Losing my Religion (or Maybe not): Religion and Fertility Patterns in Africa

Phoebe W. Ishak[*]
Aix-Marseille School of Economics

Mark Gradstein
Ben Gurion University

November 2022[+]

Abstract

How much does religion shape fertility in Africa? We document significant differences in fertility between Christians and Muslims in Africa, with the latter group having on average larger families than the former. Focusing on religious groups belonging to the same ethnicity, and employing ancient trade routes as an IV for the historical spread of Islam in Africa, we argue that historically determined religious affiliation has a persistent effect on contemporary fertility, as well as on additional associated traits pertaining to gender equality. We show that the effect of religion on fertility is largely explained by the position of women within households. In particular, the effect of religion tends to dissipate once we take into account women's age at first marriage and birth, and degree of empowerment, suggesting that women's socioeconomic status overruns the effect of religion. We also document that living in Muslim-majority communities increases fertility indicating the presence of spatial sorting, while the level of religiosity and ancient ethnic characteristics play a limited role.

Keywords: Religion; fertility; Africa

[*]Contact details: phoebe.ishak@univ-amu.fr (corresponding author); grade@bgu.ac.il

[+]We thank Nicolas Berman, Lorenzo Rotunno, and Yannick Dupraz for their helpful comments. We also thank Stelios Michalopoulos for sharing his data.

1 Introduction

> Islam is not only a religion: it is a comprehensive way of life, catering for all the fields of human existence… provides guidance for all aspects of life - individual and social, material and moral, economic and political, legal and cultural, national and international
>
> *UNESCO General History of Africa Vol III, 20*

Africa is projected to account for most of global population growth in the near future, and it is expected to triple by the year 2060 (United Nations, Population Division, 2019). This is primarily due to exceptionally high fertility rates, with five and more children per household being quite common – which stand in stark contrast to the fertility decline experienced in developed countries. Africa is also home to two of the major world religions, Christians and Muslims, whose fertility rates have differed, the latter religious group exhibiting a considerably higher fertility rate – and hence, population growth – than the former, in Africa and elsewhere (Hackett et al., 2015). But little is known about the role played by religion in affecting fertility, espicially if it constitutes an inextricable part of African society. In particular, religious cleavages often times are significant in Africa, shaping political identities and affecting violent conflicts, making the demographics of religious groups is of particular interest. Further, religious doctrines tend to take ideological stands on the issue of fertility – either explicitly through encouraging it, albeit differing in the strength of this position – or implicitly by affecting its driving factors such economic growth, girls education, and age of marriage (Barro and McClearly, 2003; McCauly, 2014; Platas, 2018).[1]

Therefore, in this paper, we undertake to explore the relationship between religion and fertility in Africa. Utilizing the DHS (Demographic and Health Survey) dataset, we begin by documenting a robust correlation between belonging to a religious group – being a Muslim or a Christian – and fertility in a sample of 24 African countries. In particular, Muslims have, on average, significantly higher fertility rates than the latter. To quantify, a simple OLS analysis with some basic controls reveals hat ten average Muslim women in our sample are bound to have five more children in their lifetime than ten Christian women.

While these observations are intriguing and hint to the role of religion as a cultural factor determining fertility, we then go further and tackle a potential causal interpretation of our findings, namely, that historically determined belonging to a religious group in Africa is affecting current fertility differentials. To do so, we estimate the effect of belonging to a

[1] Schooling is one important instrument for religious indoctrination, see Arold et al., 2022, and references therein for supporting evidence.

religious group on fertility using an instrumental variable as constructed in Micholopoulos et al., 2019, that is distance to ancient roman trade routes (i.e., pertaining to the period prior to the expansion of Islam). The logic behind using this instrument lies in the fact that early Muslims, especially merchants, made use of preexisting trade routes in the dissemination of the new religion. Merchants converting to Islam used to have clear advantages including cooperation within the Muslim trading network, building contacts to expand their trade, and rules governing commercial activities favor Muslims over non-Muslims. This is also well consistent with the historical accounts of the spread in Islam in Africa, cf., "The association of Islam and trade in sub-Saharan Africa is a well-known fact... The explanation of this phenomenon is to be found in social and economic factors. Islam is a religion born in the commercial society of Mecca and preached by a Prophet who himself had for a long time been a merchant, provides a set of ethical and practical prescripts closely related to business activities. This moral code helped to sanction and control commercial relationships and offered a unifying ideology among the members of different ethnic groups, thus providing for security and credit, two of the chief requirements of long-distance trade." (UNESCO General History of Africa, Vol. III, 39)

Employing the distance to ancient trade routes in the context of 2SLS reduces the estimated effect of belonging to a religious group on fertility, yet it remains statistically significant; in particular, we find that ten average Muslim women are bound to have one more child in their lifetime relative to ten Christian women. This result does have a causal interpretation, indicating that historical circumstances that resulted in belonging to a religious group have an effect on current fertility. The fact that the OLS estimation results in a much larger magnitude of the estimated relationship indicates that it is biased –because of a selection bias. Whereas the above results withstand the inclusion of geographic, ethnicity based, and individual controls, large shocks may have a biasing effect. We, therefore, consider European colonization alongside with the spread of Christianity. However, this doesn't change our findings in a meaningful way; in particular, our baseline result indicating that the historical accident of becoming a Muslim positively affects current fertility stands.

Additionally, we explore potential heterogeneous effects. One is derived from what is dubbed in Meon and Tojerow, 2019, the "context-dependence hypothesis" – which distinguishes between households belonging to a religious groups depending on their minority status in a region (basically, the share of the region's population). Unlike Meon and Tojerow, 2019, which are interested in a different outcome – educational attainment – across the globe, and find that differences across religious groups become insignificant when a group constitute a small minority in a region, we do discern a similar effect here. In particular, we find that our results are highly driven by Muslim majority communities. Specifically, they could be an important factor mediating the effect of a religious group's minority status on fertility and other outcomes, households with a better background being able to culturally assimilate, in particular, by self-selecting into a small minority status. An- other potentially important heterogeneous effect concerns differences in religious adherence between the two religions. To explore this, we turn to a measure of such from the World Values Surveys. Our analysis reveals, again, that there is no heterogeneous effect of strong religious adherence and fertility between Muslims and Christians – which is consistent with the cultural assimilation through self-selection interpretation above: household with a better background "self select" into being less religious. To further investigate that, we examine the fertility choices of Muslim women migrating from Muslim-minority regions to Muslim-majority regions and find that the religion positive significant effect is mostly driven by destination regions hosting a higher percentage of Muslims compared to origin regions. This agian constitutes evidence of new residents adhering to the norms of the population majority group.

We then study the channels enabling this relationship. We first document that mother's education, work status, age of first marriage and birth, child mortality and a measure of women empowerment– are all significantly correlated with the religion indicator. Specifically, Muslim women have a much lower level of education, they marry out younger, and are less prone to work compared to Christians. They also suffer from lower degree of women empowerment, lower likelihood that the partner is employed and a lower education level of the partner. Additionally, child mortality is higher, and the health indicators of Muslim women is worse relative to Christians. Consequently, when these variables are introduced as controls into our estimation framework, the regression coefficient of "being a Muslim" on fertility is reduced and becomes less statistically significant. Zooming in on additional proxies for women empowerment related to the gender roles within a household as gauged from the DHS, we find evidence that lends strong support to this channel, Muslim women having less bargaining power, differently measured, relative to the Christian ones. Finally, we do not report any evidence for effect of historical ancestral ethnic characteristics regarding gender roles on fertility.

The paper contributes to a large literature on the determinants of fertility, the majority of which has focused on the relationship between income and fertility (e.g. Becker, 1960, 1965; Galor, 2011; De la Croix and Doepke, 2003; Chatterjee and Vogl, 2018). More recently, some studies has started exploring the role of cultural factors as a determinant of fertility (e.g. Baudin, 2015; De la Croix and Perrin, 2018; La Ferrara et al., 2012; Spolaore and Wacziarg, 2021), but none have examined exiplicity the role of religion. More specifically, our paper relates to the thin line of literature that features religion as a determinant of economic growth (Barro and McClearly, 2003) and educational attainment (McCauly, 2014; Platas, 2018).[2] Others have looked at the related relationships between religion and education, espicially in the context of colonial missionaries (e.g. Becker and Ludger Woessmann, 2009; and Arold et al., 2022). We also contribute to the scarce literature on the role of religion in women empowerment in Africa (Njoh and Akiwumi, 2012). More generally, the African continent,

[2] Iannacone (1998) provides a broad survey of the effects of religious attitudes, and Becker et al. ()2021) is a survey that entertains historical perspectives on the effects of religions.

populated by multiple ethnic groups, sometimes in conflict, has attracted growing interest in this regard (see Michalopoulos and Papaioannou, 2020, for a detailed literature review). Ethnic divide in Africa, from which we explicitly abstract here in order to focus on the role of religion, has loomed large in scholarly research (e.g., Easterly and Levine, 1997; Franck and Rainer, 2012; Kramon and Posner, 2016; Miguel and Gugerty, 2005), albeit mostly in the context of conflict, and social cohesion, with less emphasis on fertility differentials or household structure. Finally, an emerging line of work studies, specifically, the effect of Islam on aspects of economic performance is particularly relevant in focusing on Islam in Africa (Kuran, 2018; Platas, 2018, 2021).

The rest of the paper is organized as follows, next section gives a historical overview of religious denominations in Africa and presents our conceptual framework. Sections 3 and 4 describe our data and empirical specification. Section 5 presents our results and heterogeneous effects based on minority status, religiosity and spatial sorting. Section 6 investigates potential mechanisms and finally, section 7 concludes.

2 Background

2.1 The spread of early Islam and Christianity in Africa

The initial spread of Islam in Africa followed the conquest of North Africa by Muslim Arabs in the 7th century CE. In the process, Islam by and large replaced Christian communities in the region that became marginalized. Islam then spread throughout West Africa mainly by peaceful means, via merchants, traders, scholars, and missionaries (often Islamized Berbers who had been variously coerced or enticed to convert) in the 8th century CE along the trade routes which crisscrossed West Africa, moving from the east coast into the interior of central Africa, finally reaching Lake Chad. The religion also spread down through Egypt and westwards through the Sudan region below the Sahara Desert. Another wave brought the religion to Africa's eastern shores, the Horn of Africa and the Swahili Coast, directly from Arabia and the Persian Gulf. Even afterwards, when the religion arrived in East Africa, and in all for at least six centuries, despite several military encounters, Islam spread largely peacefully and gradually wherever there were trade connections. As more people were converted, so more Muslim clerics were attracted from abroad, and the religion was spread further, mainly across West Africa. Figure 1 shows that modern Muslims majority regions (countries) are clustered close to historical trade routes, while Christians are .

Islam had profound effects on all aspects of daily life and society, but also blended with local traditions and customs. Of a particular relevance here is the fact that men and women's roles sometimes changed, some African communities having previously given women a more equal status with men than was the case under Muslim laws. Some African societies, initially matrilineal, changed to a patrilineal system. Polygamy and the acceptable age of girls' marriage as inspired by Islamic traditions have left their mark in those societies. It is also important to note that the spread of Islam in Africa occurred well before the demographic transition took place anywhere and in context of societies, whether originating or receiving, where gender roles were skewed against women.

European Christianity reached Africa during the Great Age of exploration, beginning with the Portuguese king Joao II. As Europeans established outposts on the coast of Africa in the 15-18 centuries, they brought along missionaries, who settled among indigenous populations. At first, the introduction of Christianity was limited to coastal areas. With a few exceptions, missionaries did not carry Christianity into the interior of the continent until relatively recently, the 1800s. Only in 19th century, did the Europeans venture inland, mainly in Sub-Saharan Africa. Most of the effort was conducted in the context of local populations who had not converted to Islam, and in areas populated by non-Muslim indigenous population. In those areas, colonial powers typically allowed missionaries to run school systems – which were viewed by local Muslims as a potential threat to their identity, see Platas, 2018, for a detailed account of Malawi in this regard. Nigeria is another good case in point. The North (mostly Muslim)-South (Christian) divide led to the proliferation of Christian schools in the latter, leading to sharp cultural divides across the regions. More generally, as is convincingly argued in Bauer et al., 2022, Christianity tended to spread in areas in which Islam had a weak following, outside of the province of Islamic kingdoms – which were located alongside ancient trade routes. This, in turn, had a profound effect on the divergence in cultural traditions between Muslim and Christian regions in Africa.

2.2 Conceptual framework

Whereas common narrative often suggests that cultural values and outcomes such as fertility may be correlated, standard economic modeling typically abstracts away from these relationships. One straightforward way to incorporate cultural values into the determination of fertility outcomes would be to assume, in the standard model of fertility choice, that individuals place different values on the number of their descendants. A somewhat more interesting and nuanced path – that is consistent with this paper's empirical approach - is to assume that individual families differ in size initially and that the preference for the number of children increases with the number of own siblings: bigger families beget larger families.

This creates a propagation mechanism that leads to persistent differences in family sizes across families over time. A variation of this thinking stipulates that the preferred number of children hinges upon the average number of siblings in own cultural group. This, too, creates persistence over time and potentially different across groups family sizes in the steady state. Further details that incorporate cultural values in a very simple and standard economic model that features a choice over fertility and schooling are in the Appendix; this is just meant to create a framework linking cultural values and their persistence to fertility patterns. This

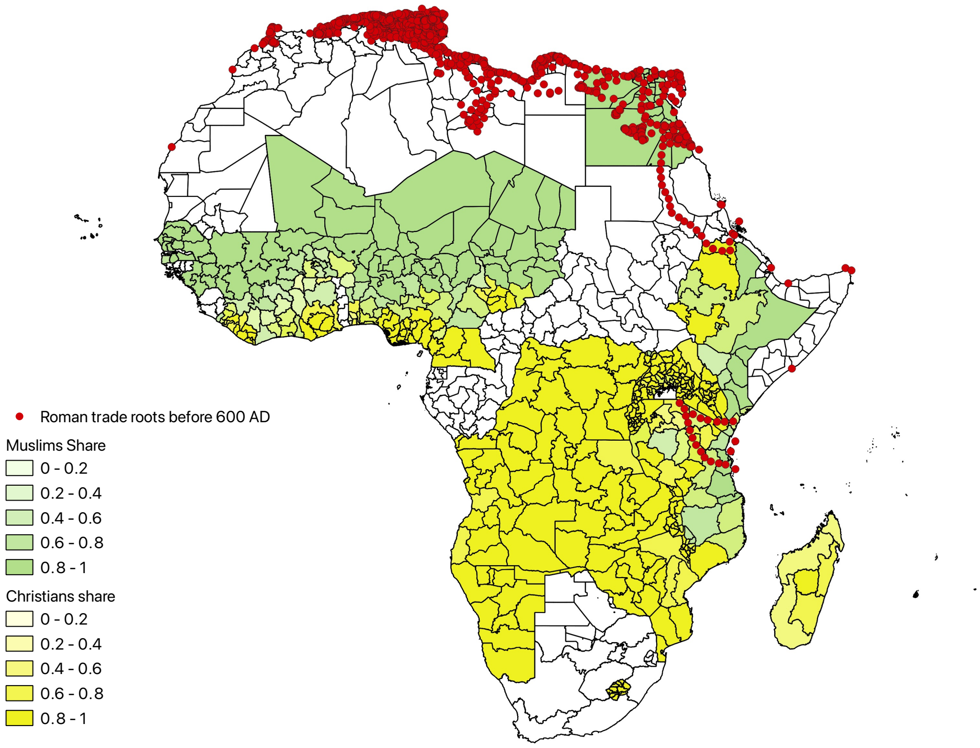


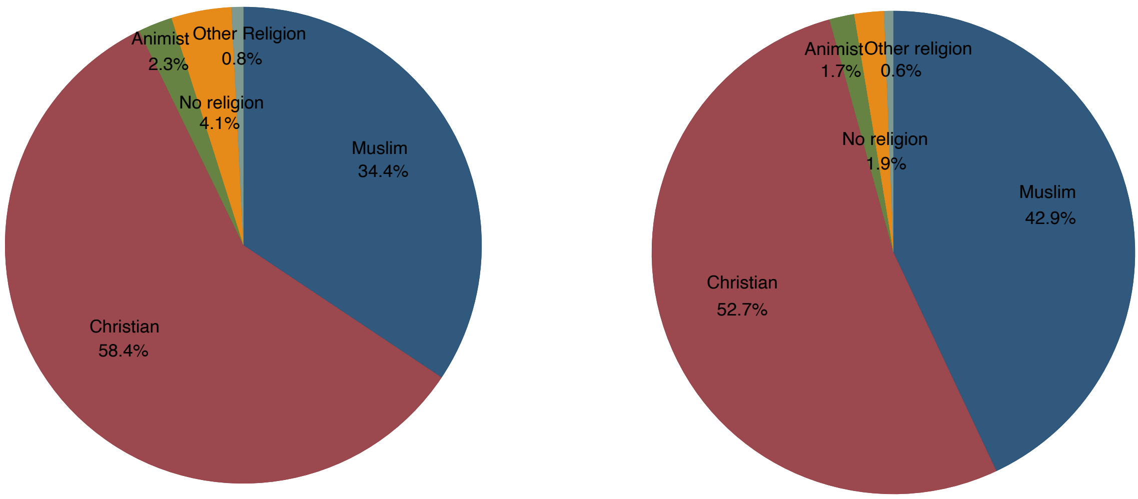
Figure 1: Historical trade routes and religious distribution

follows an emerging literature that incorporates cultural traits in economic modeling, see summaries of early work in Bisin and Verdier, 2011, and Guiso et al., 2006. Guiso et al., 2016, Giavazzi et al., 2020, and Giuliano and Nunn, 2020, address, among other things, the issue of cultural propagation, arguing that cultural traits tend to exhibit intergenerational persistence which is consistent with the formal model in the Appendix. Cultural persistence is also at the core of Platas' 2018, related work that focuses on the difference in schooling between Muslims and Christians in Africa, arguing that historical proliferation of Christian schools in Africa's regions has had a long standing effect on the motivation to puruse education across the two religious groups.

3 Data

3.1 Sample

Our data on religion, fertility, and individual characteristics come from the individual census records conducted by Demographic and Health Survey (DHS) program. The DHS household surveys provide nationally representative sample interviews of women aged 15–49 years old. Our analysis is restricted to country survey waves providing GPS information on the location of the interviewed households. This is because the geocoded data allows us to assign individ-



(a) Unweighted

(b) Population-weighted

Figure 2: Respective religions shares

Note: The Christian category includes different Christian denominations such catholic, protestant, orthodox among others. Same for the animist category, it includes traditional religions and animist adherers. No specific denominations are reported for Muslims.

uals to their ethnic regions as explained below. This leaves us with 70 surveys over the period 1990-2018 in 24 countries, namely: Angola, Burundi, Benin, Burkina Faso, Cameroon, Chad, Democratic Republic of Congo, Egypt, Ethiopia, Ghana, Kenya, Lesotho, Madagascar, Mali, Malawi, Namibia, Mozambique, Nigeria, Niger, Rwanda, Tanzania, Uganda, Zambia and Zimbabwe. The DHS provides detailed information on the respondent's religious denomination, which is standard across surveys and countries. We group religions into 5 categories: Christians, Muslims, Animists, No religion and Other religion. Figure 2 reports the (population weighted) shares of different religions in our sample. Christians represent the largest group with a share 58.4% of total sample, followed by Muslims with 34.4% and individuals adhering to no religion with 4.1%. Animists constitute 2.3% and that share of other religions is 0.8%. Using population weighted shares, we also find Christians and Muslims to present the largest two religious group with a Christians share of 52.7% and a Muslims share of 42.9%.

3.2 Fertility

To measure fertility, we make use of DHS-births history survey, which records a complete list of all children the woman has ever given birth. Birth histories include all live births, including children who later died, but omit stillbirths, miscarriages, or abortions. The survey covers all women aged 15-49 years old. Our measure of fertility reports the total number of children ever born (i.e., the stock measure). We restrict our sample to women aged 45 years and above at the time of the survey to allow us to capture complete fertility records. This leaves us with a total sample of 56,142 women born between 1945 and 1973. To illustrate, Figure 3 shows the average number of children born to women aged at least 45 years old in our sample, net of country, ethnicity and year fixed effects. We observe that, on average, the number of children born to Muslim women is generally higher than to Christian women by 15-20%. It is also noteworthy that there has been a fertility decline over time for both religious groups, more pronounced for Muslims. Still, even in recent years the average number of Muslim children is higher than the number of Christian children.

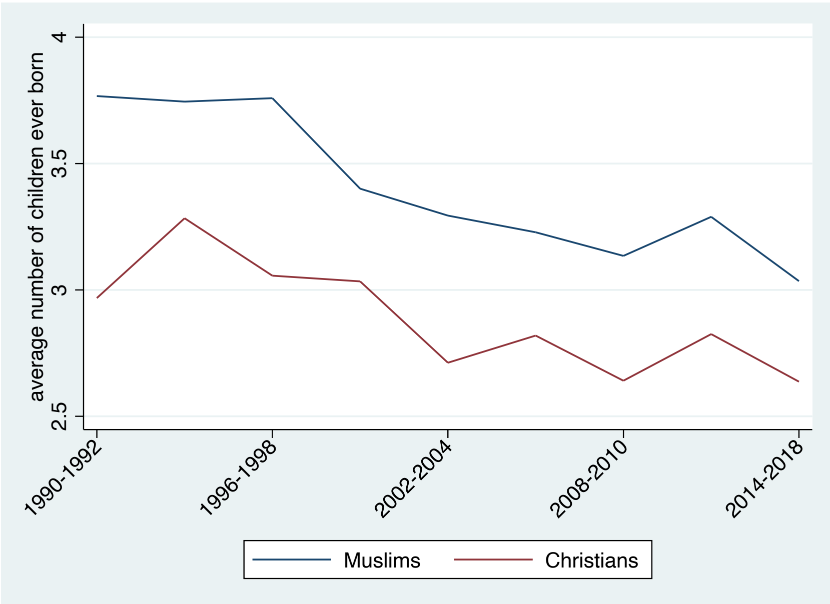


Figure 3: Number of children born by religious group

Note: Figure 3 plots both the average number of children born for women aged at least 45 years old over different time periods for available survey years. To smooth fluctuations, each bin in x-axis corresponds to a 3-year average, net of country, ethnicity and year fixed effects.

3.3 Controls

Ethnicity: Data on ethnicity has been retrieved from George Peter Murdock's Ethnographic Map of Africa, which spatially maps the distribution of around 800 African ethnicities at the time of colonization. To do so, we project the map on current regional boundaries (ADM1 level), and hence, allocating historical ethnic homeland to their respective current local regions. We then use the GPS information provided by the DHS surveys to allocate individuals to their respective local regions and ethnic homelands. From the 631 ethnicities, 412 are partitioned across 2 to 22 regions with 374 ethnicities are partitioned in regions with different religious denominations. Figure 4 plots the ethnic homelands boundaries over their respective current African regions and shows the location of ethnicities with information on religions.

Geographic controls: For our estimation strategy, it is important to supplement our individual- level data on religion with geographic information. To do so, we follow Michalopoulos and Papaioannou (2013) and partition African regions into equal grid units of 12.5 x 12.5 km. For each grid, we compute elevation, land suitability for agriculture, distance to capital, distance to coast, distance to region border, distance to protestant and catholic missions, night-time lights, climate zones, presence of water surfaces, diamond and oil. We then use the GPS information on households' locations to assign households to their corresponding grids.

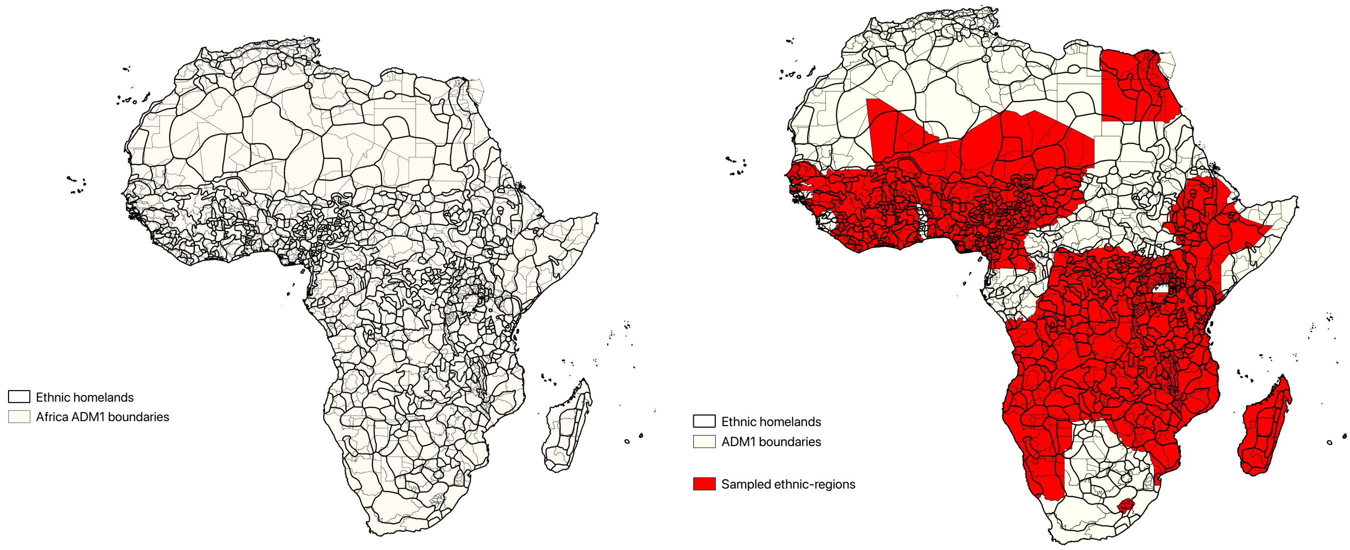


Figure 4: Ethnic homelands and African region boundaries and sampled ethnic regions

Note: Murdock's ethnographic Map of Africa projected on current regional boundaries (ADM1 level).

Individual-level controls: The DHS surveys contain a set of demographic and household-level information for women, which we use to control for individual characteristics and investigate potential mechanisms. These include age, birth cohort, place of residence (i.e., urban vs. rural), marital status, education, female occupation, household wealth index, health indicators, husband's education and occupation, use of contraceptives, age at first marriage and birth, and module on women empowerment.[4] The latter provides information on how household decisions regarding, for instance the purchase of household chores or ability of women to visit her parents, is taken in the household. Table A1 provides summary statistics of our main variables of interest.

[4] To measure household's wealth, we construct a wealth index (WI) using information on household characteristics such as household's possession of consumer durable goods, access to basic services, and housing condition. These indicators are then entered into a factor analysis -using the Principal Component Analysis- from which the first factor is selected to derive the asset weights and consequently the wealth index. The obtained WI is rescaled, so that it ranges from 0 to 100, with 0 representing households having no assets and living in lowest quality housing, and 100 representing households possessing all assets and living in highest quality housing.

4 Empirical specification

Estimation: Our identification strategies looks at differences in fertilty choices among religious adherents belonging to the same ethnicity. The main empirical specification takes the following form:

$$\begin{aligned} Y_{ivrcet} = \beta_1 Muslim_{ivrcet} + \beta_2 X_{vrce} + \beta_3 X_{ivrcet} + \\ \eta_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ivrcet} \end{aligned} \tag{1}$$

where the outcome Y_{ivrcet} is the outcome of woman i residing in cluster v region r, country c, ethnic homelnad e in the survey year t. $Muslim_{ivrcet}$ is a dummy variable takes the value of 1 if the respondant is muslim. The omitted category is christian. X_{vrce} and X_{ivrcet} are vectors of individual and region-ethnic-geographic controls, respectively. These include age, cohort, month of the survey, marital status, dummy for urban residency, household welath, distance to capital, distance coast, elevation,...etc. η_e is ethnic fixed effect, θ_r is a region fixed effect to control for unobservable regional characteristics and $\delta_c \times \phi_t$ is country-survey year fixed effects to capture shocks that are common across countries. Standard errors are clustered at the country and ethnicity level.

While the above specification makes use of religious differences within the same ethnicity to control for various cultural and traditional norms, and includes different geographic and economic controls that might affect fertility in the long run, we nevertheless cannot rule out the possibility that the presence of religious denominations in a certain region is not exogenous. Several factors feed this doubt including migration, self-sorting, and historical accounts among others. To address that, we employ an instrumental variable strategy, and following a recent contribution by Michalopoulos et al. (2019), we instrument our indicator for Muslim adherence with spatial information regarding the cluster proximity to the pre-600 CE Roman trade routes.[5] The logic behind using this instrument is rooted in several distinct features of the spread of Islam, pointed out in Micholopoulos et al., 2019. One is that early Muslims, especially merchants, made use of preexisting (thus, exogenous from the perspective of the spread of Islam) trade routes in the dissemination of the new religion. Further, Islam had an appeal among converts because of potential economic advantages associated with trade networks. Merchants converting to Islam used to have clear advantages including cooperation within the Muslim trading network, building contacts to expand their trade, and rules governing commercial activities favor Muslims over non-Muslims. Finally, proselytization and intermarriages constituted a third factor that enhanced the spread of

[5] Previous empirical work in religion have often used distance to various religious capitals as an instrumental variable for religious concentration. For instance, Cantoni (2012) shows that the distance to Wittenberg is significant determinant of the adoption of Protestantism. Coşgel et al. (2018, 2019) rely on the distances to various "religious capitals", including Rome, Nepal and Mecca, as instrumental variables for countries' levels of religious pluralism.

Islam across distant. These factors provide an indication that distance to ancient trade routes may be a valid instrument for professing Islam. An additional, fourth factor has to do with the fact that, following Islam's expansion in Africa, Islamic kingdoms were established in close proximity to ancient trade routes. This, in turn, had implications for the subsequent proliferation of Christianity in the continent, serving as an impeding factor for the interest, willingness and determination of colonial European powers and their agencies to move into areas heavily populated by followers of Islam (Bauer et al., 2022). Figure 2 shows that indeed Muslims majority regions (countries) are clustered close to historical trade routes.

Consequently, we implement an instrumental variable strategy (IV), with first stage can be represented by the following:

$$\hat{Muslim}_{ivrcet} = \alpha_1 log(\textit{Distance to trade routes})_{vrce} + \alpha_2 X_{vrce} + \alpha_3 X_{ivrcet} + \eta_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ivrcet} \tag{2}$$

where $log(\textit{Distance to trade routes})_{vrce}$ is the distance to the pre-600 CE roman trade routes. To construct the instrument, we calculated the minimum distance from the centroid of the cluster's grid to the closest historical trade route or port before 600 CE.[6] The corresponding exclusion restriction is that distance to historical trade routes only affects fertility through its affect on percentage share of muslims. It follows that in the second stage, we use the predicted level of Muslim presence to estimate the effect of religious denomination on fertility as shown below:

$$Y_{ivrcet} = \gamma_1 \hat{Muslim}_{ivrcet} + \gamma_2 X_{vrce} + \gamma_3 X_{ivrcet} + \eta_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ivrcet} \tag{3}$$

Specification: For our identification purposes, two assumptions need to be satisfied. The first requires that there are no systemic differences between religiously partitioned ethnic groups with respect to (un)observable characteristics that may affect contemporaneous fertility. Besides working at the ethnic-region level, which already controls for some of the relevant cultural, traditional, and region specific norms, using instrumental variable (IV) strategy corrects for reverse causality and omitted variable bias. To meet the exclusion restriction of the IV, we further supplement our specification with a wide range of cluster geographic and economic controls, which include distance to capital, distance coast, elevation, land suitability for agriculture, night-time lights, climate indicators, presence of water surfaces, diamond and oil. We also control for the presence of Animists in the region and distance to Protestant and Catholic missions. The latter two controls are important to take into account because the religious composition of regions has potentially changed over time during colonial time

[6]Shape files of the locations of historical trade route and ports was provided by Michalopoulos et al. (2019). More information of data sources is provided in their study.

and in the course of establishment of Christian missionaries.

Second, using historical ethnic homelands assumes that border regions have witnessed little migration over time. In other words, ethnic groups should be expected to reside in their historical locations. Following Alesina et al. (2016), we aggregate the individual-reported DHS ethnicity data at the cluster level by assigning the ethnicity of the majority group. We then compare this contemporaneous ethnic information with the one computed from the Ethnographic Map of Africa. Figure A1 in the Appendix shows that most of the clusters remain in their historical ethnic locations and that ethnic settlements have remained fairly stable over time. The results are very close to Nunn and Wantchekon (2011) and Anderson (2019) findings that 50% of Afrobarometer surveyed respondents within a given ethnic group and 60% of DHS surveyed respondents reside in their respective historical ethnic homeland, respectively.

5 Results

5.1 Main findings

. We begin by documenting some basic correlations. Thus, in column 1 of Table 1 we exhibit a statistically significant and positive correlation between the indicator of being a Muslim and fertility. Quantitatively, this simple OLS analysis with some basic, individual-level and geographic controls, reveals that ten average Muslim women are bound to have five more children in their lifetime that ten Christian women.

Column 2 and 3 in Table 1 contains the first stage two-state least squares (2SLS) estimate for the effect of religious denomination on fertility. We instrument Muslim indicator with a cluster's (log) distance to historical trade routes. The first stage results show that the instrument is a strong predictor for the share of Muslims in a given region with first stage F-statistic close to 20, hence well exceeding the recommended threshold by Staiger and Stock (1997). We see that the estimated coefficient of Muslims is positive and statistically significant at the 1% significance level. The 2SLS estimated coefficient is roughly 0.1, which is smaller in magnitude compared to the baseline OLS estimates, indicating an upward bias of the latter. This upward bias can be plausibly due to omitted variables or to classical measurement error in the Muslim indicator. Using the instrumental strategy corrects for these factors, so that the finding of the larger fertility among Muslims can be given a causal interpretation. Even though the resulting coefficient of 0.1 is smaller in size, not only is it statistically significant but also economically meaningful indicating that ten Muslim women tend to have on average one more child in the course of their lives than do ten Christian women.

In Table 2 we conduct a number of robustness checks, such as with respect to women's age; an alternative outcome variable (children born alive as opposed to the total number of children in the family); limiting the sample to households without polygamy; and including other religions. In all these variants, the results barely change. While the exhibited specifications control for a number of geographic variables, there is still concern as to potential location heterogeneity of the households that may have a bearing on the results. Indeed, Alesina et al. detect substantial regional differences with respect to their outcome of interest, educational mobility. In particular, they observe differences across religious groups in terms of places of residence, Muslim households tending to reside in less urbanized areas. In the light of these facts and in order to alleviate the concern of confounding location factors, therefore, in Table 3 we carry out our main IV estimation while limiting the sample of households to varying degrees of their residence proximity to their region's border. In all such regressions, the Muslim coefficient remains significant; in fact, for as long as the distance to the border is at least 30 kilometers, its value is higher than for the entire sample, so that generally, the relationship is stronger for households in the vicinity of the border than for the entire sample. Based on these results, we restrict our sample below to households residing within 40 kilometers of the region's border – which makes sure that their respective locations are geographically alike. In the appendix, however, we also exhibit the results for the entire sample; to summarize in brief, they are similar to albeit somewhat weaker than those presented in the text.

To further test the sensitivity of the results to omitted variable bias, we follow Oster (2019) approach, based on earlier work by Altonji, et al. (2005), in providing bounds for impacts of religious adherence on fertility based on the degree of selection on unobservables relative to observablesbased. To calculate that, Oster (2019) approach takes into account the estimated coefficient (β) and R^2 after the inclusion of all controls. These calculations depends on two assumptions: (1) a value of (δ), that is an estimator for the degree of selection on unobservables relative to observables and (2) a hypothesized value of R-squared (R_{max}). In this regard, Oster (2019) proposes two approaches. The first approach assumes a value for R_{max} and calculate the value of δ for which β=0. This can be interpreted as the degree of selection on observed and unobserved variables that would be necessary to overturn the results under the hypothesized value of R_{max}. The second approach to develop a set of bounds for β, using bounds on R_{max} and δ. If the bounds exclude zero, the results can be interpreted as robust to omitted variable bias under the stated assumptions. Following Oster (2019) recommendation, we employ a less restrictive hypothesis on R_{max} being equal to 1.3 times R^2 value of the demanding specification that includes all the controls.

We find a value of δ = 2.28, which is greater than the threshold of 1 for robustness (Altonji et al. 2005). This indicates the unobservables must be at least δ times greater than the observables to overturn the results. Considering bounds for β, one side of the bound is provided by the estimated Muslim coefficient using equation (3) which includes the full battery of controls in Table 1 under the assumption of zero selection on unobservables. The other side is calculated assuming that δ=1 and R_{max} of 1.3 $\times$ R^2. Our estaimated bound is equal to [0.012, 0.081] which excludes zero and therefore can be viewed as robust to omitted variable bias under the stated assumptions.

Athought our IV -*log(distance to trade routes)*- is found to be a strong predictor for Muslim coefficient based on first-stage *F-statistic*, the instrument validility condition requires the full excludibility of the IV. To test that, we folllow Conley et al. (2012) approach that relaxs the exclusion restriction and examine the bounds on the true effect of being Muslim on fertility. The analysis assumes that log(distance to trade routes) appear in equation (1):

$$Y_{ivrcet} = \beta_1 \hat{Muslim}_{ivrcet} + \hat{\gamma} log(Distance\ to\ trade\ routes)_{vrce} + \beta_2 X_{vrce} + \beta_3 X_{ivrcet} + \eta_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ivrcet}$$

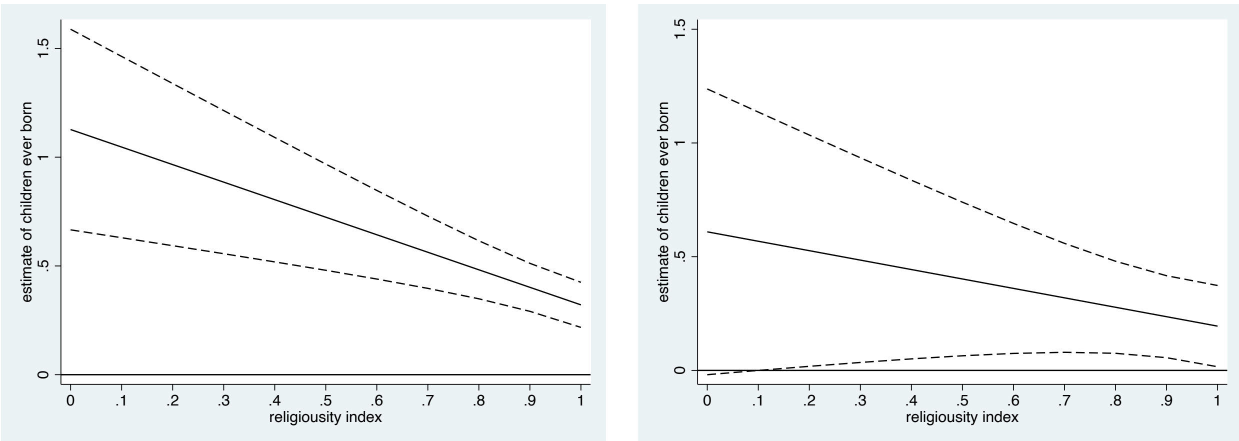
If the true value of $\hat{\gamma}$=0, we can estimate the above equation using 2SLS, however, the true value of $\hat{\gamma}$ is unknown. To address that, Conley et al. (2012) makes assumptions on $\hat{\gamma}$ that allows it to be close to, but not necessarily equal to, zero. This asssumption can take the form of the support of $\hat{\gamma}$ to calculate bounds on the estimated coefficients on β_1 (the coefficient of interest). We use the "union of confidence interval (UCI)" conservative approach to calculate these bounds, which allows for assuming the maximum and minimum value that $\hat{\gamma}$ can take. According to cenvention, the value $\hat{\gamma}$ is recommended to be no more than 0.3, or 30% of the value of β_1 estimated from the IV regression. Nevertheless, we estimate these bounds based on serveral values for $\hat{\gamma}$ of up to 100% of the value of β_1. Hence, $\hat{\gamma}$ can take all values in $[\hat{\gamma}\beta_1, \hat{\gamma}\beta_1]$, where $\hat{\gamma} = 1$, and β_1 is the coefficient of interest estimated in column (3) of Table 1. Figure A2 in Appendix A presents the results, plotting the bounds for β_1 at different magnitudes of $\hat{\gamma}$ within the 95% confidence interval. Estimates show that the all intervals exclude zero at all values of $\hat{\gamma}$.

5.2 Minority vs. majority

To develop a better understanding of the heterogeneous effect of religion on fertility, we now distinguish between households living among their co-religionists as opposed to those where co-religionists constitute a small minority in a cluster; this is in line with an approach dubbed in Meon and Tojerow (2019), the "context-dependence hypothesis" that is also pursued in Platas, 2018. This is particularly pertinent in Africa, where religious groups are generally quite segregated, see Figure 2. Existing literature well represented by the above cited work has documented significant differences in education differentials, depending on whether households reside among their coreligionists or not. Both Meon and Tojerow (2019) in their cross country analysis and Platas, 2018, in Africa's, specifically, in Malawi's context, present this finding, the latter paper also arguing that cultural channels are responsible for it.

To address this issue with our data, we distinguish between our sample's clusters according to the percentage of their Muslim households and run our baseline regressions separately for such cluster bins. We observe, see Table 4, that the Muslim indicator is insignificant in its fertility effect for as long as Muslims constitute a minority in a cluster. In contrast, whenever there is a Muslim majority, especially when it is large, then the Muslim indicator becomes highly significant and large in magnitude. For example, our results show that when the share of Muslim households in a cluster exceeds 80 percent of the cluster's sample population, ten Muslim women tend to have almost 3 more children on average than ten Christian women. As in all these regressions multiple geographic controls are included, we tend to interpret these results in causal terms, i.e., being a Muslim household among Muslim majority enhances fertility.

This finding is open to at least two possible interpretations. One is that Muslims, when in a minority, absorb the fertility norms of the majority. Another interpretation is that Muslim households without strong attachments to religion (or without strong preference to reside with co-religionists) choose to reside among households professing other religions – and also smaller household size. We are unable to distinguish between these two interpretations, but we explore below the extent to which religiosity and migration affects fertility.



(a) All sample

(b) Female only

Figure 5: Marginal estimates of children ever born at different level of muslim's relogiousity index - Net of region, ethnic and county-survey year fixed effects

5.3 Religion vs. religiosity

Our results have shown that Muslim women tend to have more children compared to Christians, but less is known on whether this result is driven by inherited Islamic teachings or contextual factors. Unfortunately, the DHS surveys do not contain information on religious adherence. Therefore, we use the World Values Survey (WVS), which includes individual level data on fertility, ethnicity and religiosity in eight African countries in four waves covering the period between 1994 and 2020.

Following our baseline strategy, we focus on two religious groups, namely Muslims and Christians, which still represents the largest two religious groups (Figure A3 in Appendix A). WVS includes questions on women's age and number of children ever born, which we make use of to construct our total fertility indicator. To measure religiosity, we include questions regarding religious values and behavior, previously used as indicators of religiosity (e.g., Bentzen, 2019; McCleary and Barro, 2006; Norris and Inglehart, 2004; Cosgel et al., 2019). These are those that asked respondents whether they (a) consider religion important in life, (b) have membership in a church or religious organization, (c) attend services regularly, (d) are a religious person, and (e) believe in god. We standardized and recoded the responses for each question, so that higher values would indicate stronger believes. We then construct an index for religiosity by taking the weighted average score of the responses to these five questions. We also make use of information available on women's ethnicity, marital and socioeconomic status such work, and income level. To get a first snapshot, Figure A4 in Appendix A shows that Muslims tend to have more children than Christians confirming our previous results, but at the same time, they are less religious.

To formally investigate that, we run the following regression to examine the effect of religion on fertility taking into the degree of religiosity. Due to unavailability of data on households' geographic location, we cannot employ an instrumental variable strategy, but we are able to control for ethnicity. Hence, the WVS equation reads as such:

$$\begin{aligned} Y_{ircet} = \beta_1 Muslim_{ircet} + \beta_2 religiosity\ index_{ircet} + \beta_3 Muslim_{ircet} \\ \times religiosityindex_{ircet} + \beta_4 X_{ircet} + \gamma_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ircet} \end{aligned} \tag{4}$$

where the outcome Y_{ircet} is the outcome of individual i belonging to ethnicity e residing in region r, country c, in the survey year t. $Muslim_{ircet}$ is a dummy variable takes the value of 1 if the respondent is Muslim. The omitted category is Christian. X_{ircet} is a vector of individual and region controls such as age, cohort, ethnicity, and wave of the survey. θ_r is a region fixed effect to control for unobservable regional characteristics, γ_e is an ethnicity fixed effect, and $\delta_c \times \phi_t$ is country-survey year fixed effects to capture shocks that are common across countries. Standard errors are clustered at the region and ethnicity level. We run this regression twice, once on the full sample of surveyed individulas and second on women only. Because of the small sample of surveyed women, we include all women irrespective of their age. Table 5 reports the results and shows that again Muslims exhibits higher fertility patterns compared to Christians. Taking into account the religiosity index, we observe a positive effect of religiosity on fertility in the full sample, but not in the sample of women (columns 2 and 5) and that the interaction term enters negative, but statistically insignificant (columns 3 and 6). To further investigate that, figure 5 plots the marginal estimates of fertility at different levels of religiosity for both full and women sample. Again, we do observe that highly religious Muslims tend to have fewer children compared to highly religious Christians.

The pattern is however relatively weaker among Muslim women.

5.4 Spatial sorting

We have shown that our baseline results are driven by Muslim majority regions, but we cannot determine whether this has to do with new residents adhering to the majority population norms, or because of spatial sorting, meaning that Muslims with high fertility preference select themselves into Muslim majority regions. One way to investigate this is to examine the fertility choices of Muslim women migrating from Muslim minority regions to Muslim majority regions. To do that, we make use of African national surveys available from IPUMS International, which report detailed migration information including the birthplace (i.e. origin region), the current place of residence (i.e. destination region), and the age of movement. This information is only available for 49 regions in Benin, Rwanda, Ghana and Guinea. Unfortunately, these surveys don't include either ethnicity data nor households' geographic location, and therefore we cannot employ an instrumental variable strategy.

As in our baseline regression, we focus on Muslim and Christian women aged at least 45 years old and construct our total fertility indicator using the question on the number of children ever born. Our equation reads as follows:

$$Y_{ikrct} = \beta_1 Muslim_{ikrct} + \beta_2 \mathit{Difference}_{ikrct} + \beta_3 X_{ikrct} + \delta_r \times \gamma_k + \delta_r \times \phi_t + \epsilon_{ikrct} \qquad (5)$$

where the outcome Y_{ircet} is the fertility of woman i belonging to cohort k residing in region r, country c, in the survey year t. $Muslim_{ikrct}$ is a dummy variable takes the value of 1 if the respondent is Muslim. The omitted category is Christian. $Difference_{ikrct}$ is the difference in the percentage of Muslims between origin and destination regions. X_{ikrct} is a vector of individual and region controls such as age, cohort, marital status, urban, years of schooling, age at move and wave of the survey. $\delta_r \times \gamma_k$ is region-cohort fixed effects to control for unobservable regional characteristics that differ by cohort, and $\delta_r \times \phi_t$ is region-survey year fixed effects to capture shocks that are common across regions. Standard errors are clustered at the region level.

The results are reported in Table 6, and they indicate that Muslims exhibits higher fertility patterns compared to Christians (column 1). Yet, the effect is mostly driven by destination regions hosting a higher percentage of Muslims compared to origin regions or where the difference in the share of Muslims between origin and destination regions is positive (column 2). This constitutes evidence of new residents adhering to the norms of the population majority group.

5.5 Spatial regression discontinuity

So far, our empirical strategy looks at fertility differences among Muslims and Christians belonging to the same ethnicity irrespective of regional proximity. African ethnicities are not confined to particular region within a country, but are rather split across regional borders and across countries. To exploit this setting, we follow a regression discontinuity design (RDD) in which we exploit split ethnicities between regional borders to examine the effects of religion on fertility among women belonging to the ethnic group. Hence, we compare the number of children born to Muslim women with those born to Christian women from the same ethnicity group, but located in two different bordering regions. As a precondition for RDD, there must be a discontinuity in the percentage of muslims (christians) around the region's border. To confirm that, Figure 2 plots the distribution of muslims across borders and shows that the neighbouring regions are religiousely segerated. Furthermore, Figure 7 shows that the precentage share of muslims indeed jumps at the region's border. Similar to equation (3), the RDD specification takes the following form:

$$\begin{aligned} Y_{ivrcet} = \omega_1 \hat{Muslim}_{ivrcet} + f(BD_{vrce}) + \omega_2 X_{vrce} + \omega_3 X_{ivrcet} + \\ \eta_e + \theta_r + \delta_c \times \phi_t + \epsilon_{ivrcet} \end{aligned} \tag{6}$$

Where $f(BD_{vrce})$ is the RD polynomial, which controls for smooth functions of geographic location of cluster v. We specify the function multidimensionally using the cluster's longtitude, latitude and distance to region border and employ a local linear polynomial for our baseline specification (e.g., Eleonora and H. Rainer, 2021). For robustness checks, we allow for using polynomials of higher orders (quadratic and cubic). Because in the first stage the Muslim coefficient is instrumented with (log) distance to historical trade routes and other controls, our specification therefore follows a fuzzy RDD. An important feature of the RDD approach is the choice of bandwidth used to delimit the clusters within a given distance to the region boarder. In doing so, there is a typical trade-off between choosing a low bandwidth, hence specifying more similar units, and between having a bigger sample. We follow the recent procedure of Calonico et al. (2020) and choose the optimal bandwidth based on a data-driven approach that uses our dependent variable, RD polynomial and other covariates. We use two bandwidth selectors, the Coverage Error Rate (CER) and the Mean Square Error (MSE), separately for both Muslims in comparison to Christians. For Muslims-Christians border analysis, the CER-optimal bandwidth procedure yields 24.9 km, while the MSE-optimal bandwidth reports 34.9 km. Based on this, we choose a window of 30 km around the boundary for our preferred specification.

Results are shown in Table 7, where column 1 uses a local linear polynomial, and columns 2 and 3 employ the quadratic and cubic forms of cluster's longitude, latitude and distance to border, respectively. Throughout all the modela, we still find a statitiscal significant evidence that Muslim women tend to give birth to more children compared to Christians. In Table 8,

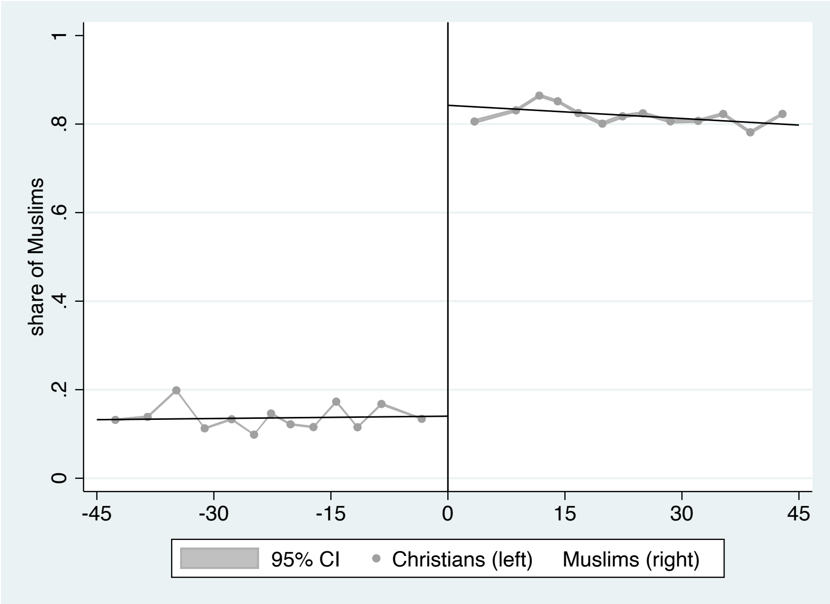


Figure 6: Religious segergation at regional borders

we run several models to check, whether the results are robust to different distance cutoffs. Indeed, our results remain stable in sign and statistical significance.

6 Mechanisms

Contemporary factors.—We next turn to potential driving forces behind the documented associations. To do so, we display, in Tables 9-11, the estimated coefficients of the effects of religious denominations on additional variables from the DHS. This is done using our preferred IV specification, with the distance cutoff of less than 40 km to the border. We observe that being a Muslim is associated with having less (women's) education; not working; being younger at the age of first marriage as well as when the first child was born. We, in particular, note the education effect, which is well consistent with previous findings, e.g., Platas, 2018. Further, it is associated with a lower degree of women empowerment; with a lower education level of the partner and lower likelihood that the partner is employed.[7]. Additionally, child mortality is higher, and the health indicators of Muslim women is worse relative to Christians.

All these are variables that are associated with fertility and have a bearing on it. When these are introduced one by one into regressions, the religion coefficient is typically reduced

[7]To measure women empowerment, we make use of DHS module questions on women's decision making within the household. Specifically, in this module, they ask women: "Who usually decides about a) major household purchases, b) health care for the wife, and c) visits to the wife's family and relatives? The responses are: "respondent," "husband," or "respondent and husband jointly". For each question, we create a dummy variable equal to 1 if the woman makes any decision within the household -whether alone or jointly with her husband- and equal to 0 if her husband decides alone. We then sum up these three dummies and create our additive women empowerment index with higher values indicating higher empowerment

in size, as illustrated in Table A2 and figure 7. In particular, when the index of women's empowerment is introduced, the Muslim coefficient is cut by a half relative to the baseline estimate. And when all controls are introduced, the religion coefficient is reduced by some 90 percent, becoming statistically insignificant. This is a further strong indication that these are relevant channels through which the fertility effect of religion manifests itself.

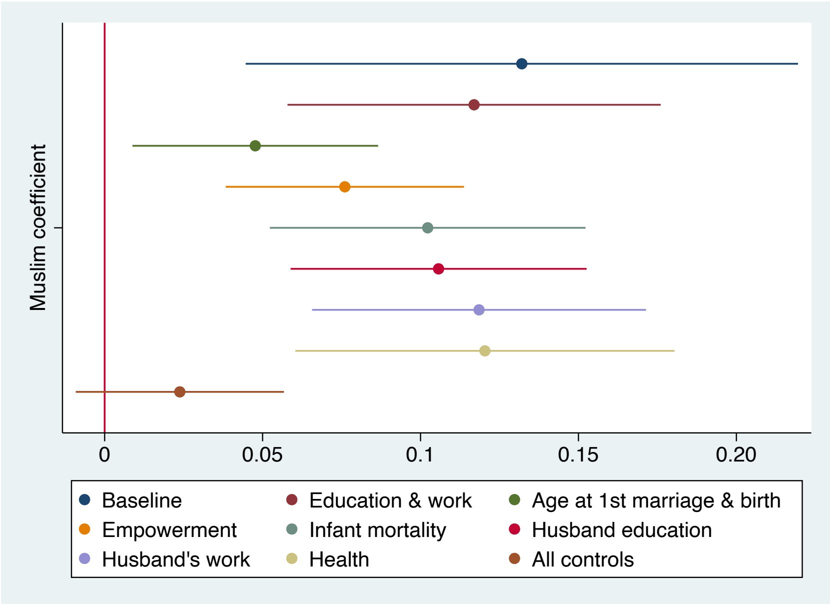


Figure 7: Estimated muslim coefficients for equation (3) after adding controls one by one

To further zoom in on the women empowerment aspect, we develop several additional measures. One is related to the difference between the actual and the desired number of children as reported in the DHS. In column 1 of Table 12, we observe that for Muslim women this disparity is significantly higher than for Christian ones, with muslim women tending to have more children than desired. Column 2 of the same table, however, indicates that there is no such significant difference among men of the two religions. Column 3 indicates a higher level of disagreement between spouses in Muslim households regarding the desired number of children; and the next two columns indicate, specifically, that, in Muslim households, a husband wants more and a wife wants less children relative to Christian households. Taken together, these findings provide evidence of a lower level of women empowerment within Muslim relative to Christian households.

Selection of covariates.—Our results indicate that Muslims tend to have more children compared to Christians, however the effect of religion is seen to have decreased once we controlled for individual characteristics. Yet, adding more variables to our model has drawbacks in terms of prediction accuracy and model interpretability. For the former, additional covariates can lead to much variability in the least squares fit if the number of observations is less than the number of covariates, resulting in overfitting and consequently poor predictions. As for the latter, it is more likely to occur if some or many of the variables used are (not) associated with the outcome. Hence, the inclusion of such variables leads to complexity in the resulting model and make it more difficult interpret the estimated coefficients. This is of a particular concern in the presence of multicollinearity, that is if these additional variables are correlated with each other. If this is the case, then we are less sure about which variables (if any) truly are predictive of the outcome, and we can never identify the best coefficients for use in the model. It also affects the reliability of estimated coefficients for religious in terms of statistical inference.

Because our IV estimation is extremely unlikely to drop coefficients, we have applied a machine learning method that allows for variable selection—that is, for excluding irrelevant variables from a multiple regression mode. We employ the double selection LASSO (Least Absolute Shrinkage and Selection Operator) technique for model inference. The idea is that in choosing the list of covariates, LASSO minimizes the sum of squared errors for a given value of penalty (λ) that increases with the complexity of the model. With $\lambda= 0$, the model boils down to an OLS model with maximum complexity, and hence no covariates are dropped. With higher the values of λ, the estimated coefficients are allowed to become zero and consequently dropped of the model. If the latter happens, it either means that a given covariate does not belong to the true model or it is correlated with covariates already selected.

For the double selection LASSO, the estimation follows three steps. In the first step, we run a lasso model regressing the Muslim and Animist indicators on the set of the controls and select the covariates that are predictive to the religion's indicators. The second step runs a similar lass model, except that we regress fertility on the set of controls and select the relevant covariates. The third step runs a linear regression model that regress fertility on religious indicators and the union of selected covariates from step 1 and 2. Results are presented in column 1 in Table 13. In choosing the optimal λ, we use the plugin method to achieve an optimal sparsity rate, that is allowing for fewer variables relative to the number of observations. The estimated coefficient indicates that religion matters less for fertility, with the Muslim indicator being statistically insignificant and the estimated coefficient is much smaller in magnitude compared to the baseline IV model. The selected covariates for muslim indicator are education, work, women empowerment, and partner's education, while the fertiliy selected covariates were wealth, age at first birth, infant mortality and education.

In short, we find that there are significant differences across the two religious groups in fertility behavior, being a Muslim households resulting in having more children on average. However, the effect of religion dissipates, once we take into account individual and household characteristics. The primary contributing factors seem to be women's education, age at first marriage, and age of giving first birth, working status, and women's empowerment within the household.

Historical ethnic characteristics.—The fact that ancient ethnic beliefs and practerises might have entagled with religious beliefs, may make it hard to distinguish the effect of religion from the effect of belonging to a certain ethnic group. Therefore, the use of ethnic fixed efffects is highly demanding and have the advantage of controlling for all ethnic characterics that might have been a priroy to a prticular religious denomination. Nevertheless, ancestral ethnic featuress have been documented to have an effect in shaping contemporary outcomes, in particular related to development (Michalopoulos et al. (2016); Michalopoulos and Papaioannou (2013)). Particular relevant to our interest is how some of these ethnic dimensions matter also for contemporary gender roles and women empowerment. For example, Alesina et al. (2016) show that part of the variation in domestic violence in Africa can be explained by differential pre-colonial ethnic features such as dependence on female-versus male-dominated economic activities, endogamy versus exogamy, brideprice versus dowry, the practice of polygamy, and the use of the plough.

To this end, we supplement our main specification in equation (3) with a set of ethnic group-level ancestral covariates which includes intensity of women's historical participation in agriculture; ancestral polygyny; ancestral bride price; ancestral plough use; ancestral pastoralism; ancestral presence of clans; indicators of ancestral settlement patterns; indicators of ancestral juridictional hierarchies beyond local communities; ancestral reliance on hunting; ancestral reliance on fishing; ancestral reliance on gathering; ancestral reliance on animal husbandry; ancestral reliance on agriculture; ancestral presence of large domesticated animals; indicators of intensity of ancestral agriculture; and year of observation of the ethnic group in the Ethnographic Atlas. The results are reported in column 2 in Table 11 and show that the estimated coefficient of muslim indicator remain positive, stable in mahnitude and statitical significant indicating that the effect persists when controlling for ancestral ethnic characteristics.

7 Conclusion

Africa has the largest fertility rates in the world and its population is projected to experience rapid growth in the near future, with potentially dramatic economic, social, and political consequences. In particular, there have been consistent differences in fertility rates across religious groups, in particular, Muslims having larger families than Christians. As cultural, specifically, religious identity in Africa is important and often times shapes the nature of political discourse and conflicts, understanding the roots of such differences is obviously important.

In this paper, we make a step in this direction arguing that the historic spread of Islam in the continent has cast a long shadow on the variation in contemporary fertility behavior. Utilizing the DHS dataset that compares women belonging to the same ethnic group, we provide evidence that not only are Muslim families larger than Christian ones, but also that the gender roles within Muslim households are tilted against the women, resulting is lower education levels, age of marriage, and a lesser say in intrafamily decision making. We also find that these effects are especially pronounced in the context of households residing in areas with a Muslim majority, whereas in areas with a Muslim minority these result cease to uphold. More generally, the paper's contribution can be seen as providing support for the view that culture has a persistent effect on gender roles and associated outcomes.

References

Alesina, A., Hohmann, S., Michalopoulos, S., and Papaioannou, P., 2020, "Religion and Educational Mobility in Africa," mimeo.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005). "Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools," Journal of political economy 113 (1), 151–184.

Anderson, S., "Legal Origins and Female HIV" American Economic Review, 2018, 108 (6): 1407-39.

Arold, B.W., L. Woessmann, and L. Zierow, 2022, "Can Schools Change Religious Attitudes? Evidence from German State Reforms of Compulsory Religious Education," CESifo WP.

Barro, Robert J. and Rachel M. McClearly. 2003, "Religion and Economic Growth across Countries." American Sociological Review 68 (2):760-781.

Baudin, T., 2015, "Religion and Fertility: The French Connection," Demographic Research, 32(13) : 397-420.

Bauer, Vincent, Melina Platas, and Jeremy Weinstein. 2022. "The Historical Legacies of Islamic Rule in Africa," World Development, 152.

Becker, G. (1960) "An Economic Analysis of Fertility," in Ansley J. Coale, ed., Demographic and economic change in developed countries. Princeton, NJ: Princeton University Press.

Becker, Gary. 1965. "A Theory of the Allocation of Time," Economic Journal, 75(299), 493–517.

Becker, Sascha O., and Ludger Woessmann. 2009. "Was Weber wrong? A human capital theory of Protestant economic history." Quarterly Journal of Economics 124 (2): 531-596.

Becker, Sascha O., Jared Rubin, and Ludger Woessmann. 2021. "Religion in economic history: A survey." In The Handbook of Historical Economics, edited by Alberto Bisin and Giovanni Federico, 585-639. London: Academic Press.

Bisin, A. and T. Verdier, 2011, “The Economics of Cultural Transmission and Socialization,” in Jess Benhabib, Alberto Bisin, and Matt Jackson, eds., Handbook of Social Economics, New York: North-Holland, 339–416.

Chatterjee, S., and T. Vogl. 2018. “Escaping Malthus: Economic Growth and Fertility Change in the Developing World” American Economic Review, 108(6): 1440–1467.

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012). ”Plausibly exogenous,” Review of Economics and Statistics 94 (1), 260–272.

Coşgel, M., Hwang, J., Miceli, T. and Yıldırım, S. (2019). “Religiosity: Identifying the effect of pluralism”. Journal of Economic Behavior & Organization, 158, 219-235,

De la Croix, D. and Doepke, M. (2003), “Inequality and Growth: Why Differential Fertility Matters”, American Economic Review, 93, 1091–1113.

De la Croix, D. and Perrin, F. (2018). ‘French fertility and education transition: rational choice vs. cultural diffusion’, European Economic Review, 108, pp. 221–45.

Easterly, William and Ross Levine (1997). “Africa’s Growth Tragedy: Policies and Ethnic Divisions”. Quarterly Journal of Economics, 112, 1203–1250.

Fage, J.D. (ed). The Cambridge History of Africa, Vol. 2. Cambridge University Press, 2001.

Franck, Raphael and Ilia Rainer (2012). “Does the Leader’s Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa”. American Political Science Review 106, 294–325.

Galor, Oded. 2011. Unified Growth Theory. Princeton, NJ: Princeton University Press.

Giavazzi, F., I. Petkov, and F. Schiantarelli, “Culture: Persistence and Evolution,” 2020, Journal of Economic Growth, forthcoming.

Giuliano, P. and N. Nunn, 2020, ”Understanding Cultural Persistence and Change”, Review of Economic Studies, forthcoming.

Guiso, L., P. Sapienza, and L. Zingales (2006). “Does Culture Affect Economic Outcomes?” Journal of Economic Perspectives, 20(2): 23–48.

Hackett, C., Connor, P., Stonawski, M., Skirbekk, V., 2015, The future of world religions: Population growth projections, 2010-2050, Pew Research Center.

Hrbek, I. (ed). UNESCO General History of Africa, Vol. III, Abridged Edition. University of California Press, 1992.

Iannaccone, Laurence R. 1998. "Introduction to the economics of religion." Journal of Economic Literature 36 (3): 1465-1496.

Kuran, T., 2018. "Islam and Economic Performance: Historical and Contemporary Links." Journal of Economic Literature 56 (4):1292-1359.

Kramon, E. and D.N Posner (2016). "Ethnic Favoritism in Education in Kenya". Quarterly Journal of Political Science 11, 1–58.

Kuran, T., 2018. "Islam and Economic Performance: Historical and Contemporary Links." Journal of Economic Literature 56 (4):1292-1359.

La Ferrara, E., Chong, A. and Duryea, S. (2012). "Soap operas and fertility: evidence from Brazil," American Economic Journal: Applied Economics, vol. 4(4), pp. 1–31.

McCauley, J. F. 2014. "Identity, Inequality, and Legitimacy: Religious Differences in Primary School Completion in Sub-Saharan Africa." American Political Science Review 108, 801-816.

Meon, P.-G. and I. Tojerow, 2019, "The minority ethic: rethinking religious denominations, minority status, and educational achievements across the globe," Journal of Comparative Economics, 47, 196-224.

Michalopoulos, S. and E. Papaioannou, 2020. "Historical Legacies and African Development." Journal of Economic Literature 58 (1):53-128.

Michalopoulos, S., A Naghavi, and G Prarolo, 2019, "Trade and Geography in the Spread of Islam," Economic Journal 128 (616), 3210-3241.

Miguel, E. and M.K. Gugerty (2005). "Ethnic diversity, social sanctions, and public goods in Kenya". Journal of Public Economics 89, 2325–2368.

Njoh, A.J. and Akiwumi, F.A. 2012, "The Impact of Religion on Women Empowerment as a Millennium Development Goal in Africa," Social Indicators Research 107(1):1-18.

Nunn N, Wantchekon L. The Slave Trade and the Origins of Mistrust in Africa. American Economic Review. 2011; 101 (7): 3221-3252.

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

Platas, M. R. 2018. "Culture and the Persistence of Educational Inequality: Lessons from the Muslim-Christian Education Gap in Africa."

Platas, M. R. 2021, "Islam and Economic Development in Sub-Saharan Africa", Oxford Handbook on Politics in Muslim Societies, edited by M. Cammett and P. Jones (forthcoming)

Spolaore, E. and R. Wacziarg, 2021, "Fertility and modernity," forthcoming, Economic Journal. United Nations, Population Division, 2019, World Population Prospects, New York, NY.

Tables

Table 1: Effect of religion on fertility - Children ever born for women $\geq$ 45 years old

	(1) OLS	(2) First stage - OLS	(3) Second stage 2SLS
	Children ever born	Muslim	Children ever born
Muslim	0.459***		0.081***
	(0.124)		(0.027)
Distance to roman trade roats (log)		-0.012***	
		(0.003)	
Underidentification test (p-value)			0.21
First stage F-statistic (Kleibergen-Paap rk F-statistic)			19.46
Anderson-Rubin Wald test (p-value)			0.00
Anderson-Rubin Wald test (p-value)			0.00
Stock-Wright LM S statistic (p-value)			0.00
Observations	47,601	47,601	47,601
Adjusted R-sqaured	0.241		
Controls	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes

Dependent variable is cildren ever born for women over 45 years old. The method of estimation in column 3 is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 2: Effect of religion on fertility - Robustness checks - All sample

	(1) children born alive	(2) age $\geq$ 47	(3) age $\geq$ 49	(4) No Polygamy	(5) Other religions
Muslim	0.052**	0.106***	0.169**	0.091***	0.096**
	(0.024)	(0.028)	(0.074)	(0.032)	(0.035)
Animist					0.031***
					(1.072)
No religion					0.254
					(0.280)
Other religion					0.817**
					(0.390)
Underidentification test (p-value)	0.214	0.292	0.398	0.220	0.232
First stage F-statistic (Kleibergen-Paap rk F-statistic)	19.477	15.494	5.480	16.915	14.359
Anderson-Rubin Wald test (p-value)	0.01	0.00	0.00	0.00	0.00
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.00	0.00	0.00
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00
Observations	47,642	25,822	8,141	38,359	49,174
Controls	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes

Dependent variable is children ever born, except in column 1 in which the dependent variable is children born a live. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 3: Effect of religion on fertility - Different distance cutoffs

	(1) $\leq$ 20 km	(2) $\leq$ 30 km	(3) $\leq$ 40 km	(4) $\leq$ 50 km	(5) $\leq$ 60 km	(6) $\leq$ 70 km	(7) $\leq$ 80 km	(8) $\leq$ 90 km	(9) $\leq$ 100 km	(10) $\leq$ 150 km
Muslim	0.011	0.101**	0.132**	0.118**	0.124*	0.128**	0.105**	0.092***	0.092***	0.088***
	(0.013)	(0.038)	(0.052)	(0.042)	(0.066)	(0.058)	(0.038)	(0.030)	(0.029)	(0.029)
First stage F-statistic (Kleibergen-Paap rk F-statistic)	25.243	16.303	11.731	12.515	4.684	8.130	14.732	18.849	19.413	24.664
Underidentification test (p-value)	0.226	0.302	0.352	0.323	0.363	0.323	0.290	0.295	0.288	0.248
Anderson-Rubin Wald test (p-value)	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Anderson-Rubin Wald test (p-value)	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Stock-Wright LM S statistic (p-value)	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	10,463	18,672	24,669	29,124	32,376	34,794	37,303	39,208	40,659	44,935
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable is cildren ever born for women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 4: Minority status - restricted 40 km

	(1) $\leq$ 10%	(2) $\leq$ 20%	(3) $\leq$ 30%	(4) $\leq$ 40%	(5) $\leq$ 50%	(6) $\geq$ 50%	(7) $\geq$ 60%	(8) $\geq$ 70%	(9)
Muslim	2.071	0.554	-0.731	-0.104	-0.095	0.148**	0.178*	0.187**	0.257***
	(3.409)	(0.989)	(2.017)	(0.224)	(0.154)	(0.060)	(0.088)	(0.079)	(0.074)
Underidentification test (p-value)	0.573	0.405	0.526	0.211	0.224	0.206	0.203	0.151	0.194
First stage F-statistic (Kleibergen-Paap rk F-statistic)	0.451	0.688	0.238	0.908	1.136	13.030	6.661	8.553	18.853
Anderson-Rubin Wald test (p-value)	0.09	0.26	0.31	0.47	0.32	0.00	0.00	0.00	0.00
Anderson-Rubin Wald test (p-value)	0.06	0.22	0.27	0.44	0.29	0.00	0.00	0.00	0.00
Stock-Wright LM S statistic (p-value)					0.10	0.00	0.00	0.00	0.00
Observations	14,865	16,054	16,764	17,200	17,596	7,031	6,686	6,428	6,114
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Dependent variable is children ever born for women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 5: Religiosity - World Value Surveys

	All sample			Women only		
	(1)	(2)	(3)	(4)	(5)	(6)
Muslim	0.385**	0.391**	1.127*	0.230**	0.230**	0.609
	(0.084)	(0.070)	(0.280)	(0.109)	(0.108)	(0.382)
Religiousity index		0.391*	0.606*		-0.004	0.117
		(0.115)	(0.194)		(0.160)	(0.176)
Muslim $\times$ Religiousity index			-0.806			-0.415
			(0.278)			(0.383)
Observations	16,605	16,605	16,605	8,496	8,496	8,496
No. of regions	142	142	142	142	142	142
Adjusted R-sqaured	0.469	0.470	0.470	0.530	0.530	0.530
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at ethnicity and region level. Individual controls include mother's age, birth cohort and month of survey. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 6: Spatial sorting

	(1) All	(2) Diff. > 0	(3) Diff. < 0
Muslim	0.184*	0.299**	0.076
	(0.094)	(0.144)	(0.135)
difference in % of muslims	78.031*	311.300**	-17.983
	(46.257)	(128.279)	(62.526)
Observations	68,202	36,927	31,267
Adjusted R-sqaured	0.171	0.143	0.193

Standard errors clustered at region level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 7: Spatial RDD

	(1) Baseline	(2) square RD	(3) cubic RD
Muslim	0.201***	0.154***	0.111***
	(0.058)	(0.051)	(0.037)
Underidentification test (p-value)	0.356	0.297	0.249
First stage F-statistic (Kleibergen-Paap rk F-statistic)	11.650	10.606	14.814
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.00
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.00
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00
Observations	26,267	26,267	26,267
Controls	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes

Standard errors clustered at region level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 8: Spatial RDD - robustness checks

	(1) $\leq$ 20 km	(2) $\leq$ 30 km	(3) $\leq$ 40 km	(4) $\leq$ 50 km	(5) $\leq$ 70 km	(6) $\leq$ 90 km	(7) $\leq$ 110 km
Muslim	0.093**	0.196***	0.222***	0.172***	0.154**	0.109***	0.122**
	(0.036)	(0.043)	(0.056)	(0.045)	(0.066)	(0.038)	(0.044)
Underidentification test (p-value)	0.237	0.257	0.326	0.301	0.331	0.305	0.306
First stage F-statistic (Kleibergen-Paap rk F-statistic)	19.433	36.497	19.339	16.530	5.461	11.175	11.390
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Observations	12,486	22,252	29,242	34,361	40,880	45,928	49,087
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors clustered at region level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 9: Driving mechanisms - restricted 40 km

	(1) Education	(2) Work	(3) Age 1st Marriage	(4) Age 1st Birth	(5) Child mortality	(6) Urban	(7) Wealth
Muslim	-0.341***	-0.019***	-0.524***	-0.368***	0.055**	-0.039***	-0.638**
	(0.103)	(0.005)	(0.141)	(0.092)	(0.026)	(0.006)	(0.241)
Underidentification test (p-value)	0.337	0.341	0.336	0.334	0.337	0.337	0.357
First stage F-statistic (Kleibergen-Paap rk F-statistic)	20.177	21.795	20.805	21.298	20.129	20.144	14.042
Anderson-Rubin Wald test (p-value)	0.01	0.00	0.00	0.00	0.06	0.33	0.00
Anderson-Rubin Wald test (p-value)	0.01	0.00	0.00	0.00	0.05	0.31	0.00
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Observations	29,241	28,409	28,794	28,348	29,176	29,245	24,931
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country $\times$ Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 10: Driving mechanisms - restricted 40 km

	(1) Married	(2) Former married	(3) Never married	(4) Women empowerment	(5) Polygamy	(6) Current contracept.	(7) Previous contracept.	(8) Never contracept.
Muslim	0.969***	-0.945***	-0.024	-0.043**	0.742	0.053	0.310	-0.363
	(0.310)	(0.292)	(0.056)	(0.017)	(1.972)	(0.577)	(0.501)	(0.461)
Underidentification test (p-value)	0.337	0.337	0.337	0.362	0.522	0.338	0.338	0.338
First stage F-statistic (Kleibergen-Paap rk F-statistic)	20.144	20.144	20.144	16.658	1.574	19.917	19.917	19.917
Anderson-Rubin Wald test (p-value)	0.03	0.03	0.69	0.00	0.60	0.92	0.50	0.42
Anderson-Rubin Wald test (p-value)	0.02	0.01	0.68	0.00	0.58	0.93	0.48	0.44
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.29	0.00	0.00	0.70	0.00	0.03
Observations	29,245	29,245	29,245	19,011	16,078	29,062	29,062	29,062
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 11: Driving mechanisms - restricted 40 km

	(1) BMI	(2) Rohre	(3) Access to health facility	(4) Partner's education	(5) Partner's work	(6) Vaccination (child)
Muslim	-0.880***	-0.762**	-0.546	-0.415***	-0.612***	0.065
	(0.285)	(0.275)	(0.363)	(0.093)	(0.092)	(0.201)
Underidentification test (p-value)	0.328	0.328	0.343	0.322	0.330	0.452
First stage F-statistic (Kleibergen-Paap rk F-statistic)	22.809	22.809	21.808	37.100	29.007	0.676
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.21	0.00	0.00	0.74
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.18	0.00	0.00	0.72
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00	0.00
Observations	17,239	17,239	28,077	25,874	25,679	3,573
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes

The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 12: Other indicators for women empowerment - restricted 40 km

	(1) Desired Women	(2) Desired men	(3) Wants same number	(4) Husband want more	(5) Want another child	(6) Undecided
Muslim	0.028**	0.089	-0.027***	0.006*	-0.0004	0.0002
	(0.024)	(0.076)	(0.007)	(0.004)	(0.002)	(0.049)
Underidentification test (p-value)	0.357	0.204	0.351	0.351	0.363	0.363
First stage F-statistic (Kleibergen-Paap rk F-statistic)	10.633	19.705	14.997	14.997	13.549	13.549
Anderson-Rubin Wald test (p-value)	0.00	0.23	0.00	0.18	0.80	0.62
Anderson-Rubin Wald test (p-value)	0.00	0.18	0.00	0.14	0.79	0.60
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00	0.00
Observations	22,050	2,847	16,392	16,392	23,283	23,283
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes

The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table 13: LASSO and ethnic characteristics - restricted 40 km

	(1) LASSO	(2) 2SLS
Muslim	0.024	0.094***
	(0.033)	(0.029)
Lambda	0.121	
R-sqaured	0.3716	
Underidentification test (p-value)		0.12
First stage F-statistic (Kleibergen-Paap rk F-statistic)		5.32
Anderson-Rubin Wald test (p-value)		0.00
Anderson-Rubin Wald test (p-value)		0.00
Stock-Wright LM S statistic (p-value)		0.00
Observations	21,550	19,792
Controls	Yes	Yes
Ethnicity FE	Yes	Yes
Region FE	Yes	Yes
Country × Survey Year FE	Yes	Yes

Dependent variable is children ever born. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include longtitude, latitude, distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Appendix

A. Model

Consider an economy that consists of a continuum of households indexed i that operates in discrete time t and is divided into two cultural communities, $j = A, B$. For simplicity, suppose that the communities are initially internally homogenous, which will imply that the same holds in future periods and let y_{jt} denote their respective income; we assume that initially these incomes are identical, $y_{A0} = y_{B0}$.[1] In each period, parents in each community divide income between family consumption c_{jt}; the number of children, $n_{(j,t+1)}$; and their (common) education level, $s_{(j,t+1)}$, respecting the budget constraint:

$$y_{jt} = c_{jt} + C(y_{jt})n_{(j,t+1)} + s_{(j,t+1)} \tag{A1}$$

where $C(y_{jt}), C' > 0$ is the opportunity cost of child rearing.

Education is translated into human capital, which is used to generate future income:

$$y_{(j,t+1)} = s_{(j,t+1)}^{\gamma} \tag{A2}$$

With $0 < \gamma < 1$. The parents derive utilities from family consumption, the number of children, and their (common) education level; the difference between the two communities is in the relative weight on the last two components. We write:

$$U_{jt} = log(c_{jt}) + \alpha_j log(n_{(j,t+1)}) + (1 - \alpha_j)log(s_{(j,t+1)}), 0 < \alpha_j < 1 \tag{A3}$$

where, for concreteness, $\alpha_A > \alpha_B$; this parameter captures the differential tradeoff between fertility and schooling. Parents face the budget constraint (A1) and the production function defined in (A2)-(A3). In each period, they choose household income allocations, taking the economy's history as given. The equilibrium is defined as an optimal sequence of one-period individual choices $c_{jt}, n_{(j,t+1)}, s_{(j,t+1)}$. Utility maximization subject to the constraints (A1)-(A2) in each period yields:

$$c_{jt} = \frac{y_{jt}}{2} \tag{A4}$$

$$n_{(j,t+1)} = \frac{y_{jt}}{C(y_{jt})} \frac{\alpha_j}{2}$$

$$s_{(j,t+1)} = y_{jt} \frac{1 - \alpha_j}{2}$$

[1] With a slight abuse of notation, we skip the index i.

And so

$$y_{(j,t+1)} = (y_{jt}\frac{1-\alpha_j}{2})^{\gamma} \tag{A5}$$

Suppose that $\frac{y_{jt}}{C(y_{jt})}$ decreases with income; this would be the case when the opportunity cost increases with income steeply enough, or when the elasticity of $C(y_{jt})$ is greater than unity. This then implies that fertility decreases in income. We define a steady state as a sequence of one-period equilibria, (c_j, n_j, s_j, y_j), that is constant over time. It follows from (A5) that

$$y_j = (\frac{1-\alpha_j}{2})^{\frac{\gamma}{1-\gamma}} \tag{A6}$$

And so:

$$n_j = \frac{y_j}{C(y_j)}\frac{\alpha_j}{2} \tag{A7}$$

$$s_j = y_j\frac{1-\alpha_j}{2}$$

It then follows that in the steady state community A's members, which have higher preferences for fertility, have higher fertility levels and lower income and education levels than community B's members. While the above result is a direct consequence of the assumed preference differences across the two population groups, a different interpretation is possible. Thus, suppose that all households' preferences are identical and are given as follows:

$$U_{jt} = log(c_{jt}) + \alpha(n_{jt})log(n_{(j,t+1)}) + (1-(n_{jt}))log(s_{(j,t+1)})$$

where $\alpha(n_{jt})$ is an increasing function. This implies that fertility preferences exhibit inertia and increase in the number of siblings. We can then proceed as in the above analysis, replacing in all derivations α_j with $\alpha(n_{jt})$. We will then, in particular, obtain the following in the steady state:

$$y_j = (\frac{(1-\alpha(n_j)}{2})^{\frac{\gamma}{1-\gamma}} \tag{A8}$$

$$n_j = \frac{y_j}{C(y_j)}\frac{\alpha(n_j)}{2} \tag{A9}$$

In general, the system (A8)-(A9) admit multiple solutions, convergence to which depends on initial conditions, specifically, on initial fertility levels. Hence, another interpretation of our result is that initial conditions, in particular, initial fertility differentials across the population groups can be persistent.

Table A1: Summary statistics - women $\geq$ 45 years old

	N	Mean	SD	Min	Max
Total children ever born	47601	6.13	2.99	0	21
Muslim	47601	0.35	0.41	0	1
Christian	47601	0.62	0.39	0	1
Animist	47601	0.02	0.12	0	1
Age	47601	46.82	1.47	45	49
Cohort	47601	1,963.36	5.13	1947	1973
Never married	47601	0.01	0.12	0	1
Currently married	47601	0.76	0.42	0	1
Formerly married	47601	0.22	0.42	0	1
Urban	47601	0.32	0.47	0	1
Rural	47601	0.68	0.47	0	1
Latitude	47601	3.44	15.13	-30.52	31.58
Longtitude	47601	22.64	13.49	-4.24	50.28
Distance to trade routes (log)	47601	6.44	1.56	.357	7.92
Distance to protestant missions (log)	47601	3.66	1.39	-1.72	7.175
Distance to catholic missions (log)	47601	4.99	1.67	-.09	7.59
Distance to coast	47601	509.88	387.10	1.76	1750.92
Distance to capital	47601	283.57	273.65	3.98	1910.54
Water area	47601	0.34	0.47	0	1
Elevation	47601	750.95	675.88	-494.43	3311.88
Soil suitability	47601	0.42	0.26	.001	.99
Diamond region	47601	0.01	0.08	0	1
Oil region	47601	0.26	0.44	0	1
Ecological zones					
Cropland	47601	0.04	0.20	0	1
Desert/Boreal	47601	0.03	0.16	0	1
Intenstive Cropland	47601	0.09	0.28	0	1
Steep terrain	47601	0.01	0.08	0	1
SubTropical/Warm/nohumid	47601	0.00	0.00	0	0
SubTropical/cool/nohumid	47601	0.02	0.15	0	1
Tropical/humid	47601	0.28	0.45	0	1
Tropical/nohumid	47601	0.38	0.49	0	1
Urban	47601	0.15	0.35	0	1
Water	47601	0.00	0.05	0	1
Wetland	47601	0.01	0.09	0	1
cold/wet	47601	0.00	0.01	0	1
Region indicator					
North	47601	0.14	0.35	0	1
South	47601	0.19	0.39	0	1
West	47601	0.32	0.46	0	1
East	47601	0.26	0.44	0	1
Central	47601	0.09	0.29	0	1

Table A2: Adding mechanisms - IV - restricted 40 km

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Muslim	0.132**	0.117***	0.048**	0.076***	0.102***	0.106***	0.119***	0.120***	0.024
	(0.051)	(0.034)	(0.023)	(0.022)	(0.029)	(0.027)	(0.031)	(0.035)	(0.019)
years of schooling		-0.041							-0.020**
		(0.025)							(0.009)
work		0.217							-0.001
		(0.149)							(0.068)
Age at 1st marriage			-0.010						-0.007
			(0.013)						(0.011)
Age at 1st birth			-0.202***						-0.153***
			(0.009)						(0.008)
women_empowerment1				-0.012					-0.026
				(0.045)					(0.019)
Child mortality					0.895***				0.761***
					(0.040)				(0.032)
husband's years of education						-0.012			-0.005
						(0.028)			(0.006)
husband work statuts							0.274*		0.181***
							(0.133)		(0.042)
health								0.111	0.123
								(0.193)	(0.133)
Underidentification test (p-value)	0.357	0.367	0.349	0.359	0.362	0.341	0.340	0.311	0.285
First stage F-statistic (Kleibergen-Paap rk F-statistic)	12.462	11.788	11.709	14.105	11.263	18.602	20.100	21.742	22.348
Observations	24,669	23,826	23,664	17,944	24,669	21,564	21,444	15,568	11,326
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Outcomes are women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table A3: Driving mechanisms - All sample

	(1) Education	(2) Work	(3) Age 1st Marriage	(4) Age 1st Birth	(5) Child mortality	(6) Urban	(7) Wealth (log)
Muslim	-0.161	-0.033***	-0.194**	-0.134**	0.036*	0.011	-0.055**
	(0.104)	(0.009)	(0.087)	(0.061)	(0.018)	(0.009)	(0.019)
Underidentification test (p-value)	0.166	0.199	0.173	0.170	0.166	0.166	0.196
First stage F-statistic (Kleibergen-Paap rk F-statistic)	11.181	12.690	11.607	10.502	11.027	11.171	20.319
Anderson-Rubin Wald test (p-value)	0.01	0.00	0.00	0.00	0.06	0.33	0.00
Anderson-Rubin Wald test (p-value)	0.01	0.00	0.00	0.00	0.05	0.31	0.00
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.00	0.00	0.00	0.01	0.00
Observations	56,169	54,986	55,425	53,731	55,329	56,185	47,946
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Outcomes are women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table A4: Driving mechanisms - All sample

	(1) Married	(2) Former married	(3) Never married	(4) Women empowerment	(5) Polygamy	(6) Current contracept.	(7) Previous contracept.	(8) Never used contracept.
Muslim	0.528*	-0.481	-0.046	-2.901**	0.464	0.063	-0.191	0.128
	(0.308)	(0.296)	(0.047)	(1.076)	(0.645)	(0.251)	(0.295)	(0.305)
Underidentification test (p-value)	0.166	0.166	0.166	0.274	0.200	0.165	0.165	0.165
First stage F-statistic (Kleibergen-Paap rk F-statistic)	11.171	11.171	11.171	24.162	7.357	11.129	11.129	11.129
Anderson-Rubin Wald test (p-value)	0.04	0.05	0.31	0.00	0.25	0.49	0.80	0.47
Anderson-Rubin Wald test (p-value)	0.02	0.04	0.28	0.00	0.25	0.47	0.79	0.45
Stock-Wright LM S statistic (p-value)	0.00	0.02	0.19	0.00	0.25	0.15	0.58	0.42
Observations	56,185	56,185	56,185	37,316	35,388	55,190	55,190	55,190
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Outcomes are women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

Table A5: Driving mechanisms - All sample

	(1) BMI	(2) Rohre	(3) Access to health facility	(4) Partner's education	(5) Partner's work	(6) Children vaccination
Muslim	-0.012***	-0.011***	-0.320	-0.234**	-0.897***	-0.015
	(0.004)	(0.003)	(0.553)	(0.106)	(0.298)	(0.105)
Underidentification test (p-value)	0.279	0.279	0.208	0.221	0.229	0.723
First stage F-statistic (Kleibergen-Paap rk F-statistic)	21.024	21.024	13.637	13.443	13.844	0.135
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.52	0.00	0.00	0.89
Anderson-Rubin Wald test (p-value)	0.00	0.00	0.51	0.00	0.00	0.88
Stock-Wright LM S statistic (p-value)	0.00	0.00	0.14	0.00	0.00	0.83
Observations	33,571	33,571	53,513	50,402	50,375	8220
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Ethnicity FE	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Country × Survey Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Outcomes are women over 45 years old. The method of estimation is 2SLS with muslim indicator instrumented by (log) distance to roman trade routes before 600 A.D. Individual and houshold controls include mother's age, birth cohort, marital status, month of survey, urban dummy and household wealth. Geographic controls at the region-ethnic level include distance to the capital, distance to coast, (log) distance to protestant and catholic missions, elevation, soil suitability, night-time lights, climate indicators, presence of water, diamond and oil. Standard errors clustered at ethnicity and country level. Significantly different from zero at *10% significance, **5% significance level, ***1% significance level.

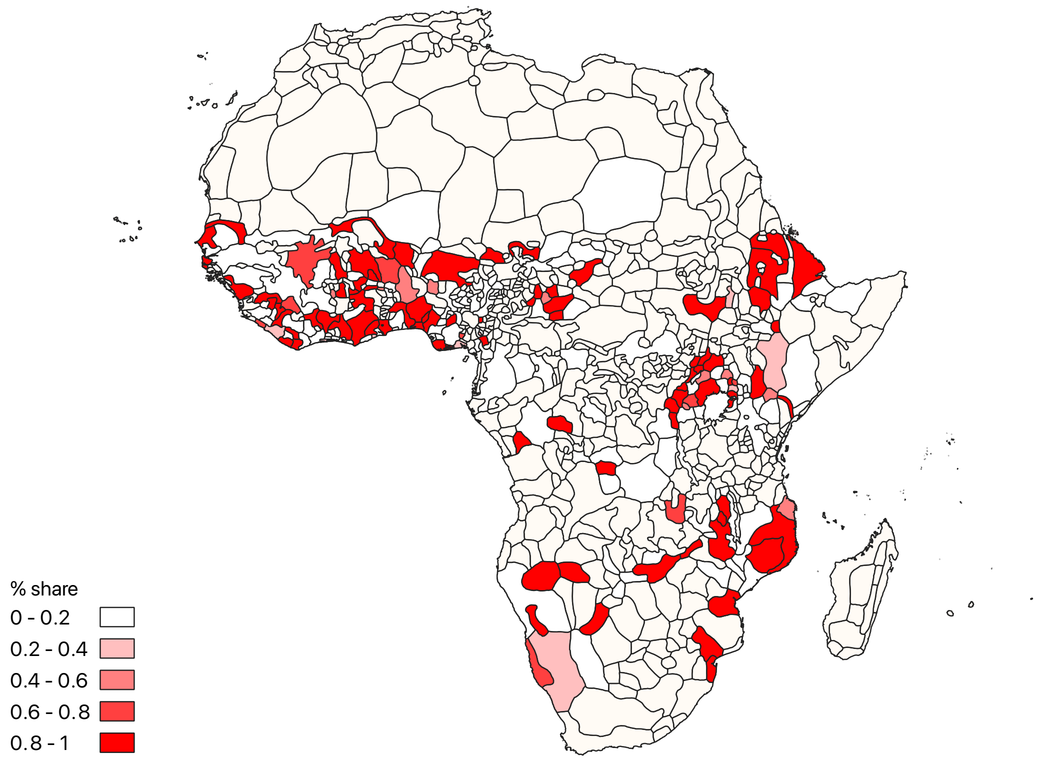


Figure A1: Share of DHS reported ethnicities residing in their historical ethnic homelands

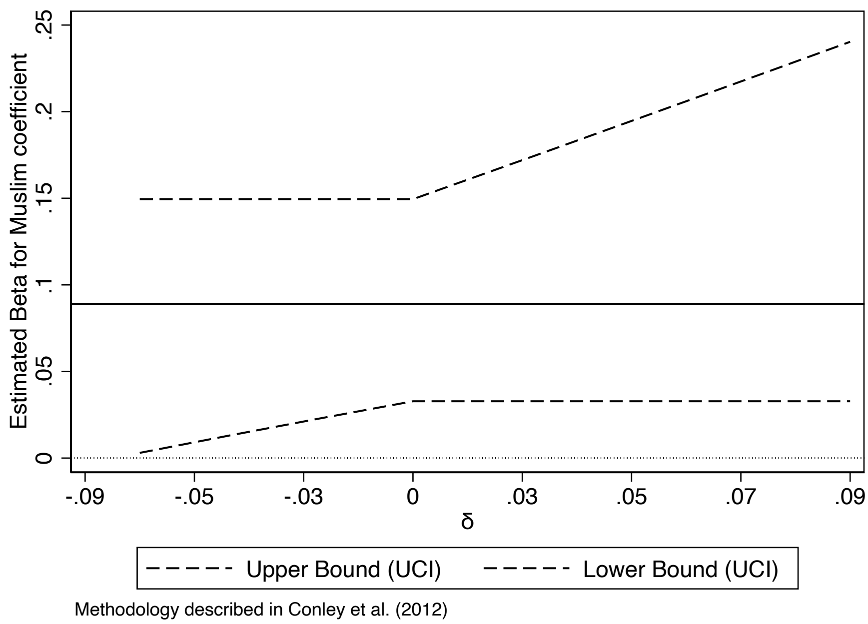


Figure A2: The union of confidence intervals (UCI) approach of Conley et al. (2012).

Note: The solid line is the estimated Muslim coefficient from column 3 in Table(1) and dashed lines are the 95% confidence intervals.

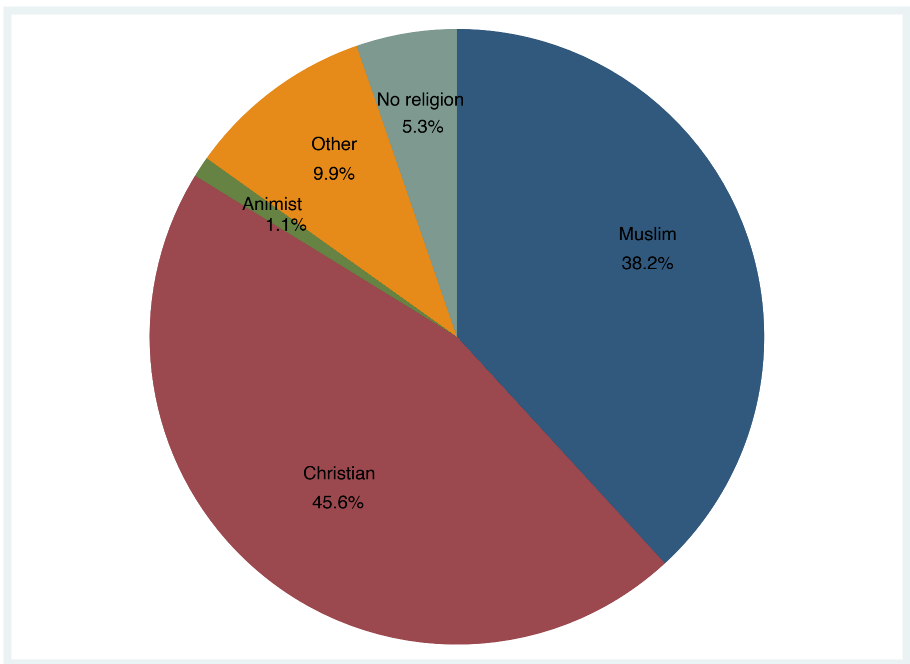


Figure A3: Shares of religious denominations - WVS

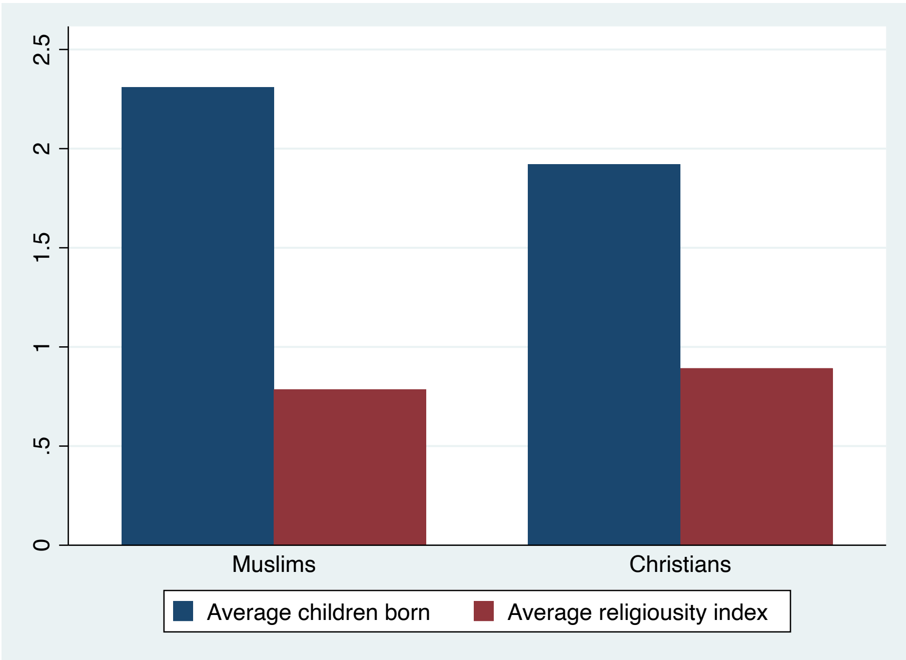


Figure A4: Average number of children and religiosity index - WVS