Explaining inter-ethnic marriage in Sub-Saharan Africa

Sanghamitra Bandyopadhyay1 | Elliott Green2

1School of Business and Management, Queen Mary, University of London, London, UK
2Department of International Development, London School of Economics, London, UK

Correspondence
Elliott Green, Department of International Development, London School of Economics, Houghton Street, London WC2A 2AE, UK.
Email: e.d.green@lse.ac.uk

Abstract
We use Demographic and Health Survey data to undertake the largest ever examination into the causes of inter-ethnic marriage in contemporary Sub-Saharan Africa and document a number of novel findings. First, we show that inter-ethnic marriage rates are high, at 19.4% on average, vary significantly across countries and are rising over time. Second, we show that individual variables associated with modernization such as urbanization, literacy/education and declines in polygamy and agricultural employment are correlated with inter-ethnic marriage. Third, we show that inter-ethnic marriage is quadratically correlated with ethnic group size and negatively correlated with both educational homogamy and country-population size.

KEYWORDS
DHS data, ethnicity, inter-marriage, modernization, Sub-Saharan Africa

1 INTRODUCTION

Inter-cultural marriages have long been of great interest to social scientists who wish to examine how ethnic, religious, racial, linguistic, national, caste and tribal identities form and change over time. There has been a great deal of literature on inter-racial marriages in the United States over the past few decades, as well as on inter-religious and inter-national marriages in other contexts. However, there remains a relative absence of quantitative evidence on contemporary inter-ethnic marriage in Sub-Saharan Africa despite a long-standing qualitative focus on the subject from anthropologists and historians, as well as evidence from multiple other contexts that demonstrate several common causes of intermarriage. This subject is particularly important inasmuch as inter-ethnic marriage has the...
potential to reduce levels of ethnic diversity in Africa, which has long been seen as an impediment to economic development (Easterly & Levine, 1997; Gören, 2014).

As such, we undertake the largest ever examination into the causes of inter-ethnic marriage in contemporary Sub-Saharan Africa. We use Demographic and Health Survey (DHS) Couples data for up to 26 countries and employ multi-level models to uncover several novel and important findings. First, we show that inter-ethnic marriage rates are high across Africa, with inter-ethnic marriages comprising one out of every five marriages (19.4%) in our sample. Second, as expected from the literature on inter-cultural marriages in other contexts, we show that a variety of individual measures of modernization such as urbanization, literacy/education, a decline in polygamy, employment outside agriculture and higher age of marriage are significantly correlated with rising levels of inter-ethnic marriage. In support of recent evidence on the effects of educational expansion across the continent, we find a negative relationship between educational homogamy and inter-ethnic marriage, such that couples with similar levels of education are less likely to be in an inter-ethnic marriage. Finally, we show that ethnic group size is quadratically correlated with inter-ethnic marriage and that country-population size is negatively correlated with inter-ethnic marriage. All in all our results suggest that, as Africa continues to modernize via higher levels of education, urbanization, non-polygamous marriage and non-agricultural employment, inter-ethnic marriage will continue to rise into the foreseeable future but that these trends may be mitigated by higher population growth and processes of ethnic assimilation.

The rest of the paper is organized as follows. In Section 2, we discuss the previous literature on inter-marriage and provide an overview of inter-ethnic marriages in Africa. Section 3 describes the broader historical context of inter-ethnic marriage in Africa. In Section 4, we describe our data, present estimates from our multi-level model and cross-national panel data regressions and discuss our results. Finally, in Section 5, we conclude.

2 | LITERATURE REVIEW

The vast majority of literature on intermarriages has been on inter-racial marriages in the United States (see Qian & Lichter, 2011, for an overview), with some assorted literature on other countries such as the Netherlands (Kalmijn & Van Tubergen, 2006; Van Tubergen & Maas, 2007), Singapore (Lee, 1988), Sweden (Dribe & Lundh, 2008, 2011) and the United Kingdom (Muttarak & Heath, 2010). In the developing country context, there is a small literature on inter-caste and inter-religious marriage in India (Ahuja & Ostermann, 2016; Goli, Singh, & Sekher, 2013; Ray, Chaudhari, & Sahai, 2017), inter-ethnic marriage in China (Mamet, Jacobson, & Heaton, 2005) and mixed marriages by nationality in Qatar (Alharahsheh, Mohieddin, & Almeer, 2015). Contemporary cross-national examinations of intermarriage are extremely small in number, with Hou, Zheng, Schimmele, and Myles (2015)’s study of Canada and the United States the only recent such work of note.

In contrast to much of the aforementioned literature, the scholarship on inter-marriage in Africa is almost entirely qualitative in nature and focused on single field sites within countries such as Ghana, Nigeria, Tanzania and Uganda (cf. Arens & Arens, 1978; Peil, 1975; Smith, 2005). What quantitative literature on Africa exists is mostly at the individual country level, such as in Kenya (Bratton & Kimenyi, 2008), Ghana and Nigeria (Stewart, 2009) and South Africa (Jacobson, Amoateng, & Heaton, 2004), and largely avoids discussion of the determinants of inter-ethnic marriage. The only cross-national quantitative paper that exists on the subject of the causes of inter-ethnic marriage across Africa is Crespin-Boucaud (2020), which also uses DHS data to analyse intermarriage patterns for 15 African countries and finds a positive correlation between inter-ethnic marriage and both urbanization and education. However, her analysis is different from ours for several reasons. First, hers is limited to countries which have had two or more DHS surveys each, which leaves out large and important countries such as Ethiopia, Mozambique and Nigeria that are included in our analysis; as such, our analysis covers countries with 54.9% of the population of Sub-Saharan Africa compared to only 24.4% in her analysis. Second, Crespin-Boucaud (2020) runs a linear probability model with country fixed effects rather than the multi-level model we use here, which allows us to assess the role of ethnic-group and country-level effects. Third and finally, she focusses on both inter-religious and inter-ethnic
marriage, whereas ours is only on the latter, for reasons explained below; as such, we uniquely examine the effects of occupational choice on inter-ethnic marriage.¹

The literature on intermarriage has focused a great deal of attention on several important correlates of modernization as a factor that generates incentives for individuals to marry spouses from different backgrounds. Perhaps, the foremost such correlate is education (Chiappori, Oreffice, & Quintana-Domeque, 2016; Choi & Tienda, 2017; Dribe & Lundh, 2008; Furtado, 2012; Furtado & Theodoropoulos, 2011; Hou, Zheng, Schimmele, & Myles, 2015; Kalmijn & Van Tubergen, 2006; Muttarak & Heath, 2010; Qian, Lichter, & Tumin, 2018; Ray, Chaudhari, & Sahai, 2017; Van Tubergen & Maas, 2007), with a twofold mechanism: education should promote greater tolerance of individuals from different backgrounds, whereas more educated individuals are more likely to come into contact with members of other groups. Similarly, any factor that might increase inter-cultural contact before marriage could lead to greater intermarriage, with strong evidence for the positive role of urbanization (Choi & Tienda, 2017) and service in the armed forces (Arens & Arens, 1978; Chiswick & Houseworth, 2011; Fryer, 2007; Peil, 1975; Smith, 2005). Finally, we would expect a similar relationship to hold for other social norms that are associated with low historic levels of education, such as polygamy (Fenske, 2015). As such, we test for the role of education, urbanization, occupation and polygamy in our analysis below.

Education also potentially plays a role in promoting intermarriage in societies where education becomes more important in spousal selection than other factors, such that marriage between spouses of the same educational level—or educational homogamy—will rise over time while racial, ethnic or religious homogamy will decline. This trend has been documented for Protestants and Catholics in the 20th-century United States (Kalmijn, 1991), although more recent evidence has questioned the degree to which educational homogamy increases over time in the United States (Rosenfeld, 2008). In Africa, there is recent evidence for the opposite trend, such that the expansion of education across the continent has led to declining educational homogamy as both educational hypergamy (marrying someone with a higher level of education) and hypogamy (marrying someone with a lower level of education) are on the rise (Lopus & Frye, 2020). In our analysis below, we examine evidence for educational homogamy as well as its relationship with inter-ethnic marriage in Africa.

In other contexts where there is a clear distinction in status between ethnic or racial groups—such as among races in the United States or castes in India—there is evidence of ‘social status exchange’, such that potential spouses from lower status groups have a higher probability of intermarriage as their education or earnings goes up, whereas those from a higher status increase their probability of inter-ethnic marriage as their education/earnings goes down (seminally Merton, 1941; more recently Chiappori, Oreffice, & Quintana-Domeque, 2016; Kalmijn & Van Tubergen, 2006). Yet, in most cases in Africa—with the obvious exceptions of apartheid South Africa and Liberia before 1980—ethnic groups are non-hierarchical or ‘unranked’ (Horowitz, 2000, p. 22), in part because of a history of shifting ethnic power relations across the colonial and post-colonial periods. As such, it is not surprising that we find no evidence for a status exchange relationship in our African data.²

Finally, there is a robust finding in the literature for a negative role for group size (Chiswick & Houseworth, 2011; Choi & Tienda, 2017; Fryer, 2007; Furtado & Theodoropoulos, 2011; Hou, Zheng, Schimmele, & Myles, 2015; Kalmijn & Van Tubergen, 2010; Mamet, Jacobson, & Heaton, 2005; Van Tubergen & Maas, 2007), due to the fact that members of smaller groups are forced to interact more with non-group members than members of larger groups and that this effect is decreasing with size (Furtado & Theodoropoulos, 2011). As such, we add a quadratic measure of group size to our analysis.

## 3 | OVERVIEW OF INTERMARRIAGES IN AFRICA

Inter-cultural marriages have a long history in Africa. Most obviously, whole ethnic or racial groups have developed around a history of mixed marriage, most prominently the Coloured groups in southern Africa who emerged from centuries of mixing between men of European descent and women of African and/or Asian descent. In Cape Verde,
the majority ethnic group is the Creoles of mixed African and European descent, a group which features prominently in other former Portuguese colonies as well, whereas in Tanzania and Comoros, members of the Shirazi ethnic group claim to have mixed African-Persian ancestry.

Africa's high level of ethnic diversity has contributed to a long history of inter-ethnic marriages between groups as a means of developing extensive kinship ties and political alliances while still maintaining group identities, such as in Cameroon, the eastern Congo, Ethiopia, Kenya, Malawi and Nigeria (Forrest, 2004, pp. 38–41; Lonsdale, 2012, pp. 93–94). Although inter-ethnic marriage was discouraged in colonial Africa as colonial officials attempted to define tribes as distinct units (Lonsdale, 2012, p. 102)—and in apartheid South Africa white/non-white marriage was actually banned—there is nonetheless evidence that inter-ethnic marriage rates were 30% or higher in colonial towns in Sierra Leone, South Africa and Zambia (Banton, 1957; Hellmann, 1948; Wilson, 1942). Moreover, anecdotal evidence exists that inter-ethnic marriage continued to remain high in post-colonial Botswana (Werbner, 2002), Ghana (Schildkrout, 1973) and Tanzania (Arens & Arens, 1978). Indeed, there are numerous examples of post-colonial African heads of state from such varied countries as Angola, Botswana, Côte d'Ivoire, Gabon, Malawi and Senegal who married spouses from different ethnic or racial backgrounds before they took office, demonstrating the degree to which such marriages are considered socially acceptable. Thus, African inter-marriage rate trends differ from the United States, for instance, whose history of high social costs associated with inter-racial/inter-ethnic marriages (Fryer, 2007) is incomparable to the African experience (outside South Africa).

4 | RESULTS

4.1 | Data description

We use Demographic and Health Survey (DHS) Couples data, where each observation is a married couple living in the same household. The DHS programme has surveyed individuals from low- and middle-income countries since the late 1980s with the purpose of collecting data on fertility, family planning, health and nutrition; its data on ethnic identification have been used for a number of years in the political science literature (cf. Franck & Rainer, 2012; Kramon & Posner, 2013) but only recently for the purposes of understanding inter-ethnic marriage. The structure of the surveys is always the same across countries and years, such that individuals are asked first about basic personal data (age, education, access to public goods and asset ownership) before being asked about their religious and ethnic identity. For all questions (other than those about location, length of interview, etc.)—and across all surveys—the enumerator enters the response given by the respondent, and enumerators are explicitly told never to suggest answers to respondents or change the wording or sequence of questions (ICF, 2019, p. 10). We used the most recent survey per country with a full set of data, with a focus on making sure that they were regionally representative as conflict has occasionally prevented the DHS from covering certain areas (as in the 2012 survey in Mali). In some cases, the most recent survey to ask a question about ethnic identification was more than two decades old; as such, we decided to drop data for three countries (namely the Central African Republic, Rwanda and Tanzania) from our main analysis as they date to 1992–1996.

The DHS usually (but not always) collects data from both men and women in their surveys, whereupon men are interviewed randomly in every second or third household surveyed and the couples data are created by matching data from the information provided in the individual surveys; as such, the couples data are a subset of both the women’s and men’s surveys as each observation consists of one husband and one wife. The women in the couples data are always aged 15–49; the age range of the men is more variable but always includes those between the ages of 20 and 49 (and usually aged 20–54). Each observation contains both individual- and household-level data; the former include age, ethnicity, religion, literacy and education, whereas the latter include urban residence and asset ownership/wealth. To calculate data on intermarriage, we compared individual-level data from both spouses on ethnic identity and coded marriages as inter-ethnic where the ethnic identities of the spouses differed.
There are three notable limitations to the DHS data. First, we do not have either separate datasets or even separate questions for women and men on the ethnic or religious identity of their spouse which would allow us to track differential reported rates of inter-marriage by gender such as in Chiappori, Oreffice, and Quintana-Domeque (2016), although evidence from South Africa, for instance, suggests that in at least one African context the gender difference in the inter-marriage rate is close to zero (Bangstad, 2004, p. 355). Second, the data on age at marriage and date of marriage are only captured for the first marriage per individual, rather than the current marriage, which means we cannot track inter-marriages by marriage date for the full sample. (Across the dataset, 14.9% of women have been married more than once, compared to 32.9% of all men.) The third and possibly most notable limitation to the DHS data is that we do not have data on individual characteristics before and after marriage to control for the effects of assimilation into the spouse's ethnic group or religious conversion. In the former case, it is thus possible that we are underestimating inter-ethnic marriage as a consequence, although it is unclear how this problem would affect the different rates of change over time discussed below. In the latter case, there are widespread cultural norms which prohibit Muslims from marrying non-Muslims unless the non-Muslim partner converts to Islam, which would lead the marriage of someone who converted to Islam in order to get married to be incorrectly coded as intra-religious. As such, we decided to drop any examination of inter-religious marriage from our analysis, as well as the use of Muslim identification as a co-variate in explaining inter-ethnic marriage.

We are thus left with 23 country surveys from 2006 through 2018 across the continent with data on ethnic identity for both women and men; in our additional analysis discussed below, we use the full set of 26 surveys dating back to 1992. We only use one survey per country for two reasons: (1) there is wide variation in the number of surveys per country, ranging from one in many cases to seven in Senegal, and (2) ethnic categories very commonly change across surveys. (Thus, to take one extreme example, the Nigeria 2008 DHS survey lists 10 ethnic categories, whereas the 2013 survey lists 307.) As such, any pooled dataset would be unbalanced and would also limit the scope of the analysis to countries with consistent ethnic categories.

As regards the data on ethnic identification, in many cases, the data are either missing or they are given by language groups (i.e., Namibia) or race (South Africa and Zimbabwe). Due to the well-known methodological difficulties of defining and measuring inter-ethnic marriages (Craig-Henderson & Lewis, 2015), we are particularly careful to include only countries where the DHS ethnic categories correspond to popularly accepted ethnic identities in each country, as recorded in other data on ethnic fractionalization by Alesina, Devleeshauwer, Easterly, Kurlat, and Wacziarg (2003) and Fearon (2003). As seen when comparing the DHS ethnic categories to two other datasets in Table S1, in most cases, the DHS does a good job of capturing socially relevant ethnic categories, such that an inter-ethnic marriage recorded in the survey would most likely be perceived as such by other researchers and in the spouses’ community. However, for this reason, we exclude data from the Democratic Republic of Congo, where the DHS ethnic categories are given as geographically defined super-ethnicities such as Cuvette Centrale, Uele-Lac Albert and Ubangi/Itimbiri-Ngiri, whereas in another case, we used the older 2011 survey for Cameroon, whose nine ethnic categories are similar to the nine in Alesina, Devleeshauwer, Easterly, Kurlat, and Wacziarg (2003) and 11 in Fearon (2003), rather than the more recent 2018 survey with 144 categories.

We recognize that, even in countries where the DHS surveys capture local conceptions of ethnicity accurately, there often still exists a multiplicity of ways in which to measure ethnic identity. For instance, the 2011 Cameroon survey's ethnic categories include large groupings like ‘Bantu of the South-West’ and ‘Grassfields’, whereas the 2018 survey contains much smaller groups, many of which include only one respondent. Arguably, both surveys are valid in the way they capture locally understood ethnic categories, which leads us to use questions on native language as a secondary measure of ethnicity. Indeed, in much of Sub-Saharan Africa, one's native language is synonymous with ethnic identity, as seen in the use of language data as a proxy for ethnicity in much of the literature (cf. Fearon, 2003). We thus use data from the 11 countries in which surveys asked about native language usage to construct a rate of inter-language marriage as a robustness measure of inter-ethnic marriage. (We exclude countries with a lingua franca that dominates local language use like Rwanda, Burundi and Tanzania). As can be seen in Table S1 in the online supplement, the number of ethnic groups per country roughly corresponds to the number of
language groups; the three exceptions are Nigeria, Uganda and Zambia, where over 30% of respondents listed ‘other’ as their home language in each case.

4.2 | Descriptive statistics

We first begin our analysis with a series of descriptive statistics. The average percentage of marriages that are inter-ethnic across the entire dataset are 19.4%, compared to 13.6% that are inter-language. To put these numbers in perspective, inter-marriage rates in other ethnically mixed countries include inter-caste marriage rates of 6.1% in India (Goli, Singh, & Sekher, 2013) and inter-racial marriage rates in the United States of roughly 10% (Livingston & Brown, 2017). The country averages for inter-ethnic and inter-language marriage rates are given in Table 1 alongside survey size and year, and Figure 1 lists all countries in order of their inter-ethnic marriage rates at the time of their survey. One way to verify the accuracy of this data is to see if they correspond to previous survey or census evidence on inter-ethnic marriage across individual African countries. Previous academic literature on the topic has found rising rates of inter-ethnic marriage in Kenya (Bratton & Kimenyi, 2008) and higher rates of inter-ethnic marriage in Ghana than in Nigeria (Stewart, 2009), all of which correspond exactly to our findings. Crespin-Boucaud (2020) also finds comparable rates of inter-ethnic marriage of 20.4% across her smaller sample of 15 countries. In other words, we can say with confidence that our findings match other data on this topic.

Some obvious reasons help to explain much of the cross-country variation across Africa. For instance, countries with many small groups such as Zambia (54 ethnic groups with one or more members in the DHS survey, with the largest comprising only 24% of the population) will naturally see higher inter-ethnic marriage rates due to the composition of the marriage market than countries with small numbers of ethnic groups such as Guinea and Togo (with only six ethnic groups each). Another difference between countries on either end of the spectrum is the level of urbanization, which creates more possibilities for people from different ethnic groups to interact with each other: UN data at the time of the DHS survey records levels of urbanization at 16.9% in Niger versus 40% in Zambia and 86.4% in Gabon. Finally, we also observe an inverse correlation between country population size and inter-ethnic marriage, with low levels of inter-ethnic marriage in large countries such as Nigeria (with a population of 196 million people in 2018) and Kenya (44.9 million) and higher levels of inter-ethnic marriage in smaller countries such as Gabon (1.6 million) and Congo-Brazzaville (4.2 million). The theoretical basis for this correlation is that increasing scale does not necessarily lead to more ethnic fractionalization but could instead lead to a decline in social cohesion, inasmuch as each ethnic group becomes large enough that its members need rely less upon members of other groups for resources (Gerring & Veenendaal, 2020).

We next examine the times series trend for inter-ethnic marriage using date of birth in Figure 2. We use 1962 and 2000 as end dates as those are the first and last dates where there is a minimum of 200 observations per year. It is clear that the inter-ethnic marriage rate in Africa has been steadily increasing, from an average of 15% for those born in the early 1960s to 20% for those born in the early-1990s. The decline in the late 1990s can be understood as an artefact of the changing composition of the dataset, such that the last years included are increasingly dominated by the 2016 survey from Ethiopia and 2018 survey from Nigeria, which have two of the lowest rates of inter-ethnic marriage across our dataset. (This point is easily seen in Figure S1 in the supporting information, which breaks down the trend across time according to the year of survey.) The rise in the inter-ethnic (and inter-language) marriage rate corresponds to rising rates of intermarriage in other contexts such as India, Qatar and the United States (Alharahsheh, Mohieddin, & Almeer, 2015; Goli, Singh, & Sekher, 2013; Qian & Lichter, 2011).4

In Table 2, we break down the inter-ethnic marriage rate by individual and couple characteristics, with binary variables listed with their difference-of-mean test results. Although the results are similar with men, we use data for women because 5.8% of men are duplicated due to being polygamous and thus are listed twice or more in the couples data. Table 2 shows that multiple different measures of development—declining prevalence of polygamy, urban residence and increasing levels of literacy and education—are correlated with inter-ethnic marriage. It also
| Country              | Year | Survey size | Inter-ethnic Marriage rate | Inter-language Marriage rate |
|----------------------|------|-------------|---------------------------|----------------------------|
| Benin                | 2017 | 3,660       | 0.144                     | 0.142                      |
| Burkina Faso        | 2010 | 5,088       | 0.084                     |                             |
| Cameroon             | 2011 | 2,973       | 0.126                     |                             |
| Chad                 | 2014 | 3,043       | 0.116                     |                             |
| Congo-Brazzaville    | 2011 | 2,372       | 0.287                     |                             |
| Côte d'Ivoire        | 2011 | 6,745       | 0.239                     |                             |
| Ethiopia             | 2016 | 6,141       | 0.117                     |                             |
| Gabon                | 2012 | 1,946       | 0.315                     |                             |
| Gambia               | 2013 | 1,388       | 0.198                     |                             |
| Ghana                | 2014 | 1,828       | 0.200                     | 0.266                      |
| Guinea               | 2018 | 2,483       | 0.119                     | 0.109                      |
| Kenya                | 2014 | 4,211       | 0.088                     | 0.108                      |
| Liberia              | 2013 | 1,844       | 0.274                     |                             |
| Malawi               | 2015 | 3,806       | 0.330                     | 0.178                      |
| Mali                 | 2018 | 2,642       | 0.259                     | 0.156                      |
| Mozambique           | 2011 | 2,141       | 0.254                     |                             |
| Namibia              | 2013 | 1,249       | 0.195                     |                             |
| Niger                | 2006 | 2,226       | 0.121                     | 0.049                      |
| Nigeria              | 2018 | 5,582       | 0.087                     | 0.080                      |
| Senegal              | 2016 | 1,400       | 0.153                     |                             |
| Sierra Leone         | 2013 | 3,725       | 0.183                     |                             |
| Togo                 | 2013 | 2,270       | 0.109                     |                             |
| Uganda               | 2016 | 2,479       | 0.250                     | 0.180                      |
| Zambia               | 2013 | 7,198       | 0.434                     | 0.161                      |
| Average              | 2013 | 2,840       | 0.194                     | 0.133                      |

### Panel B: Pooled Data

| Variable                  | Obs. | Mean   | St. Dev. | Min. |
|---------------------------|------|--------|----------|------|
| Interethnic Marriage Rate | 68,493 | 0.193  | 0.395    | 0    |
| Age (Women)               | 76,829 | 30.711 | 8.101    | 15   |
| Age (Men)                 | 77,012 | 38.342 | 9.343    | 15   |
| Literate (Women)          | 77,012 | 0.407  | 0.491    | 0    |
| Literate (Men)            | 77,012 | 0.595  | 0.491    | 0    |
| Urban residence           | 77,012 | 0.321  | 0.467    | 0    |
| Polygamy                  | 76,006 | 0.249  | 0.432    | 0    |
| Educational Homogamy      | 77,012 | 0.595  | 0.491    | 0    |
| Ethnic group size (Women) | 75,763 | 0.173  | 0.129    | 0    |
| Ethnic group size (Men)   | 75,763 | 0.166  | 0.130    | 0    |
| Country Population (log)  | 77,012 | 16.822 | 1.196    | 14.294 |
shows that spouses with the same levels of education have lower rates of inter-ethnic marriage than those with differing levels of education, which is the opposite of what one might expect given evidence from Kalmijn (1991). This finding suggests that greater similarity between spouses in their levels of education is actually correlated with lower, not higher, levels of inter-ethnic marriage, which could be due to declining educational homogamy that itself is a result of rising levels of education across the continent (Lopus & Frye, 2020).

### 4.3 Regression results

We now turn to our regression results; we focus only on inter-ethnic marriages due to the greater availability of data. Due to the nature of the datasets available, much of the literature in the field has employed a wide range
of methods which suit the data structure, for example, linear probability models (Crespin-Boucaud, 2020; Furtado & Theodoropoulos, 2011), log-linear models (Qian & Lichter, 2011) and binomial logistic regression (Van Tubergen & Maas, 2007). However, as with other recent papers in the field (Choi & Tienda, 2017; Hou, Zheng, Schimmele, & Myles, 2015; Kalmijn & Van Tubergen, 2010), we used a multi-level logistic regression approach because the individual DHS data are nested within ethnic groups and countries. There are many clear benefits of using a multi-level approach for nested groups. Instead of estimating several models for each group (in our case often over a hundred), we obtain individual estimates by allowing for residual components at each level of the hierarchy. These individual estimates are valuable for the purposes of the study. Each of these groups are treated as a random sample from a population of groups, thus allowing us to make inferences about the groups. The multi-level models thus allow us to specifically estimate the extent to which ethnic-group and country-level differences individually explain variation in inter-ethnic marriage rates, unlike models with ethnic-group and country fixed-effects.

In order to justify the use of a multi-level model statistically, we used a variance components test to generate an intra-class correlation for each level of the analysis (ethnic group and country). The general rule of thumb is that a given level should be included in a multi-level model if 5% or more of the variance is explained at that level (Bacikowski, 1981). We found that the ethnic group level was far above the 5% threshold, at 20.0%, whereas the country level explained 8.5% of the variance. As such, we ran a three-tiered multi-level logit model, with levels 1–3

| TABLE 2 | Inter-ethnic marriage rates by female/couple characteristics |
|---------|-------------------------------------------------------------|
|         | Inter-Ethnic Marriage | Difference     |
| All     | 0.194               |                |
| Urban Residence |                |                |
| Rural   | 0.151               | 0.138***       |
| Urban   | 0.289               |                |
| Polygamous Marriage |                |                |
| Yes     | 0.142               |                |
| No      | 0.209               | 0.066***       |
| Literacy |                |                |
| Illiterate |                |                |
| Literate | 0.259               | 0.110***       |
| Level of Education |                |                |
| No Education |                | 0.127         |
| Some Primary |                | 0.223         |
| Some Secondary |                | 0.283         |
| Some Higher   |                | 0.318         |
| Educational Homogamy |              |                |
| Same level of education |    | 0.178         |
| Different levels of education | | 0.216 | 0.037*** |

Note. The level of statistical significance is given for a mean-comparison test between two variables and is listed after the second observation. The results are nearly identical for men.

* $p \leq 0.1$.

** $p \leq 0.05$.

*** $p \leq 0.01$ (two-tailed tests).
accounted by the individual, ethnic group and country. We included age, age squared and literacy as individual-level variables for both women and men. We did not include the wealth variable, as it is potentially post-treatment, but did include urban residence and polygamy, which are measured at the household level but are arguably individual characteristics inasmuch as individuals may choose to live in urban areas and/or marry polygamously before they consider the ethnic background of their spouse. At the group level, we included group size and size squared to account for size of the marriage market. At the country level, we included the country (logged) population, which was the only country-specific variable that was consistently statistically significant. (We tried adding a wide variety of country-level variables that might be correlated with ethnic diversity and inter-ethnic relations such as historical slave exports, the level of democracy, GDP per capita and spatial ethnic segregation [Robinson, 2020] in Table S4, none of which was statistically significant.) Finally, we included the year of the survey as a time trend and as a test of the validity of our specification. In other words, if we have included the correct co-variates in our model, then the year of the survey—which functions as a measure of a cohort effect after controlling for age—should not be statistically significant in either direction.

Our first set of results is listed in Table 3; we first list the results for women and then men, with polygamous men only included once (leading to a lower number of observations in Column 2). (In Table S3, we present results from the 26-country sample discussed above, which includes data from the Central African Republic [DHS survey in 1994], Rwanda [1992] and Tanzania [1996], with results almost identical to those of Table 3.) At the individual level, literacy is always positively and statistically associated: the increase in the odds for being in an inter-ethnic marriage when literate is 44% for women and 31% for men. As expected, for both men and women inter-ethnic marriage is positively associated with urbanization and negatively with both polygamy and educational homogamy, which conform with our findings in Table 2.

### TABLE 3 Inter-ethnic marriage, main results (dependent variable: inter-ethnic marriage)

| Individual Characteristics | Women            | Men            |
|----------------------------|------------------|----------------|
| Age                        | 0.048*** (0.010) | 0.036*** (0.009) |
| Age (squared)              | −0.0008*** (0.0002) | −0.0005*** (0.0001) |
| Literate                   | 0.433*** (0.026) | 0.356*** (0.028) |
| Urban                      | 0.832*** (0.024) | 0.89*** (0.025) |
| Polygamy                   | −0.112*** (0.029) | −0.132*** (0.032) |
| Educational Homogamy       | −0.137*** (0.022) | −0.060*** (0.023) |
| Group size                 | −11.747*** (1.564) | −7.761*** (1.150) |
| Group size (squared)       | 19.222*** (4.234) | 19.229*** (3.073) |
| Country Population (log)   | −0.340*** (0.103) | −0.241** (0.122) |
| Year of Survey             | −0.014 (0.030)   | −0.042 (0.033)   |
| Constant                   | 32.728 (59.571)  | 87.159 (66.122)  |
| Wald Chi squared           | 2202.09***       | 1972.79***      |
| Countries                  | 23               | 23             |
| Ethnic Groups              | 375              | 366            |
| Observations               | 67,727           | 62,901         |

Note. The results above are multi-level logit model estimates with random intercepts at the ethnic group and country levels. 
p ≤ 0.1. 
**p ≤ 0.05. 
***p ≤ 0.01 (two-tailed tests).
The group-level results show in both columns a very strong U-shaped quadratic relationship with intermarriage, which again concords with previous literature and suggests a large role for group size in determining inter-group marriages. Finally, at the country level, there is a consistently negative and significant relationship between inter-ethnic marriage and population size, such that more populous countries have lower rates of inter-marriage. This novel finding is possibly driven by the scale effects discussed above (Gerring & Veenendaal, 2020), inasmuch as more populous ethnic groups are more self-reliant and thus their members may interact less with members of other groups. One obvious implication of this finding is that population growth could have a dampening effect on inter-ethnic marriage within Africa; however, this relationship remains a topic for further research.

As an alternative measure of modernization, we next examine the results according to occupation in Table 4, using the raw occupational identities given in the DHS surveys (and unemployed status as the null category); we lose data from a number of countries which did not include a question on clerical work or membership of the armed forces. As discussed above, we would expect that any profession that increased inter-cultural contact would be positively associated with inter-ethnic marriage, with previous evidence finding a positive relationship with service in the armed forces in particular. Here, our results support this finding, with a positive and statistically significant coefficient for men in the armed forces. (We do not include a variable for armed service for women in Column 1, as only four women in the entire sample listed the army as their occupation, vs. 492 for men.) For women, more modern professions such as clerical work, professional/technical/managerial positions and employment in the sales/services industry are associated positively with inter-ethnic marriage, whereas agricultural work is negatively correlated for both men and women. We should note that these variables are potentially endogenous as they might have been chosen post-marriage; unfortunately, the DHS does not ask how long individuals have been in their given occupation which might allow for a more causal explanation.

In our final table, we examine the determinants of inter-ethnic marriage at the country-survey level, using the full universe of DHS surveys with data on ethnicity dating back to the early 1990s. As noted above, this sample

| TABLE 4 Inter-ethnic marriage and occupation (dependent variable: inter-ethnic marriage) |
|---------------------------------|---------------------------------|-----------------|
| **Individual Characteristics**  | **Women**                       | **Men**         |
|                                 |                                 |                 |
|                                 | 1                               | 2               |
| **Controls from Table 3**       | Yes                             | Yes             |
| **Job: Agriculture**            | –0.305*** (0.032)               | –0.559*** (0.041) |
| **Job: Army**                   |                                 | 0.559** (0.119) |
| **Job: Clerical**               | 0.634*** (0.135)                | 0.122 (0.103)   |
| **Job: Professional/Technical/Managerial** | 0.471*** (0.061) | 0.037 (0.052)   |
| **Job: Sales and Services**     | 0.197*** (0.033)                | –0.082 (0.053)  |
| **Job: Skilled Manual**         | –0.083 (0.065)                  | –0.018 (0.045)  |
| **Job: Unskilled Manual**       | 0.139** (0.063)                 | 0.036 (0.053)   |
| **Constant**                    | –21.920 (62.984)                | 51.961 (45.254) |
| **Wald Chi squared Countries**  | 2236.41*** 20                   | 2011.95*** 18   |
| **Ethnic Groups**               | 342                             | 308             |
| **Observations**                | 58,437                          | 45,129          |

Note. The results above are multi-level logit model estimates with random intercepts at the ethnic group and country levels. For the women’s sample, there is no data on clerical occupation from Benin and Senegal, and on unskilled labour from Benin and Chad; for the men, there are no data on army occupation from Benin, Burkina Faso, Ghana, Mozambique and Senegal.

*p ≤ 0.1.

**p ≤ 0.05.

***p ≤ 0.01 (two-tailed tests).
includes wildly varying numbers of ethnic categories per survey, especially in Cameroon (between 9 and 144 categories), Congo-Brazzaville (11 to 50) and Nigeria (10 to 307). As such, we regress the country-survey average level of inter-ethnic marriage on the (logged) number of ethnic categories alongside the full set of variables that were statistically significant in Table 3, plus round fixed effects. More specifically, we included country-survey averages of (female) literacy and urbanization alongside the (logged) country population and the country-level female ethno-linguistic fractionalization index (ELF) as a composite measure of the relative size of ethnic groups. We also included the correlation coefficient between the highest level of education achieved by husbands and wives in each couples country survey, which captures educational homogamy in the sample. More specifically, a higher correlation coefficient means that there is more similarity in educational levels between the sexes; if Kalmijn (1991)’s evidence on the positive relationship between educational homogamy and inter-ethnic/racial/religious marriage holds in Africa, then we would thus expect there to be a positive correlation between the education correlation coefficient and inter-ethnic marriage.

We use both a random effects model and pooled OLS for robustness due to the small number of observations; the random effects model with its specific within and between modelling allows the researcher more flexibility in modelling, making it more suitable than a fixed-effects model. (As before, no other country-year variables such as the absolute number of ethnic groups, the level of democracy, GDP per capita and spatial ethnic segregation, alongside other country time-invariant variables like British colonization, absolute latitude and historical slave exports, were ever statistically significant.) We cluster our standard errors at the country level; whereas the number of country clusters is low, there is ample evidence that the standard errors are not largely biassed due to the use of lower numbers of clusters (Hansen, 2007).

As we show in Table 5, our analysis confirms our previous findings on ethnic group size and country population, whereas literacy and urbanization are positive and statistically significant in the random-effects model but not when using OLS. We also find a negative and statistically significant coefficient on the education variable in both models, which matches our results in Tables 2 and 3 above.

**Table 5** Inter-ethnic marriage, country-round observations (dependent variable: inter-ethnic marriage)

| Estimation Method | RE | Pooled OLS |
|-------------------|----|------------|
|                   | 1  | 2          |
| Number of Ethnic Categories (log) | 0.035** (0.017) | 0.050** (0.024) |
| Literacy (Women)  | 0.159** (0.065) | 0.124 (0.082) |
| Urbanization      | 0.140** (0.055) | 0.034 (0.145) |
| ELF               | 0.239** (0.094) | 0.244* (0.139) |
| Country Population (log) | −0.042*** (0.011) | −0.053*** (0.018) |
| H/W Education Correlation Coefficient | −0.124* (0.067) | −0.333*** (0.133) |
| Round Fixed Effects | Yes | Yes |
| Constant          | 0.529*** (0.173) | 0.835*** (0.248) |
| R² (within)        | 0.637 |          |
| R²                 | 0.656 |          |
| Countries          | 26  | 26        |
| Observations       | 76  | 76        |

Note. The standard errors are clustered at the country level.

*p ≤ 0.1.

**p ≤ 0.05.

***p ≤ 0.01 (two-tailed tests).
5  |  DISCUSSION AND CONCLUSIONS

In this paper, we analysed the determinants of inter-ethnic marriage in contemporary Sub-Saharan Africa, using DHS couples data from up to 26 countries whose combined population comprises the majority of the region’s population. We found that the prevalence of inter-ethnic marriage is high at 19.4%, with strong evidence that individual measures of modernization such as literacy/education, urbanization, non-polygamous marriages and non-agricultural employment are correlated with inter-ethnic marriage and could help to explain why inter-ethnic marriage rates have been steadily increasing in Africa since the 1980s. We also found a negative relationship between educational homogamy and inter-ethnic marriage, which may be a result of educational expansion across the continent. A U-shaped relationship with group size and a novel finding of a negative relationship with country population size—which was only possible due to the relatively large number of countries used in the study—help to explain variance across ethnic groups and countries within Sub-Saharan Africa. An analysis at the country-round level confirmed these findings.

The implications of our findings are several. On the one hand, inter-ethnic marriage should continue to grow in prevalence in Africa as the continent’s population continues to become more educated and urbanized, and less polygamous and employed in agriculture, especially as we do not find evidence of fewer children from inter-ethnic marriages. On the other hand, there are three mitigating factors that could limit the growth of inter-ethnic marriages in the future, namely, country population, ethnic structure (i.e., ethnic group size and ELF) and educational homogamy. In the first case, the negative relationship between country population and inter-ethnic marriage is an important one in a continent with high population growth rates as a consequence of high fertility and low mortality rates, particularly in the Sahel and Central/East African regions. Second, the U-shaped relationship with ethnic group size (and positive relationship with ELF) is also notable in a context where there is a history of ethnic assimilation from smaller groups to larger groups, whether via industrialization as individuals assimilate from smaller to larger ethnic groups (Gellner, 1983; Green, 2019), the provision of public goods due to the desire to identify with politically dominant ethnic groups (Wimmer, 2018) or other factors. In other words, processes of ethnic assimilation that lead ethnic groups to grow in size and ELF measures to diminish could have a negative effect on inter-ethnic marriage across the continent. However, if the offspring of inter-ethnic marriages identify as a separate group instead of with either of their parents’ groups—as historically happened in the ethnogenesis of the Coloured, Creole and Shirazi groups, as noted above—then these processes could again be mitigated. Third and finally, although educational homogamy is currently on the decline in Africa, this trend is likely to reverse in coming decades as the educational gap between men and women closes and norms about educational hypergamy change (Lopus & Frye, 2020); thus, if the negative relationship between educational homogamy and inter-ethnic marriage continues to hold, rising levels of educational homogamy could contribute to lower levels of inter-ethnic marriage in the future.

There are multiple potential avenues of future research on this topic. With more detailed personal histories, it would be possible to test the theory that exposure to inter-ethnic violence could be correlated with lower levels of intermarriage (Horowitz, 2000, pp. 61–62). With data on the parents of spouses, it might also be possible to test Ray, Chaudhari, and Sahai (2017)’s argument about the effect of parental education on inter-caste marriage in India in the African context. Similarly, it would be interesting to use DHS data in other developing country contexts outside Africa. For instance, preliminary analysis of DHS couples data on inter-ethnic marriage in the Philippines shows an inter-ethnic marriage rate of 26% and a positive correlation with urbanization but surprisingly no correlation with literacy, education or group size.

Finally, tracking data on the children of inter-marriages could allow for better understanding of how such children identify ethnically once they become adults, which would greatly add to our understanding of how inter-marriages can affect ethnic demography across time. Indeed, preliminary analysis from Dulani, Harris, Horowitz, and Kayuni (2020) drawing from survey data in Kenya and Malawi suggests that children of inter-ethnic marriages are less likely to engage in ethnic voting, which suggests that inter-ethnic marriage can have important political implications as well. Similarly, Le Bas (2017)’s examination of the effect of inter-ethnic marriage on trust in local officials in Nigeria could be replicated elsewhere.
ACKNOWLEDGEMENTS
For comments and suggestions we would like to thank Ernestina Coast, Arjan Gjonça, Tiziana Leone and seminar participants at the LSE. All errors remain our own.

DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID
Sanghamitra Bandyopadhyay https://orcid.org/0000-0003-4299-9599
Elliott Green https://orcid.org/0000-0002-0942-5756

ENDNOTES
1 We should note as well that by including data from 2018 DHS surveys in Guinea, Mali and Nigeria, our analysis is more up to date.
2 We used two different measures of high status for African ethnic groups—the co-ethnics of the President and members of the Staatsvolk (the historically politically and socially dominant ethnic group; cf. Green, 2020)—and examined whether there was a trade-off between status and literacy for those in inter-ethnic marriages and found no evidence for either men or women (results available from authors upon request).
3 Indeed, it could be possible that spouses differ on whether they view their marriage as inter-ethnic, depending on how they define their ethnicity and that of their spouse.
4 We also examined intermarriage between individuals from matrilineal and patrilineal groups, who have different systems of inheritance; using data from Rossi and Rouanet (2015) across 13 of the countries in our dataset, we similarly find an upward sloping trend by birth cohort from 5% to 9% between the 1960s and 1990s.
5 The results with inter-language marriage are given in Table S2 and largely correspond to those for inter-ethnic marriage.
6 We included data on literacy rather than level of education as the former is much easier to compare across different contexts than the latter, given the large heterogeneities that exist in the quality of education within Africa. Literacy is measured directly by the enumerator, who ask respondents to read a sentence in their chosen language. We code respondents as literate when they can read parts of the sentence or the whole sentence and code them as illiterate when they cannot read the sentence at all.
7 In the former case evidence, we reran our analysis using a sub-sample of 18,382 individuals who had always lived in the same location, with a positive and statistically significant coefficient for urban residence, whereas in the latter case, there is evidence that contemporary decisions to engage in polygamous marriage are historically rooted in the colonial period (Fenske, 2015).
8 For polygamous men, we use the couple where they are first listed in the dataset, which amounts to a random sample according to the rank of the wife within polygamous relationships.
9 We only include one country-survey per round.
10 Education is measured ordinally, with no education, primary, secondary and tertiary as the categories.
11 In our analysis, we did not include country-survey averages of polygamy, as it is highly correlated with literacy at the country-survey level ($r = −0.81$); it is not, however, ever statistically significant when included in an alternative model (available from the authors upon request).

REFERENCES
Ahuja, A., & Ostermann, S. L. (2016). Crossing caste boundaries in the modern Indian marriage market. Studies in Comparative International Development, 51(3), 365–387. https://doi.org/10.1007/s12116-015-9178-2
Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. Journal of Economic Growth, 8(2), 155–194. https://doi.org/10.1023/A:1024471506938
Alharahsheh, S. T., Mohieddin, M. M., & Almeer, F. K. (2015). Marrying out: Trends and patterns of mixed marriage amongst Qataris. International Journal of Social Science Studies, 3(6), 211–225. https://doi.org/10.4000/cy.4399
SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Bandyopadhyay S, Green E. Explaining inter-ethnic marriage in Sub-Saharan Africa. J. Int. Dev, 2021;33:627–643. https://doi.org/10.1002/jid.3535