



# Interethnic and interfaith marriages in sub-Saharan Africa

Juliette Crespín-Boucaud

## ► To cite this version:

Juliette Crespín-Boucaud. Interethnic and interfaith marriages in sub-Saharan Africa. 2018. halshs-01834808v1

**HAL Id: halshs-01834808**

**<https://shs.hal.science/halshs-01834808v1>**

Preprint submitted on 11 Jul 2018 (v1), last revised 30 Sep 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## **WORKING PAPER N° 2018 – 37**

### **Interethnic and interfaith marriages in sub-Saharan Africa**

**Juliette Crespin-Boucaud**

**JEL Codes: J15, J12, Z12 , N37.**

**Keywords : Ethnicity, Marriage, Religion, Sub-Saharan Africa**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS

TÉL. : 33(0) 1 80 52 16 00=

[www.pse.ens.fr](http://www.pse.ens.fr)

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES  
ÉCOLE DES PONTS PARISTECH – ÉCOLE NORMALE SUPÉRIEURE  
INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE – UNIVERSITÉ PARIS 1

# Interethnic and interfaith marriages in sub-Saharan Africa

Juliette Crespín-Boucaud \* †

July, 2018

## Abstract

How does the salience of ethnic cleavages vary with time in Africa? This study documents patterns of inter-ethnic marriages and contrasts them with patterns of inter-faith marriages. Using DHS data that spans marriages contracted between 1960 and 2010 in 13 countries, I find significant discrepancies between countries: the share of intermarriage is as low as 10% in DRC and as high as 48% in Zambia. Looking at changes in the extensive margin, I find that the share of interethnic marriages increases with time in half of the countries while inter-faith marriage share does not. At the intensive margin, there are no changes: inter-ethnic marriages happen mostly between partners belonging to groups that are linguistically close to each other. Evidence suggests that urbanization and changes in social norms about marriage are what drive the increase in the share of interethnic unions. Inter-faith marriages remain rare, which indicates that religious cleavages are either extremely salient or that partners convert before marrying.

**JEL Classification:** J15, J12, Z12 , N37.

**Keywords:** Ethnicity, Marriage, Religion, Sub-Saharan Africa.

---

\*Paris School of Economics (PSE); Address: Paris School of Economics, office R6-01, 48 boulevard Jourdan, 75014 Paris ; E-mail: juliette.crespinboucaud@psemail.eu

†The author is grateful to Denis Cogneau, Sylvie Lambert and Alexander Moradi for insightful discussions on previous versions of this work; and to participants to seminars at the Paris School of Economics (PSE) and to the Monday economics workshop at the University of Sussex. Remaining errors are the author's sole responsibility.

# 1 Introduction

Ethnic fractionalization has been studied extensively over two decades, and deemed to be a key factor to explain economic and political outcomes. Yet little attention has been paid to the mechanisms at play behind “ethnic fractionalization”. How is ethnic identity transmitted from parents to children? Do people live ethnically homogeneous towns or block? Does ethnicity matter on the marriage market?

Ethnolinguistic fractionalization and political competition measures suppose that most individuals only belong to one group. Starting with the seminal paper of Easterly and Levine [1997], many scholars have pointed out the detrimental effects of ethnic diversity (Eifert et al. [2010], Miguel [2004]). A few surveys report that ethnic diversity at local level has a positive effect on economic growth (Montalvo and Reynal-Querol [2016], Gisselquist et al. [2016]). Recent papers (Gershman and Rivera [2018]) have emphasized the importance of deep cleavages to explain the provision of public goods<sup>1</sup>. However, while the literature on the impact of ethnic divisions is vast and growing, quantitative evidence about micro-level interactions is scarce. The Implicit Association Test (IAT) was used to show that individuals hold implicit biases against members of other ethnic groups (Habyarimana et al. [2007] in Uganda; Lowes et al. [2015], in DRC) and it is not clear that such results can be generalized to other countries.

This paper studies household formation and marriage patterns across two sets of identities: ethnicity and religion. Studying interethnic marriages allows us to describe the current state of societies as well as to understand the mechanisms by which identity cleavages gain and lose significance within a society. Ethnic identity differs from other subsets of identity (e.g. religious denomination, language) with respect to the rule for membership, as “membership is determined by attributes associated with, or believed to be associated with, descent” (Chandra [2006]). Bisin and Verdier [2000] show that mixed marriages result in parents from minority groups having access to a weaker socialization technology. Marrying within one’s group is hence a way to transmit one’s identity. While the empirical context differs<sup>2</sup> we nonetheless expect the theoretical result put forward to hold: high shares of intermarriage should be associated with weaker ethnic affiliation for both parents and children.

This paper contributes to two strands of the literature. First, it complements the literature on ethnic fractionalization by stressing that the salience of ethnic cleavages is likely to vary with time, and that ELF measures may need to be updated. Second, it is linked to the literature on social

---

<sup>1</sup>See de la Cuesta and Wantchekon [2016] for a recent review of the state of research on ethnic cleavages in sub-Saharan Africa.

<sup>2</sup>Bisin and Verdier [2000] empirically study the French aristocracy and orthodox Jews which both are cases of minorities in a majority minority setting. In Sub-Saharan Africa, not all intermarriages can be classified as minority-majority unions.

stratification and marriage.

I study interethnic marriage patterns in 13 sub-Saharan countries as a way to understand the role played by ethnic affiliations. Studying unions is a means to measure the degree of segregation of societies, as it combines a measure of segregation (who may meet whom and where) and a measure of who is thought to be an acceptable spouse. If intermarriage shares are low, it means that people either hold strong preferences for marrying within their group, or that they interact mostly with people from their own group - and hence that marriage markets are segregated by ethnicity (Furtado [2012]). I describe patterns across space and time. I then study factors that could explain these patterns: I focus on education and urban residence. Building on recent research on how to measure ethnolinguistic diversity (Desmet et al. [2016], Gershman and Rivera [2018] ), I compute linguistic distance measures for each interethnic union. I use these measures to understand better intermarriages: Do people who marry outside their ethnic group typically marry into a group that is close to their own group, in terms of culture and language? This allows me to look at the extensive margin (marrying outside one's ethnic group) and at the intensive margin (how far outside one's ethnic group?). Finally, I contrast interethnic marriage patterns with interfaith marriage ones.

Using DHS data that spans marriages and unions contracted between 1960 and 2010, I find significant discrepancies between countries: the share of intermarriage is as low as 10% in DRC and as high as 48% in Zambia. These differences cannot be solely attributed to differences in the ethnic structure across countries. Looking at time trends, I show that intermarriage shares increased in half of the countries and remained stable in the other half. Education and urban residence are factors that drive interethnic marriage shares up. However, evidence points at changes in preferences for ethnic homogamy across generations. This change in norms is likely to explain the part of the trend that is not explained by education or increased urbanization. When looking at linguistic distance, I do not find changes in across time.

Interfaith marriages decrease in most countries. This change can be attributed to the fact that followers of traditional religions are fewer and fewer over time, a fact consistent with the well-documented idea that Africa is experiencing a religious resurgence. Muslim/Christian intermarriage is rare and remains so. I do not find evidence that norms around interfaith marriage changed over, though norms and religious practices at the individual level are themselves changing.

The rest of the paper is organized as follows: The next section presents the data and discusses how ethnicity is measured. Section 3 reports main results on intermarriage patterns across countries. Section 4 presents results on how marriage patterns changed, as well as tests of factors that may explain these changes. Section 5 provides robustness analyses of the findings. Section 6 con-

cludes. Appendices present detailed information on data used and explanations on how linguistic distances are computed.

## 2 Data and measure

### 2.1 DHS data

I use Demographic and Health Surveys (DHS) that were implemented between 1992 and 2016 in sub-Saharan Africa. Table 14 (in Appendix A.1) lists all countries in which DHS were conducted, and whether they are part of my sample.

DHS collect two-stage stratified samples of households. All women aged 15-49 are surveyed in all households. Men are surveyed in a fraction of surveyed households<sup>3</sup>. Cohabiting spouses and partners<sup>4</sup> are then matched. Individual survey weights are provided. It is important to note that sample size is not proportional to population size and that the number of waves differs from country to country.<sup>5</sup> DHS are the only surveys spanning the African continent that provides information on ethnicity and on marital history. Afrobarometer surveys include information about ethnic identity but not marital history. However, DHS aim primarily at providing information on demography and health, hence their name and focus on reproductive history rather than on marital history<sup>6</sup>.

In order to present descriptive result across space, I use data from 25 countries for which respondents' ethnic and religious identification is available.

To study changes in time, I use a subsample of the 25 countries. The main sample is made up of 13 countries. Criteria for inclusion in that sample are as follows: First, countries must have implemented at least two survey waves that include ethnic information. Second, ethnic classifications must be comparable across waves. Third, ethnic groups must be ethnolinguistic groups that can be matched to linguistic groups<sup>7</sup> using Ethnologue (Simons and Fennig [2017]). Fourth, surveys must include women born between 1950 and 1991. DHS surveys are not a balanced panel of countries. While surveys are designed to be implemented roughly every 5 years, surveys are

---

<sup>3</sup>The lowest sampling rate of men is 25% (Malawi 2000) . The highest one is 100% (DHS Ghana 2014, DHS Zambia 2013/2014) . Men aged 15-59 are surveyed. I adjust the weights to take into account these differences in survey rates of men.

<sup>4</sup>"Marriage" is used to refer to married couples as well as to cohabiting couples.

<sup>5</sup>DHS analysts advice to use men's weight to analyze couples file as it corrects for men's non-reponse which is slightly higher than women's non-response. However as the sample was designed to be a nationally representative sample of the population of women aged 15-49, I reweight the observations using the women's weights. Reweighting consist in making sure the weights used in a given survey sum up to a set number - and not up to sample size. I then correct for the different number of survey waves, so that each country's weights sum up to its population.

<sup>6</sup>The reason to include ethnicity and religion in DHS-III is listed as follows: "There is considerable evidence that reproductive behavior is influenced by normative attitudes associated with religious values and ethnicity. This information is also of potential pragmatic value in identifying particular groups that may have specific needs."

<sup>7</sup>For instance, DRC and Chad list groups that correspond to geographic areas (e.g. "cuvette central" and "uele lac albert" in DRC). These places are heterogenous in terms of ethnic groups, thus leading me to exclude Chad and DRC from the main sample.

discontinued, or started earlier or later depending on countries. This imbalance is the reason why I restrict the sample to countries where I have observations for all the (birth) years between 1950 and 1991.

The final sample includes Benin, Burkina Faso, Cote d'Ivoire, Gabon, Ghana, Guinea, Kenya, Malawi, Mali, Niger, Senegal, Togo and Zambia.

There are two main limitations to a comparative analysis of DHS data across time and space: First, not all countries gather data on ethnicity. Second, marital history information is limited to the characteristics of the current marriage.

This main sample of 13 countries is not a random sample of African countries, as including a question on the respondent's ethnic identity is not a decision made independently from whether ethnicity matters in a given country. For instance, the questionnaire for DHS Tanzania does not include these questions, consistently with the idea that ethnicity is "depoliticized" in Tanzania (Weber [2010]). DHS Rwanda 1992 included a question on ethnicity, but the survey waves conducted after the 1994 genocide do not include this question. In other countries, the question about ethnicity was included only once (e.g. Central African Republic) or the classification changed across survey waves (e.g. Nigeria).

The second caveat is that information on marital history is limited. To study marriages and what drives matches on the marriage market, one would want to access information on characteristics of both partners at the time of the search and match. DHS gather information on current partnerships and on the current characteristics of both partners (ethnicity, education, occupation...) but, apart from the year of first cohabitation, it does not include retrospective information on marital history. Some characteristics are unlikely to change after marriage: The education level at the time of the survey should be a good proxy of the level of education at the time of the union was formed<sup>8</sup>. However, the current place of residence is unlikely to be the same as the place of residence around the time of marriage. Information on migration history is lacking: we do not have information on where the respondents were born, nor on the places they lived before marrying. The older survey waves listed whether the respondent had grown up in an urban or rural area, but the question was dropped in more recent questionnaires.

A stronger limitation is the fact that, especially in West Africa, women are likely to have entered several (successive) unions while DHS allows us to study only the current union, as no questions about the characteristics of previous partners are included. For women who remarried, I know neither how many unions they entered, nor whether they were divorcees or widows when they

---

<sup>8</sup>Men are expected to be employed in order to provide for their wife (or wives), so they are unlikely to still be students when they marry. It is unlikely women can stay in school after getting married. In my sample, under the assumption that girls start school at 8, 90% of women started their first cohabitant union after they had completed their schooling, 4% at the same time as they completed it, and 6% before that.

remarried<sup>9</sup>. I discuss the impact in subsection 4.3 and in section 5.

## 2.2 Identifying changes in time

### 2.2.1 Measuring ethnic and religious identities

The first issue is to make sure survey waves are comparable across time. The number of interethnic unions varies according to the number of ethnic groups listed in the survey. Imagine a country where two survey waves are implemented, while everyone's marital situation remains the same across waves. In period 1, the survey lists groups A (e.g. Kikuyu), B (e.g. Luhya) and O ("other") and in period 2, it lists groups A, B, C (e.g. Luo) and O ("other"). This change in classification results in a higher share of interethnic marriages in period 2 than in period 1, despite all unions remaining the same. Marriages between A and B, B and O, A and O are counted as interethnic unions in both surveys, but marriages between C and O that were previously counted as intraethnic<sup>10</sup>, are now counted as interethnic marriages. Therefore, in order to make comparisons across time when using several survey waves, classifications must be held constant<sup>11</sup>.

This common classification includes only the groups that are listed in all surveys. I recode ethnic classifications under the assumption that individuals have a preferred answer to the question "what is your ethnic group?", and that this answer is not affected by changes in the classification. If individuals pick an answer that is not in the list (e.g. Maasai), this answer is recoded in "other". I hence assume that a Maasai individual would have been coded as belonging to the "other" group in DHS surveys that do not include this group. Based on that assumption, I assign to all Maasai individuals the identity "other", as the common classification for Kenya does not include the group Maasai.

Members of groups listed in only some of the waves are recoded as belonging to the "other" group. I assume subgroups are recoded into the corresponding group in the classification<sup>12</sup>. Subgroups which are split in some waves and not in others fall under a common category, as we cannot distinguish two groups in surveys with a common entry for both.<sup>13</sup>

I check that the share of each ethnic group within the total population (women) remains roughly constant across cohorts and survey waves, alleviating concerns about misidentification<sup>14</sup>.

---

<sup>9</sup>Divorced women are likely to have more of a say about whether they remarry, and to whom (Thiriat and Locoh [1995]).

<sup>10</sup>As C was part of the group "other", both partners were considered as being part of the same group.

<sup>11</sup>In that country, the common classification would be "A, B, and others".

<sup>12</sup>Early DHS in Ghana list ethnic groups as "1 Asante, 2 Akwapim, 3 Fante, 4 other Akan 5. Ga/Adangbe (etc.)", later waves list only "1. Akan 2. Ga/Adangbe (etc)". I recode all Akan answers into a single category.

<sup>13</sup>Having similar enough classifications across survey waves was one criterion for inclusion in the sample.

<sup>14</sup>Figure 5 (Appendix A.1) shows the share of women that do not belong to one of the groups listed in the common classification



Similar issues arise with respect to religious nomenclatures. Religious classifications change faster than ethnic nomenclatures, perhaps because ethnic groups formation and recognition requires work by local ethnic entrepreneurs while new religious groups/sects may be created, or “imported” from other countries. For instance, the well-documented surge of Pentecostalism in West Africa (Mayrargue [2004]) results in a change in nomenclature. In the case of Benin, the 1996 nomenclature lists 7 groups (Traditional, Moslem, Catholic, Protestant, Other Christian, Other, No religion) whereas the 2011 one lists 10 groups (Vodoun, Other traditional, Islam, Catholic, Methodist, Other Protestant, Celestes (Celestial Church of Christ), Other Christian, Other religion, No religion). Harmonization of nomenclatures across waves requires to use a high level of aggregation: Christian, Muslim, Other (which includes followers of traditional religions, atheists, and members of new religious movements that cannot be linked to Christianity or Islam).

### 2.2.2 Selection into marriage

Sample selection is an additional issue to take into account. For the younger cohort, not all women within this cohort are married by the time of the latest wave of the survey. For the older cohort, selective attrition is a concern. Older women who are still married when they were surveyed are those who survived and whose husband survived<sup>15</sup>. The bias would play against finding a positive trend in interethnic and interfaith marriage shares. Women from older cohorts, in case of selective mortality, would be more likely to be in an interethnic marriage than their counterparts who died or become widows. Women from younger cohorts that are married are more likely to have married younger, and hence to be in more traditional unions than their counterparts that will marry later. This would “flatten out” increases in intergroup marriages.

Additionally, in order to mitigate selection concerns, I use controls for age at the time of the survey. I built my sample in order to include only a time period (birth years 1950-1991) which appear in at least two survey waves in each country.

## 3 Heterogeneity across space

### 3.1 Methodology

Ethnic classifications reflect local choices on what the relevant groups are within a country. Choices made include the level of aggregation chosen<sup>16</sup>. The ethnic structure - the number of groups listed

<sup>15</sup>Indicating that either they were both from a more well-off background, or that the woman could remarry after her husband’s death.

<sup>16</sup>For instance Cote d’Ivoire DHS lists around 40 ethnic groups in two survey waves, but only 5 “aggregated” groups in another. Crespin-Boucaud [2015] studies interethnic patterns in Ghana and shows the impact of classification changes, using GLSS Ghana (groups and subgroups listed) and DHS (groups listed) surveys that were conducted around the same

as well as their respective sizes - of a country also influences the share of interethnic marriages that is computed. Hence, cross-country comparisons must be done carefully, as intermarriage measures depend upon the ethnic structure. In order to neutralize this effect, I compute structural intermarriage rates for each country. These rates are “structural” because they describe the share of intermarriages that would be observed if individuals were matched randomly with each other in a national marriage market. It is equivalent to computing the ethnolinguistic fractionalization index (ELF).

$$ELF_c = 1 - \sum_{k=1}^n (P(X = k))^2$$

This structural share is computed for each country ( $c$ ).  $n$  is the number of ethnic groups in the survey (including “other ethnicity/nationality” as a group).  $P(X = k)$  is the likelihood that the ethnic group to which belongs the individual  $X$  is the group  $k$  (it is also each ethnic group population share). As we suppose that this individual is drawn randomly from the population, this likelihood is equal to the share of the ethnic group  $k$  in the population. Empirically, we identify the share of ethnic group  $k$  with the share of women that belong to the group  $k$ <sup>17</sup>. This measure allows me to compare intermarriage rates with structural rates, thus making cross-country comparisons more meaningful.

This measure does not include any information on how similar ethnic groups are. Since the seminal paper by Greenberg [1956], the literature has suggested adjusting measures of fractionalization by a measure of similarity between groups. Yet, in that case, it is not a concern as I treat the country-level ELF as a benchmark of how the share of intermarriages we observe compares to random matching. I discuss linguistic distance measures in section 4.4.

## 3.2 Results

Figure 1 shows the share of interethnic marriages by country. Striking differences between countries appear: In the Congo, Uganda and Zambia, more than 40% of married women are in an interethnic marriage, whereas this share is lower than 10% in DRC, Namibia and Nigeria. The results do not change much when this share is expressed as the share of the random interethnic marriage we would observe if matches were realized randomly at the country level: The map does not get much darker.

Figure 2 shows the same maps, but for interfaith marriages. In stark contrast to interethnic

---

period.

<sup>17</sup>This hypothesis does not hold if there are discrepancies in the share of married women from a given ethnic group and the share of married men from the same ethnic group. However, these discrepancies are small and for a “structural” measure, this hypothesis remains plausible.

marriage patterns, the share of interfaith marriages is low. However this low share results from the religious make-up of the countries. Most countries have very little religious diversity (Niger or Zambia), hence the much darker shades of the map on the right panel. In Nambia and Zambia, the observed share of interfaith marriages is higher than the share that would be observed under random matching. This indicates that members of religious minorities may experience difficulties in finding a partner within their faith. It is important to note that religious groups may be very small groups (e.g. 0.62% of Nigerien women identify as Christians in 2006) while an ethnic group of the same size is unlikely to be listed in the ethnic classification.

## 4 Time trends

### 4.1 Methodology

First, I study time trends for my sample as a whole, using country fixed effects. Second, I run the same specifications separately for each country, in order to study whether trends and patterns are homogenous across countries.

My baseline specification is a linear probability model<sup>18</sup> at the country level. I define an inter-marriage dummy that describe each observed union.

In that model, individuals (i) are women. As polygyny is prevalent in some countries<sup>19</sup>, it means that while a woman appears only one time, a man appear matched with all his cohabiting wives. Results do not change if I study intermarriage collapsing the dataset to keep one observation per man<sup>20</sup>.

$$Intermarriage_i = \beta_1 * BirthYear_i + \beta_2 * Age_i + \beta_3 * Age2_i + X_i' * \delta + \gamma_c + \epsilon_i \quad (1)$$

$$Intermarriage_i = \alpha_c + \beta_{1c} * BirthYear_i + \beta_{2c} * Age_i + \beta_{3c} * Age2_i + X_i' * \delta_c + \epsilon_i \quad (2)$$

In equations 1 and 2  $intermarriage_i$  is a dummy variable, and  $X_i'$  is a vector of variables at the individual level. These variables include dummies for the highest completed education level (No education/Primary/Secondary and higher education<sup>21</sup>), the type of place of residence (Rural/Urban) and a dummy that indicates that the current marriage is not the respondent's first union.

<sup>18</sup>A quadratic term for birth year is never significant. Results available from the author on request.

<sup>19</sup>In the pooled sample, 20% of women declared having at least one co-wife, and 14% of women appear with at least a co-wife in the dataset. This average hides discrepancies between countries: While (cohabiting) polygyny is rare in Kenya and Zambia, its prevalence is high in West African countries (e.g. 35.5% of Guinean women)

<sup>20</sup>In that case I consider a man to be in an interethnic (interfaith) union if at least one of his wives is not a member of his ethnic (religious) group. Results available from the author.

<sup>21</sup>DHS classification distinguishes between secondary and higher education. However, only 2% of married women in my sample have completed university, so I aggregate secondary education and higher education into a single category.

Equation 2 is a regression that is run at the country-level, allowing me to study each case separately. Equation 1 is a regression on the pooled sample, and includes country-level fixed effects.

An important issue is that several outcomes are jointly determined with marriage (or with marriage projects). Current place of residence is an outcome of marriage, but preferences about residence places also determine one's partner choice. The same is true of education: In many cases, parents choose between sending their daughters to school or marrying them. Hence, the age at marriage (or the year of marriage) is not exogenous to education choices. In order not to bias my estimates, I focus on birth year rather than on marriage year for my main estimates. Marriage year is used as robustness check in section 5.

## 4.2 What could drive time trends?

Two mechanisms could drive the increase in intermarriage shares. The first mechanism rests on a change in norms and values, allowing people to marry outside their group (or allowing people to select their spouse rather than accepting elders' decisions about a potential spouse or being "given to marry"). The second mechanism is a decrease in segregation, or the fact that people from different groups have more opportunities to meet (migrations to cities, longer studies, preferences for similar education levels). Education may play a role in both mechanisms: either it could change agents' preferences, or if the preference for educational homogamy trumps the one for ethnic endogamy, it could also increase the share of interethnic marriages.

## 4.3 Extensive margin - Interethnic marriages

Table 1 shows that interethnic marriage have increased over time<sup>22</sup>. Younger cohorts are more likely to marry outside their ethnic group than older cohorts, even when controlling for age at the time of the survey. This increase can partly be explained by the fact that younger cohorts are more educated and more likely to live in urban areas. The trend remains positive and significant, even once changes in the composition of the population (education, urbanization) are controlled for.

Figure 3 present a visual representation of the change in the share of interethnic marriage between the cohort of women born in the 1960s and the cohort of women born in the 1980s. The share of interethnic marriages is higher for the 1980s cohort than for the the 1960s in all countries, except for Benin, Mali and Niger. Turning to regression analysis, table 2 lists the coefficient  $\beta_1$  of regressions including different subsets of controls. The share of interethnic marriages has significantly increased over time in Burkina Faso, Cote d'Ivoire, Ghana, Guinea, Kenya and Togo.

---

<sup>22</sup>Figure 4 shows the coefficients associated to birth year when this variable is introduced in a non-linear way.

In Burkina Faso and in Kenya, this increase can be attributed to changes in the composition of the population. The increase in the education level of the population explains the trend away in Burkina Faso. In Kenya, the higher share of interethnic marriages among younger generations is explained by urbanization. Yet, in Gabon and Senegal, when controlling for these factors, younger generations are more likely to marry outside of their ethnic group, with most of the effect coming from controlling for remarriage. The reverse happens in Malawi: When controlling for those factors, the trend becomes negative and significant.

Table 3 shows the role played by each factor on the likelihood to be in an interethnic union. Having any kind of formal schooling, living in an urban area at the time of the survey and having remarried are all positively correlated with the likelihood to be in an interethnic union. For women, having completed at least primary education has a positive and significant impact on the likelihood to have a husband who is not a co-ethnic, except in Kenya, Chad and Ghana. Having completed at least secondary school has a positive and significant impact on the likelihood to be married outside of one's ethnic group in all countries. This variable is likely to capture several factors that could favor interethnic unions: Students that attend secondary education are more likely to come from better-off backgrounds, they have learnt the vehicular language used in secondary school. Moreover, especially for students that attended university, schools can be where one meets their partner.

The association of these characteristics with a higher likelihood to in an interethnic union is consistent with several channels. High schools and universities are where many interethnic couples meet<sup>23</sup>. School is likely to be one of the places where a vehicular language can be learnt. Urban areas are more mixed<sup>24</sup> and traditional norms are less likely to hold in these spaces than in the countryside.

When assessing these time trends, a concern is that while comparing older cohorts with younger cohorts, I contrast unions that more likely to be second unions with unions that are more likely to be first unions. If older women who remarry, after divorce or widowhood, marry as the cohort of younger women, then I would be underestimating change across years. Hence, if social norms relax across time, women from a given birth cohort would be more likely to remarry outside their ethnic group than they were when they entered their first union, then this would bias my estimates downward. Another possibility is that interethnic unions are short-lived, hence leading to a higher share of divorce among women who had first marry outside of their group, and these women could then marry within their group, leading us to overestimate the change in time.

<sup>23</sup>Qualitative interviews in Cote d'Ivoire in June 2016, discussions with informants.

<sup>24</sup>I compute ELF indices using geolocalised GPS data. ELF indices are computed for squares of 10km<sup>2</sup>, hence restricting concerns about cluster size. The average urban cluster has an ELF of 0.64 whereas the average rural cluster has an ELF of 0.34.

When comparing the share of interethnic unions among women in their first union and other women currently in union, I find that either the share of interethnic unions is similar in both groups, or that women who remarried are more likely to be in an interethnic union (table 9). One plausible explanation is that older women are the ones who reenter a union and they do not have to follow strict social rules about choosing a spouse. Another is that entering a second union is more difficult than entering a first one, so these women cannot find partners in their own ethnic group, so they marry outside their group. Both are competing explanations that I cannot test for. Yet, when I split the sample between women still in their first union and women who remarried, I do not find a trend for women that remarried, except in Ghana and Guinea (table 4). Coefficients for Gabon, Ghana, Guinea and Kenya are all higher for remarried women.

#### 4.4 Intensive margin - Linguistic distance

The ELF has been criticized for not taking into account the depth of cleavages or differences across groups. Classifying unions into two categories (intermarriage or not) has the same caveats as using the ELF. It lumps together very different situations: In Zambia some ethnic groups speak the same language<sup>25</sup> while some groups (for instance the Somali and the Kikuyu in Kenya) are very distant, according to the linguistic tree. Following recent research on ethnolinguistic diversity (Desmet et al. [2016], Gershman and Rivera [2018]), I compute linguistic distance measures. The methodology is detailed in Appendix B.

Table 5 shows how the linguistic distance (conditional on being in an interethnic union) changes across cohorts. In most countries, the mean linguistic distance has not changed across cohorts. When the mean linguistic distance has varied across time, it does not seem to follow general patterns, or to be associated to the increase in the share of interethnic marriages.

In Ghana and Togo, both countries where the share of interethnic unions has not significantly increased across time, linguistic distance has increased. While interethnic unions have become more frequent in Kenya, the mean linguistic distance of such unions has decreased: Women are more likely to marry outside their group, but they also marry closer to their group than the older cohort did. Finally, in Burkina Faso, the share of interethnic unions has not changed over time, but linguistic distance decreased over time.

Unfortunately, DHS do not include questions on the (main) language spoken at home, contrary to other surveys such as the World Values Survey (for instance used by Desmet et al. [2016]). An extension of this study would be to examine the mechanisms by which ethnicity is passed on to the next generation, especially in the case of children born from parents who do not belong to the

---

<sup>25</sup>This is a well-know example, notably documented by Posner [2005].

same ethnic group.

#### 4.5 Interfaith marriages

Figure 3 represent the change in the share of interethnic and interfaith marriage between the cohort of women born in the 1960s and the cohort of women born in the 1980s, without controls. Whereas interethnic marriages are more frequent, the reverse is true for interfaith unions. However, as discussed below, most of this effect is unlikely to be due to changes in marital preference, or to people becoming more intolerant. This decrease is explained by the fact that traditional religions (classified as “other”) are losing followers across time. As these people were the ones who married outside of their religious group, the share of interfaith marriages is decreasing over time.

Moreover, if ethnicity can be considered to be exogenous<sup>26</sup>, religion is not. It is likely that some people convert in order to get married, which may explain the low share of interfaith unions. Interfaith unions in DHS are hence unions in which spouses kept different faith, or converted after being married. It is hence a lower bound on the share of interfaith unions.

Table 6 shows the coefficient  $\beta_1$  (equation 2). In columns (1) and (2), I consider three faith groups: Muslim, Christian and others. I find that the share of interfaith unions increased in Niger and Togo, while it decreased in Benin, Burkina Faso, Ghana, Malawi and Zambia. The share of interfaith unions remained stable in the other countries.

However, countries in which the share of interreligious unions decreased are also countries in which religious diversity decreased over time, mainly due to traditional religions not having members in the youngest generations in West African countries. Table 8 shows how the likelihood to identify as a member of a religion which is neither Islam or Christianity correlates with birth year and age at the time of the survey. Older cohorts are less likely to be Muslim or Christian. More surprisingly, age effects are also at play: Older women are more likely to be Muslim or Christian than younger women. Put together, these two effects corroborates anthropological works on the religious resurgence: Younger generations are more likely to identify as Muslim or Christian (and hence do not belong/identify as members of traditional religions), and women from older cohorts convert<sup>27</sup> to Islam or Christianity over the course of their lives.

In columns (3) and (4), only Muslim/Christian intermarriages are counted as interfaith marriages. There does not seem to be a change in the share of Muslim/Christian unions: they are rare and remain so. Table 9 shows that the highest share of Muslim/Christian marriages is found in Gabon, where it reaches 3.9%. Only in Ghana does that share decrease with time.

---

<sup>26</sup>I discuss this hypothesis in section 5.

<sup>27</sup>Convert or start identifying as Muslim/Christian. Traditional religion members could have had a proximity to Islam or Christianity and have had practices belonging to both religions.

Table 7 shows how individual characteristics correlate with the likelihood to be in an interethnic union. Strikingly, education variables are negatively correlated with the likelihood to be in an interfaith union in almost half of the sample (Benin, Gabon, Kenya, Malawi, Togo and Zambia). There are positively correlated with such an outcome in Burkina Faso, Cote d'Ivoire and Ghana, countries that have sizable population shares of both Muslims and Christians.

## 5 Robustness analyses

### 5.1 Year of cohabitation

Table 11 shows patterns using year of the first union. Any variation in the share of interethnic unions across birth cohorts is also found using year of the first union. Yet, for Benin, Mali, Senegal and Zambia there is an increase in the share of interethnic marriages only when considering the year of the first union.

The fact that trends are stronger, and exist in more countries, when using marriage year rather than birth year is consistent with the hypothesis that norms change and that women and men have more freedom in choosing their partners, and with the hypothesis that the countries are becoming less ethnically segregated: Using marriage year is a better proxy for norms and for the possibilities to meet partners outside one's group at the time marriage decisions when taken by women (or their older relatives).

Concerning interfaith marriages, using birth year or marriage year yields similar results. In Senegal, the trend for interfaith marriages is positive and significant using marriage year. This effect is driven by a very high share of interfaith marriages in the most recent years. Senegal being one of the countries with the most recent survey wave, this may indicate that people started cohabiting but have yet to convert, or will convert before getting married.

### 5.2 Ethnic identity and time in union

In order to identify changes in time, ethnic identification must be exogenous to union-related characteristics. Especially, in order to be able to make cross-cohorts comparisons and to study variations over time within a given country, ethnic identification must not be correlated to the length of the union. My concern is that women in younger cohorts may declare they are from a different ethnic group from their partner's while women in older cohorts are older and hence have spent more time in their (first) union than younger women, and may more likely to declare the same ethnic group as their husband as they may have "assimilated".

Exploiting the fact that I have several survey waves for each country, I can study whether



women that married for the first time the same year and were born the same year are more (less) likely to report having the same ethnic group as their husband with time spent in union. However, the identification ultimately rests on differences across survey waves - so this will also capture any effect linked to survey wave. If the length of the union is negatively correlated with the likelihood to be in an interethnic union, then it means that either women (or their husbands) start assimilating into the other's ethnic group or that interethnic marriages are more likely to end in divorces<sup>28</sup>.

Table 10 shows that the coefficient of birth year is not affected by this additional control. The coefficient for the number of years spent in union is negative and significant in Ghana, Malawi and Zambia. Yet it is positive in Burkina Faso, Guinea and Senegal<sup>29</sup>: This indicates that it is likely that I am mainly capturing differences between waves that are not due to the length of union in itself, but to measurement error and differences in survey quality across time.

### 5.3 Uncertainty: Others and remarriage

In table 12 (ethnicity) and table 13 (religion), I relax two of the assumptions I made. Column (1) and (2) present results for the main specification. In columns (3) and (4) (table 12), I consider that unions between spouses that both declare they belong to the “other” ethnic (religious) group are intraethnic (intrareligious) unions. This change the results only for Cote d’Ivoire, Gabon and Malawi. In Cote d’Ivoire and Gabon, the share of members of an unclassified ethnic groups is large (see figure 5). Results are consistent with the fact that the share of “others” changed over time in these countries. In columns (3) and (4) (table 13), results do not change when relaxing the “other-other” assumption, except for Cote d’Ivoire.

In columns (5) and (6) of tables 12 and 13, I consider that women that are now remarried were previously in an interethnic (interreligious) union and I hence consider them as being in a interethnic (interreligious) union. This assumption turns the birth year coefficient negative (or changes its magnitude) in most countries. In order to understand why this assumption changes the patterns, further work is required on divorce and widowhood in sub-Saharan Africa. The fact that results change in the same direction for both interethnic and interreligious marriages, and that the effect is stronger for countries in which the difference between share of remarried women across cohorts is the greatest<sup>30</sup> indicates that the effect captured is mostly a change in the share of remarried women across time. If that change in shares of remarried women is due to differences

<sup>28</sup>Several explanations could support this hypothesis. For instance, women who enter interethnic unions are more likely to be educated than their counterparts, and educated women are more likely to ask for divorce.

<sup>29</sup>I do not have an hypothesis that could explain why length of union increases the likelihood to belong to a different group as one's spouse.

<sup>30</sup>In Niger, the share of remarried women (among women in union) is 32% for the 1960s cohort, while it is around 8% for the 1980s cohort.

in mortality patterns across time, then the introduction of age effects would not be able to fully control for these changes.

## 6 Concluding remarks

This paper documents patterns of intermarriage in sub-Saharan Africa, focusing on ethnicity and on religion. I use data from Demographic and Health Surveys that contain information on marital history, education and geographic location. I find the share of interethnic marriages varies greatly between countries, but that such unions are not rare: 21.7% of women are married to a man who is not from the same ethnic group as them, contrasting with 11% of women being married to someone who does not share their faith, and 1.7% if I look only at Muslim-Christian unions.

Looking at time trends, I find that interethnic marriages are becoming more common in half of the sample and that their share remained constant in the other half. This suggests that ethnic cleavages may lose salience in some parts of sub-Saharan Africa. The study concludes that increased access to education and widespread urbanization contributed to that change, but that it cannot explain all of it. There is suggestive evidence that norms around marriage are changing across generations. On the contrary, interfaith marriages are becoming less common, a fact that can mostly be attributed to the decline in traditional religions in West Africa.

Further research could focus on understanding better the channels by which education and urbanization impact the likelihood to marry outside own's ethnic group : Are cities different from rural areas because they are more heterogenous in terms of ethnicity? Or is that people living in cities have more agency to chose a partner? Related work would consist on investigating whether marrying outside own's ethnic group could be the result of strategic behavior (Luke and Munshi [2006]). Are urban-dwellers turning away from membership in their ethnic group to benefit from their membership in a religious group? Faith groups or networks related to attending the same church or mosque could also be an opportunity to access jobs and support. Deepening our understanding on whether marriage decisions reinforces or changes identity affiliations would bring important contributions to the political economy literature on conflict, as well as to the literature on networks.

## References

- Alberto Bisin and Thierry Verdier. “Beyond the melting pot”: cultural transmission, marriage, and the evolution of ethnic and religious traits. *The Quarterly Journal of Economics*, 115(3):955–988, 2000. 1
- Cécile Canut. Langues et filiation en Afrique. *Les Temps Modernes*, (4):410–440, 2002. 37
- Kanchan Chandra. What is ethnic identity and does it matter? *Annual Review of Political Science*, 9(1):397–424, June 2006. ISSN 1094-2939, 1545-1577. doi: 10.1146/annurev.polisci.9.062404.170715. URL <http://www.annualreviews.org/doi/10.1146/annurev.polisci.9.062404.170715>. 1
- Juliette Crespín-Boucaud. Interethnic marriages and homogamy patterns in Kenya and Ghana. *Mimeo, Paris School of Economics*, 2015. 6
- Brandon de la Cuesta and Leonard Wantchekon. *Is Language Destiny? The Origins and Consequences of Ethnolinguistic Diversity in Sub-Saharan Africa*, pages 513–537. Palgrave Macmillan UK, London, 2016. ISBN 978-1-137-32505-1. doi: 10.1007/978-1-137-32505-1\_19. URL [https://doi.org/10.1007/978-1-137-32505-1\\_19](https://doi.org/10.1007/978-1-137-32505-1_19). 1
- Klaus Desmet, Ignacio Ortuno-Ortín, and Romain Wacziarg. Linguistic cleavages and economic development. In *The Palgrave Handbook of Economics and Language*, pages 425–446. Springer, 2016. 2, 11
- William Easterly and Ross Levine. Africa’s Growth Tragedy: Policies and Ethnic Divisions. *The Quarterly Journal of Economics*, 112(4):1203–1250, 1997. ISSN 0033-5533. URL <http://www.jstor.org/stable/2951270>. 1
- Benn Eifert, Edward Miguel, and Daniel N. Posner. Political competition and ethnic identification in Africa. *American Journal of Political Science*, 54(2):494–510, 2010. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-5907.2010.00443.x/full>. 1
- Delia Furtado. Human capital and interethnic marriage decisions. *Economic Inquiry*, 50(1):82–93, January 2012. ISSN 00952583. doi: 10.1111/j.1465-7295.2010.00345.x. URL <http://doi.wiley.com/10.1111/j.1465-7295.2010.00345.x>. 2
- Boris Gershman and Diego Rivera. Subnational diversity in Sub-Saharan Africa: Insights from a new dataset. *Journal of Development Economics*, 133:231–263, July 2018. ISSN 03043878. doi: 10.1016/j.jdeveco.2018.01.003. URL <http://linkinghub.elsevier.com/retrieve/pii/S0304387818300051>. 1, 2, 11, 36

- Rachel M Gisselquist, Stefan Leiderer, and Miguel Niño-Zarazúa. Ethnic heterogeneity and public goods provision in Zambia: Evidence of a subnational “diversity dividend”. *World Development*, 78:308–323, 2016. 1
- Joseph H Greenberg. The measurement of linguistic diversity. *Language*, 32(1):109–115, 1956. 7
- James Habyarimana, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. Why does ethnic diversity undermine public goods provision? *American Political Science Review*, 101(4):709–725, 2007. URL <https://www.cambridge.org/core/journals/american-political-science-review/article/why-does-ethnic-diversity-undermine-public-goods-provision/36C285D07DC0DE604BFC123F53060DF6>. 1
- Sara Lowes, Nathan Nunn, James A. Robinson, and Jonathan Weigel. Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT). *American Economic Review*, 105(5):340–345, May 2015. ISSN 0002-8282. doi: 10.1257/aer.p20151075. URL <http://pubs.aeaweb.org/doi/10.1257/aer.p20151075>. 1
- Nancy Luke and Kaivan Munshi. New roles for marriage in urban Africa: Kinship networks and the labor market in Kenya. *The Review of Economics and Statistics*, 88(2):264–282, 2006. URL <http://www.mitpressjournals.org/doi/abs/10.1162/rest.88.2.264>. 15
- Cédric Mayrargue. Trajectoires et enjeux contemporains du pentecôtisme en Afrique de l’Ouest. *Critique internationale*, 22(1):95, 2004. ISSN 1290-7839, 1777-554X. doi: 10.3917/cii.022.0095. URL <http://www.cairn.info/revue-critique-internationale-2004-1-page-95.htm>. 6
- Edward Miguel. Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania. *World Politics*, 56(03):328–362, April 2004. ISSN 0043-8871, 1086-3338. doi: 10.1017/S0043887100004330. URL [http://www.journals.cambridge.org/abstract\\_S0043887100004330](http://www.journals.cambridge.org/abstract_S0043887100004330). 1
- José G. Montalvo and Marta Reynal-Querol. Ethnic Diversity and Growth: Revisiting the Evidence. *Manuscript, Universat Pompeu Frabra-ICREA*, 2016. URL [http://84.89.132.1/~reynal/Ethnic\\_Diversity\\_Development\\_160712.pdf](http://84.89.132.1/~reynal/Ethnic_Diversity_Development_160712.pdf). 1
- Daniel N Posner. *Institutions and ethnic politics in Africa*. Cambridge University Press, 2005. 11
- Gary F Simons and Charles D Fennig. Ethnologue: Languages of the world. *SIL International*, 20, 2017. 3

Marie-Paule Thiriat and Thérèse Locoh. Divorce et remariage des femmes en afrique de l'ouest. le cas du togo. *Population*, 50(1):61–93, 1995. ISSN 0032-4663. doi: 10.2307/1533793. URL [https://www.persee.fr/doc/pop\\_0032-4663\\_1995\\_num\\_50\\_1\\_5902](https://www.persee.fr/doc/pop_0032-4663_1995_num_50_1_5902). 5

Anke Weber. The causes of politicization of ethnicity-a comparative case study of Kenya and Tanzania. 2010. 4

## 7 Figures and Tables

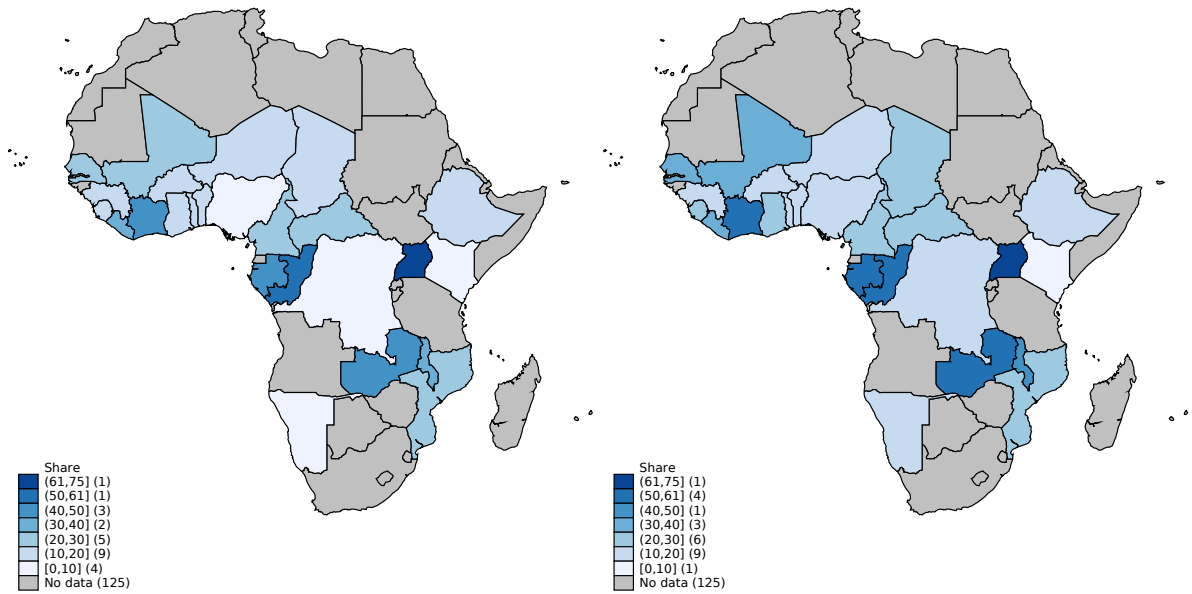


Figure 1: Left panel: Share of interethnic marriages. Right panel: Share of interethnic marriages, as a share of the interethnic marriage share we would observe under random matching.

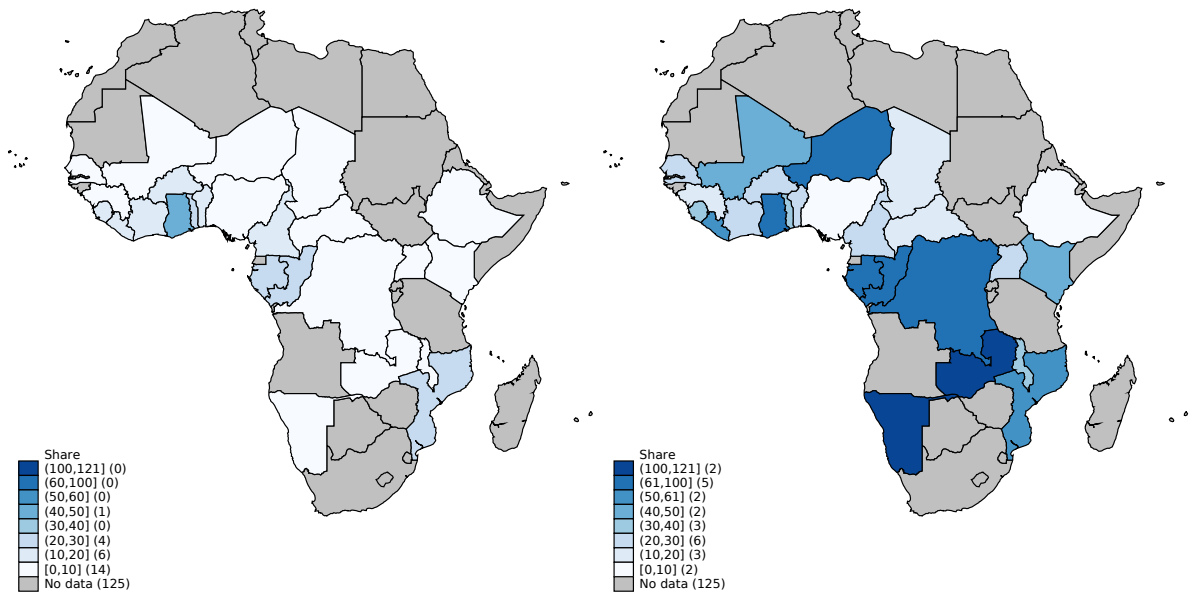


Figure 2: Left panel: Share of interfaith marriages. Right panel: Share of interfaith marriages, as a share of the interfaith marriage share we would observe under random matching.

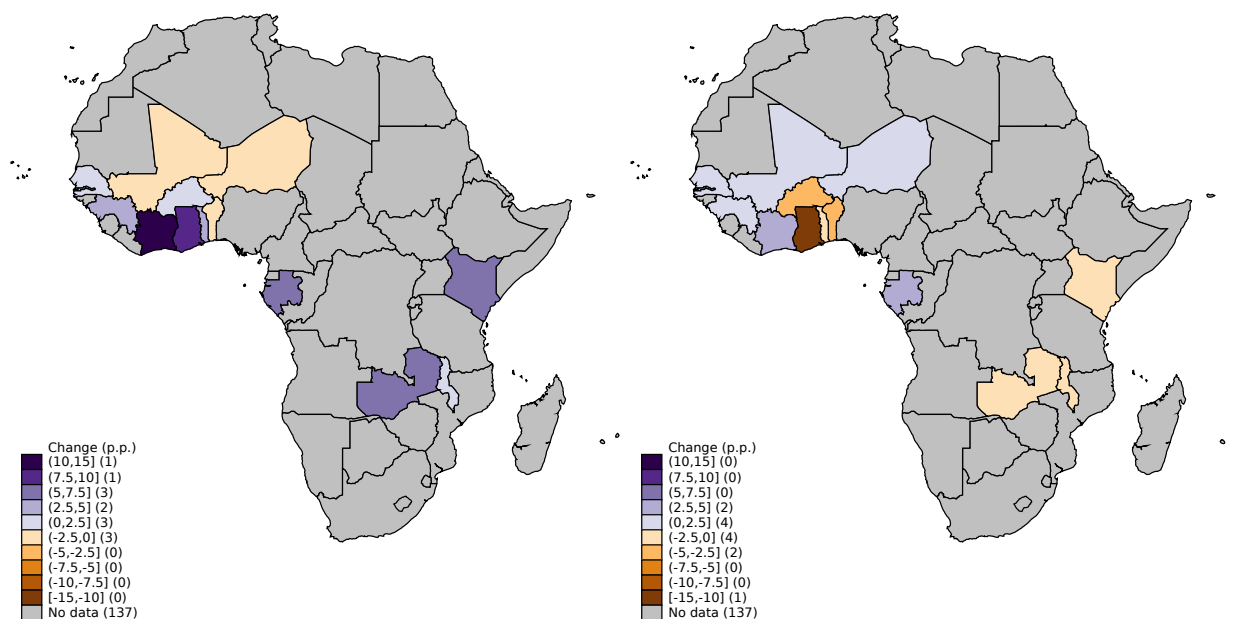


Figure 3: Left panel: Differences (in percentage points) between the share of interethnic marriage for the women born in the 1960s and for those born in the 1980s. Right panel: Differences (in percentage points) between the share of interfaith marriage for the women born in the 1960s and for those born in the 1980s.

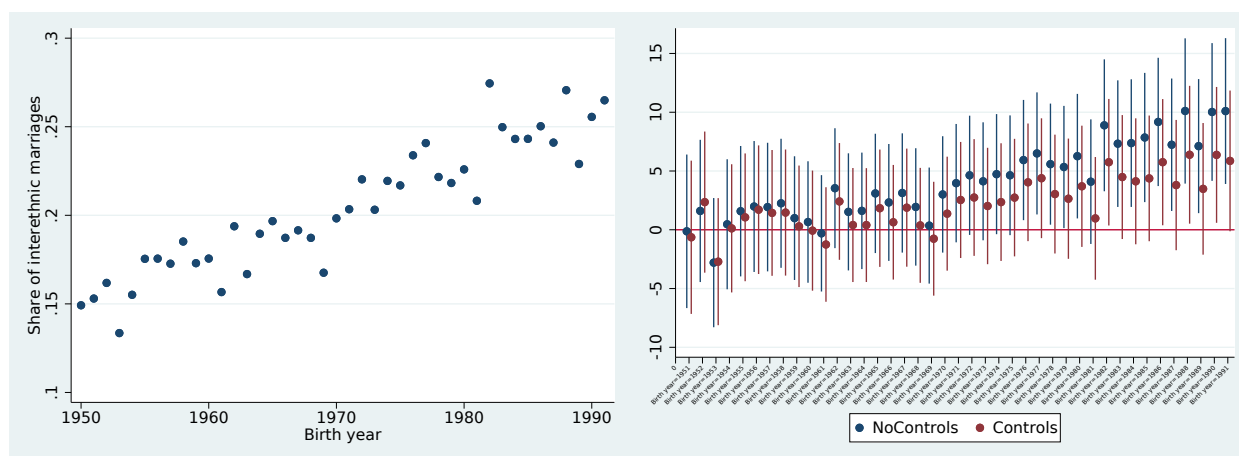


Figure 4: Left panel: Share of interethnic marriages by birth year. Right panel: Coefficient from a non-parametric regression of birth year on a dummy that indicates whether the respondent is in an interethnic union. “No controls” shows results for a regression on that dummy on age controls and country fixed effects. “Controls” shows results for equation 1.

Not all women know their age accurately, hence leading to age heaping around “round numbers”. As this correlates with education, some discontinuities on the left panel can be explained by the fact that the relative weights of countries changes when the prevalence age heaping is high.

Table 1: Trends - Intermarriage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Interethnic marriage	Interethnic marriage	Distance	Distance	Interfaith marriage	Interfaith marriage	Muslim/Christian marriage	Muslim/Christian marriage
Birth year	0.182*** (0.0229)	0.156*** (0.0359)	-0.00172 (0.00322)	-0.0126** (0.00514)	-0.136*** (0.0184)	-0.352*** (0.0330)	0.0115 (0.00730)	-0.00471 (0.0119)
Age		-0.0294 (0.183)		-0.0436* (0.0247)		-0.568*** (0.143)		-0.0557 (0.0602)
Age squared		0.000524 (0.00278)		0.000395 (0.000385)		0.00269 (0.00221)		0.000457 (0.000912)
Primary		4.070*** (0.596)		-0.0771 (0.0802)		1.571*** (0.524)		0.726*** (0.199)
Secondary/Higher		8.364*** (0.755)		0.0460 (0.0961)		-1.061 (0.648)		0.822*** (0.240)
Urban		13.71*** (0.624)		0.124 (0.0767)		-2.905*** (0.472)		0.974*** (0.191)
Remarried		5.053*** (0.558)		0.0976 (0.0642)		3.579*** (0.543)		0.952*** (0.199)
Observations	89097	89097	15072	15072	88412	88412	88412	88412
R-squared	0.272	0.297	0.791	0.791	0.174	0.182	0.023	0.026
Controls								
Country-fixed effects	✓	✓	✓	✓	✓	✓	✓	✓

OLS regression. Columns (1) and (2) the dependant variable is a variable that equals 0 if the woman is married within her ethnic group, 100 if is married outside of her group. Columns (3) and (4) the dependant variable is a variable that equals the linguistic distance between spouses, if the union is interethnic. Columns (5) and (6) the dependant variable is a variable that equals 0 if the union is an intrafaith one, 100 if it is an interfaith one. Faith groups are defined (for all countries) as follows: Muslim, Christian, Other religion/No religion.

Results for colums (1), (2), (5) and (6) should be read as percentage points.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.



Table 2: Trend - Respondent does not declare the same ethnicity as her husband

	(1)	(2)	(3)	(4)	(5)	(6)
	Intermarriage	Intermarriage	Intermarriage	Intermarriage	Intermarriage	Observations
Benin	0.0947 (0.0936)	0.0159 (0.0909)	0.0314 (0.0934)	0.114 (0.0933)	0.00971 (0.0908)	8864
Burkina Faso	0.232* (0.134)	0.165 (0.133)	0.186 (0.134)	0.237* (0.134)	0.160 (0.133)	9792
Cote d'Ivoire	0.420*** (0.129)	0.404*** (0.127)	0.451*** (0.125)	0.424*** (0.130)	0.447*** (0.123)	2894
Gabon	0.419 (0.277)	0.413 (0.269)	0.433 (0.273)	0.461* (0.279)	0.439 (0.268)	2417
Ghana	0.279** (0.127)	0.217* (0.127)	0.219* (0.125)	0.293** (0.126)	0.202 (0.126)	6791
Guinea	0.790*** (0.209)	0.750*** (0.213)	0.775*** (0.212)	0.794*** (0.209)	0.754*** (0.213)	4949
Kenya	0.264*** (0.0552)	0.222*** (0.0542)	0.111** (0.0521)	0.263*** (0.0553)	0.0973* (0.0527)	9815
Malawi	-0.182 (0.185)	-0.351* (0.186)	-0.200 (0.180)	-0.157 (0.186)	-0.286 (0.183)	7120
Mali	0.120 (0.146)	0.0614 (0.143)	0.210 (0.144)	0.147 (0.145)	0.174 (0.143)	8952
Niger	-0.138 (0.113)	-0.169 (0.114)	-0.136 (0.112)	-0.0908 (0.115)	-0.102 (0.116)	6016
Senegal	0.580*** (0.101)	0.420*** (0.0967)	0.452*** (0.0969)	0.612*** (0.101)	0.415*** (0.0965)	6196
Togo	0.390*** (0.104)	0.214** (0.105)	0.301*** (0.102)	0.407*** (0.103)	0.209** (0.105)	3976
Zambia	-0.00980 (0.123)	-0.158 (0.121)	-0.0242 (0.118)	0.00789 (0.124)	-0.0649 (0.119)	11315
Controls						
Age & Age <sup>2</sup>	✓	✓	✓	✓	✓	
Education		✓			✓	
Urban			✓		✓	
Marital status				✓	✓	

OLS regression. Dependant variable is a variable that equals 0 if the union is intraethnic, 100 if the union is interethnic. Results can be interpreted as changes in percentage points.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

”

Table 3: Individual characteristics and interethnic marriage

	(1) Benin	(2) Burkina Faso	(3) CI	(4) Gabon	(5) Ghana	(6) Guinea	(7) Kenya
Birth year	0.00971 (0.0908)	0.160 (0.133)	0.447*** (0.123)	0.439 (0.268)	0.202 (0.126)	0.754*** (0.213)	0.0973* (0.0527)
Age	0.173 (0.465)	-0.462 (0.459)	0.571 (0.987)	-0.308 (1.579)	-1.177* (0.616)	0.972* (0.503)	-1.023* (0.563)
Age squared	-0.00312 (0.00729)	0.0120* (0.00726)	-0.00729 (0.0152)	0.00591 (0.0233)	0.0152* (0.00907)	-0.00489 (0.00721)	0.0148* (0.00832)
Primary	8.142*** (1.440)	6.970*** (1.910)	2.622 (2.706)	17.27*** (5.928)	1.960 (1.664)	3.624* (2.130)	1.641 (1.201)
Secondary/Higher	15.06*** (2.112)	21.43*** (2.843)	17.85*** (3.933)	20.95*** (5.799)	3.265** (1.509)	8.390*** (3.002)	2.605* (1.348)
Urban	11.54*** (1.250)	6.469*** (1.899)	22.31*** (2.729)	-0.945 (2.932)	5.087*** (1.452)	2.725 (1.883)	12.03*** (1.614)
Remarried	3.230** (1.382)	2.324 (1.487)	4.901* (2.632)	13.76*** (3.428)	3.480** (1.445)	2.974* (1.724)	7.428*** (2.078)
Constant	-12.85 (181.5)	-297.4 (265.5)	-880.6*** (247.0)	-844.9 (534.1)	-363.2 (251.0)	-1501.2*** (426.6)	-170.5 (105.0)
Observations	8864	9792	2894	2417	6791	4949	9815
R-squared	0.053	0.026	0.099	0.034	0.014	0.022	0.042
	Malawi	Mali	Niger	Senegal	Togo	Zambia	Pooled sample
Birth year	-0.286 (0.183)	0.174 (0.143)	-0.102 (0.116)	0.415*** (0.0965)	0.209** (0.105)	-0.0649 (0.119)	0.156*** (0.0359)
Age	-0.399 (0.644)	0.901 (0.612)	-0.345 (0.514)	1.236 (0.777)	-0.435 (0.737)	-0.447 (0.542)	-0.0294 (0.183)
Age squared	0.00000375 (0.0103)	-0.0107 (0.00936)	0.00399 (0.00836)	-0.0133 (0.0120)	0.00843 (0.0112)	0.000299 (0.00821)	0.000524 (0.00278)
Primary	3.563** (1.587)	6.871*** (2.406)	2.913 (2.228)	11.52*** (2.502)	4.350** (1.704)	3.418** (1.662)	4.070*** (0.596)
Secondary/Higher	13.20*** (2.898)	19.43*** (3.390)	10.68*** (3.260)	15.19*** (3.325)	11.34*** (2.717)	11.51*** (2.014)	8.364*** (0.755)
Urban	17.38*** (2.603)	17.55*** (2.536)	11.28*** (1.813)	15.33*** (2.052)	8.873*** (2.022)	25.80*** (1.508)	13.71*** (0.624)
Remarried	6.131*** (1.563)	4.139* (2.359)	6.485*** (1.581)	5.953*** (2.278)	4.651** (1.900)	8.185*** (1.415)	5.053*** (0.558)
Constant	600.2 (364.9)	-334.2 (285.7)	217.3 (230.9)	-831.9*** (195.4)	-400.9* (207.2)	172.7 (237.2)	
Observations	7120	8952	6016	6196	3976	11315	89097
R-squared	0.032	0.047	0.028	0.084	0.044	0.088	0.297

OLS regression. Dependant variable is a variable that equals 0 if the union is intraethnic, 100 if the union is interethnic. Results can be interpreted as changes in percentage points.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 4: Trend - Respondent does not declare the same ethnicity as her husband

	(1) All	(2) 1st union	(3) Remarried	(4) Share remarried women
Benin	-0.00954 (0.0939)	0.0307 (0.0974)	-0.117 (0.221)	14.50
Burkina Faso	0.188 (0.132)	0.196 (0.129)	0.179 (0.251)	10.77
Cote d'Ivoire	0.453*** (0.130)	0.523*** (0.137)	0.0851 (0.273)	14.74
Gabon	0.489* (0.264)	0.376 (0.275)	0.836 (0.525)	25.48
Ghana	0.208 (0.131)	0.134 (0.138)	0.514** (0.251)	20.76
Guinea	0.770*** (0.216)	0.771*** (0.223)	0.762* (0.399)	15.41
Kenya	0.0941* (0.0543)	0.0807 (0.0552)	0.220 (0.303)	5.67
Malawi	-0.343* (0.184)	-0.379* (0.198)	-0.134 (0.371)	20.15
Mali	0.148 (0.141)	0.155 (0.142)	0.371 (0.353)	9.87
Niger	-0.202 (0.173)	-0.159 (0.161)	-0.114 (0.436)	21.69
Senegal	0.411 (0.302)	0.486 (0.300)	0.583 (0.739)	15.72
Togo	0.174 (0.110)	0.210* (0.112)	0.123 (0.235)	17.48
Zambia	-0.115 (0.121)	-0.0973 (0.130)	-0.0953 (0.246)	16.66
Controls				
Age & Age <sup>2</sup>	✓	✓	✓	
Education	✓	✓	✓	
Urban	✓	✓	✓	

OLS regression. Dependant variable is a variable that equals 0 if the union is intraethnic, 100 if the union is interethnic. Results can be interpreted as changes in percentage points.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 5: Trend - Linguistic distance between spouses

	(1) Distance - only intermarried couples	(2) Distance - only intermarried couples	(3) Observations
Benin	0.00154 (0.0129)	0.00209 (0.0126)	917
Burkina Faso	-0.0471** (0.0237)	-0.0445* (0.0238)	586
Cote d'Ivoire	-0.0346 (0.0219)	-0.0271 (0.0212)	461
Gabon	0.00285 (0.00935)	0.00899 (0.00933)	471
Ghana	0.0220** (0.00991)	0.0200** (0.0101)	984
Guinea	0.0362 (0.0289)	0.0388 (0.0287)	468
Kenya	-0.0626*** (0.0238)	-0.0463* (0.0246)	723
Malawi	0.00365 (0.00348)	0.00295 (0.00346)	1695
Mali	-0.0108 (0.0110)	-0.0142 (0.0109)	2033
Niger	0.00664 (0.00989)	0.00760 (0.00993)	780
Senegal	-0.0280* (0.0167)	-0.0314** (0.0159)	860
Togo	0.0265* (0.0137)	0.0246* (0.0138)	293
Zambia	0.00229 (0.00424)	0.000620 (0.00407)	4801
Controls			
Age & Age <sup>2</sup>	✓	✓	
Education		✓	
Urban		✓	
Marital status		✓	

OLS regression. Dependant variable is the linguistic distance (measure defined in Appendix B) associated to each interethnic union. Resultats can be interpreted as changes in percentage points.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 6: Trend - Respondent does not declare the same religion as her husband

	(1)	(2)	(3)	(4)	(5)
	Intermarriage	Intermarriage	Intermarriage	Intermarriage	Observations
	All	All	Muslim/Christian	Muslim/Christian	
Benin	-0.438*** (0.0955)	-0.378*** (0.0960)	0.0252 (0.0415)	0.0223 (0.0429)	8864
Burkina Faso	-0.268*** (0.0857)	-0.254*** (0.0857)	0.0193 (0.0405)	-0.00244 (0.0408)	9792
Cote d'Ivoire	0.0770 (0.124)	0.0816 (0.122)	0.0669 (0.0417)	0.0731* (0.0426)	2894
Gabon	-0.556** (0.223)	-0.415* (0.222)	0.0549 (0.101)	0.0348 (0.105)	2417
Ghana	-2.168*** (0.113)	-2.086*** (0.115)	-0.176*** (0.0539)	-0.173*** (0.0535)	6791
Guinea	-0.0155 (0.113)	-0.0100 (0.114)	0.00839 (0.0264)	-0.000788 (0.0269)	4949
Kenya	-0.179*** (0.0434)	-0.0921** (0.0445)	0.0285 (0.0197)	0.0163 (0.0208)	9815
Malawi	-0.0202 (0.114)	0.0794 (0.115)	0.0229 (0.0632)	0.0463 (0.0616)	7120
Mali	0.00103 (0.0683)	-0.00976 (0.0684)	0.0184 (0.0230)	0.0170 (0.0224)	8952
Niger	0.0889* (0.0536)	0.0929* (0.0532)	-0.00156 (0.0145)	-0.00254 (0.0147)	6016
Senegal	-0.0399 (0.0834)	-0.0388 (0.0831)	-0.0350 (0.0660)	-0.0322 (0.0639)	5511
Togo	-0.232** (0.106)	-0.129 (0.110)	0.0799*** (0.0263)	0.0642** (0.0271)	3976
Zambia	-0.167*** (0.0426)	-0.145*** (0.0429)	-0.00258 (0.0157)	-0.00232 (0.0159)	11315
Controls					
Age & Age <sup>2</sup>	✓	✓	✓	✓	
Education		✓		✓	
Urban		✓		✓	
Marital status		✓		✓	

OLS regression. Dependant variable is a variable that equals 0 if the union is intrafaith, 100 if the union is interfaith. Results can be interpreted as changes in percentage points.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 7: Individual characteristics interfaith marriage

	(1) Benin	(2) Burkina Faso	(3) CI	(4) Gabon	(5) Ghana	(6) Guinea	(7) Kenya
Birth year	-0.378*** (0.0960)	-0.254*** (0.0857)	0.0816 (0.122)	-0.415* (0.222)	-2.086*** (0.115)	-0.0100 (0.114)	-0.0921** (0.0445)
Age	-1.084** (0.488)	-0.394 (0.403)	-0.178 (0.978)	0.0901 (1.064)	-2.561*** (0.635)	-0.366 (0.385)	-0.358 (0.376)
Age squared	0.00929 (0.00752)	0.00333 (0.00615)	0.00228 (0.0151)	-0.0136 (0.0155)	0.00689 (0.00933)	0.00545 (0.00599)	0.00175 (0.00569)
Primary	-2.197* (1.292)	0.762 (1.344)	10.40*** (2.538)	7.064 (4.360)	8.395*** (1.719)	2.785** (1.354)	-7.564*** (1.438)
Secondary/Higher	-2.554 (1.610)	6.395*** (2.043)	11.87*** (3.858)	-0.873 (3.698)	2.287 (1.509)	0.684 (1.194)	-10.19*** (1.432)
Urban	-1.653 (1.113)	-3.693*** (1.130)	-8.812*** (2.176)	-4.566* (2.436)	-3.936*** (1.374)	-3.570*** (0.861)	-1.254* (0.739)
Remarried	6.044*** (1.345)	4.032*** (1.268)	8.623** (3.630)	4.692 (2.943)	5.355*** (1.512)	2.095* (1.224)	-0.374 (1.293)
Constant	786.0*** (192.5)	522.1*** (172.1)	-140.3 (245.8)	850.8* (444.2)	4215.2*** (229.5)	30.73 (227.5)	205.6** (87.86)
Observations	8864	9792	2894	2417	6791	4949	9815
R-squared	0.009	0.007	0.030	0.024	0.098	0.006	0.019
	Malawi	Mali	Niger	Senegal	Togo	Zambia	Pooled sample
Birth year	0.0794 (0.115)	-0.00976 (0.0684)	0.0929* (0.0532)	-0.0388 (0.0831)	-0.129 (0.110)	-0.145*** (0.0429)	-0.352*** (0.0330)
Age	-0.378 (0.453)	-0.484 (0.335)	-0.0131 (0.169)	0.218 (0.214)	-0.0143 (0.721)	-0.0468 (0.214)	-0.568*** (0.143)
Age squared	0.00780 (0.00738)	0.00603 (0.00507)	0.00169 (0.00287)	-0.00408 (0.00329)	-0.00377 (0.0109)	-0.00127 (0.00328)	0.00269 (0.00221)
Primary	-4.476*** (1.284)	0.664 (1.007)	-0.491 (0.721)	-0.813 (0.780)	0.105 (1.785)	-4.373*** (0.950)	1.571*** (0.524)
Secondary/Higher	-7.671*** (1.540)	0.940 (1.405)	-0.763 (0.698)	1.851 (1.720)	-4.884** (2.282)	-5.872*** (1.101)	-1.061 (0.648)
Urban	1.893 (1.400)	-4.209*** (0.703)	-0.182 (0.475)	0.148 (0.875)	-5.282*** (1.907)	0.722 (0.562)	-2.905*** (0.472)
Remarried	2.694** (1.116)	1.605* (0.969)	0.302 (0.519)	0.849 (0.920)	4.729** (2.108)	0.254 (0.618)	3.579*** (0.543)
Constant	-142.3 (229.1)	34.61 (136.7)	-182.4* (105.1)	75.98 (165.9)	280.5 (218.2)	296.3*** (85.56)	
Observations	7120	8952	6016	5511	3976	11315	88412
R-squared	0.011	0.006	0.002	0.003	0.012	0.010	0.182

OLS regression. Dependant variable is a variable that equals 0 if the union is intrafaith, 100 if the union is interfaith. Results can be interpreted as changes in percentage points.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 8: Other and traditional religions

	(1) Other religion	(2) Other religion	(3) Other religion
Birth year	-0.246*** (0.0214)	-0.522*** (0.0422)	-0.438*** (0.0416)
Age		-0.855*** (0.143)	-0.678*** (0.141)
Age squared		0.00632*** (0.00213)	0.00419** (0.00210)
Primary			-6.350*** (0.579)
Secondary/Higher			-8.209*** (0.669)
Urban			-5.915*** (0.494)
Remarried			2.396*** (0.540)
Observations	88412	88412	88412
R-squared	0.216	0.221	0.239
Controls			
Country-fixed effects	✓	✓	✓
Turning point (age)			79.2

OLS regression. The dependant variable is a variable that equals 100 if a woman is a member of the group "other religion/no religion", 0 otherwise. Most people in that category identify with traditional religion.

Results for colums (1), (2) and (3) should be read as percentage points.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 9: Descriptive statistics: Intermarriage patterns and linguistic distance

	Interethnic marriages				Interreligious marriages			Muslim/Christian marriages			
	All	1st union	Remarried	Linguistic distance	All	1st union	Remarried	All	1st union	Remarried	
Benin	15.4	15.0	17.5	3.3		17.0	16.0	22.3	2.7	2.6	3.7
Burkina Faso	18.9	18.7	20.5	4.7		11.9	11.3	16.1	3.5	3.5	3.8
Cote d'Ivoire	24.7	24.6	25.2	3.9		19.6	18.2	27.4	2.7	2.4	4.1
Gabon	39.5	36.1	49.1	2.3		17.4	16.0	21.5	3.9	3.2	5.9
Ghana	19.8	19.4	21.0	3.2		26.0	24.3	32.5	2.5	2.3	3.2
Guinea	14.5	14.2	16.1	5.9		5.0	4.7	6.9	0.9	0.8	1.2
Kenya	10.5	10.1	17.0	4.5		6.8	6.8	7.1	1.1	1.0	2.0
Malawi	31.3	30.6	34.2	2.8		7.9	7.2	10.8	2.4	1.8	4.5
Mali	33.1	32.6	37.6	6.7		6.0	5.9	7.1	1.1	1.1	1.2
Niger	12.9	11.5	17.7	4.4		1.9	1.9	2.0	0.4	0.4	0.3
Senegal	22.2	21.5	25.6	3.6		1.4	1.4	1.6	0.9	0.9	1.3
Togo	14.7	14.3	17.0	3.3		19.1	18.2	23.4	1.4	1.2	2.0
Zambia	47.1	46.4	50.2	2.4		3.9	3.7	4.6	0.4	0.3	0.5

Descriptive statistics: Shares of interethnic and interfaith marriages. The linguistic distance measure is defined in Appendix B.

Weighted sample.



Table 10: Ethnic identification and time in union

	(1) Benin	(2) Burkina Faso	(3) CI	(4) Gabon	(5) Ghana	(6) Guinea	(7) Kenya
Birth year	0.0933 (0.0972)	0.230* (0.130)	0.477*** (0.133)	0.267 (0.295)	0.180 (0.130)	0.778*** (0.214)	0.192*** (0.0572)
Number of years since cohabitation	-0.0672 (0.112)	0.265* (0.141)	-0.0103 (0.184)	-0.410 (0.372)	-0.250* (0.143)	0.641*** (0.217)	-0.0908 (0.0749)
Age at cohabitation	0.802*** (0.157)	1.484*** (0.254)	1.179*** (0.304)	0.796* (0.438)	0.352* (0.192)	1.259*** (0.267)	0.580*** (0.146)
Constant	-183.0 (194.8)	-463.7* (261.2)	-939.7*** (265.0)	-502.6 (590.4)	-339.7 (260.4)	-1552.1*** (428.4)	-379.9*** (113.0)
Number of observations	7,595	8,597	2,444	1,664	5,355	4,164	9,172
	Malawi	Mali	Niger	Senegal	Togo	Zambia	Pooled sample
Birth year	-0.204 (0.195)	0.0706 (0.148)	-0.144 (0.117)	0.385*** (0.108)	0.377*** (0.105)	-0.0819 (0.130)	0.214*** (0.0376)
Number of years since cohabitation	-0.417** (0.206)	-0.0330 (0.170)	-0.160 (0.122)	0.276* (0.153)	0.110 (0.130)	-0.548*** (0.141)	-0.0271 (0.0427)
Age at cohabitation	-0.0204 (0.317)	0.938*** (0.244)	0.534** (0.248)	1.771*** (0.266)	1.219*** (0.229)	1.180*** (0.232)	0.917*** (0.0721)
Constant	437.9 (389.7)	-122.3 (295.0)	288.6 (233.1)	-773.1*** (216.4)	-756.3*** (209.9)	193.6 (260.1)	
Number of observations	5,586	7,945	4,578	5,286	3,279	9,288	74,953

Note: OLS regression of a dummy variable (100 if the union is interethnic) on year of birth (linear variable), age at first union and number of years since wedding. The sample is made up of women who are still in their first union (as only the year of the first union is recorded).

Weighted data, standard errors are clustered at the DHS cluster level, strata are accounted for. Standard errors are in parentheses. Significance levels are denoted as follows: \* p<0.10,

\*\* p<0.05, \*\*\* p<0.01.

Table 11: Trend - Year of marriage

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Interethnic marriage				Interfaith marriage				Observations
	Marriage year		Birth year		Marriage year		Birth year		Observations
Benin	0.328*** (0.0776)	0.0891 (0.0791)	0.124 (0.0984)	0.0241 (0.0960)	-0.280*** (0.0788)	-0.245*** (0.0803)	-0.398*** (0.102)	-0.381*** (0.102)	7595
Burkina Faso	0.378*** (0.118)	0.227* (0.118)	0.253* (0.131)	0.183 (0.130)	-0.148* (0.0794)	-0.155* (0.0800)	-0.225** (0.0873)	-0.224** (0.0872)	8597
Cote d'Ivoire	0.660*** (0.120)	0.534*** (0.117)	0.539*** (0.133)	0.533*** (0.127)	0.177 (0.113)	0.139 (0.111)	0.0587 (0.127)	0.0459 (0.125)	2444
Gabon	0.160 (0.141)	0.112 (0.138)	0.404 (0.301)	0.349 (0.284)	-0.284*** (0.101)	-0.231** (0.0998)	-0.522** (0.241)	-0.347 (0.241)	1664
Ghana	0.344*** (0.0986)	0.191* (0.101)	0.240* (0.131)	0.115 (0.129)	-1.477*** (0.0952)	-1.457*** (0.101)	-2.103*** (0.128)	-2.049*** (0.130)	5355
Guinea	0.711*** (0.153)	0.637*** (0.168)	0.773*** (0.215)	0.730*** (0.220)	0.0146 (0.0915)	0.0405 (0.0995)	-0.0363 (0.119)	-0.0315 (0.121)	4164
Kenya	0.312*** (0.0473)	0.132*** (0.0496)	0.243*** (0.0576)	0.0785 (0.0539)	-0.186*** (0.0364)	-0.0654* (0.0380)	-0.162*** (0.0440)	-0.0681 (0.0445)	9172
Malawi	-0.00796 (0.157)	-0.363** (0.164)	-0.208 (0.197)	-0.356* (0.197)	-0.128 (0.0968)	-0.0154 (0.0967)	-0.0662 (0.124)	0.0278 (0.123)	5586
Mali	0.289** (0.115)	0.205* (0.112)	0.129 (0.146)	0.165 (0.142)	0.0118 (0.0553)	0.00938 (0.0567)	0.0253 (0.0711)	0.00640 (0.0710)	7945
Niger	0.0231 (0.0949)	-0.0857 (0.0965)	-0.106 (0.114)	-0.121 (0.115)	0.0696 (0.0575)	0.0771 (0.0619)	0.0788 (0.0654)	0.0806 (0.0659)	4578
Senegal	0.694*** (0.0964)	0.324*** (0.0922)	0.563*** (0.110)	0.351*** (0.106)	0.148** (0.0691)	0.161*** (0.0624)	0.00375 (0.0890)	-0.000406 (0.0896)	5286 /4750
Togo	0.549*** (0.0941)	0.314*** (0.0990)	0.442*** (0.107)	0.236** (0.110)	-0.232** (0.0986)	-0.124 (0.108)	-0.175 (0.115)	-0.0882 (0.119)	3279
Zambia	0.530*** (0.107)	0.125 (0.106)	0.00727 (0.131)	-0.0612 (0.127)	-0.153*** (0.0391)	-0.113*** (0.0397)	-0.157*** (0.0487)	-0.135*** (0.0489)	9288
Controls									
Age & Age <sup>2</sup>	✓	✓	✓	✓	✓	✓	✓	✓	
Education		✓		✓		✓		✓	
Urban		✓		✓		✓		✓	
Marital status		✓		✓		✓		✓	

OLS regression. Dependant variable is a variable that equals 100 if the union is an interfaith or interethnic union. Results can be interpreted as changes in percentage points.  
The sample is made up of women that are still in their first union.

Reported coefficients are the coefficient associated to the birth year and marriage year variables.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 12: Bounds - Respondent does not declare the same ethnicity as her husband

	(1)	(2)	(3)	(4)	(5)	(6)
	Main specification		Robustness - Other		Robustness - Remarriage	
Benin	0.0947 (0.0936)	-0.0124 (0.0911)	0.0947 (0.0936)	-0.0124 (0.0911)	-0.490*** (0.111)	-0.575*** (0.109)
Burkina Faso	0.232* (0.134)	0.154 (0.133)	0.232* (0.134)	0.154 (0.133)	-0.0154 (0.130)	-0.0752 (0.129)
Cote d'Ivoire	0.420*** (0.129)	0.437*** (0.123)	0.120 (0.193)	0.145 (0.188)	0.292** (0.137)	0.301** (0.134)
Gabon	0.419 (0.277)	0.426 (0.266)	0.927*** (0.286)	0.848*** (0.277)	0.0985 (0.260)	0.170 (0.250)
Ghana	0.279** (0.127)	0.190 (0.127)	0.177 (0.153)	0.119 (0.151)	-0.170 (0.136)	-0.200 (0.135)
Guinea	0.790*** (0.209)	0.749*** (0.213)	0.960*** (0.223)	0.919*** (0.228)	0.508** (0.203)	0.486** (0.205)
Kenya	0.264*** (0.0552)	0.100* (0.0529)	0.284*** (0.0698)	0.243*** (0.0709)	0.242*** (0.0670)	0.108 (0.0661)
Malawi	-0.182 (0.185)	-0.312* (0.182)	-0.329* (0.195)	-0.455** (0.193)	-0.558*** (0.182)	-0.587*** (0.181)
Mali	0.120 (0.146)	0.154 (0.143)	0.120 (0.146)	0.154 (0.143)	-0.226 (0.141)	-0.190 (0.136)
Niger	-0.138 (0.113)	-0.154 (0.113)	-0.171 (0.113)	-0.188* (0.113)	-0.790*** (0.132)	-0.794*** (0.133)
Senegal	0.580*** (0.101)	0.381*** (0.0959)	0.666*** (0.101)	0.448*** (0.0964)	-0.00112 (0.121)	-0.166 (0.121)
Togo	0.390*** (0.104)	0.190* (0.106)	0.332** (0.130)	0.181 (0.130)	-0.00559 (0.122)	-0.166 (0.125)
Zambia	-0.00980 (0.123)	-0.0856 (0.118)	0.00907 (0.123)	-0.0731 (0.118)	-0.175 (0.111)	-0.216** (0.110)
Controls						
Age & Age <sup>2</sup>	✓	✓	✓	✓	✓	✓
Education		✓		✓		✓
Urban		✓		✓		✓

OLS regression. Dependant variable is a variable that equals 100 if the union is an interethnic union. Resultats can be interpreted as changes in percentage points.

Specification (1) is my main specification (as reported in table 2). In specification (2), the outcome variable is set to be 100 (interethnic marriage) if both spouses belong to the group "other" (0 in the main specification). In specification (3), the outcome variable is set to be 100 (interethnic marriage) in case a woman has remarried.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 13: Bounds - Respondent does not declare the same religion as her husband

	(1)	(2)	(3)	(4)	(5)	(6)
	Main specification		Robustness - Other		Robustness - Remarriage	
Benin	-0.438*** (0.0955)	-0.419*** (0.0953)	-0.919*** (0.163)	-0.796*** (0.161)	-0.932*** (0.103)	-0.902*** (0.103)
Burkina Faso	-0.268*** (0.0857)	-0.266*** (0.0857)	-0.823*** (0.152)	-0.764*** (0.150)	-0.455*** (0.0945)	-0.444*** (0.0939)
Cote d'Ivoire	0.0770 (0.124)	0.0633 (0.122)	-0.375* (0.211)	-0.405** (0.204)	-0.121 (0.146)	-0.138 (0.141)
Gabon	-0.556** (0.223)	-0.419* (0.223)	-0.574** (0.230)	-0.385* (0.227)	-0.614** (0.276)	-0.337 (0.273)
Ghana	-2.168*** (0.113)	-2.105*** (0.115)	-2.814*** (0.130)	-2.532*** (0.133)	-1.930*** (0.129)	-1.823*** (0.130)
Guinea	-0.0155 (0.113)	-0.0133 (0.114)	-0.114 (0.216)	-0.0993 (0.219)	-0.190 (0.157)	-0.176 (0.157)
Kenya	-0.179*** (0.0434)	-0.0923** (0.0444)	-0.211*** (0.0520)	-0.0842 (0.0552)	-0.136** (0.0591)	-0.0250 (0.0583)
Malawi	-0.0202 (0.114)	0.0680 (0.115)	-0.0154 (0.117)	0.0849 (0.118)	-0.573*** (0.171)	-0.375** (0.171)
Mali	0.00103 (0.0683)	-0.0177 (0.0683)	-0.209** (0.0900)	-0.232*** (0.0897)	-0.458*** (0.0851)	-0.461*** (0.0847)
Niger	0.0889* (0.0536)	0.0906* (0.0541)	0.103* (0.0572)	0.105* (0.0578)	-0.731*** (0.119)	-0.716*** (0.119)
Senegal	-0.0399 (0.0834)	-0.0454 (0.0844)	-0.0293 (0.0840)	-0.0346 (0.0850)	-0.778*** (0.196)	-0.766*** (0.195)
Togo	-0.232** (0.106)	-0.149 (0.110)	-1.113*** (0.171)	-0.471*** (0.161)	-0.482*** (0.115)	-0.407*** (0.120)
Zambia	-0.167*** (0.0426)	-0.145*** (0.0427)	-0.179*** (0.0442)	-0.156*** (0.0443)	-0.439*** (0.0889)	-0.352*** (0.0876)
Controls						
Age & Age <sup>2</sup>	✓	✓	✓	✓	✓	✓
Education		✓		✓		✓
Urban		✓		✓		✓

OLS regression. Dependant variable is a variable that equals 100 if the union is an interfaith union. Resultats can be interpreted as changes in percentage points.

Specification (1) is my main specification (as reported in table 6). In specification (2), the outcome variable is set to be 100 (interfaith marriage) if both spouses belong to the group "other" (0 in the main specification). In specification (3), the outcome variable is set to be 100 (interfaith marriage) in case a woman has remarried.

Reported coefficients are the coefficient associated to the birth year variable.

Weighted data. Standard errors are clustered at the DHS-cluster level. Significance levels are denoted as follows: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

## A Appendix

### A.1 Data - detailed

Table 14: Databases - DHS in sub-Saharan Africa

Descriptive sample	Main sample	Country	Survey years	Nb ethnic groups
		Angola		Not asked
✓	✓	Benin	1996 2001 2006 2011	9
✓	✓	Burkina Faso	1993 1998 2003 2010	10
		Burundi		Not asked
✓		Cameroon	1998 2004 2011	49; 10
✓		Central African Republic	only one wave	
✓		Chad	1996 2004	13
		Comoros		Not asked
✓		Congo (Brazzaville)	2005 2011	96; 13
✓		Congo Democratic Republic	2007 2013	10
✓	✓	Cote d'Ivoire	1994 2011	46
✓		Ethiopia	2000 2004 2010	classifications do not match
✓	✓	Gabon	2000 2012	9
✓	✓	Ghana	1993 1998 2003 2008 2014	8
✓	✓	Guinea	1999 2005 2012	7
✓	✓	Kenya	1993 1998 2003 2008 2014	11
		Lesotho		Not asked
✓		Liberia	only one wave	
		Madagascar		Not asked
✓	✓	Malawi	2000 2004 2010	9
✓	✓	Mali	1995 2001 2006 2012	10
✓		Mozambique	1997 2011	7; 21
✓		Namibia	only one wave	
✓	✓	Niger	1992 1998 2006	9
✓		Nigeria	2008 2013	11; 400
		Rwanda		Only one wave
		Sao Tome and Principe		Not asked
✓	✓	Senegal	1992 1997 2005 2010 2012-2013-2014	7
✓	✓	Sierra Leone	2008 2013	9
		South Africa		Not asked
		Sudan		Not asked
		Swaziland		Not asked
		Tanzania		Not asked
✓	✓	Togo	1998 2013	6
✓		Uganda	1996 2011	34; 19
✓	✓	Zambia	1996 2001 2007 2013	45
		Zimbabwe		Not asked

Note: Number of ethnic groups: Number of ethnic groups with at least one (female) member in the sample, including the group "others". There is always only one "other" group in each country, which includes foreigners and all the nationals that did not identify with any of the listed ethnic groups (in the common classification). When there are several values for a give country, it indicates the number of ethnic groups in each waves - in that case, due to too huge changes in the nomenclature, these databases were not used. "Not asked" indicates that the question about ethnicity was not included in the DHS questionnaire. The number of waves corresponds to the number of waves which include the ethnicity question.

Selection criteria were that there were at least two waves including the ethnicity question and that the ethnic group classification must not have changed enough not to be able to be unified across waves.

Country	Share of “others”	Country	Share of “others”
BF2	7.19	KE5	4.69
BF3	11.42	KE6	6.03
BF4	8.07	ML3	10.72
BF6	13.15	ML4	4.54
BJ3	4.10	ML5	5.54
BJ4	4.26	ML6	6.80
BJ5	4.33	MW3	8.28
BJ6	1.73	MW4	7.61
CI3	29.03	MW5	6.43
CI6	25.42	NI2	1.89
GA3	24.74	NI3	1.55
GA6	31.16	NI5	0.39
GH2	4.81	SN2	4.56
GH3	4.30	SN4	7.93
GH4	8.44	SN5	9.06
GH5	4.61	SN6	4.92
GH6	3.56	SN7	7.27
GN3	1.07	TG3	8.63
GN4	1.14	TG6	6.60
GN6	3.64	ZM3	2.42
KE2	4.69	ZM4	3.32
KE3	4.76	ZM5	3.18
KE4	7.00	ZM6	2.73

Figure 5: Note: Share of women that do not identify with any listed ethnic group (hence identified as “others”), across survey waves.

## B Linguistic and religious diversity

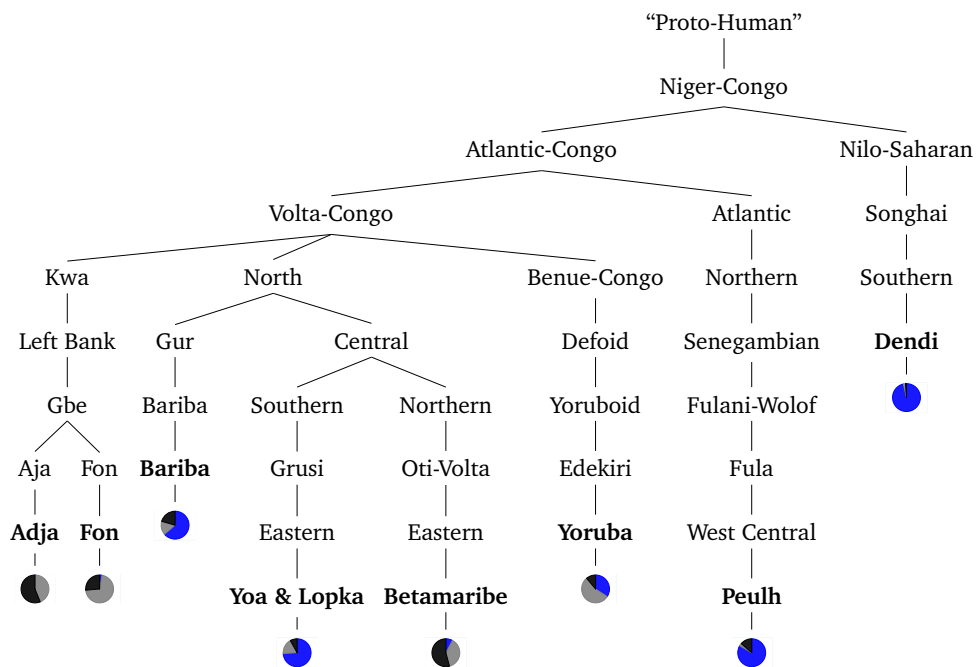


Figure 6: Ethnolinguistic tree for groups listed in DHS Benin. Pie charts represent the religious composition of each group (all women, DHS). Blue: Muslim; Gray: Christian; Black: Other religions (traditionalist, animist, none etc.)

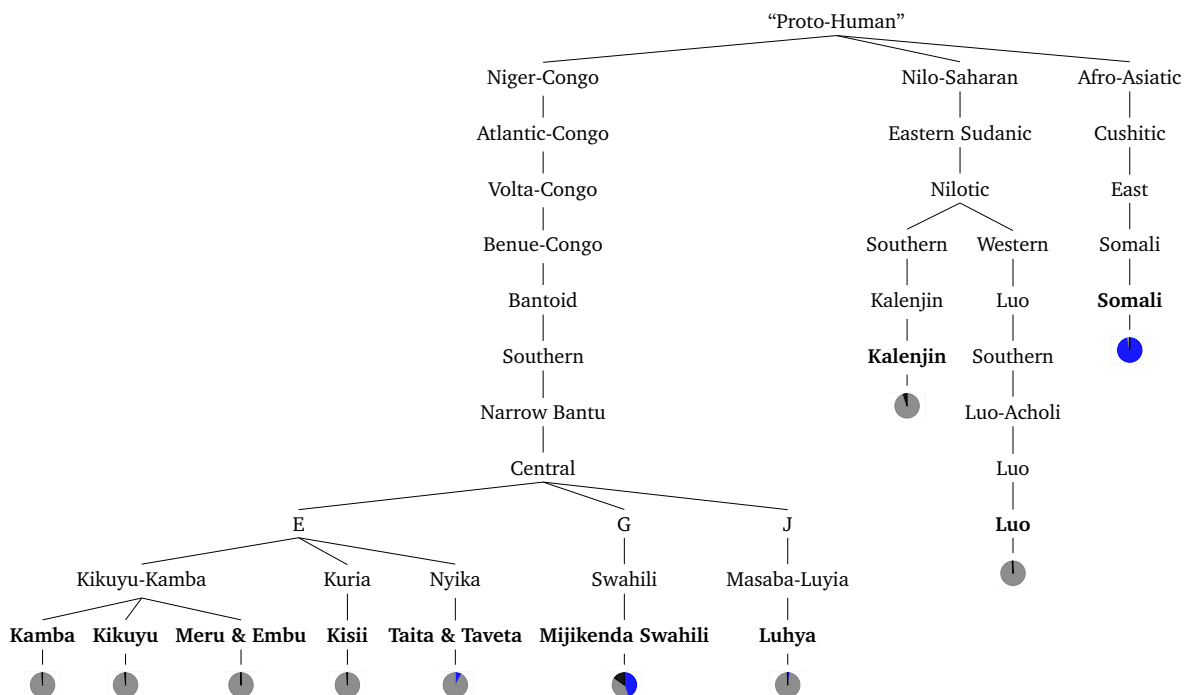


Figure 7: Ethnolinguistic tree for groups listed in DHS Kenya. Pie charts represent the religious composition of each group (all women, DHS). Blue: Muslim; Gray: Christian; Black: Other religions (traditionalist, animist, none etc.)

Gershman and Rivera [2018] describe the specificity of linguistic trees in the case of ethnolinguistic groups in sub-Saharan Africa. I define linguistic distance between two groups as the mean number of nodes to their first common subfamily. While languages are not always associated with ethnicity (Canut [2002]), there is currently a strong association between languages (or language groups) and ethnic groups. For instance, the linguistic distance between Kamba and Kikuyu is 1 (one step needed to go to the last common subgroup - Kikuyu-Kamba); the linguistic distance between Kalenjin and Luo is 4.5 (the average of the distance Kalenjin-Nilotic (3) and Luo-Nilotic (6)). When one spouse does not belong to any identified ethnic group, but that their partner does, I define the linguistic distance between them as the mean linguistic distance of intermarried couples in which the woman was born in 1950-1960.