity-level. We include a measure of distance from ethnic group's centroid to Massawa. We also include ethnic group's longitude, absolute latitude, which controls for differences between the eastern and the western part of the continent, altitude and proportion of land in tropical and sub-tropical capture differences in the agricultural and ecological areas. Finally, we include a measure of agricultural suitability from the Food and Agricultural Organization (FAO).

2.3.3 Contemporary Data

The second set of our analysis examines the long-run economic impacts of rinderpest among the descendants of cattle-based societies. The contemporary individual-level data stem from the *Demographic and Health Surveys* (DHS). The DHS are cross-sectional

7. Since the ethnicity names from the Murdock's Map do not always coincided with the names from the Ethnographic Atlas, Fenske (2014) and Michalopoulos and Papaioannou (2013) constructed a concordance rule that allows to link the groups from the Atlas to the groups in the Map.

nationally representative household surveys which cover over 90 developing countries. They consist in individual-level datasets which provide information on a range of topics including households wealth, individual human capital, migration and level of trust in medicine.

We use data from 84 surveys, covering 22 Sub-Saharan African countries over the period 1992-2016. We include in the analysis all the Sub-Saharan African surveys that have data on ethnicity of the respondent. We link the *Ethnographics atlas* data with the DHS using ethnic group's names. The final data consist of 527,748 individuals.

For the contemporary analysis, we focus on three main outcomes : wealth, education and migration. Wealth is a categorical variable coded 1 to 5, and measures household's cumulative living standard. [8] Education is a dummy variable for respondent ever attended school. We measure migration by a dummy variable for whether the respondent is not longer living in her ancestral homeland. Additionally, we examine the effect on fertility and the level of trust in medicine. Fertility is measured by the number of children ever born, and on the level of trust in medicine is captured by a dummy variable for whether the respondent refused blood test.

Our last empirical analysis explores the long-run social impact of the Rinderpest Epizootic. We examine whether rinderpest caused a culture of mistrust to develop within Africa. Data on trust are taken from the 2005 Afrobarometer surveys. They are a series of nationally representative individual-level surveys. The sample consists of 21,822 respondents in 17 sub-Saharan countries from which information on ethnicity names are available. We link the Afrobarometer data with the *Ethnographics atlas* and index of rinderpest exposure based on ethnicity names. The resulting sample consists of 20,500 individuals.

The Afrobarometer asks respondents how much they trust their relatives, neighbours, people from other ethnic groups, and those from the same ethnic groups. We construct a measure of each trust as variables that take on the value of 0, 1, 2 and 3 : where 0 corresponds to "not at all" ; 1 to "just a little" ; 2 to "somewhat" and 3 to "a lot".

8. The DHS wealth index is constructed as a composite based on household ownership of various assets (televisions, bicycles), housing material, and access to water facilities.

2.4 Empirical Strategy

To estimate the effects of rinderpest on development in sub-Saharan Africa, we follow a difference in differences framework. Our estimation strategy compares outcomes across cattle-based and non-cattle-based societies that were more or less exposed to rinderpest. We estimate the following equation :

$$Y_i = \alpha + \beta CSI_i \times Drought_i + \gamma CSI_i + \delta Drought_i + X_i^{'} \Omega + \epsilon_i \tag{2.1}$$

where Y_i represents the historical outcomes for ethnic groups i. For the long-run analysis, outcomes are denoted by the subscript (j, i, c) to identify individuals j, who reside in country c, and are descendants of historical ethnic group i. The term CSI_i measures the average suitability of the ethnic homeland for cattle-rearing according to local agronomic conditions. $Drought_i$ is an indicator for homelands that experienced a drought the year preceding the outbreak between 1890 and 1896. The vector X_i includes a number of geographic conditions of ethnic homelands that might simultaneously influence development and rinderpest exposure. We control for a measure of market access to Massawa - the entry point of the 1890s outbreak. We also include an index for agricultural suitability and the proportion of the ethnic homeland that falls in tropical and subtropical zones from the FAO, that take in account environmental conditions and soil fertility. Finally, we control for ethnic homeland altitude, longitude and absolute latitude, that capture differences for the eastern and western part of the continent, as well as differences across agro-ecological zones.[9]

The main variable of interest is the interaction term, $CSI_i \times Drought_i$. The difference-in-difference coefficient, β, captures the differential in outcomes between cattle-based and non-cattle based societies that experienced a drought during the rinderpest outbreak, relative to the differential in regions that did not experience a drought during the outbreak. The coefficient captures the rinderpest exposure as reflected in societies that were *both* likely to raise cattle *and* vulnerable to the disease

9. In the long-run analysis, we also include individual covariates : a quadratic in respondents age, gender of the household head, survey year fixed effects, and (in some specifications) country fixed effects.

because of contemporaneous environmental conditions.

Our identification assumption is that the relative gap in outcomes between cattle-based and non-cattle-based societies would have been similar across regions, absent the arrival of rinderpest to the continent. In particular, there should be no differences in relative outcomes systematically related to locations of drought in the 1890s. This assumption is supported by the fact our measure of CSI, which depends on the specific interaction of three agronomic factors, is uncorrelated to a range of other geoclimatic and historical conditions relevant for long-run development (Talba, 2019). One potential concern with our approach is serial correlation in drought probability, so that locations that experienced a drought in the 1890s were more drought-prone in general. As a result, there may systematic differences in economic outcomes in areas that experienced a contemporaneous drought, independent of rinderpest. The inclusion of the main effect, $Drought_i$, controls common differences between more or less drought-prone regions. Nevertheless, the ability to cope with underlying drought risk may differ systematically in areas with access to cattle (e.g. Marshall and Hildebrand, 2002), a violation of our identifying assumption. To address this concern and validate our empirical approach, we conduct a series of placebo test, that examine the interaction between CSI and droughts in the decades prior to 1890. The results show that only the interaction with contemporaneous droughts have relevance for subsequent economic outcomes (see section).

Two final estimating details are worth noting. First, historical regressions are unweighted and long-run regressions use survey weights. Second, we cluster the standard errors at the level of cultural provinces to capture both spatial and genealogical potential correlation in outcomes across ethnic groups. These provinces have been grouped by Murdock (1959) based on written, archeological records and linguistic similarities (Alsan 2015).

2.5 Results

2.5.1 Impacts of the Rinderpest Epizootic on Early 20th Century Ethnic Outcomes

Cattle Ownership

Table 2.1 reports the effects of rinderpest exposure on cattle presence, based on various specification of equation (1). Column (1) reports the baseline difference-in-difference estimates. In column (2), we add controls for distance to the port of Massawa, which may have influence the timing of rinderpest arrival. In column (3), we add controls for the quality of agricultural land, which may have mitigated the local impact of rinderpest by facilitating the transition from cattle herding to agricultural production. The remaining columns (4)-(8) include additional controls for local geoclimatic factors that may also have contributed to the spread and severity of the disease.

The main coefficient estimates for rinderpest exposure – $CSI \times Drought$ – are large, negative, and statistically significant. These estimates are quantitatively meaningful : a one standard deviation increase in rinderpest exposure among ethnic groups on land suitable for cattle-rearing decreases by 13% the likelihood to have cattle present in the early 20th century. This effect corresponds to a 21.7% decrease relative to the average cattle presence rate among ethnic groups on land suitable for cattle-rearing that was unaffected by rinderpest. The estimates are also remarkably stable across the different specifications, and unaffected by the inclusion of additional covariates. Combining these estimates with the R-squared from the regression models, we follow the procedure outlined by Oster (2019) to evaluate the potentially omitted variables to account for the observed effects. We calculate that selection on unobservables must be 10 times larger than selection on observables to account for the $CSI \times Drought$ effect. [10] Intuitively, the

10. This calculation is derived based on the following equation $\beta^* - \tilde{\beta} \approx \delta[(\hat{\beta} - \tilde{\beta})\frac{(R_{max}-\tilde{R})}{\tilde{R}-\hat{R}}]$, where $(\tilde{\beta}, \tilde{R}, \hat{\beta}, \hat{R})$ denote the coefficient estimates and R-squared from the simple model (col. 1) and the fully specified model (col. 8), respectively ; where $\beta^* = 0$ denotes the true coefficient estimate assuming no effect ; and where δ denotes the proportional selection on unobservables relative to observables. Following the guidelines in Oster (2019), we calculate relative selection assuming $R_{max} = 1.3 \times \tilde{R}$. Even under the extreme scenario of $R_{max} = 1$, in which there is no measurement error in the dependent variable and

fact that the coefficient estimates remain stable, even as controls that have explanatory power for cattle ownership are included, suggests that they cannot be solely attributed to some other omitted factors.

Table 2.2 reports results from modified versions of equation (1) that include the interaction of lagged drought measure with CSI. Comparing the CSI interaction with three drought periods – 1870-1876, 1880-1886, and 1890-1896 – only drought exposure that coincided with the spread of the virus throughout Africa affected cattle ownership. These results provide strong support for our identification assumption that the specific drought conditions the year before the outbreak contributed to the decline in cattle ownership, rather than some other unobservable characteristic that was more common in drought-prone regions.

In Table 2.3, we explore the effects of rinderpest on the reliance on cattle at the intensive margin. We estimate regressions for cattle dependence, a variable that ranges from 0 to 9 based on the degree to which the society is dependent on cattle for subsistence, based on information reported in Murdock (1967). The results are similar in sign and significance to the patterns reported in Table 2.1.

Population Density, Political Centralization, and Economic Activity

Having established the impact of rinderpest exposure on historical cattle ownership, we turn the broader effects of the outbreak on historical economic and institutional outcomes. Table 2.4, col. 1 reports the effects on population densities. We find that societies that were relatively more exposed to rinderpest had lower population densities in the early 20th century, suggesting that cattle-based did not experience rapid economic recovery from the shock. The coefficient estimate is large and negative, though statistically insignificant.[11]

We find no effect on the level of political centralization (col. 2) and intensive agriculture (col. 4). This pattern suggests that the shock did not cause a fundamental reorganization of ethnic institutional arrangements. Our results provide a counterpoint

---

potential covariates can completely explain cattle ownership, proportional selection on unobservables would still have to be 1.7 times greater to account for the findings.

11. Note that the estimate in col. 1 is not significant because of the small sample size.

to Acemoglu et al. (2012), who argues that large environmental shocks can lead to major institutional upheaval, as societies shift from one equilibrium to another. Instead, they are consistent with much of the historiography, which argues that despite the substantial economic and social costs of the Rinderpest Epizootic, the event did not cause widespread institutional upheaval (Spinage 2003).

Spinage(2003, p.617) argues the social disintegration due to rinderpest led to the sale of children by their parents and self-enslavement. The results in col. 3 is an empirical support for a positive correlation between rinderpest and the practice of indigenous slavery. A one standard deviation increase in rinderpest exposure is associated with a statistically significant 11.7% increase in the probability an ethnic group used slave labor.

Why did societies affected by the Rinderpest Epizootic experience relative population decreases into the 20th century ? To explore this question, we study how this episode reshaped local economic activity across ethnic groups. Table 2.5, col. 1-2 show the relative decrease in cattle-herding was offset by a rise in shepherding. There may have been economic costs associated with this shift, since cattle offer a number of economic advantages relative to sheep and goats. For example, cattle contribute to agricultural productivity by providing a source of draft power and manure for fertilizer (Schneider, 1957). Their milk provides a better supply of nutrients and they are a particularly reliable source of food during droughts (Marshall and Hildebrand, 2002).

We also find some evidence that the Rinderpest Epizootic caused cattle-based societies to shift to agricultural production (Table 2.5, col. 4). The coefficient estimate is large and positive, albeit statistically insignificant. Meanwhile, the estimates for gathering, hunting, and fishing are all small and statistically insignificant. There may have been short-run economic costs associated with the transition to agriculture, given the lack of knowledge of farming practices. Future research will explore the extent to which cattle-based societies that were less specialized and had previously engaged in agricultural production were better able to withstand the rinderpest shock.

2.5.2 Impacts of the Rinderpest Epizootic on Contemporary Outcomes

Wealth, Education, Fertility, and Migration

In this section, we explore the long-run economic effects of rinderpest among descendants of cattle-based societies. The analysis is based on an individual-level regression variant of equation (1), using data on contemporary economic outcomes from the DHS. In addition to the historical ethnic group controls, we add a vector of individual controls that include a quadratic in respondents age, gender of the household head, survey year fixed effects, and (in some specifications) country of residence fixed effects. To identify historical exposure to rinderpest, we link individuals in DHS to the Murdock ethnic groups, following the procedure described in Talba (2019). Overall, we link 85 percent of the DHS sample to the *Ethnographic Atlas* to one of 215 ethnic groups in the *Atlas*.

Table 2.6 reports the long-run impacts of rinderpest on household wealth, based on a composite index designed to capture household living standards.[12] The estimates in columns 1-3 are consistently negative and statistically significant, implying that rinderpest imposed significant long-run economic costs that persisted for more than a century after the outbreak. In columns (4) and (5), we add controls for whether the household resides in a rural area, and whether the household still resides within the historical ethnic boundaries of their ancestors.[13] The inclusion of these controls reduces the main effect sizes by one fourth, the estimates remain significant at 5 percent. Taken together, the results suggest that the long-term decline in household wealth may be partly explained by differential migration patterns, an issue we explore in more detail below.

The decreases in household wealth cannot be attributed to changes in human capital or fertility decisions among descendants of cattle-based societies. We find no evidence that the outbreak had persistent effects on educational or fertility decision among descendants

12. The DHS wealth index is constructed as a composite based on household ownership of various assets (televisions, bicycles), housing material, and access to water facilities.

13. Given potential endogeneity in these explanatory variable, the estimates from these regressions should be interpreted with caution.

of cattle-based societies (Table 2.7, cols. 1-6). The coefficient estimates are all small in magnitude and statistically insignificant.

To further examine the role of migration in driving long-run economic outcomes, we estimate a series of regressions that explore the relative mobility across societies that were more or less exposed to rinderpest. Specifically, we construct a dependent variable – 'mover – that identifies whether contemporary descendants no longer reside within the boundaries of their historical ethnic homeland. [14] We also identify whether 'movers reside a region that has lower mean wealth levels than the historical area of their ancestors and whether they currently live in an urban or rural area.

The results show systematic differences in the migration patterns of ethnic groups that were more severely affected by rinderpest. Descendants of exposed ethnicities were significantly more likely to live outside their historical ethnic homeland (Table 2.8, cols. 1-2), and *also* more likely to have migrated to rural areas with lower average wealth levels (cols. 3-6). Together, these patterns are consistent with historical distressed migration. The Rinderpest Epizootic led widespread outmigration from cattle-based tribes to neighboring areas (Spinage 2003). Informal barriers to mobility, resulting from either interethnic conflict or colonial forces, may have hampered the ability of displaced workers to seek out new economic opportunities. To the extent that subsequent generations faced mobility constraints, the consequences of these migration patterns may have persisted to present day. These findings contrast with the historical U.S. context, where high levels geographic mobility substantially mitigate the long-run consequences of local shocks (e.g., Hornbeck, 2012).

The Cultural Legacy of Rinderpest

"I never saw all these different diseases until after the 'white men came : my cattle were all right until everything was changed by them."
– Fulani elder

To conclude the analysis, we explore the lasting effects of Rinderpest Epizootic on reported levels of trust among descendants of cattle-based ethnicities. There are

14. All regression models include controls for ethnic homeland area.

number of channels through which this episode may have shaped cultural attitudes. The outbreak led to a widespread breakdown of many social and economic relationship that were undergirded by cattle (Sunseri, 2018). The marriage market shut down, since bride prices could not be paid. Without the dominant source of transferable wealth, economic relationships ceased functioning. Additionally, displacement of workers to new territory and widespread cattle raiding contributed to inter-group hostilities (Spinage 2003). Finally, the outbreak reinforced animus towards white colonizers. There was deep resentment of policies that required the slaughtering of infected herds, and rumours circulated that the disease had been introduced by colonists in efforts suppress workers in cattle-based societies. These views were exacerbated by the response of a number of colonial mine companies, who cut worker wages in response to the outbreak (Phoofolo, 1993).

We use self-assessed levels of trust from the 2005 Afrobarometer to study the long-term cultural impact of the Rinderpest Epizootic. We estimate individual-level regression variants of equation (1), using linking individuals to ancestors who were more or less affected by the outbreak. Table 2.9 reports the results for several different measures of trust. We estimate large negative effects across a number of different measures of trust. Descendants of ethnic groups exposed to rinderpest reported lower levels of trust in neighbors, and lower levels of both inter- and intra-group trust. In contrast, we find no statistical difference in trust of relatives. We also find that these groups were more likely to refuse to take the DHS blood test, a response that has been linked to lower levels of trust in the medical system (Lowes, 2018). Overall, these patterns suggest that the Rinderpest Epizootic had persistent negative effects on cultural attitudes. These results support an emerging literature documenting the role of major historical episodes in shaping contemporary cultural attitudes in Sub-Saharan Africa (e.g., Nunn and Wantchekon, 2011 ; Lowes, 2017, 2018).

2.6 Conclusion

The 1890s Rinderpest Epizootic is considered to be the worst natural disaster in Africas history. This paper studies the immediate economic consequences of this episode across tribes in early 20th century and tracks the long-run impacts among modern descendants. In the short-run, we find that the cattle plague led to a substantial decrease in cattle holdings and a substitution towards shepherding and agricultural production. Societies exposed to rinderpest experienced significant decreases in population density, although the outbreak did not cause broader institutional reorganization. Rinderpest had long-run impacts on economic and social outcomes among descendants of cattle-herding ethnic groups. We find persistent negative effects on household wealth, more than a century after the outbreak. These losses can partly be attributed to distressed migration. Finally, the episode appears to have had a long-run impact on levels of trust that originated from the strains the disease imposed on inter- and intra-group interactions.

This paper provides new insights into the ability of households to cope with environmental catastrophe in a context in which they face informal barriers to mobility. These findings may have relevance for many modern developing country settings, in which both government policy and inter-ethnic conflict may limit agents ability to cope with severe shocks. In ongoing research, we will explore how ethnic conflict and colonial policy interacted with the shock to shape long-run outcomes.

Appendix

A Figures

Figure 2.1 : Cattle Suitability Index

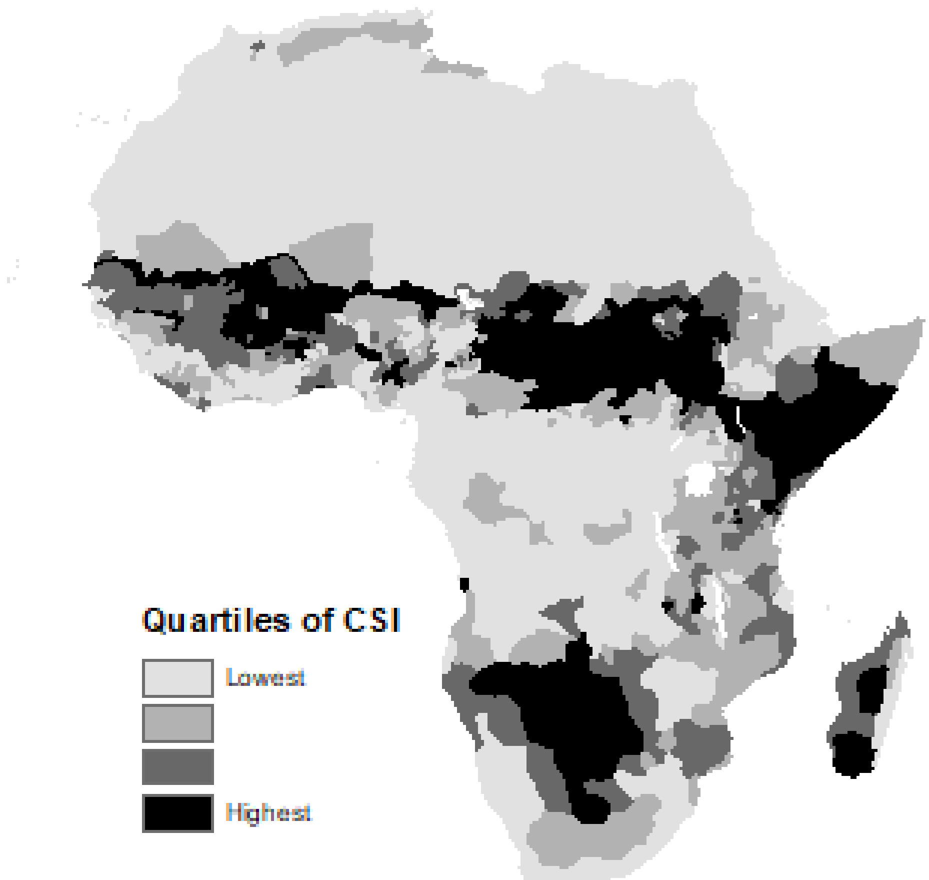


Figure 2.2 : Historical drought exposure

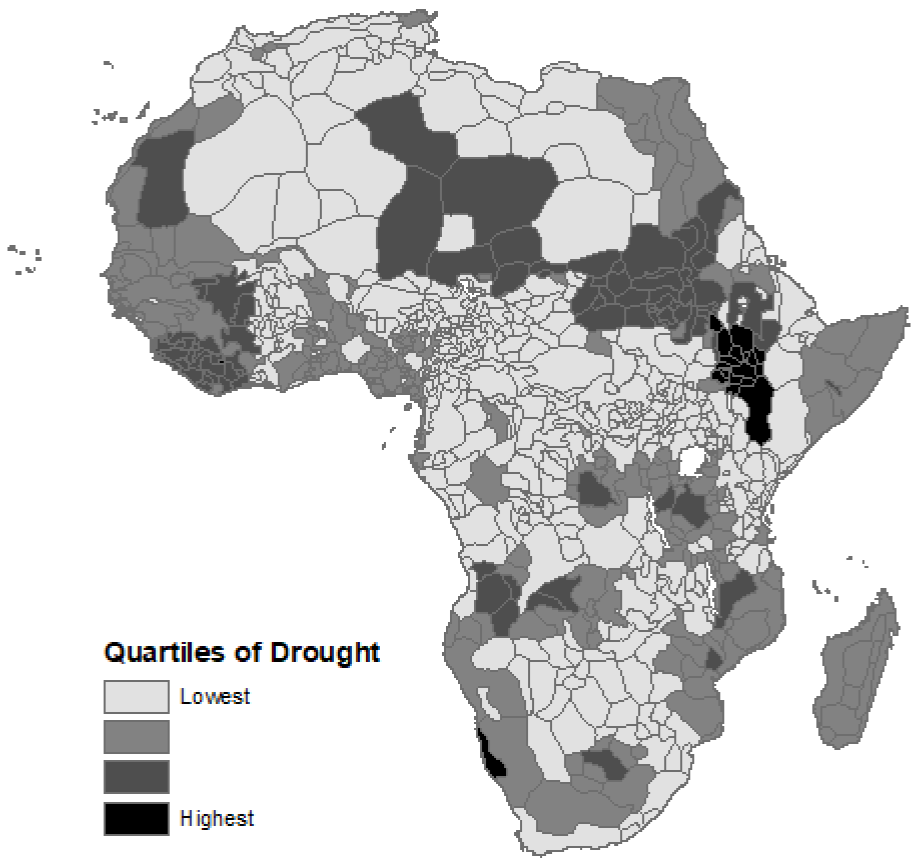
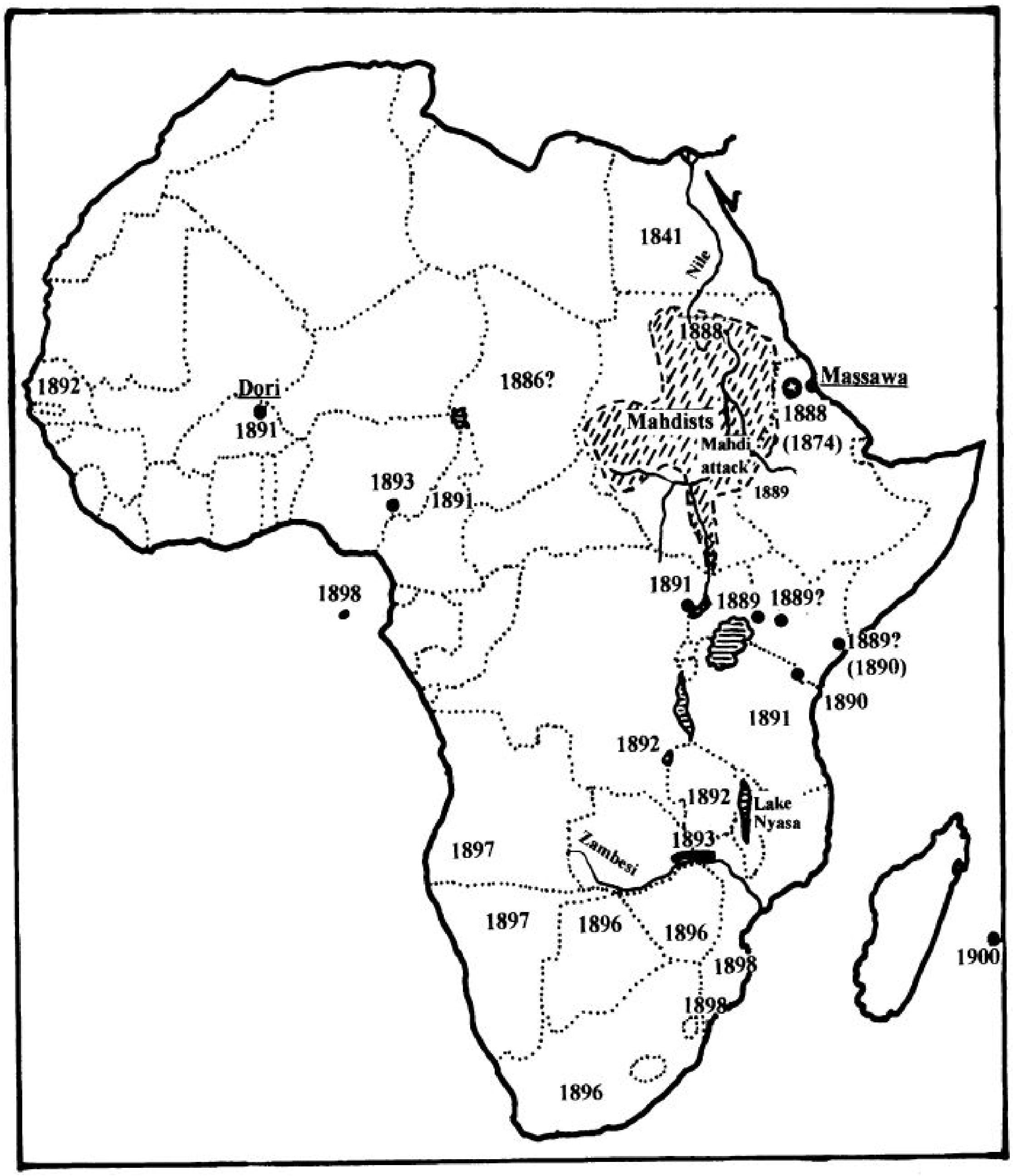


Figure 2.3 : The spread of rinderpest in Africa 1888-1896



Source : Spinage (2003).

Table 2.1 : Reduced-form estimates of the relationship between precolonial cattle presence and cattle plague

	Cattle (1)	Cattle (2)	Cattle (3)	Cattle (4)	Cattle (5)	Cattle (6)	Cattle (7)	Cattle (8)
$CSI \times Drought$	-0.536***	-0.529***	-0.507***	-0.515***	-0.515***	-0.528***	-0.530***	-0.525***
	(0.177)	(0.186)	(0.186)	(0.184)	(0.182)	(0.183)	(0.179)	(0.176)
Drought	0.210*	0.190	0.180	0.184	0.184	0.199	0.188	0.193
	(0.123)	(0.127)	(0.129)	(0.129)	(0.128)	(0.128)	(0.126)	(0.123)
CSI	0.897***	0.855***	0.879***	0.877***	0.876***	0.848***	0.871***	0.862***
	(0.107)	(0.121)	(0.114)	(0.117)	(0.115)	(0.109)	(0.111)	(0.113)
Distance to Massawa Port		-0.138**	-0.113*	-0.106*	-0.114	-0.126	-0.170*	-0.108
		(0.057)	(0.061)	(0.062)	(0.084)	(0.087)	(0.091)	(0.108)
Agricultural Suitability Index			-0.338**	-0.321*	-0.314*	-0.221	-0.236	-0.204
			(0.150)	(0.163)	(0.162)	(0.161)	(0.159)	(0.158)
Mean Altitude				0.049	0.065	0.058	0.047	-0.000
				(0.095)	(0.096)	(0.098)	(0.097)	(0.090)
Longitude					-0.000	-0.000	-0.002	-0.002
					(0.004)	(0.004)	(0.004)	(0.004)
Absolute Latitude						0.008	0.001	0.001
						(0.007)	(0.008)	(0.009)
Proportion in Tropics							-0.492**	-0.292
							(0.239)	(0.298)
Ruggedness								0.157*
								(0.086)
Observations	442	442	442	442	442	442	442	442
Number clusters	43	43	43	43	43	43	43	43
R-squared	0.139	0.167	0.182	0.183	0.184	0.190	0.200	0.210
$CSI \times Drought$ std.dev.	0.254	0.254	0.254	0.254	0.254	0.254	0.254	0.254
Dep. var. mean unaffected	0.599	0.599	0.599	0.599	0.599	0.599	0.599	0.599

Notes. OLS estimates. Standard errors in parenthesis, clustered at the cultural province level. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *Cattle* is an indicator variable that takes the value one if ethnic group had cattle in precolonial times. $CSI \times Drought$ is the index of rinderpest exposure by the ethnic group. *CSI* stands for Cattle Suitability Index, it is a continuous variable that lies between zero and one. *Drought* is an indicator variable that takes the value of one if an ethnic group experienced drought anomaly the year preceding the outbreak.
***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.2 : Reduced-form estimates of the relationship between cattle plague and cattle presence

	Cattle (1)	Cattle (2)	Cattle (3)	Cattle (4)
$CSI \times Drought$	-0.525***	-0.607***	-0.575***	-0.698***
	(0.176)	(0.194)	(0.187)	(0.206)
$CSI \times Drought_{1878-1887}$		0.143		0.222
		(0.233)		(0.238)
$CSI \times Drought_{1868-1877}$			0.193	0.298
			(0.286)	(0.280)
Historical Controls	Yes	Yes	Yes	Yes
Observations	442	442	442	442
Number Clusters	43	43	43	43
R-squared	0.210	0.225	0.217	0.236
$CSI \times Drought$ std.dev.	0.254	0.254	0.254	0.254
Dep. var. mean unaffected	0.599	0.599	0.599	0.599

Notes. OLS estimates. Standard errors in parenthesis, clustered at the cultural province level. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *Cattle* is an indicator variable that takes the value one if ethnic group had cattle in precolonial times. $CSI \times Drought$ is the index of rinderpest exposure by the ethnic group. CSI stands for Cattle Suitability Index, it is a continuous variable that lies between zero and one. *Drought* is an indicator variable that takes the value of one if an ethnic group experienced drought anomaly the year preceding the outbreak.

***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.3 : Reduced-form estimates of the relationship between cattle plague and presence of cattle, other measure of cattle

	Cattle (1)	Cattle (2)	Cattle (3)	Cattle (4)	Cattle (5)	Cattle (6)	Cattle (7)	Cattle (8)
$CSI \times Drought$	-1.970**	-1.931**	-1.762*	-1.829**	-1.837**	-1.876**	-1.885**	-1.891**
	(0.941)	(0.935)	(0.874)	(0.888)	(0.903)	(0.900)	(0.881)	(0.889)
Drought	0.724	0.596	0.517	0.556	0.551	0.596	0.550	0.546
	(0.478)	(0.479)	(0.455)	(0.465)	(0.463)	(0.454)	(0.446)	(0.444)
CSI	3.299***	3.023***	3.209***	3.188***	3.189***	3.106***	3.202***	3.210***
	(0.717)	(0.648)	(0.508)	(0.508)	(0.512)	(0.556)	(0.563)	(0.562)
Distance to Massawa Port		-0.892***	-0.704***	-0.648**	-0.615*	-0.650*	-0.831**	-0.889**
		(0.253)	(0.239)	(0.248)	(0.345)	(0.341)	(0.362)	(0.372)
Agricultural Suitability			-2.563***	-2.417***	-2.412***	-2.138***	-2.200***	-2.231***
			(0.581)	(0.646)	(0.649)	(0.664)	(0.657)	(0.682)
Mean Altitude				0.420	0.418	0.399	0.353	0.397
				(0.376)	(0.354)	(0.370)	(0.366)	(0.379)
Longitude					0.002	0.003	-0.005	-0.005
					(0.014)	(0.014)	(0.014)	(0.014)
Absolute Latitude						0.023	-0.005	-0.005
						(0.030)	(0.035)	(0.034)
Proportion in Tropics							-2.016*	-2.203*
							(1.007)	(1.138)
Ruggedness								-0.148
								(0.390)
Observations	442	442	442	442	442	442	442	442
Number clusters	43	43	43	43	43	43	43	43
R-squared	0.127	0.206	0.263	0.268	0.268	0.272	0.284	0.284
$CSI \times Drought$ std.dev.	0.254	0.254	0.254	0.254	0.254	0.254	0.254	0.254
Dep. var. mean unaffected	1.812	1.812	1.812	1.812	1.812	1.812	1.812	1.812

Notes. OLS estimates. Standard errors in parenthesis, clustered at the cultural province level. The unit of observation is an ethnic group from the *Ethnographic Atlas*. The dependent variable *Cattle* measures the degree to which an ethnic group depended on cattle for subsistence ranged 0−9. $CSI \times Drought$ is the index of rinderpest exposure by the ethnic group. CSI stands for Cattle Suitability Index, it is a continuous variable that lies between zero and one. *Drought* is an indicator variable that takes the value of one if an ethnic group experienced drought anomaly the year preceding the outbreak.
***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.4 : Reduced-form estimates of the relationship between cattle plague and historical Africa development

	Log Population Density (1)	Political Centralization (2)	Indigenous Slavery (3)	Intensive Agriculture (4)
$CSI \times Drought$	-0.484	0.177	0.462**	-0.029
	(0.703)	(0.168)	(0.217)	(0.186)
Drought	0.702	-0.069	-0.194*	-0.109
	(0.487)	(0.096)	(0.114)	(0.135)
CSI	0.437	-0.208*	-0.386**	0.035
	(0.584)	(0.105)	(0.143)	(0.194)
Historical Controls	Yes	Yes	Yes	Yes
Observations	361	428	409	442
Number clusters	42	43	43	43
R-squared	0.182	0.117	0.218	0.089
$CSI \times Drought$ std.dev.	0.254	0.254	0.254	0.254
Dep. var. mean unaffected	1.408	0.333	0.856	0.348

Notes. OLS estimates. Standard errors in parenthesis, clustered at the cultural province level. The unit of observation is an ethnic group from the *Ethnographic Atlas*. $CSI \times Drought$ is the index of rinderpest exposure by the ethnic group. CSI stands for Cattle Suitability Index, it is a continuous variable that lies between zero and one. *Drought* is an indicator variable that takes the value of one if an ethnic group experienced drought anomaly the year preceding the outbreak.

***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.5 : Reduced-form estimates of the relationship between cattle plague and historical economic activity

	Cattle	Sheep and Goats	Husbandry	Agriculture	Gathering	Hunting	Fishing
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$CSI \times Drought$	-0.525***	0.391**	-1.270	0.991	-0.062	0.146	0.163
	(0.176)	(0.167)	(0.776)	(0.716)	(0.300)	(0.329)	(0.513)
Drought	0.193	-0.112	0.262	-0.298	0.019	0.050	-0.014
	(0.123)	(0.113)	(0.434)	(0.393)	(0.118)	(0.195)	(0.284)
CSI	0.862***	-0.601***	1.154	-0.569	0.360*	-0.312	-0.658
	(0.113)	(0.113)	(0.704)	(0.641)	(0.189)	(0.211)	(0.430)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	442	442	479	479	479	479	479
Number clusters	43	43	43	43	43	43	43
R-squared	0.210	0.221	0.365	0.199	0.097	0.086	0.091
$CSI \times Drought$ std.dev.	0.254	0.254	0.254	0.254	0.254	0.254	0.254
Dep. var. mean unaffected	0.599	0.307	2.532	5.566	0.391	0.782	0.722

Notes. OLS estimates. Standard errors in parenthesis, clustered at the cultural province level. The unit of observation is an ethnic group from the *Ethnographic Atlas*. $CSI \times Drought$ is the index of rinderpest exposure by the ethnic group. CSI stands for Cattle Suitability Index, it is a continuous variable that lies between zero and one. *Drought* is an indicator variable that takes the value of one if an ethnic group experienced drought anomaly the year preceding the outbreak.
***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.6 : Reduced-form estimates of the relationship between historical cattle plague and current wealth index

	Wealth (1)	Wealth (2)	Wealth (3)	Wealth (4)	Wealth (5)
$CSI \times Drought$	-1.078**	-1.047**	-0.950*	-0.754**	-0.739**
	(0.463)	(0.487)	(0.527)	(0.379)	(0.372)
Drought	1.001***	1.002***	0.998***	0.767***	0.757***
	(0.272)	(0.274)	(0.266)	(0.202)	(0.198)
CSI	0.140	0.128	-0.0549	-0.199	-0.194
	(0.305)	(0.308)	(0.306)	(0.244)	(0.244)
Rural				-1.727***	-1.730***
				(0.0590)	(0.0564)
Mover					-0.0382
					(0.0939)
Historical Controls	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
Country FE	No	No	Yes	Yes	Yes
Observations	185,048	185,048	185,048	185,048	185,048
Number Clusters	111	111	111	111	111
R-squared	0.036	0.058	0.068	0.417	0.417
$CSI \times Drought$ std.dev.	0.244	0.244	0.244	0.244	0.244
Dep. var. mean unaffected	2.860	2.860	2.860	2.860	2.860

Notes. OLS estimates. *Individual controls* include age, age squared, gender of the household head, and survey year fixed effects. *Historical controls* include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absolute latitude coordinates of the ethnic group, ruggedness. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level.
***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.7 : Reduced-form estimates of the relationship between historical cattle plague, education and fertility

	Education	Education	Literacy	Literacy	Number of children	Number of children
	(1)	(2)	(3)	(4)	(5)	(6)
$CSI \times Drought$	-0.197	-0.117	0.039	-0.005	0.035	0.195
	(0.152)	(0.130)	(0.176)	(0.132)	(0.404)	(0.489)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	No	Yes	Yes	Yes
Country FE	No	Yes	No	Yes	No	Yes
Observations	402,626	402,626	238,147	238,147	292,253	292,253
Number Clusters	150	150	115	115	149	149
R-squared	0.190	0.201	0.273	0.298	0.578	0.581
$CSI \times Drought$ std.dev.	0.226	0.226	0.244	0.244	0.224	0.224
Dep. var. mean unaffected	0.195	0.195	0.300	0.300	3.225	3.225

Notes. OLS estimates. *Individual controls* include age, age squared, a dummy for the respondent gender, and survey year fixed effects. *Historical controls* include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absolute latitude coordinates of the ethnic group, ruggedness. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level.
***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.8 : Reduced-form estimates of the relationship between historical cattle plague and modern educational and migration

	Mover	Mover	Mover to rural area	Mover to rural area	Mover to lower inc. region	Mover to lower inc. region	Mover to urban area	Mover to urban area	Mover to higher inc. region	Mover to higher inc. region
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$CSI \times Drought$	0.268	0.447	0.228	0.274	0.495*	0.528**	0.0398	0.173	-0.227	-0.0812
	(0.308)	(0.281)	(0.226)	(0.232)	(0.271)	(0.257)	(0.148)	(0.147)	(0.338)	(0.341)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	402,639	402,639	402,639	402,639	402,639	402,639	402,639	402,639	402,639	402,639
Number Clusters	150	150	150	150	150	150	150	150	150	150
R-squared	0.139	0.180	0.055	0.074	0.063	0.077	0.063	0.079	0.117	0.165
$CSI \times Drought$ std.dev.	0.226	0.226	0.226	0.226	0.244	0.244	0.244	0.244	0.224	0.224
Dep. var. mean unaffected	0.562	0.562	0.342	0.342	0.188	0.188	0.220	0.220	0.373	0.373

Notes. OLS estimates. *Individual controls* include age, age squared, a dummy for the respondent gender, and survey year fixed effects. *Historical controls* include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absolute latitude coordinates of the ethnic group, ruggedness. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level. ***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.9 : Reduced-form. Historical cattle plague and Trust

	Trust of relatives (1)	Trust of neighbors (2)	Intra-group trust (3)	Inter-group trust (4)	Refused blood test (5)
$CSI \times Drought$	-0.140	-0.289**	-0.389***	-0.406***	0.017***
	(0.116)	(0.117)	(0.115)	(0.092)	(0.005)
Historical Controls	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes
Observations	13,305	13,274	13,226	13,086	69,870
R-squared	0.119	0.128	0.146	0.119	0.018

Notes. OLS estimates. *Individual controls* include age, age squared, a dummy for a respondent being a female, and survey year fixed effect. *Historical controls* include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absloute latitude coordinates, and average altitude of the ethnic group. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level.

***, **, * indicate significance at the 1%, 5% and 10% level.

Table 2.10 : Reduced-form. Historical cattle plague and Health

	Height (1)	Height (2)	BMI (3)	BMI (4)	Anemia (5)	Anemia (6)
$CSI \times Drought$	-16.61*	-14.24*	-106.3	-78.29	0.013	-0.008
	(8.555)	(7.952)	(71.83)	(90.45)	(0.039)	(0.045)
Historical Controls	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	Yes	No	Yes
Observations	110,203	110,203	178,759	178,759	66,369	66,369
R-squared	0.022	0.029	0.100	0.106	0.030	0.033

Notes. OLS estimates. *Individual controls* include age, age squared, a dummy for a respondent being a female, and survey year fixed effect. *Historical controls* include the suitability for the ethnic group location for agriculture, the proportion of the location that is tropical or subtropical, longitude and absloute latitude coordinates, and average altitude of the ethnic group. Coefficients are reported with robust standard errors in parenthesis clustered at the ethnicity level.

***, **, * indicate significance at the 1%, 5% and 10% level.

Chapitre 3

Sibling Gender and Marriage Patterns in Sub-Saharan Africa

3.1 Introduction

Women's age at first marriage differ widely across sub-Saharan Africa. While late marriage is most prevalent in Southern African societies (South Africa, Namibia and Bostwana), early marriage is common in other developing countries (Cameroon, Mali, Senegal)[1]. It has proven that early marriage has negative effects on women's socioeconomic outcomes including education, health and fertility (Tertilt 2005, Field and Ambrus 2008, Parsons et al. 2015). To design policies aiming at reducing early marriage, economists have become interested in identifying the determinants of women's age at first marriage, research has focused on the role of economic, demographic, and cultural factors[2]. However, little attention has been given to family composition that plays a tremendous role in early-adulthood outcomes.

This paper tests whether sibling gender affects women's age at first marriage in sub-Saharan Africa, where bride price marriage is prevalent. Having a sister can increase family size and lower family resources, which may accelerate the timing of older sister's marriage. For instance, parents of two first daughters are more likely to go on to have an

1. See UNEP DESA 2014, Singh and Samara 1996, Althaus 1991
2. See Tertilt 2005, Field and Ambrus 2008

additional child, raising the family size, and decreasing the average child investment and age at marriage (Becker and Lewis 1973, Busfield 1972, Bumpass and Mburugu 1977, Tashakkori et al. 1987, Angrist and Evans 1996, Westoff and Potter 2015). In the same way, as bride price is decreasing with bride's age, parents may feel pressured to marry their older daughter early to obtain the resources from daughter's bride price, which they can use in response to the high family size generated by the presence of second daughter. In this case having a younger sister relative to brother accelerates women's marriage. Lastly, in societies with arranged marriage custom, parents tend to marry their daughters in order of birth. The framework predicts in this case that, the presence of younger sister hastens the marriage of older sister, and the presence of older sister delays the marriage of younger sister.

My analysis is based on the IPUMS-international census data from all sub-Saharan countries covering the period 1960-2013. To establish the causal impact of sibling gender, I exploit the random assignment of the second-born's gender, and I compare the age at first marriage for first-born women with a second-born sister to those with a second-born brother[3]. I do not use all siblings because the final sibling gender composition in a household is endogenous, and will produce biased estimates.

The results show consistent evidence that women with second-born sister are significantly more likely to get married earlier than their counterparts with second-born brother. The coefficient estimates are large in magnitude, implying that first-born women with second-born sister get married about a month younger than those with second-born brother. Consistent with early marriage evidence, I also find that having a second-born sister decreases women's likelihood to complete primary school, and raises their levels of fertility. In addition, compared to women with next-youngest brother, those with next-youngest sister live in poorer households, and have less educated husbands. By contrast, I find different effects for men. Having next-youngest sister lower men's age at first marriage, but does not affect their education.

Next, I explore the mechanisms underlying the relationship between sibling gender

3. This empirical approach has been used by Brenoe (2018) who shows that sibling gender composition affects women's occupational choice.

and age at first marriage. The results can be partly explained by family size and marriage payments patterns. Controlling for family size reduces the estimates for sibling gender by one third [4]. In addition, an analysis based on twin births shows that first-born women with younger sister have more siblings than women with younger brother. Taken together, these findings suggest that younger sister increases family size, and decreases households resources, which ultimately accelerate women's marriage.

The effects of sibling gender are stronger for women in countries where bride price custom is the prevailing norm. Whereas, bride price appears to have no effect on women that belong to ethnic groups that do not practice bride price marriage. These results are consistent with the family size and resource constraints patterns. In response to resource constraints generated by large family size, parents with two first daughters, who belong to ethnic groups that practice bride price marriage, are more likely to marry their daughters younger in exchange for money.

This paper contributes to two strands of the literature. First it is directly related to the literature on determinants of marriage. Previous studies have raised the importance of economic factors (Tertilt 2005, Field and Ambrus 2008, Corno, Hildebrandt and Voena 2020). While others emphasized social and cultural factors including arranged marriage (Vogl 2013, Pesando and Abufhele 2019), bride price and dowry (Corno, Hildebrandt and Voena 2020, Ashraf et al. 2018), and patrilineal and matrilineal inheritance norms (Fafchamps and Quisumbing 2007). I contribute to this literature by showing that, sibling gender plays a tremendous role in determining women's age at marriage, specially in countries where bride price is the norm. I show that having two first daughters significantly increases family size, which led to lower family resources and accelerate the marriage of first-born women.

Second, the paper also contributes to the literature on sibling rivalry. Previous works have demonstrated the impact of household composition on childhood and adulthood outcomes, including education (Conley and Glauber 2006, Yu, Su and Chiu 2012, Lei et al. 2017, Hauser and Kuo 1998), health (Garg and Morduch 1996 and 1998, Morduch

4. Given potential endogeneity in these explanatory variable, the estimates from these regressions should be interpreted with caution

2000), fertility (Lyngstad and Prskawetz 2010, Morosow and Kolk 2019, Whitworth and Stephenson 2002), labor market (Cools and Pattacchini 2017, Peter et al 2018, Dammert 2010), gender roles and competitivity (Brenoe 2018, Kornreich et al. 2003, Galambos 2004), and marriage (Vogl 2013). I contribute to the literature by showing how sibling gender affects the timing of marriage of first-born women.

The paper proceeds as follows : Section 3.2 discusses the Conceptual framework ; Section 3.3 describes the data ; Section 3.4 presents the empirical strategy ; Section 3.5 presents the results ; and Section 3.6 concludes.

3.2 Conceptual framework

The most salient theory on the relationship between sibling gender and age at marriage is the concepts of resource dilution and household resource constraints. The resource dilution model posits that family resources are finite, and sibship characteristic shapes the amount of family resources accessible for each child (Downey 2001, Steelman et al. 2002). As suggested by Angrist and Evans 1996, parents of same-sex siblings are more likely to go on to have an additional child, which increases family size. Moreover, according to Becker and Lewis (1973), the number of children influences investment in each child, affecting child's health and education. This hypothesis is supported by empirical evidence (Conley and Glauber 2006, Bhai 2015, Blake 1985). Given that family resources matter for parental and individuals decisions about marriage in sub-Saharan Africa, the resource dilution model should explain the relationships between sibling sex composition and the transition to marriage (Yu, Su and Chiu 2012, Pesando and Abufhele 2019).

Moreover, in sub-Saharan Africa where bride price is prevalent, having a younger sister accelerates marriage of older sister. First, according to the dilution resources models, having a younger sister may increase family size and lower family resources, which may accelerate the timing of older sister's marriage. Second, as bride price is decreasing with a bride's age, parents may feel pressured to marry their older daughter early, to obtain the resources from daughter's bride price. This can help them to response to family resource constraints generated by the presence of second daughter. Lastly, in societies with arranged marriage custom, parents tend to marry their daughters in order of birth. The framework predicts in this case that, the presence of younger sister hastens the marriage of older sister, and the presence of older sister delays the marriage of younger sister.

3.3 Data

To examine how sibling gender composition can affect the timing of marriage, I use the IPUMS-International census data. The IPUMS-international data is composed of microdata, which provides information at the individual-level, and covers several subjects on a range of population characteristics including household composition, nuptiality, education, fertility, labor force participation, and religion. I consider all the available census from sub-Saharan countries over the period 1960-2013. I restrict the sample to first-born women who have at least one younger sibling, and I exclude all women who are twins.

The advantage of the IPUMS census data set over the Demographic and Health Surveys (DHS) is that, in contrast to DHS, the IPUMS census data provide predetermined characteristics of respondents such as parents' age, education, and religion. More importantly, information on siblings in the DHS are measured with errors [5]. Thus using DHS data may produce biased estimates due to measurement errors. Fortunately, the IPUMS census data registers less such drawbacks. However, a disadvantage of the IPUMS census data is the absence of health information such as height and body mass index are not available.

I use the following outcome variables to analyse the effect of sibling gender. The first outcome is age at first marriage, the age at which the respondent first get married formerly or traditionally. The information on age at first marriage is collected retrospectively during the interview. Respondents were asked to recall the month and year when they

5. Some respondents report siblings' information many time, and have difficulties to remember age of their siblings.

were first married. I also examine the effect on education. The IPUMS census data provide information on education attainment, years of schooling, and literacy. The variable *Education attainment*, is ranged between 1-4, with categories "less than primary completed ", "primary completed ", "secondary completed", and "university completed". The variable *Literacy* is a dummy variable for the respondent being literate. The fourth variable is the *Number of children* measured by the number of respondent's children ever born. Lastly, I explore the effect on labor force participation using the variable *FLFP*, which is a dummy for respondent being in the labor force.

Table 3.1 reports descriptive statistics on the outcome variables and on predetermined characteristics for women and men. As expected, women get married younger than men on average. They get married by age of 19 years old, while male marry latter when the aged 23 years. However, there are any differences between women and men in education, fertility and labor force participation. Women have on average completed primary school, and they have an average of fourth years of education. Among women, only 32% have higher levels of literacy skills, and 37% participate in the workplace. On average, age gap to next sibling is 5.5 years, and mothers are more educated than fathers.

3.4 Empirical Strategy

To establish the causal impact of sibling gender composition on women's age at first marriage, I exploit the random assignment of the second-born's gender, and I compare the age at first marriage for first-born women with a second-born sister to those with a second-born brother. [6] I estimate the following equation :

$$Y_i = \alpha + \beta Sister_i + X_i' \delta + \epsilon_i \tag{3.1}$$

where Y_i represents the outcomes for first-born woman i. The explanatory variable $Sister_i$ equals to 1 if the second-born sibling is a sister, and zero if the second-born sibling is a brother. The vector X_i includes respondent's year of birth, age gap between the two first

6. This empirical approach has been used by Brenoe (2018) who shows that sibling gender composition affects women's occupational choice.

sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects and (in some specifications) country of residence fixed effects. ϵ_i is the error term.

The coefficient of interest, β, will capture the causal effect of having a younger sister relative to younger brother conditional on having at least one sibling.

I do not use all siblings because the final sibling gender composition in a household is endogenous, and will produce biased estimates. The identification strategy requires that conditional on the first child's gender, and conditional on having two children, the gender of the second-born child is random.

3.5 Results

3.5.1 The effects on age at marriage

Table 3.2 presents the OLS estimates of the effect of having a next-youngest sister on women's age at marriage. In column (1) I add a vector of individual controls that include respondent's year of birth, age gap between respondent and her next sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. In column (2), I add controls for whether the respondent resides in a rural area. In column (3) I include country of residence fixed effects. The coefficient estimates on younger sister in columns 1-3 are consistently negative and statistically significant. Having a younger sister relative to brother reduces women's age at first marriage. First-born women with a second-born sister get married 0.104 years earlier than first-born women with a second-born brother, representing an average effect of about a month. In column (4), I add control for respondent's family size measured with number of siblings. The inclusion of this control reduces the main effect sizes by one third, the estimates remain significant at 5 percent. Taken together, the results suggest that the effect of younger sister on women's transition to marriage may be partly explained by differential family size patterns, an issue I explore in more detail below.

To further examine the role of family size in driving marriage outcomes, I estimate a series of regressions relying on twin births [7]. Specifically, in columns 4-6 of Table 3.3, I examine the effect of twins on the number of siblings. Column (4) shows that having a younger sister relative to brother increases the number of siblings by 0.136, consistent with the hypothesis that family size may explain the results. The twins analysis confirm this explanation. The estimates in column (5) show that having a twin sisters increases the number of siblings relative to having a singleton sister. Meanwhile, the estimates in column (6) show that younger twins brother decreases the number of siblings.

In addition, columns 1-3 of Table 3.3 show that the estimates in Table 3.2 reflect a positive effect of younger brother on women's age at first marriage. Column (2) shows that there is no substantial differences between women with second-born twin sisters and women with second-born singleton sister. The coefficient is positive but statistically insignificant [8]. In contrast, the estimates in column (3) show that women with second-born twin brothers get married one year earlier than women with second-born singleton brother. Thus, increasing the number of younger brothers delays women's marriage, whereas increasing the number of younger sisters does not.

Taken together, the results confirm that the negative effect of having a second-born sister on women's age at first marriage reflects the positive effect of an additional brother, which ultimately decreases family size. Implying that family size substantially drives the results. These findings contrast with the South Asia context, where family size does not play a role in explaining women's marriage outcomes (e.g., Vogl 2013).

Heterogeneity of the effects by historical mode of marriage

To examine whether sibling's gender has heterogeneous effects by mode of marriage, I exploit historical heterogeneity in marriage payments across different countries and ethnic groups. Data on traditional marriage norms comes from the Atlas of pre-colonial societies (Muller, Marti, Shiedt, and Arpagaus, 2010). This data is an extension of Murdock's Ethnographic Atlas (1967), and has been used in the literature subjects such

7. I make the assumption that twin births are perfectly random in sub-Saharan Africa.
8. The standard errors are large due to the small number of twin girls in the sample.

as marriage and colonialism (Corno, Hildebrandt and Voena 2020, Ziltener and Walter 2017). In addition to ethnographic information, the Atlas of pre-colonial societies provides information on the exact number of population of each ethnic group, as well as the historical mode of marriage payment. Using these information, I build a measure of the prevalence of the bride price marriage at the country level. Bride price is the prevalent norm in a country if more than 50 percent of the population of this country traditionally used bride price custom.

Table 3.4 reports the effect of sibling gender composition by bride price prevalence across countries. The estimates in columns 1-2 are small and statistically insignificant, implying that sibling's gender does not affect women's age at first marriage in countries where the bride price marriage is not the norm. Conversely, having a younger sister lower age at first marriage for women in countries where the bride price custom is the prevailing norm. Thus, in countries where bride price is the traditional mode of marriage, women with younger sister get married younger than women with younger brother, while there is any relationship between sibling's gender and women's age at marriage.

Heterogeneity of the effects by demographic group

Next, I explore the effects of sibling gender composition on women's transition to marriage by several other characteristics. Table 3.5 presents the results by place of residence, age gap to next sibling, and parents' education. The estimates in columns 1-2 show that the effects are stronger for women in rural areas than for their urban counterparts. In addition, Columns 4-5 examine heterogeneity by age gap to next sibling. I find stronger effects of having younger sister when the age gap is high, but the coefficients are statistically insignificant. Finally, column 5-8 report the estimates by parents' years of schooling. The results show that the effects are stronger among women whose parents have above six years of education. Taken together, the findings suggest that the effects may be driven by economic constraints among poor households.

3.5.2 The effects on education, fertility and labor force participation

There are number of channels through which sibling gender composition may affect education. In the standard framework, siblings compete for limited households resources, so that having a sister compared to brother increases the number of children and decreases average child investment (Becker and Lewis 1973, Busfield 1972, Bumpass and Mburugu 1977, Tashakkori et al. 1987, Angrist and Evans 1996, Westoff and Potter 2015). To test for this, I use three measures of education, and I examine the effect of having younger sister on education attainment, years of schooling and literacy.

Table 3.6 reports large negative effects of the effect of having younger sister on education. First-born women with younger sister have lower levels of education attainment, fewer number of years of schooling, and lower levels of literacy. These patterns are consistent with Vogl (2013), who finds that women with younger sisters have significantly lower literacy skills in South-Asia.

Next, I explore whether these findings reflect the marriage age effects. This is the case only if sibling's gender affects education only at the marriage age group. To test for this, I examine the effects more in details by age group and education. Table 3.7 reports the effects of younger sister by group of education. Columns 1-2 present the effects of having a younger sister on the likelihood to attend preschool among girls aged between three and six years old. I find a positive relationship between sibling's gender and the likelihood for girls under age six to attend preschool. Girls with younger sister are 0.7% more likely to attend preschool. Columns 3-4 report the effects on the likelihood to attend primary school among girls in age to attend primary school (6 to 12 years old). The results show any association between having sister and attending primary, the coefficient is zero and is statistically insignificant. In column 5-6, and 7-8, I examine the effect on the likelihood to addend secondary school and college respectively. The results show that having a sister has consistent negative effects on the likelihood to attend secondary and college. The effects on college are stronger than the effects on attending secondary school. These findings suggest that the effect on education is driven by the differential drop out at secondary and college level. Given that girls attend college when they are about 18 years old, which is equivalent to the average age at marriage, I can conclude that the findings for education reflect the marriage age effects.

I investigate the effect on fertility and labor force participation in Table 3.8. Columns 1-2 report the effect on female labor force participation. The estimates are negative and statistically significant. Women with younger sister are 1.2% less likely to participate in the workplace. Columns 3-4 show the results for fertility. Women with younger sister have fewer children. Compared to women with younger brother, those with younger sister have 0.06 fewer children ever born.

3.5.3 Comparison with men

In this section, I examine the effect of sibling's gender on men's outcomes. Table 3.9 presents the effects of having younger sister for both men and women. I report the effects for women in Columns 1-3. As shown before, women with next younger sister get married younger, are less educated, and have lower levels of literacy. I estimate the effects for men in columns 4-6. I find consistent negative effects on men's age at first marriage, the estimates are large and statistically significant at 1% level : Having a younger brother delays men's marriage by 0.155 years. Importantly, the effect for men is stronger than the effect for women.

Columns 5-6 examines the effects on men's education and literacy. The estimates are positive and statistically significant. In contrast to women, men with next younger sister have higher education attainment and levels of literacy skills. These findings suggest that parents may have reallocated resources to first son when the second child is a girl, while they may invest equally between children when the second sibling is a boy.

3.5.4 The effects of having an older sister

This section investigates the effect of older sister. Table 3.10 presents the results for both men and women. Compared to women with older brother, those with older sister get married younger, have more schooling and fewer number of children (cols. 1-4). Taken together with the results of younger sister on women suggest that in households with two first daughters, the eldest sister is substantially worse off : they marry earlier, and are less educated. Whereas, the younger sister benefits more in education. Moreover, the negative effect of older sister on younger sister's timing of marriage is consistent with the dilution resources models, and family size patterns. In addition, the results rule out the hypothesis of order birth marriage custom.

These patterns are different for men. Having an older sister delays men's marriage and increases their education and fertility. Thus, the presence of any sister has a positive impact on men's education, consistent with the siblings competition frameworks.

3.6 Conclusion

This paper examines how gender of younger sibling can affect women's transition to marriage in sub-Saharan Africa. The analysis rely on both information on historical ethnic norms from the Ethnographic Atlas with data from IPUMS-International census data on modern women's outcomes. To establish the causal effect of sibling gender composition on women's age at first marriage, I exploit the random assignment of the second child's gender, and I compare the age at first marriage for women with a second-born sister to those with a second-born brother.

The paper's findings show that women with younger sister get married earlier than their counterparts with younger brother. Moreover, I find that those women are less educated, have lower levels of literacy and higher levels of fertility. The effects are stronger in countries where the bride price marriage is the norm. The results can partly be explained by differential family size patterns, where having a sister increases the number of children, which ultimately reduces family resources and increase the likelihood for first-born women to get married younger.

The paper adds to our understanding of the determinants of age at marriage in sub-Saharan Africa. In addition to establishing the role of sibling's gender, the results demonstrate a tremendous role of the tradition of bride price. Designing policies to delay women's age at first marriage necessitates to understand the combined effects of demographic, economic and cultural factors.

Table 3.1 : Descriptive statistics, First-born Sample

	Women			Men		
	Mean (1)	Std. Dev. (2)	Obs. (3)	Mean (4)	Std. Dev. (5)	Obs. (6)
Outcome variables						
Age First Marriage	19.30	5.59	72,500	22.64	6.49	42,586
Education Attainment	1.42	0.64	4,978,230	1.41	0.64	5,748,727
Years of schooling	4.15	4.20	4,358,541	4.17	4.18	4,915,228
Literacy	0.32	0.47	5,007,952	0.36	0.48	5,789,775
Number of Children	0.22	0.78	6,191,927	0.05	0.42	7,151,025
Labor Force Participation	0.37	0.48	3,191,506	0.43	0.50	3,876,691
Respondent characteristics						
Age	13.94	10.30	6,542,339	14.42	9.75	7,502,318
Age gap	5.49	5.82	6,357,289	5.45	5.70	7,352,922
Number of Siblings	1.88	1.95	6,555,882	2.08	2.09	7,517,386
Rural	0.69	0.46	5,904,350	0.72	0.45	6,771,362
Muslim	0.12	0.33	6,525,838	0.15	0.35	7,487,617
Christian	0.38	0.49	6,525,838	0.36	0.48	7,487,617
Mother Educattion	1.21	0.84	6,552,259	1.18	0.83	7,512,220
Father Educattion	0.97	0.96	6,555,718	0.97	0.93	7,517,206
Mother Muslim	0.11	0.31	6,502,362	0.13	0.33	7,422,761
Mother Christian	0.36	0.48	6,536,971	0.34	0.47	7,491,179
Father Muslim	0.10	0.30	6,555,820	0.12	0.33	7,517,283
Father Christian	0.27	0.44	6,555,527	0.26	0.44	7,516,856

Figure 3.1 : Age at First Marriage

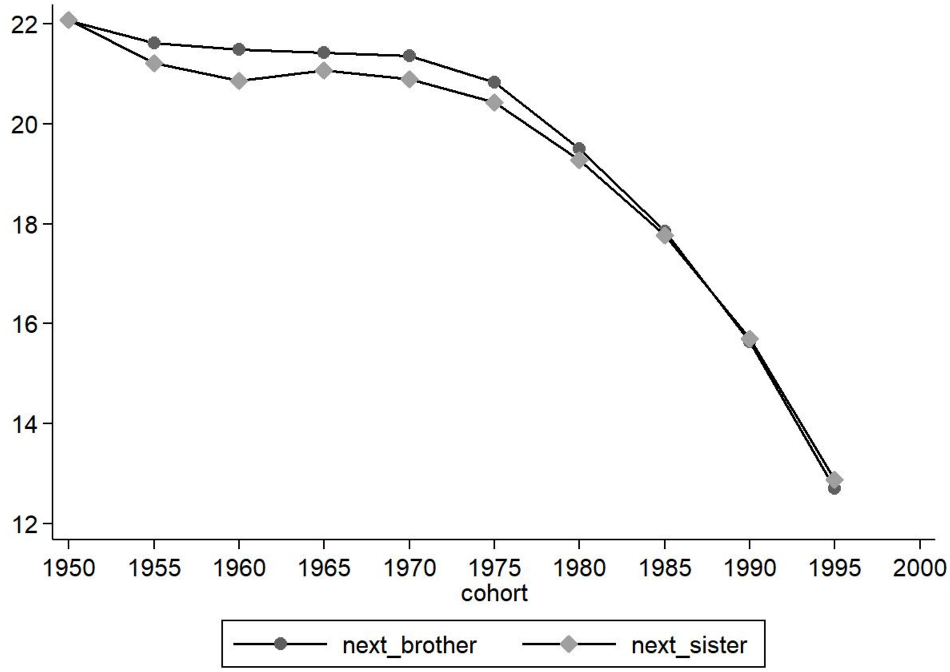


Figure 3.2 : Education Attainment

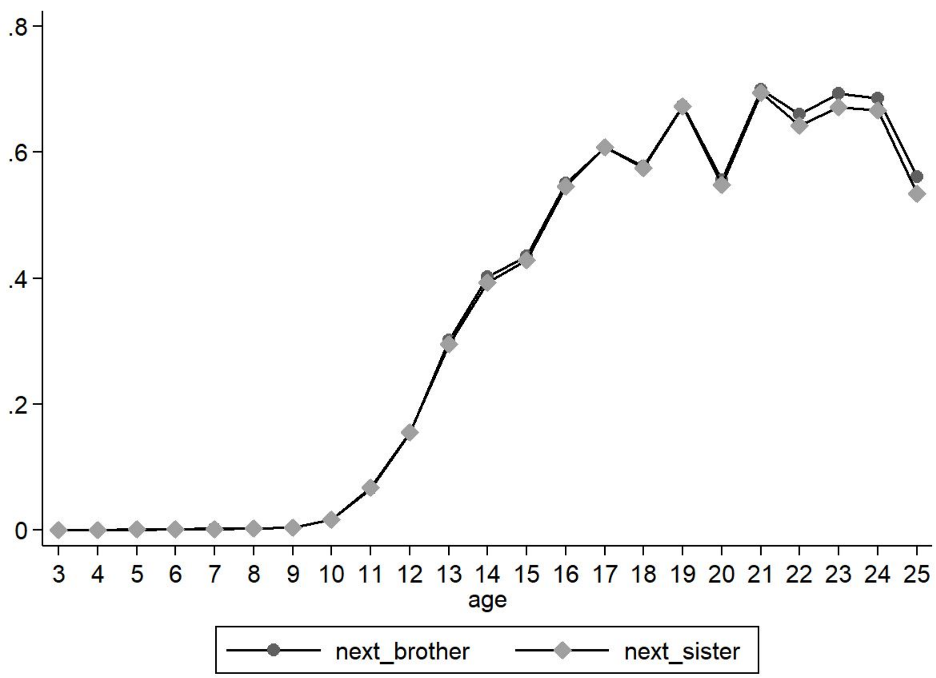


Table 3.2 : Younger sister effects on women age at marriage

	(1) Age First Marriage	(2) Age First Marriage	(3) Age First Marriage	(4) Age First Marriage
Younger sister	-0.104***	-0.097**	-0.089**	-0.075*
	(0.037)	(0.042)	(0.042)	(0.042)
Rural		-0.741***	-0.754***	-0.726***
		(0.055)	(0.055)	(0.055)
Baseline Controls	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes
Country FE	No	No	Yes	Yes
Family Size	No	No	No	Yes
Observations	63,621	54,896	54,896	54,896
R-squared	0.216	0.226	0.233	0.235
Dep. var. mean unaffected	18.97	19.05	19.05	19.05

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.3 : Effects on family size

	(1) Age First Marriage	(2) Age First Marriage	(3) Age First Marriage	(4) Number Siblings	(5) Number Siblings	(6) Number Siblings
Younger sister	-0.097**			0.136***		
	(0.042)			(0.019)		
Younger sister is a twin relative to singleton sister		0.280			0.940***	
		(0.196)			(0.123)	
Younger brother is a twin relative to singleton sister			1.013***			-1.752***
			(0.107)			(0.056)
Observations	54,896	23,579	31,317	54,896	23,579	31,317
R-squared	0.226	0.241	0.217	0.247	0.087	0.383

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.4 : Younger sister effects by marriage payment

	Bride price≤50		Bride price≥50	
	(1) Age First Marriage	(2) Age First Marriage	(3) Age First Marriage	(4) Age First Marriage
Younger sister	0.023	0.035	-0.476***	-0.432***
	(0.051)	(0.050)	(0.046)	(0.046)
Country FE	No	Yes	No	Yes
Observations	17,920	17,920	58,449	58,449
R-squared	0.149	0.158	0.336	0.346

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.5 : Younger sister effects by demographic groups

	Place of residence		Age gap to next sibling		Father years education		Mother years education	
	(1) Urban	(2) Rural	(3) ≤5	(4) >5	(5) ≤6	(6) >6	(7) ≤6	(8) >6
Younger sister	-0.132	-0.079*	0.001	-0.055	-0.113**	0.007	-0.120***	0.092
	(0.104)	(0.044)	(0.051)	(0.095)	(0.044)	(0.126)	(0.044)	(0.137)
Observations	11,620	43,276	26,086	9,661	49,507	5,425	50,315	4,617
R-squared	0.220	0.210	0.309	0.221	0.218	0.243	0.216	0.284

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.6 : Younger sister effects on education

	(1) Education attainment	(2) Education attainment	(3) Years schooling	(4) Years schooling	(5) Literacy	(6) Literacy
Younger sister	-0.007***	-0.007***	-0.006*	-0.027***	-0.008***	-0.011***
	(0.0005)	(0.0005)	(0.003)	(0.003)	(0.0003)	(0.0003)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Survey year FE	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	No	Yes	No	Yes	No	Yes
Observations	4,286,126	4,286,126	3,754,811	3,754,811	4,241,623	4,241,623
R-squared	0.382	0.394	0.460	0.461	0.447	0.448
Dep. var. mean unaffected	1.406	1.406	5.112	5.112	0.549	0.549

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.7 : Younger sister effects by level of education

	Preschool		Primary		Secondary		University	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Younger sister	0.007***	0.006***	-0.0008	-0.0001	-0.003***	-0.004***	-0.006***	-0.0056***
	(0.001)	(0.001)	(0.001)	(0.0006)	(0.001)	(0.001)	(0.001)	(0.001)
Country FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	444,258	444,258	1,105,049	2,243,155	1,071,514	1,071,514	1,133,990	1,133,990
R-squared	0.151	0.182	0.253	0.269	0.300	0.333	0.250	0.261

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include respondent's year of birth, age gap between the two first sibling, mother and father education attainment, a dummy for mother and father being Muslim or Christian, and survey year fixed effects. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.8 : Younger sister effects on labor and fertility

	. FLFP (1)	FLFP (2)	Number of children (3)	Number of children (4)
Younger sister	-0.012***	-0.013***	-0.057***	-0.057***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Country FE	No	Yes	No	Yes
Observations	2,713,817	2,713,817	5,387,845	5,387,845
R-squared	0.306	0.318	0.212	0.218

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include number of sibling, number of sibling squared, and a set of fixed effects for years of birth, age gap, years of survey, a dummy for the respondent being Christian, and a dummy for a respondent being Muslim. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.9 : Younger sister effects on women versus men

	Women			Men		
	(1) Age First Marriage	(2) Education Attainment	(3) Literacy	(4) Age First Marriage	(5) Education Attainment	(6) Literacy
Younger sister	-0.097**	-0.007***	-0.005***	-0.155***	0.006***	0.0052***
	(0.041)	(0.0005)	(0.0004)	(0.057)	(0.001)	(0.00003)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	54,896	4,286,126	4,241,623	34,864	4,969,506	4,932,008
R-squared	0.226	0.382	0.387	0.378	0.359	0.358

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include number of sibling, number of sibling squared, and a set of fixed effects for years of birth, age gap, years of survey, a dummy for the respondent being Christian, and a dummy for a respondent being Muslim. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Table 3.10 : Older sister effects on women versus men

	Women				Men			
	(1) Age First Marriage	(2) Education attainment	(3) Literacy	(4) Number of Children	(5) Age First Marriage	(6) Education attainment	(7) Literacy	(8) Number of Children
Older sister	-0.075	0.008***	0.007***	-0.013***	0.293***	0.023***	0.017***	0.003***
	(0.059)	(0.0005)	(0.0004)	(0.0003)	(0.084)	(0.0005)	(0.0004)	(0.0002)
Baseline Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,322	2,969,873	2,858,616	4,189,817	12,509	3,365,128	3,243,204	4,674,860
R-squared	0.329	0.431	0.427	0.173	0.463	0.403	0.407	0.040

Notes. OLS estimates. Robust standard errors in parentheses. Data are from the IPUMS-International census data, and include individuals evaluated between 1960 and 2013 with at least one sibling. Twins and women with no younger siblings are excluded. *Younger sister* is a dummy variable, equals one if respondent's next-younger sibling is a girl, zero otherwise. The baseline controls include number of sibling, number of sibling squared, and a set of fixed effects for years of birth, age gap, years of survey, a dummy for the respondent being Christian, and a dummy for a respondent being Muslim. ***, **, * indicate the level of significance at the 1, 5 and 10 percent level.

Bibliographie

Aberle, D. F. (1961). Matrilineal descent in cross-cultural comparison. *Matrilineal kinship*, 655–730.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The colonial origins of comparative development : An empirical investigation. *American economic review*, *91*(5), 1369–1401.

Acemoglu, D., Johnson, S., & Robinson, J. A. (2012). The colonial origins of comparative development : An empirical investigation : Reply. *American Economic Review*, *102*(6), 3077–3110.

Albanesi, S., & Olivetti, C. (2007). *Gender roles and technological progress.* National Bureau of Economic Research Cambridge, Mass., USA.

Alesina, A., Brioschi, B., & Ferrara, E. L. (2016). *Violence against women : A cross-cultural analysis for africa* (Rapport technique). National Bureau of Economic Research.

Alesina, A., Giuliano, P., & Nunn, N. (2011). Fertility and the plough. *American Economic Review*, *101*(3), 499–503.

Alesina, A., Giuliano, P., & Nunn, N. (2013). On the origins of gender roles : Women and the plough. *The Quarterly Journal of Economics*, *128*(2), 469–530.

Alsan, M. (2015). The effect of the tsetse fly on african development. *American Economic Review*, *105*(1), 382–410.

Althaus, F. (1991). Early marriage and childbearing remain common in much of developing world, even as fertility falls. *International Family Planning Perspectives*, 119–120.

Andersen, S., Ertac, S., Gneezy, U., List, J. A., & Maximiano, S. (2013). Gender, competitiveness, and socialization at a young age : Evidence from a matrilineal and a patriarchal society. *Review of Economics and Statistics*, *95*(4), 1438–1443.

Anderson, S. (2007). The economics of dowry and brideprice. *Journal of Economic Perspectives*, *21*(4), 151–174.

Angrist, J. D., & Evans, W. N. (1996). *Children and their parents' labor supply : Evidence from exogenous variation in family size* (Rapport technique). National bureau of economic research.

Arnold, F., Choe, M. K., & Roy, T. K. (1998). Son preference, the family-building process and child mortality in india. *Population studies*, *52*(3), 301–315.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2016). Bride price and female education.

Ashraf, N., Bau, N., Nunn, N., & Voena, A. (2018, 2019). Bride price and female education. *Journal of Political Economy*.

Bailey, D. W., Gross, J. E., Laca, E. A., Rittenhouse, L. R., Coughenour, M. B., Swift, D. M., & Sims, P. L. (1996). Mechanisms that result in large herbivore grazing distribution patterns. *Rangeland Ecology & Management/Journal of Range Management Archives*, *49*(5), 386–400.

Bau, N. (2016). *Can policy crowd out culture ?* (Rapport technique). working paper.

Becker, A. (2018). On the economic origins of female genital cutting. *Available at SSRN*.

Becker, G. S. (1974). A theory of marriage : Part ii. *Journal of political Economy*, *82*(2, Part 2), S11–S26.

Becker, G. S., & Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, *81*(2, Part 2), S279–S288.

Besley, T., & Persson, T. (2010). State capacity, conflict, and development. *Econometrica*, *78*(1), 1–34.

Bhai, M. (2015). Brothers are better : The effect of sibling sex composition on women's schooling, health, earnings, and labor supply. *Health, Earnings, and Labor Supply (February 23, 2015)*.

Blake, J. (1985). Number of siblings and educational mobility. *American Sociological Review*, 84–94.

Blanchard, O., & Lawrence, F. (1992). und l. katz (1992),regional evolutions. *Brookings Papers on Economic Activity*, *1992*(1), 1–61.

Boserup, E., Tan, S. F., & Toulmin, C. (1970). *Woman's role in economic development.* Routledge.

Botte, R. (1985). Rwanda and burundi, 1889-1930 : chronology of a slow assassination, part 1. *The International Journal of African Historical Studies*, *18*(1), 53–91.

Boustan, L. P., Kahn, M. E., & Rhode, P. W. (2012). Moving to higher ground : Migration response to natural disasters in the early twentieth century. *American Economic Review*, *102*(3), 238–44.

Boyd, R., & Richerson, P. J. (1988). *Culture and the evolutionary process.* University of Chicago press.

Brenøe, A. (2018). Origins of gender norms : sibling gender composition and women's choice of occupation and partner. *University of Zurich, Department of Economics, Working Paper*(294).

Bumpass, L. L., & Mburugu, E. K. (1977). Age at marriage and completed family size. *Social Biology*, *24*(1), 31–37.

Burchell, W. J. (1822). *Travels in the interior of southern africa* (Vol. 1). .

Busfield, J. (1972). Age at marriage and family size : Social causation and social selection hypotheses. *Journal of Biosocial Science*, *4*(1), 117–134.

Cáceres, S. B. (2011). The long journey of cattle plague. *The Canadian Veterinary Journal*, *52*(10), 1140.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, *29*(2), 238–249.

Chakraborty, T., & Kim, S. (2010). Kinship institutions and sex ratios in india. *Demography*, *47*(4), 989–1012.

Comaroff, J. L. (1990). Goodly beasts, beastly goods : cattle and commodities in a south african context. *American ethnologist*, *17*(2), 195–216.

Conley, D., & Glauber, R. (2006). Parental educational investment and childrens academic risk estimates of the impact of sibship size and birth order from exogenous variation in fertility. *Journal of human resources*, *41*(4), 722–737.

Conley, T. G. (1999). Gmm estimation with cross sectional dependence. *Journal of econometrics*, *92*(1), 1–45.

Cools, A., & Patacchini, E. (2017). Sibling gender composition and women's wages.

Corno, L., Hildebrandt, N., & Voena, A. (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, *88*(3), 879–915.

Crouzet, J. (1889). Les missions catholiques. , *21*, 62–63.

Dalton, J. T., & Leung, T. C. (2014). Why is polygyny more prevalent in western africa ? an african slave trade perspective. *Economic Development and Cultural Change*, *62*(4), 599–632.

Dammert, A. C. (2010). Siblings, child labor, and schooling in nicaragua and guatemala. *Journal of Population Economics*, *23*(1), 199–224.

Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., & Hwa-Ok, B. (2003). Why is son preference so persistent in east and south asia ? a cross-country study of china, india and the republic of korea. *The Journal of Development Studies*, *40*(2), 153–187.

Deere, C. D., & Doss, C. R. (2006). The gender asset gap : What do we know and why does it matter ? *Feminist economics*, *12*(1-2), 1–50.

Deininger, K., Goyal, A., & Nagarajan, H. (2013). Women's inheritance rights and intergenerational transmission of resources in india. *Journal of Human Resources*, *48*(1), 114–141.

DESAP, U. (2014). World urbanisation prospects : the 2014 revision, highlights (st/esa/ser. a/352). department of economic and social affairs. *Population Division, New York : United Nations*, *16*.

Deshler, W. (1963). Cattle in africa : Distribution, types and problems. *Geographical Review*.

Diamond, J. (1997). Guns. *Germs and Steel : The Fates of Human Societies, Vintage*, 186.

Dixon-Fyle, M. (1983). Reflections on economic and social change among the plateau tonga of northern rhodesia, c. 1890-1935. *The International Journal of African Historical Studies*, *16*(3), 423–439.

Downey, D. B. (2001). Number of siblings and intellectual development : The resource dilution explanation. *American psychologist*, *56*(6-7), 497.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic literature*, *50*(4), 1051–79.

Dupire, M. (1963). The position of women in a pastoral society. *Women of tropical Africa*, 47–92.

Dyson, T., & Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in india. *Population and development review*, 35–60.

Ebenstein, A. (2014). Patrilocality and missing women. *Available at SSRN 2422090*.

Edington, A. (1896). Rinderpest. *Agricultural Journal*, *9*, 253–254.

Edlund, L., & Ku, H. (2011). The african slave trade and the curious case of general polygyny.

Engels, F. (1884). *The origin of the family, private property and the state.* Foreign Languages Publishing House Moscow.

Epstein, H., & Mason, I. (1984). Cattle. in : (ed. mason il) evolution of domesticated animals. *Longman, New York.*

Fafchamps, M., & Quisumbing, A. R. (2007). Household formation and marriage markets in rural areas. *Handbook of development economics*, *4*, 3187–3247.

Feir, D., Gillezeau, R., Jones, M., et al. (2019). *The slaughter of the bison and reversal of fortunes on the great plains* (Rapport technique). Federal Reserve Bank of Minneapolis.

Fenske, J. (2014). Ecology, trade, and states in pre-colonial africa. *Journal of the European Economic Association*, *12*(3), 612–640.

Fenske, J. (2015). African polygamy : Past and present. *Journal of Development Economics*, *117*, 58–73.

Fenske, J., & Kala, N. (2015). Climate and the slave trade. *Journal of Development Economics*, *112*, 19–32.

Fernand, R. (1906). *l expansion coloniale au congo français.* Paris, Larose.

Field, E., & Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in bangladesh. *Journal of political Economy*, *116*(5), 881–930.

Fischer, G., Nachtergaele, F. O., Prieler, S., Teixeira, E., Tóth, G., Van Velthuizen, H., . . . Wiberg, D. (2012). Global agro-ecological zones (gaez v3. 0)-model documentation.

Galambos, N. L. (2004). Gender and gender role development in adolescence.

Ganskopp, D., George, M., Bailey, D., Borman, M., Surber, G., & Harris, N. (2007). Factors and practices that influence livestock distribution. *University of California Division of Agriculture and Natural Resources*, *8217*, 20.

Garg, A., & Morduch, J. (1998). Sibling rivalry and the gender gap : Evidence from child health outcomes in ghana. *Journal of Population Economics*, *11*(4), 471–493.

Garg, A., Morduch, J., et al. (1996). *Sibling rivalry, resource constraints, and the health of children*. Harvard Institute of Economic Research.

Gitungwa, H. (2018). The relationship of male and female pastoralist income with household food security and nutrition status in tanzania : Maasai, sukuma, and barabaig ethnic groups.

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition : Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637–1664.

Goldin, C. (1994). *The u-shaped female labor force function in economic development and economic history* (Rapport technique). National Bureau of Economic Research.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *American economic review*, *96*(2), 1–21.

Goldin, C., & Katz, L. F. (2002). The power of the pill : Oral contraceptives and womens career and marriage decisions. *Journal of political Economy*, *110*(4), 730–770.

Goldschmidt, W. (1974). The economics of brideprice among the sebei and in east africa. *Ethnology*, *13*(4), 311–331.

Graham, J. P., Hirai, M., & Kim, S.-S. (2016). An analysis of water collection labor among women and children in 24 sub-saharan african countries. *PloS one*, *11*(6), e0155981.

Green, E. (2013). Production systems in pre-colonial africa. *The History of African Development. African Economic History Network. Online at https ://www. aehnetwork. org/textbook/. Accessed*, *6*(27), 2017.

Greenwood, J., Seshadri, A., & Yorukoglu, M. (2005). Engines of liberation. *The Review of Economic Studies*, *72*(1), 109–133.

Grossbard, A. (1976). An economic analysis of polygyny : The case of maiduguri. *Current anthropology*, *17*(4), 701–707.

Hanotte, O., Tawah, C., Bradley, D., Okomo, M., Verjee, Y., Ochieng, J., & Rege, J. (2000). Geographic distribution and frequency of a taurine bos taurus and an indicine bos indicus y specific allele amongst sub-saharan african cattle breeds. *Molecular ecology*, *9*(4), 387–396.

Hansen, C. W., Jensen, P. S., & Skovsgaard, C. V. (2015). Modern gender roles and agricultural history : the neolithic inheritance. *Journal of Economic Growth*, *20*(4), 365–404.

Hauser, R. M., & Kuo, H.-H. D. (1998). Does the gender composition of sibships affect women's educational attainment ? *Journal of Human Resources*, 644–657.

Henderson, M., & Whatley, W. (2014). Pacification and gender in colonial africa : evidence from the ethnographic atlas.

Holden, C. J., & Mace, R. (2003). Spread of cattle led to the loss of matrilineal descent in africa : a coevolutionary analysis. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, *270*(1532), 2425–2433.

Hornbeck, R. (2012). The enduring impact of the american dust bowl : Short-and long-run adjustments to environmental catastrophe. *American Economic Review*, *102*(4), 1477–1507.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, *7*(1), 63–88.

Joshi, McLaughlin, & Phillips. (1957). Types and breeds of african cattle. *FAO Agricultural Studies*(37).

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. E., & Sachs, J. (2004). A global index representing the stability of malaria transmission. *The American journal of tropical medicine and hygiene*, *70*(5), 486–498.

Kornreich, J. L., Hearn, K. D., Rodriguez, G., & O'Sullivan, L. F. (2003). Sibling influence, gender roles, and the sexual socialization of urban early adolescent girls. *Journal of Sex Research*, *40*(1), 101–110.

La Porta, R., Lopez-de Silanes, F., & Shleifer, A. (2008). The economic consequences of legal origins. *Journal of economic literature*, *46*(2), 285–332.

Lei, X., Shen, Y., Smith, J. P., & Zhou, G. (2017). Sibling gender compositions effect on education : evidence from china. *Journal of population economics*, *30*(2), 569–590.

Lewis, J. (2018). Infant health, women's fertility, and rural electrification in the united states, 1930–1960. *The Journal of Economic History*, *78*(1), 118–154.

Lowes, S. (2017). Matrilineal kinship and spousal cooperation : evidence from the matrilineal belt. *Unpublished manuscript. URL : https ://scholar. harvard. edu/files/slowes/files/lowes_matrilineal. pdf*.

Lowes, S. (2018). *Matrilineal kinship and gender differences in competition* (Rapport technique). Working paper, Harvard University. Available online : https ://scholar .

Lowes, S., & Nunn, N. (2017). Bride price and the wellbeing of women. *Towards Gender Equity in Development*, 117.

Lowes, S. R., & Montero, E. (2018). The legacy of colonial medicine in central africa.

Lugard, L. F. (1893). *The rise of our east african empire. early efforts in nyasaland and uganda, 2 vols.* William Blackwood and Sons, London.

Lundberg, S., & Pollak, R. A. (1993). Separate spheres bargaining and the marriage market. *Journal of political Economy*, *101*(6), 988–1010.

Lyngstad, T. H., & Prskawetz, A. (2010). Do siblings fertility decisions influence each other ? *Demography*, *47*(4), 923–934.

Mann, M. E., & Jones, P. D. (2003). Global surface temperatures over the past two millennia. *Geophysical Research Letters*, *30*(15).

Manser, M., & Brown, M. (1980). Marriage and household decision-making : A bargaining analysis. *International economic review*, 31–44.

Marquardt, G. (2005). Water, wood and wild animal populations : seeing the spread of rinderpest through the physical environment in bechuanaland, 1896. *South African Historical Journal*, *53*(1), 73–98.

Marshall, F., & Hildebrand, E. (2002). Cattle before crops : the beginnings of food production in africa. *Journal of World Prehistory*, *16*(2), 99–143.

Mattioli, R. C., Pandey, V. S., Murray, M., & Fitzpatrick, J. L. (2000). Immunogenetic influences on tick resistance in african cattle with particular reference to trypanotolerant ndama (bos taurus) and trypanosusceptible gobra zebu (bos indicus) cattle. *Acta Tropica*, *75*(3), 263–277.

McElroy, M. B., & Horney, M. J. (1981). Nash-bargained household decisions : Toward a generalization of the theory of demand. *International economic review*, 333–349.

Michalopoulos, S., & Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, *81*(1), 113–152.

Michalopoulos, S., & Papaioannou, E. (2014). National institutions and subnational development in africa. *The Quarterly journal of economics*, *129*(1), 151–213.

Michalopoulos, S., Putterman, L., & Weil, D. N. (2016). *The influence of ancestral lifeways on individual economic outcomes in sub-saharan africa* (Rapport technique). National Bureau of Economic Research.

Miller, G. (2009). Contraception as development ? new evidence from family planning in colombia. *The Economic Journal*, *120*(545), 709–736.

Monteil, P. (1895). *De saint-louis a tripoli par le lac tchad. voyage au travers du soudan et du sahara accompli pendant les annees 1890-91-92.* Felix Alcan, Paris.

Morduch, J. (2000). Sibling rivalry in africa. *American Economic Review*, *90*(2), 405–409.

Morens, D. M., Holmes, E. C., Davis, A. S., & Taubenberger, J. K. (2011). Global rinderpest eradication : lessons learned and why humans should celebrate too. *Journal of Infectious Diseases*, *204*(4), 502–505.

Morgan. (1871). Systems of consanguinity and affinity of the human family. *Washington.*

Morosow, K., & Kolk, M. (2019). How does birth order and number of siblings affect fertility ? a within-family comparison using swedish register data. *European Journal of Population*, 1–37.

Mukasa-Mugerewa, E. (1989). *A review of reproductive performance of female bos indicus (zebu) cattle. international livestock centre for africa (ilca), monograph, addis ababa, ethiopia.* Ethiopia.

Muller, H.-P., Marti, C., Schiedt, E., & Arpagaus, B. (2010). *Atlas vorkolonialer gesellschaften (atlas of pre-colonial societies) : Cultural heritage and social structures in african, asian and melanesian states.* Berlin : Dietrich Reimer Verlag GmbH.

Murdock, G. P. (1967). Ethnographic atlas : a summary. *Ethnology*, *6*(2), 109–236.

Murdock, G. P., et al. (1959). Africa : its peoples and their culture history.

Murdock, G. P., & White, D. R. (1969). Standard cross-cultural sample. *Ethnology*, *8*(4), 329–369.

Murray, M., Morrison, W., & Whitelaw, D. (1982). Host susceptibility to african trypanosomiasis : trypanotolerance. In *Advances in parasitology* (Vol. 21, pp. 1–68). Elsevier.

Nicholson, S. E., Klotter, D., & Dezfuli, A. K. (2012). Spatial reconstruction of semi-quantitative precipitation fields over africa during the nineteenth century from documentary evidence and gauge data. *Quaternary Research*, *78*(1), 13–23.

Nunn, N. (2008). The long-term effects of africa's slave trades. *The Quarterly Journal of Economics*, *123*(1), 139–176.

Nunn, N. (2012). Culture and the historical process. *Economic History of Developing Regions*, *27*(sup1), S108–S126.

Nunn, N., & Wantchekon, L. (2011). The slave trade and the origins of mistrust in africa. *American Economic Review*, *101*(7), 3221–52.

Oster, E. (2019). Unobservable selection and coefficient stability : Theory and evidence. *Journal of Business & Economic Statistics*, *37*(2), 187–204.

Owens, M. K., Launchbaugh, K. L., & Holloway, J. (1991). Pasture characteristics affecting spatial distribution of utilization by cattle in mixed brush communities. *Journal of Range Management*, 118–123.

Pandey, V. (2007). *Analysis and modeling of cattle distribution in complex agro-ecosystems of south florida* (Thèse de doctorat non publiée). University of Florida.

Parsons, J., Edmeades, J., Kes, A., Petroni, S., Sexton, M., & Wodon, Q. (2015). Economic impacts of child marriage : a review of the literature. *The Review of Faith & International Affairs*, *13*(3), 12–22.

Pesando, L. M., & Abufhele, A. (2019). Household determinants of teen marriage : Sister effects across four low-and middle-income countries. *Studies in family planning*, *50*(2), 113–136.

Peter, N., Lundborg, P., Mikkelsen, S., & Webbink, D. (2018). The effect of a siblings gender on earnings and family formation. *Labour Economics*, *54*, 61–78.

Phoofolo, P. (1993). Epidemics and revolutions : the rinderpest epidemic in late nineteenth-century southern africa. *Past & Present*(138), 112–143.

Pribyl, K., Nash, D. J., Klein, J., & Endfield, G. H. (2019). The role of drought in agrarian crisis and social change : The famine of the 1890s in south-eastern africa. *Regional Environmental Change*, *19*(8), 2683–2695.

Pryor, F. L. (1985). The invention of the plow. *Comparative Studies in Society and history*, *27*(4), 727–743.

Quisumbing, A. R., Meinzen-Dick, R., Raney, T. L., Croppenstedt, A., Behrman, J. A., & Peterman, A. (2014). Closing the knowledge gap on gender in agriculture. In *Gender in agriculture* (pp. 3–27). Springer.

Refik-Bey, & Refik-Bey. (1899). La peste bovine en turquie. *Ann. Inst. Past.*, 596–611.

Ross, M. L. (2008). Oil, islam, and women. *American political science review*, *102*(1), 107–123.

Schneider, H. K. (1957). The subsistence role of cattle among the pakot and in east africa. *American Anthropologist*, *59*(2), 278–300.

Schneider, H. K. (1964). Economics in east african aboriginal societies. *Economic transition in Africa*, 108–123.

Senft, R., Rittenhouse, L., & Woodmansee, R. (1985). Factors influencing patterns of cattle grazing behavior on shortgrass steppe. *Journal of Range Management*, 82–87.

Singh, S., & Samara, R. (1996). Early marriage among women in developing countries. *International family planning perspectives*, 148–175.

Spinage, C. A. (2003). *Cattle plague : a history.* Springer Science & Business Media.

Steelman, L. C., Powell, B., Werum, R., & Carter, S. (2002). Reconsidering the effects of sibling configuration : Recent advances and challenges. *Annual review of sociology*, *28*(1), 243–269.

Sunseri, T. (2018). The african rinderpest panzootic, 1888–1897. In *Oxford research encyclopedia of african history.*

Talba, L. (2019). The historical determinants of gender roles in sub-saharan africa. *Working paper*.

Tashakkori, A., Thompson, V. D., & Mehryar, A. H. (1987). Iranian adolescents' intended age of marriage and desired family size. *Journal of Marriage and the Family*, 917–927.

Tertilt, M. (2005). Polygyny, fertility, and savings. *Journal of Political Economy*, *113*(6), 1341–1371.

Teso, E. (2018). The long-term effect of demographic shocks on the evolution of gender roles : Evidence from the transatlantic slave trade. *Journal of the European Economic Association*, *17*(2), 497–534.

Trivers, R. (1972). *Parental investment and sexual selection* (Vol. 136). Biological Laboratories, Harvard University Cambridge, MA.

Van de Walle, N. (2009). The institutional origins of inequality in sub-saharan africa. *Annual Review of Political Science*, *12*, 307–327.

Van Onselen, C. (1972). Reactions to rinderpest in southern africa 1896-97. *Journal of African history*, 473–488.

Vogl, T. S. (2013). Marriage institutions and sibling competition : Evidence from south asia. *The Quarterly journal of economics*, *128*(3), 1017–1072.

Voigtländer, N., & Voth, H.-J. (2013). How the west" invented" fertility restriction. *American Economic Review*, *103*(6), 2227–64.

Vroklage, B. (1952). Bride price or dower. *Anthropos*(H. 1./2), 133–146.

Westoff, C. F., & Potter, R. G. (2015). *Third child : a study in the prediction of fertility.* Princeton University Press.

White, B. (2015). *World development report 2015 : mind, society, and behavior, by the world bank group.* Taylor & Francis.

Whitworth, A., & Stephenson, R. (2002). Birth spacing, sibling rivalry and child mortality in india. *Social science & medicine*, *55*(12), 2107–2119.

Xia, X. (1992). Uncertainty of paternity can select against paternal care. *The American Naturalist*, *139*(5), 1126–1129.

Yu, W.-h., Su, K.-h., & Chiu, C.-T. (2012). Sibship characteristics and transition to first marriage in taiwan : Explaining gender asymmetries. *Population Research and Policy Review*, *31*(4), 609–636.

Ziltener, P., Künzler, D., & Walter, A. (2017). Research note : Measuring the impacts of colonialism : A new data set for the countries of africa and asia. *Journal of World-Systems Research*, *23*(1), 156–190.

Zonchello, A. (1917). Ii gulhai nel sahel. *Clinical Vet*, *30*, 113–132.