

## Explaining African Ethnic Diversity

Elliott Green<sup>1</sup>

February 2012

Forthcoming in the *International Political Science Review*

The consequences of ethnic diversity in Africa have been widely noted. However, the sources of this diversity remain unexplained, which is surprising in light of Africa's high level of ethnic diversity as well as its large internal variation. Here I show that ethnic diversity in Africa is a result of its tropical location, the pre-colonial slave trade, the colonial creation of large states and low levels of urbanization, and that my results are robust to various controls. I also show that, after controlling for the slave trade, Africa is no more ethnically diverse than the rest of the world.

---

<sup>1</sup> Department of International Development, London School of Economics, Houghton Street, London WC2A 2AE, UK. Email: [e.d.green@lse.ac.uk](mailto:e.d.green@lse.ac.uk).

## 1. INTRODUCTION

In recent years Sub-Saharan Africa's ethnic diversity has become almost synonymous for the continent's economic and political problems. Scholars have argued that ethnic diversity has been responsible for Africa's low economic growth (Easterly and Levine, 1997; Posner, 2004a), political instability and conflict (Buhaug, 2006; Easterly, 2001), high inequality ( Barr and Oduro, 2002; Milanovic, 2003; ) and low provision of public goods (Kimenyi, 2006; Miguel and Gugerty, 2005).<sup>1</sup> What remains unexplained, however, are the origins of Africa's ethnic diversity.

In particular what makes Africa particularly interesting in this regard is both its unusually high levels and large internal variation in ethnic diversity. As regards the former, previous attempts to explain global ethnic diversity still find an Africa dummy variable to be statistically significant (Ahlerup and Olsson, 2011; Michalopoulos, 2012). As regards the latter, Africa's standard deviation in the number of ethnic groups per country is more than 35% higher than any other region according to one recent dataset on ethnic diversity (Fearon, 2003: 204).<sup>2</sup> Moreover, as seen in Table 1, another recent dataset from Alesina et al. (2003) shows that the standard deviation in ethnic diversity per country as measured by the Ethno-Linguistic Fractionalization (ELF) index is largest in Africa,<sup>3</sup> which also contains the countries with the smallest and largest measures of ethnic diversity in the world. Indeed, in a continent with high levels of diversity in such countries as Liberia and Uganda it is striking to find largely homogenous countries like Burundi and Comoros.

[Insert Table 1 here]

In this paper I probe the origins of African ethnic diversity. I show that ethnic diversity within Africa is a result of its low latitudes, the pre-colonial slave trade, the creation of large

colonial states and low levels of urbanization. More specifically, I demonstrate that a variable capturing the number of slaves exported per country between 1400 and 1900 eliminates the statistical significance of an Africa dummy variable. I also show that low population densities led colonialists to create large states, thereby leading to higher levels of ethnic diversity within Africa. Moreover, I use a lagged measure of urbanization to demonstrate that urbanization is correlated with ethnic homogenization over time. Finally, I employ survey data from Afrobarometer to show that urbanization is associated with greater attachment to the nation than to one's ethnic group. The paper thus suggests that African ethnic diversity has pre-colonial, colonial and post-colonial roots and is continuing to change.

The paper is organized as follows. First, I examine a variety of theories for the pre-colonial and colonial origins of African ethnic diversity, and test them in both an African sample and a global sample excluding the Americas and Oceania. I then examine theories of ethnic change in post-colonial Africa and use two different empirical strategies to show that urbanization is correlated with ethnic homogenization and greater attachment to the nation. Finally I conclude with wider thoughts on African ethnic diversity.

## 2. AFRICAN ETHNIC DIVERSITY IN HISTORICAL PERSPECTIVE

The vast majority of political economy work on Africa takes ethnicity as fixed and thus uses it as an independent variable rather than a dependent variable.<sup>4</sup> This phenomenon has continued to occur despite the fact that anthropologists began to move away from a 'primordial' understanding of ethnicity towards a more 'constructivist' point of view in the late 1960s (Barth, 1969; Cohen, 1969). Indeed, of late political scientists and sociologists have continued to emphasize the dynamism of ethnicity (Chandra, 2006; Brubaker, 2004), although it is too early to say whether these scholars have made a major impact on the field.

One strong conclusion that has come out of the constructivist literature is that modern processes of economic and political development have profoundly shaped ethnic identities. Gellner (2006 [1983]), for instance, proposed that industrialization can have a strong impact on the formation of cross-ethnic national identities, while the colonial state has been acknowledged to have had a major impact on shaping modern ethnic identities in Africa (Posner, 2003; Laitin, 1994). I thus focus my attention here on a variety of potential causes of African ethnic diversity, first beginning with pre-colonial and colonial origins in the rest of section 2 before moving onto post-colonial sources in section 3.

### *2.1. Pre-Colonial and Colonial Origins*

There are several already existing hypotheses about the origins of Africa's pre-colonial ethnic diversity. Several of these are geographic in nature: for instance, many scholars have suggested that latitude and/or temperature has an inverse relationship with ethnic diversity, in that warm tropical environments are ideal for growing food and thereby create few incentives for inhabitants to migrate elsewhere or trade extensively with other human populations. This isolation thus spurs the creation of new ethnic groups in very much the same way as Darwin had proposed for finch species in the Galapagos islands (Ahlerup and Olsson, 2011; Cashdan, 2001; Collard and Foley, 2002; Nichols, 1992; Sutherland, 2003; ; ). Other researchers have found a significant positive correlation between elevation and ethnic diversity (Nichols, 1992; Sutherland, 2003), which presumably works along the same mechanism of encouraging isolation among human populations.

An alternative hypothesis for the pre-colonial era is the slave trade: as argued by Nunn (2008), the slave trade may have weakened ties between Africans and thereby inhibited the development of broader ethnic identities. Empirically there is a good deal of evidence supporting a relationship between slavery and ethnic diversity: Nunn (2008: 164), for instance, shows a

significant and strong correlation between slave exports per area and contemporary ethnic fractionalization. The names of many African ethnic groups also suggest a link, such as the Dioula of West Africa, whose name comes from the Manding word *gyo-la*, or ‘the land of the slaves captured through warfare,’ while the name of their neighbours the Gio possibly comes from the Bassa phrase *gii-o*, or ‘slave people’ (Holsoe and Lauer, 1976: 141-142). Moreover, this evidence is not limited to West Africa: data from 1856 on freed slaves in what is now north-west Mozambique lists at least 21 different ethnic backgrounds (Isaacman, 1972: 447), while the current Bantu ethnic minority in southern Somalia originated as slaves from southern Africa who were shipped along the East African coast in the 19<sup>th</sup> century by Zanzibari traders (Webersik, 2004: 525-526).

Moving closer to the present, scholars have long argued that the colonial period had a profound effect on ethnicity in Africa, for a variety of reasons. In particular the literature suggests that colonialists directly promoted ethnic diversity through ‘divide and rule’ tactics (Berman, 1998; Blanton, Mason and Athow, 2001) and indirectly by providing ‘tribal’ chiefs and missionaries with the incentives to promote ethnic differences (Laitin, 1994). Scholarship on the colonial period has also repeatedly emphasized the way colonies were governed according to cost-saving methods. Indeed, at the time of its colonization Africa was relatively peripheral to European economic interests (Young, 1994: 84-85), and the principle of colonial self-sufficiency meant that colonies had to pay for themselves, no matter how oddly shaped or small they were.

The consequences of colonial cost-saving on African ethnic diversity were threefold. First, European efforts at creating economies of scale and minimizing bureaucratic costs led to the creation of large states (Green, 2012). For instance, the average British district commissioner was responsible for an area the size of Wales (Herbst, 2000: 78), while French colonial administrators created unusually large colonies in the more under-populated areas of Central and West Africa. The result has been that the median former colony in Africa is 37% and 74% larger than the median former colony in Latin America and the Caribbean or Asia, respectively (Green, 2012). When we

recall that ethnic groups in Africa are different from caste groups in India or racial groups in the US in the way they tend to be regionally concentrated, it is thus no surprise to find that smaller African states like Burundi, Djibouti and Lesotho have lower levels of ethnic diversity than Chad, the DRC or Tanzania.

A second consequence of colonial cost-saving was the drawing of arbitrary colonial borders. More specifically, due to the costs involved in drawing accurate borders that reflected local political economies, Europeans inadvertently ‘split up ethnic groups and exacerbated pre-existing high levels of ethnic and linguistic diversity’ when dividing up the continent (Easterly and Levine, 1997: 1213). In particular Englebert, Tarango and Carter (2002: 1096) have noted that some 44% of colonial borders were straight lines, leading to as many as 177 ethnic groups split across two and sometimes three colonial borders; when added together these partitioned ethnic groups represent 43% of the average African state’s population. Indeed, the supposed arbitrariness of these borders has led some social scientists like Miguel (2004) and Posner (2004b) to use them as ‘natural’ experiments inasmuch as they were drawn exogenous to pre-existing local political conditions.<sup>7</sup>

A third consequence of colonial cost-saving was the creation of supposedly homogenous rural ‘tribal’ areas with their own ‘tribal’ chiefs as ‘decentralized despots’ (Mamdani, 1996). Labelled as ‘indirect rule,’ this policy led in areas to the creation of new ethnic identities as colonialists split up Africans into separate tribal areas with their own tribal chiefdoms, thereby encouraging separation and ethnic difference. Moreover, the way in which colonial rule restricted both migration between tribal areas and property rights to ‘natives’ from each area only helped to encourage isolation and further fractionalization.

## *2.2. Empirical Analysis*

My goal is to regress ethnic diversity on these aforementioned variables. As already noted there exists two recently compiled datasets on global ethnic diversity, namely Alesina et al. (2003)

and Fearon (2003). What makes these dataset different from previous codings of ethnic diversity is their focus on ethnic rather than ethnolinguistic diversity, whereby groups in Rwanda and Somalia who speak the same language but are ethnically distinct are coded as such. Thus I deliberately do not employ other datasets that explicitly use language as a proxy for ethnicity such as (Roeder, 2001). I utilize both of these new datasets here inasmuch as they both have advantages and disadvantages: in particular, Alesina et al. (2003)\_ has a wider coverage of countries but is slightly less accurate in certain cases than Fearon (2003), such as with Tanzania.<sup>9</sup>

To summarize my hypotheses, I have listed seven historic variables here which are hypothesized to have effects on ethnic diversity in Africa: for the pre-colonial period I have elevation, latitude, temperature and slave exports and for the colonial period I listed state size, artificial borders and indirect rule. Several variables present no problems in measurement, with uncontroversial data available for elevation, latitude, temperature and state size.<sup>10</sup> Other data, however, present certain problems. Nunn (2008), for instance, uses contemporary ethnic group identities to code historical slave exports, which he then claims led to the creation of modern ethnic groups, a problem which has been already criticized elsewhere (Austin, 2008). However, I use his data here as it is so far the only such cross-national data on slave exports.

Artificial borders could be estimated according to Englebert et al. (2002)'s estimates on the number of straight and non-straight borders in Africa. Instead, however, I utilize a new 'fractal' measure of artificial borders by Alesina, Easterly and Matuszeski (2011) which captures the degree to which a country's land borders are one or two dimensional. Thus African countries range from the straight-line borders of Somalia (fractal = 0.003) and Equatorial Guinea (0.007) to the curved and jagged borders of Sierra Leone (0.055) and Rwanda (0.072).

Finally, measuring indirect rule is also problematic since the sole available measure is one created by Lange (2004) for British colonies, leaving us with only 14 observations and a non-representative sample.<sup>11</sup> Another alternative is to use colonial identity as a proxy for indirect rule, inasmuch as there is some evidence that the British were more forthright in using indirect rule than

the French in their African colonies (Grier, 1999; Blanton, 2001). I thus introduce colonial dummies for the three main colonial powers of France, Portugal and the UK.

I begin my analysis by testing various historical variables against each other. Columns 1-3 employ Alesina et al. (2003) as the dependent variable while columns 4-6 employ Fearon (2003). Inasmuch as elevation and temperature are highly correlated with each other for obvious reasons ( $r = -0.72$ ), I include them in separate regressions. I also report the results with the *Fractal* variable separately in columns 3 and 6 as Alesina et al. (2011) do not compute their data for every African country.

Nonetheless, as reported in Table 2, I find a series of consistent results. First, absolute latitude is correlated with ethnic diversity across all six specifications while elevation, temperature and the colonial dummy variables are never significant. Second, slave exports are always positively and significantly correlated with ethnic diversity, with a consistent coefficient around 0.03, as is state size with a coefficient between 0.04 and 0.05. Moreover, in all six regressions I can explain a majority of the variation in ethnic diversity across Africa.

[Insert Table 2 here]

To examine whether these variables explain the Africa dummy variable I can thus regress Alesina et al. (2003) and Fearon (2003) on a series of variables on a global scale. However, Nunn (2008)'s data on slavery was only calculated for Sub-Saharan and North Africa. I can supplement this data by examining non-African countries where evidence exists that they exported slaves in commercial exchanges between 1400 and 1900. While by no means as well documented as the Atlantic slave trade, recent scholarship has begun to estimate the number of slaves involved and their origin in the early modern Indian Ocean, Central Asian and Crimean slave trades. As in Nunn (2008) I normalize the export figures by taking the natural log of the total number of exports divided by the size of the country (and multiplied by 1000).



My new estimates, summarized in Appendix 2, shows that Africa was by no means the only part of the world to export significant numbers of slaves in the early modern period. For instance, evidence suggests that slave traders in the Indian Ocean not only exported slaves from east Africa but also from the Coromandel coast in south India and the Arakan/Bengal coast in what is now Bangladesh and Myanmar (Carter, 2006; Vink, 2003). Moreover, there is evidence of Portuguese slave traders exporting slaves from India, Indonesia and Malaysia to Manila in the late 16<sup>th</sup> and early 17<sup>th</sup> centuries (Seijas, 2008: 22-23), while Warren (2007: xxxv) estimates that around 300,000 slaves from Indonesia and the Philippines were exported to the Sulu Sultanate (in what is now south-west Philippines) between 1768 and 1878.

As for the Central Asian slave trade, the expansion of the Mughal Empire in the 16<sup>th</sup> and 17<sup>th</sup> centuries resulted in ‘the enslavement and exportation of hundreds of thousands of individuals’ from what is now India and Pakistan (Levi, 2002: 283-284), which I estimate here as around 300,000 people. Given a ratio of 2:1 of Indo-Pakistani slaves to Afghan slaves in late 16<sup>th</sup>-century Samarkand (Levi, 2002: 284), I also estimate that roughly 150,000 Afghans were subjected to the Central Asian slave trade.<sup>13</sup> Finally, an estimate that half of the population of Khanata of Khiva was comprised of Persian slaves in 1851 suggests that Iran exported a minimum of 350,000 slaves during this period (Hopkins, 2008: 645). To compare these figures with Nunn (2008)’s African data suggests that slave exports from Afghanistan were roughly proportional to the number of exports from Congo-Brazzaville while Iran was more comparable to Madagascar and India/Pakistan to Ethiopia.

But perhaps the most striking numbers come from Kizilov (2007: 7), who suggests that there were roughly 2 million slaves exported from ten countries in Eastern Europe through the Crimean Slave trade between 1500 and 1700. Indeed, one source discusses the negative political and demographic effects of slavery in early modern Georgia in much the same way as it is often discussed for Africa:

The constant feudal warfare and the miserable life of the serfs on the individual estates served the interests of the slave traders so well that the population of the entire region was greatly reduced by the eighteenth century to the benefit of the rival Safavid and Ottoman states (Kortepeter, 1966: 105).

With this new data I can thus estimate the effect of slavery on contemporary measures of ethnic diversity on a global scale and observe whether it has any effect on the Africa dummy variable.<sup>14</sup> I exclude the Americas and Oceania from my analysis since there is no evidence of any early modern slave trade exporting slaves within either region (as opposed to importing them from elsewhere), making the Africa dummy and the slave export variable perfectly correlated within this subset. For the rest of the sample, however, the Africa dummy and the slave export variable correlate at  $r = 0.59$ .

Table 3 presents my results, again using both datasets as dependent variables. I first start out in columns 1 and 3 by regressing ELF on a set of variables already demonstrated by Ahlerup and Olsson (2011) and Michalopoulos (2012) as correlating with global ethnic diversity, namely elevation, latitude, state size and dummies for Sub-Saharan Africa and British and French colonial rule. Finally, I also include a new variable computed by Ahlerup and Olsson (2011), who model the emergence of new ethnic groups over time in response to a lack of public goods; they calculate the length of time modern humans have populated each country, measured here in millions of years, and find a positive correlation between origin time and contemporary ethnic diversity.

[Insert Table 3 here]

In columns 1 and 3 both latitude and the Africa dummy are statistically significant in the expected directions, while none of the other variables are significant. In columns 2 and 4, however, I introduce the updated slavery variable, which is not only highly statistically significant but also entirely removes the significance of the Africa dummy. Moreover, the slave export variable appears not to affect the other variable except to increase statistical significance for elevation;

interestingly, the origin time variable is not significant in any of my four regressions. Finally, it is important to note that the slavery variable produces a higher  $R^2$  as well in both regressions; indeed, in univariate regressions excluding the Americas and Oceania the slavery variable explains 32% of the variation in ELF in the Fearon dataset and 34% in the Alesina dataset despite the fact that my data only records 62 countries as having exporting any slaves at all.

### 3. CONTEMPORARY AFRICAN ETHNIC DIVERSITY

#### *3.1. Post-Colonial Origins*

So far I have examined the potential effects of historical variables on African ethnic diversity. Here I turn towards more contemporary suggestions from many scholars that socio-economic development can alter ethnic diversity. For instance, Alesina and Spolaore (2003) have suggested that ethnic heterogeneity leads to poorer public goods provision due to diverse preferences, and that as a result citizens from ethnically heterogeneous states have the incentive to secede and create new, more ethnically homogenous states (and thereby also make the original ‘mother’ state more homogenous as well). They claim that these incentives would be enhanced as countries democratize and thus give their citizens greater latitude to choose their own future. Yet there is very little evidence supporting this theory in post-colonial Africa, where there has been a remarkable lack of secessionist movements even after democratization swept the continent in the 1990s (Englebert, 2009).

There is, however, more evidence supporting the role of urbanization in promoting homogenization. Social theorists from Marx and Engels to Gellner (2006 [1983]) have invoked urbanization as a key variable in the transition from old multi-ethnic agrarian societies to modern nation-states, while historians like Weber (1976) have shown how urbanization spread nationalist ideology and homogenization in France and other western countries in the 19<sup>th</sup> and 20<sup>th</sup> centuries.

In particular urbanization has been seen as both the mechanism by which people from different but related ethnic groups come in contact with each other, realize their commonalities and form new and broader ethnic identities, and the way by which the state can enforce nationalist ideology through education and propaganda. Rural citizens, on the other hand, remain isolated from each other physically and socially, and modern nation-states have often had difficulties reaching out to their rural peripheries. Within Africa perhaps the best example of the effects of urbanization on ethnic diversity phenomenon is Botswana, which had a 7.1% annual rate of urbanization between 1960 and 2000, or the highest of any country in the world. The result of such a large-scale shift was the creation of neutral ethnic spaces which ‘provided an enabling context for minorities to join and partially “assimilate” to national life’ (Solway, 2004: 132).

In Africa the relative lack of urbanization could be one reason behind the continent’s high ethnic diversity. Indeed, colonial rulers deliberately suppressed urbanization until the mid-twentieth century, partially due to the need to utilize labour in rural areas but also because they were worried about the existence of an urban ‘detribalized’ proletariat as a fertile recruiting tool for nationalist anti-colonial movements (Mamdani, 1996: 102). Many of these policies persisted into the post-colonial era, as most African governments solved urban housing shortages by taking up ‘systematic round-ups and expulsions of urban dwellers living in self-constructed shacks’ (Freund, 2007: 148). Governments often targeted the urban unemployed or underemployed, who were then redeployed to rural areas in a direct continuation of colonial policy (albeit with ‘greater zeal’) (Brennan, 2006: 403). The result has been that, as demonstrated in Table 4, Africa remains the continent with the lowest level of urbanization, and is projected to become a majority urban region only after 2030.

[Insert Table 4 here]

Yet, despite state efforts to suppress urbanization, the dual effects of colonial pacification and economic development programs have encouraged urbanization indirectly via high population growth, local land scarcities and out-migration from tribal areas. While some of these migrants moved to rural areas, many of them moved to cities, where they formed new ethnic identities. Indeed, as noted by Freund (2007: 90-91) and Wallerstein (1960: 133), Africans have often assimilated into larger ethnic groups in order to find security and prestige in the difficult urban environment. The Ibo of Nigeria, the Jola of Senegal, Duala of Cameroon and Bangala of the DRC can thus be seen as classic examples of previously different ethnic groups amalgamating into larger ethnic identities as urban migrants found commonalities among each other and transferred these new identities back to their rural brethren as well (Eckert, 1999; Nugent, 2008; Young, 1976).

### *3.2. Empirical Analysis*

In order to see if urbanization has had any effect in promoting ethnic homogenization within Africa I could regress ethnic diversity onto a variety of variables as in Table 2, but any correlation between the post-colonial variables and ethnic diversity would not, of course, tell us anything about how changes in urbanization affect changes in ethnic diversity. Moreover, inasmuch as different countries measure urbanization according to vastly different definitions any static correlations across countries are bound to be inaccurate.<sup>17</sup> I thus employ two different empirical strategies here.

First, I regress change over time in levels of ethnic diversity on changes in urbanization in a long-difference model, thereby allowing me to control for different definitions for urbanization and other variables like international migration. Theoretically one way to do this would be to use measures of ethnic diversity from different country census reports; yet this exercise is not feasible here as most African countries no longer ask questions about ethnicity on their censuses in large part because of the controversies surrounding ethnic headcounts (Morning, 2008). There is,

however, one extant source of changes in ethnic fractionalization over time, namely Roeder (2001)'s data on ethnic diversity in 1961 and 1985 as drawn from Soviet ethnographic sources.

As regards the independent variable I use data on changes in levels of urbanization from 1960 to 1985, along with three control variables. First, I control for changes in the number of immigrants between 1960 and 1985. Second, I control for initial levels of ELF and urbanization to account for the possibility that the relationship between change in ELF and urbanization is spurious due to the effects of initial levels on subsequent change for both variables

I present my results in Table 5. Column 1 reports the result of regressing change in ELF on the four basic variables listed above. In regression 2 I add changes in GDP/capita growth between 1961 and 1985; however, due to missing data we lose six observations from column 1. In column 3 I substitute change in agricultural productivity, both because it is a proxy for economic development with fewer missing observations and because previous scholars have suggested that it is correlated with urbanization (Yuki, 2007). Finally, in regression 4 I control for changes in fertility, mortality and sex ratio inasmuch as changes in the population composition should have an effect on a given society's ethnic demographics.

[Insert Table 5 here]

The results are clear: in all four regressions the coefficient for urbanization is both statistically significant and consistently located between -0.162 and -0.195, with the exception of column 2 where the use of a smaller sample makes the urbanization coefficient drop to -0.122 and lose some of its statistical significance. Moreover, none of the other variables are consistently statistically significant at the 5% level. The results therefore suggest that higher rates of urbanization are robustly correlated with greater ethnic homogenization within Africa.<sup>19</sup>

My second empirical strategy involves the use of survey data from the Afrobarometer. Indeed, the causal effect of urbanization on nationalism has already been established by Robinson

(2009) with third-round Afrobarometer survey data from 21,937 respondents across 16 African countries, whereby logistic regressions which include age, gender, education and wealth as controls show a strong correlation between urbanization and the likelihood of feeling more strongly attached to one's national identity than one's ethnic identity. More specifically, each respondent was asked the question,

Let us suppose that you had to choose between being a [Ghanaian/Kenyan/etc.] and being a [Respondent's Ethnic Group]. Which of the following best expresses your feelings?

- 1 = I feel only [Respondent's ethnic group]
- 2 = I feel more [Respondent's ethnic group] than [Ghanaian/Kenyan/etc.]
- 3 = I feel equally [Ghanaian/Kenyan/etc.] and [Respondent's ethnic group]
- 4 = I feel more [Ghanaian/Kenyan/etc.] than [Respondent's ethnic group]
- 5 = I feel only [Ghanaian/Kenyan/etc.]

Thus a higher number indicates an attachment to a higher-order identity. While not an ideal measure of ethnic homogenization, this question at least captures the degree to which respondents identify mostly or entirely with their national identity, which I assume would be correlated with ethnic homogenization.

This question was also asked in fourth-round Afrobarometer data from 25,705 respondents across 20 countries (listed in Appendix 3), allowing us to assess whether the association of urban residence and stronger attachment to a national identity is a consistent finding. To do so I use an urban dummy variable; inasmuch as the dummy takes a value of 1 for urban areas and 2 for rural areas I would expect a negative relationship between urbanization and the nationalism dependent variable. I also utilize a series of controls, including the respondent's age and age squared, gender, self-described level of poverty, education, level of political participation (as measured by whether he/she voted in the last election), number of languages spoken, religion, ethnic identity, region, country and native language. Adding data on the frequency by which the respondent obtains access to news from radio and television, ranging from never (coded as 0) to every day (4), helps us control for the effects of the media. Finally, I use robust standard errors clustered at the regional

level to account for similar access of neighboring respondents to education, newspapers and radios; clustering at the level of the country and ethnic group give similar results.

Before proceeding, however, I should examine the country-level data for any outliers, since Robinson (2009) identified Tanzania as having an unusually high level of respondents who identify with the nation in 3<sup>rd</sup>-round survey data. Plotting the percentage of respondents who only identified with the nation on the vertical axis and the average score for the nationalism variable on the horizontal axis, Figure 1 demonstrates that Tanzania is an outlier in 4<sup>th</sup>-round data as well. Tanzania's exceptionally high level of national identity has long been a source of debate for scholars of African politics; this debate is covered in more detail in Green (2011b) and need not detain us here. As a result of Tanzania's exceptionally high levels of national identity I run two sets of regressions, one including data from Tanzania and one without.

[Insert Figure 1 here]

The results, as seen in Table 6, indicate that urbanization is consistently correlated with an adherence to national over ethnic identities, and that this relationship is stronger when the Tanzanian data are removed. It also indicates that greater use of radios and newspapers is associated with higher identification with the nation. Robinson (2009)'s findings that being male and that age has an inverted-U relationship with identification with the nation are consistent with the results here, albeit with a weaker levels of significance for the age variables when Tanzania is excluded, and all other variables are insignificant.

[Insert Table 6 here]

#### 4. INTERPRETATION



My results thus suggest that latitude, slave exports, state size and urbanization play a strong causal role in African ethnic diversity, and that other hypothesized causes have little to no significant effect. Indeed, while at first the negative results for the other hypothesized variables are counterintuitive, a closer look at the evidence suggests otherwise. As regards the pre-colonial variables, previous scholars have found contradictory evidence: Nettle (1996), for instance, finds no relationship between diversity and elevation in West Africa, while many of the studies mentioned above only examine bivariate correlations and therefore fail to check if any variables lose significance and/or coefficient size in multivariate regressions.

As for the colonial period, there is a good amount of equivocal evidence about the impact of both arbitrary borders and indirect rule on ethnic diversity. For instance, despite their lack of detailed ethnographic knowledge European colonialists often tried to avoid splitting indigenous states between colonies when they could: the Angola/Zambia, Benin/Nigeria, Ghana/Burkina Faso, Niger/Nigeria, and Rwanda/Tanganyika borders, among others, were all drawn to ensure that pre-colonial African states were not partitioned. Indeed, the British Prime Minister Lord Salisbury pushed for the Ndian river to be demarcated as the border between Nigeria and Kamerun in 1887 as it marked 'a distinct ethnographic frontier' between Bantu and non-Bantu peoples (Partridge and Gillard, 1995: 251). After the partition of the German colonial empire in 1919 the then British Colonial Secretary was similarly concerned about the initial French/British border in the Cameroons as it 'cut across tribal and administrative divisions,' and was pleased at the addition of the British Cameroons and Togoland to the Empire as it brought 'completely within our borders native Tribes which have hitherto been partly within British territory and partly outside it' (Louis, 1966: 885-886). Finally, there is evidence as well of European powers allowing for the transfer of populations caught on the wrong side of the borders between Gambia and Senegal, the Belgian Congo and Uganda, and Kenya and Italian Somaliland in the early colonial period (Griffiths, 1986; Trouval, 1966).

There is also evidence that European ‘tribal’ policies did not always increase pre-existing ethnic diversity. In many places such as central Uganda colonial policies had little effect on ethnic identities due to the strength of pre-colonial attachments (Green, 2008). In other areas ethnic groups that were too small to be formed into individual chiefdoms or kingdoms were consolidated into larger groupings and given chiefs with ‘invented traditions’ (Ranger, 1983); in this way the Abaluhya of Kenya and Ndebele of Zimbabwe, among many others, were grouped together in a single district and eventually became single ethnic groups (Mamdani, 1996: 81-82; Ndegwa, 1997;). It was not only the colonial state which promoted ethnic homogeneity through cost-saving: missionaries often ‘reduced Africa’s innumerable dialects to fewer written languages’ and thereby promoted ethnic homogeneity in large part due to budgetary constraints (Iliffe, 2007: 239; Posner, 2003).

Finally, the results with urbanization correspond to ones already obtained by Green (2011a) for both a global sample of countries and mid-twentieth century Turkey. These results thus provide further evidence of a causal effect of urbanization on ethnic homogenization.

## 5. CONCLUSIONS

In this paper I have shown that Africa’s unusually high ethnic diversity has four sources, namely its tropical location, the early modern slave trade, the creation of large colonial states and low levels of urbanization. I demonstrated that the inclusion of a variable capturing the early modern slave trade in regressions explaining global ethnic diversity makes an Africa dummy variable statistically insignificant. I also showed that these four variables are robust to various specifications and that a number of other variables such as indirect rule, artificial borders and economic growth are not robustly correlated with ethnic diversity. Finally, I also used Afrobarometer data to show that urbanization is correlated with greater attachment to the nation.

As such this paper has two broad contributions to the literature on Africa and ethnic diversity. First, this paper suggests that Africa's ethnic diversity does not have mysterious origins. Africa's unusually high levels of ethnic diversity can be explained by the early modern slave trade, and its large standard deviation in ethnic diversity is the result of internal variation in latitudinal spread, state size and urbanization. The relationship between diversity and latitude confirms the results found by Ahlerup and Olsson (2011) and Michalopoulos (2012) on a global scale, and the significance of state size confirms the role of colonialists in creating ethnically diverse states in Africa, albeit not in the divide and rule strategy suggested above. Indeed, the results suggest that the unusual lack of secessionist movements in post-colonial Africa should thus not only be seen as partially responsible for the continent's lack of state formation and consolidation but also for its ethnic diversity (Herbst, 2000; Englebert, 2009). Finally, the ethnic diversity that pervaded rural society in other regions in the pre-modern era has persisted in Africa in part because of the continent's low levels of urbanization.

Second, it suggests that African ethnic diversity is not primordial or fixed and thus cannot continue to be used as an exogenous variable in analyses of African political economy. Further research in this area should continue to interrogate the primordialist assumptions that still plague the study of African ethnicity. Moreover, the results here suggest that measures to alleviate the negative effects of ethnic diversity could concentrate less on such suggestions as the creation of locally homogenous political units (Laitin, 2007) and more on promoting urbanization as a means of ethnic homogenization.

Additional research on this topic could focus on a number of important items. First, inasmuch as pre-colonial slave exports are significantly correlated with modern ethnic diversity, it is important to develop more accurate quantitative measures of pre-colonial slavery which do not use ethnicity for identification purposes and which better account for the internal slave trade within Africa and Eurasia. Secondly, more substantial time-series data on ethnic change over time would

improve our understanding of how political and economic phenomena affect ethnicity in Africa.

Finally, similar analyses for other parts of the world would add to our knowledge on this subject.

## Bibliography

- Ahlerup P and Olsson O (2011) *The Roots of Ethnic Diversity*. Department of Economics, Göteborg University.
- Alesina A, Devleeshauer A, Easterly W, Kurlat S and Wacziarg R (2003) Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.
- Alesina A, Easterly W and Matuszeski J (2011) Artificial States. *Journal of the European Economic Association* 9(2): 246-277.
- Alesina A and Spolaore E (2003) *The Size of Nations*. Cambridge, MA: MIT Press.
- Austin G (2008) The 'Reversal of Fortune' Thesis and the Compression of History: Perspectives from African and Comparative Economic History. *Journal of International Development* 20(8): 996-1027.
- Barr A and Oduro A (2002) Ethnic Fractionalization in an African Labour Market. *Journal of Development Economics* 68(2): 355-379.
- Barth F (ed). (1969) *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. London: Allen and Unwin.
- Berman B (1998) Ethnicity, Patronage and the African State: The Politics of Uncivil Nationalism. *African Affairs* 97(388): 305-341.
- Blanton R, Mason T and Athow, B (2001) Colonial Style and Post-Colonial Ethnic Conflict in Africa. *Journal of Peace Research* 38(4): 473-492.
- Brennan J (2006) Blood Enemies: Exploitation and Urban Citizenship in the Nationalist Political Thought of Tanzania 1958-1975. *Journal of African History* 47(3): 389-413.
- Brubaker R (2004) *Ethnicity Without Groups*. Cambridge, MA: Harvard University Press.
- Buhaug H (2006) Relative Capability and Rebel Objective in Civil War. *Journal of Peace Research* 43(6): 691-708.
- Carter M (2006) Slavery and Unfree Labour in the Indian Ocean. *History Compass* 4(5): 800-813.

- Cashdan E (2001) Ethnic Diversity and Its Environmental Determinants: Effects of Climate, Pathogens and Habitat Diversity. *American Anthropologist* 103(4): 968-991.
- Chandra K (2006) What is Ethnic Identity and Does it Matter? *Annual Review of Political Science* 9: 397-424.
- Cohen A (1969) *Custom and Politics in Urban Africa: A Study of Hausa Migrants in Yoruba Towns*. London: Routledge.
- Collard I and Foley R (2002) Latitudinal Patterns and Environmental Determinants of Recent Human Cultural Diversity: Do Humans Follow Biogeographic Rules? *Evolutionary Ecology Research* 4(3): 371-383.
- Easterly W (2001) Can Institutions Resolve Ethnic Conflict? *Economic Development and Cultural Change* 49(4): 687-706.
- Easterly W Levine R (1997) Africa's Growth Tragedy: Policies and Ethnic Divisions. *Quarterly Journal of Economics* 112(4): 1203-1250.
- Eckert A (1999) African Rural Entrepreneurs and Labor in the Cameroon Littoral. *Journal of African History* 40(1): 109-126.
- Englebert P (2000) Solving the Mystery of the AFRICA Dummy. *World Development* 28(10): 1821-1835.
- Englebert P (2009) *Africa: Unity, Sovereignty and Sorrow*. Boulder, CO: Lynne Rienner.
- Englebert P, Tarango S and Carter M (2002) Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries. *Comparative Political Studies* 35(10): 1093-1118.
- Fearon J (2003) Ethnic and Cultural Diversity by Country. *Journal of Economic Growth* 8(2): 195-222.
- Freund B (2007) *The African City: A History*. Cambridge: Cambridge University Press.
- Gellner E (2006 [1983]) *Nations and Nationalism*. Oxford: Blackwell.
- Government of Tanzania (1969) *1967 Population Census*. Dar es Salaam: Government Printer.

- Green E (2008) Understanding the Limits to Ethnic Change: Evidence from Uganda's Lost Counties. *Perspectives on Politics* 6(3): 473-485.
- Green E (2011a) Endogenous Ethnicity. Department of International Development, London School of Economics.
- Green E (2011b) The Political Economy of Nation Formation in Modern Tanzania: Explaining Stability in the Face of Diversity. *Commonwealth and Comparative Politics* 49(2): 223-244.
- Green E (2012) On the Size and Shape of African States. *International Studies Quarterly* 56(forthcoming).
- Grier R (1999) Colonial Legacies and Economic Growth. *Public Choice* 98: 317-335.
- Griffiths I (1986) The Scramble for Africa: Inherited Political Boundaries. *Geographical Journal* 152(2): 204-216.
- Herbst J (2000) *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton, N.J.: Princeton University Press.
- Heston A, Summers R and Aten B (2009) Penn World Table, Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.
- Holsoe S and Lauer J (1976) Who are the Kran/Guere and the Gio/Yacouba? Ethnic Identification along the Liberia-Ivory Coast Border. *African Studies Review* 19(1): 139-149.
- Hopkins B (2008) Race, Sex and Slavery: 'Forced Labour' in Central Asia and Afghanistan in the Early 19th Century. *Modern Asian Studies* 42(4): 629-671.
- Iliffe J (2007) *Africans: The History of a Continent* (2 ed.). Cambridge: Cambridge University Press.
- Isaacman A (1972) The Origin, Formation and Early History of the Chikunda of South Central Africa. *Journal of African History* 13(3): 443-461.
- Kimenyi M (2006) Ethnicity, Governance and the Provision of Public Goods. *Journal of African Economies* 15(AERC Supplement 1): 62-99.

- Kizilov M (2007) Slave Trade in the Early Modern Crimea From the Perspective of Christian, Muslim, and Jewish Sources. *Journal of Early Modern History* 11(1-2): 1-31.
- Kortepeter C (1966) Ottoman Imperial Policy and the Economy of the Black Sea Region in the Sixteenth Century. *Journal of the American Oriental Society* 86(2): 86-113.
- Laitin D (1994) The Tower of Babel as a Coordination Game: Political Linguistics in Ghana. *American Political Science Review* 88(3): 622-634.
- Laitin D (2007) *Nations, States and Violence*. Oxford: Oxford University Press.
- Lange M (2004) British Colonial Legacies and Political Development. *World Development* 32(6): 905-922.
- Levi S (2002) Hindus Beyond the Hindu Kush: Indians in the Central Asian Slave Trade. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 12(3): 277-288.
- Louis W (1966) Great Britain and the African Peace Settlement of 1919. *American Historical Review* 71(3): 875-892.
- Mamdani M (1996) *Citizen and Subject: Contemporary Africa and the Legacy of Late Colonialism*. Princeton, N.J: Princeton University Press.
- Michalopoulos S (2012) The Origins of Ethnolinguistic Diversity. *American Economic Review* forthcoming.
- Miguel E (2004) Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania. *World Politics* 56(3): 327-362.
- Miguel E and Gugerty M (2005) Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics* 89(11-12): 2325-2368.
- Milanovic B (2003) Is Inequality in Africa Really Different? Working Paper #3169, Development Research Group, World Bank.
- Mitchell T, Carter T, Jones P, Hulme Mand New M (2004) A Comprehensive Set of High-Resolution Grids of Monthly Climate for Europe and the Globe. Working Paper #55, Tyndall Centre for Climate Change Research, Norwich.

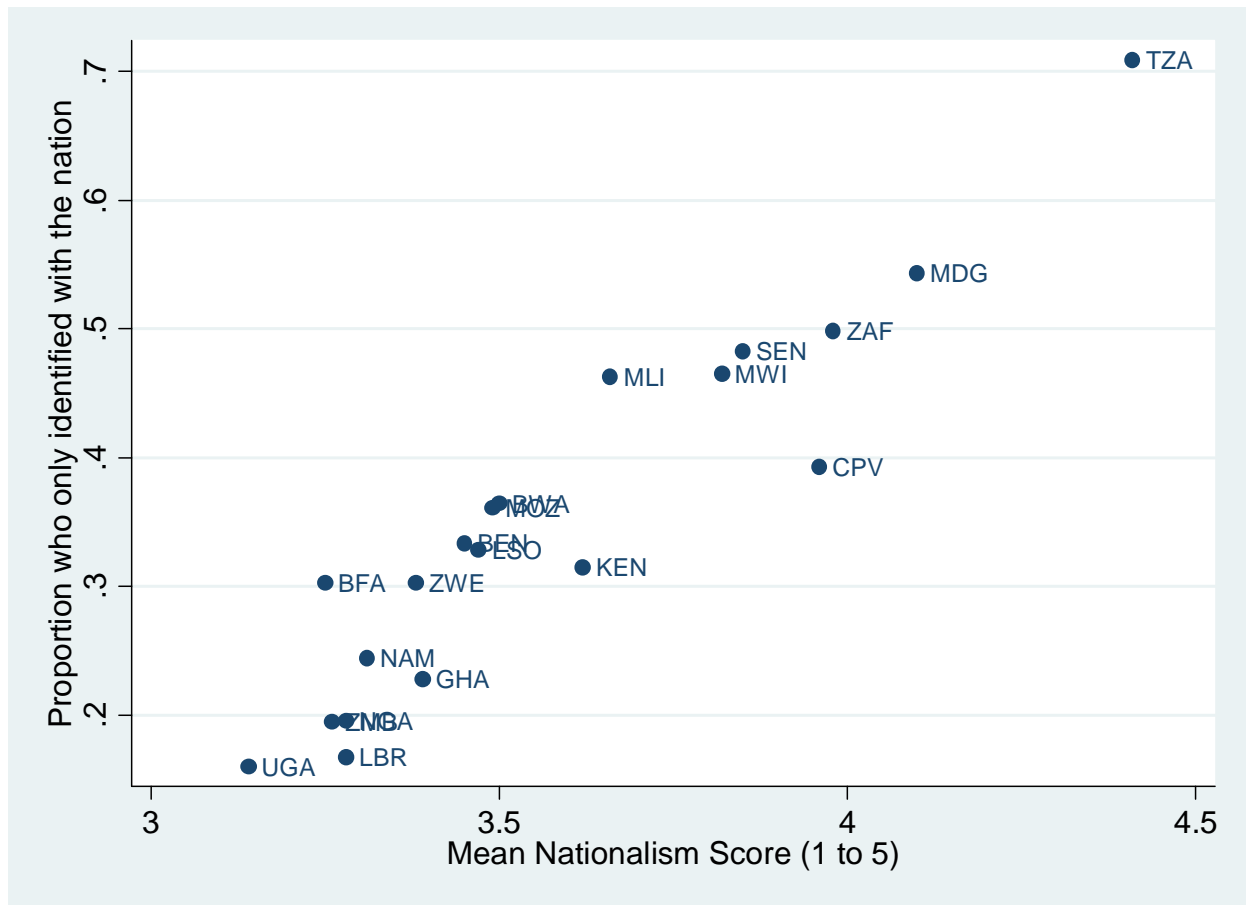


- Morning A (2008) Ethnic Classification in Global Perspective: A Cross-National Survey of the 2000 Census Round. *Population Research and Policy Review* 27(2): 239-272.
- Ndegwa S (1997) Citizenship and Ethnicity: An Examination of Two Transition Moments in Kenyan Politics. *American Political Science Review* 91(3): 599-616.
- Nettle D (1996) Language Diversity in West Africa. *Journal of Anthropological Archaeology* 15(4): 403-438.
- Nichols J (1992) *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Nugent P (2008) Putting the History Back in Ethnicity: Enslavement, Religion and Cultural Brokerage in the Construction of Mandika/Jola and Ewe/Agotime Identities in West Africa, c. 1650-1930. *Comparative Studies in Society and History* 50(4): 920-948.
- Nunn N (2008) The Long Term Effects of Africa's Slave Trade. *Quarterly Journal of Economics* 123(1): 139-176.
- Partridge M and Gillard D (eds) (1995) *British Documents on Foreign Affairs: Reports and Papers from the Foreign Office Confidential Print. Part I: From the Mid-Nineteenth Century to the First World War. Series G: Africa, 1848-1914*. (Vol. 19: West Africa: Diplomacy of Imperialism, 1868-1895). Frederick, MD: University Publications of America.
- Posner D (2003) The Colonial Origins of Ethnic Cleavages: The Case of Linguistic Divisions in Zambia. *Comparative Politics* 35(2): 127-146.
- Posner D (2004a) Measuring Ethnic Fractionalization in Africa. *American Journal of Political Science* 48(4): 849-863.
- Posner D. N. (2004b) The Political Salience of Cultural Difference: Why Chewas and Tumbukas Are Allies in Zambia and Adversaries in Malawi. *American Political Science Review* 98(4): 529-546.
- Ranger T (1983) The Invention of Tradition in Colonial Africa. In: Hobsbawm E and Ranger T (eds) *The Invention of Tradition*, pp 211-262. Cambridge: Cambridge University Press.

- Robinson A (2009) National versus Ethnic Identity in Africa: State, Group and Individual Level Correlates of National Identification. Afrobarometer Working Paper #112, Michigan State University.
- Roeder P (2001) Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985. <http://weber.ucsd.edu/~proeder/elf.htm>.
- Seijas T (2008) The Portuguese Slave Trade to Spanish Manila: 1580-1640. *Itinerario* 32(1): 19-38.
- Solway J (2004) Reaching the Limits of Universal Citizenship: 'Minority' Struggles in Botswana. In: Berman B, Eyoh D and Kymlicka W (eds) *Ethnicity and Democracy in Africa*, pp 128-147. Oxford: James Currey.
- Sutherland W (2003) Parallel Extinction Risk and Global Distribution of Languages and Species. *Nature* 423(6937): 276-279.
- Trouval S (1966) Treaties, Borders and the Partition of Africa. *Journal of African History* 7(2): 279-293.
- United Nations (2008) World Urbanization Prospects: The 2007 Revision. Department of Economic and Social Affairs/Population Division.
- United Nations (2010) World Population Policies 2009. Department of Economic and Social Affairs, United Nations Secretariat.
- Vink M (2003) "The World's Oldest Trade": Dutch Slavery and Slave Trade in the Indian Ocean in the Seventeenth Century. *Journal of World History* 24(2): 131-177.
- Wallerstein I (1960) Ethnicity and National Integration in West Africa. *Cahiers d'Études Africaines* 1(3): 129-139.
- Warren J (2007) *The Sulu Zone, 1768-1898: The Dynamics of External Trade, Slavery, and Ethnicity in the Transformation of a Southeast Asian Maritime State* (2 ed.). Singapore: NUS Press.
- Weber E (1976) *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Stanford, CA: Stanford University Press.

- Webersik C (2004) Differences that Matter: The Struggle of the Marginalized in Somalia. *Africa: Journal of the International African Institute* 74(4): 516-533.
- Young C (1976) *The Politics of Cultural Pluralism*. Madison, WI: University of Wisconsin Press.
- Young C (1994) *The African Colonial State in Comparative Perspective*. New Haven, CT: Yale University Press.
- Yuki K (2007) Urbanization, Informal Sector and Development. *Journal of Development Economics* 84(1): 76-103.

Figure 1: Afrobarometer Results, Country Level



**Table 1: Ethnic Diversity Across Continents**  
(Source: (Alesina et al., 2003))

	Obs.	Mean	Stand Dev.	Min	Max
Sub-Saharan Africa	46	0.662	0.226	0	0.930
Asia	42	0.425	0.224	0.002	0.769
Americas	35	0.420	0.215	0.095	0.864
Europe	47	0.311	0.214	0.041	0.809

**Table 2: Historical Sources of African Ethnic Diversity**

Dependent Variable	Alesina	Alesina	Alesina	Fearon	Fearon	Fearon
	(1)	(2)	(3)	(4)	(5)	(6)
Latitude	-0.045** (0.020)	-0.041* (0.020)	-0.050** (0.020)	-0.053** (0.020)	-0.047** (0.020)	-0.058*** (0.021)
Ln(Exports/Area)	0.029*** (0.010)	0.028** (0.013)	0.027*** (0.010)	0.025** (0.010)	0.028*** (0.010)	0.024** (0.010)
Km2	0.044** (0.018)	0.038** (0.016)	0.047* (0.025)	0.056*** (0.018)	0.043*** (0.014)	0.051** (0.023)
British Colony	0.040 (0.075)	0.052 (0.071)	0.037 (0.076)	0.058 (0.077)	0.079 (0.077)	0.054 (0.073)
French Colony	0.038 (0.094)	0.034 (0.079)	0.037 (0.076)	0.014 (0.081)	0.032 (0.080)	-0.002 (0.084)
Portuguese Colony	0.016 (0.093)	0.040 (0.092)	0.059 (0.097)	0.014 (0.079)	0.045 (0.086)	0.010 (0.085)
Elevation	-0.028 (0.035)		-0.029 (0.042)	-0.056 (0.018)		-0.061 (0.038)
Temperature		0.010 (0.011)			0.007 (0.010)	
Fractal			-0.324 (1.911)			-1.154 (1.918)
Constant	0.244 (0.010)	-0.090 (0.205)	0.243 (0.451)	0.289 (0.260)	-0.074 (0.255)	0.477 (0.378)
N	43	43	40	42	42	39
R <sup>2</sup>	0.563	0.570	0.578	0.604	0.589	0.619

\*  $p \leq 0.1$ , \*\*  $p \leq 0.05$ ; \*\*\*  $p \leq 0.01$ ; robust standard errors in parentheses.

**Table 3: African and Global Ethnic Diversity**

Dependent Variable	Alesina	Alesina	Fearon	Fearon
	(1)	(2)	(3)	(4)
Elevation	0.020 (0.023)	0.055*** (0.020)	0.011 (0.023)	0.041* (0.021)
Latitude	-0.065** (0.025)	-0.042** (0.021)	-0.067** (0.027)	-0.044* (0.024)
British Colony	0.043 (0.055)	0.058 (0.048)	0.037 (0.057)	0.059 (0.052)
French Colony	0.049 (0.057)	0.056 (0.050)	0.008 (0.054)	0.022 (0.050)
Km <sup>2</sup>	0.013 (0.014)	-0.011 (0.013)	0.018 (0.016)	-0.002 (0.015)
Origin Time	0.040 (1.098)	0.348 (1.061)	-0.235 (1.017)	0.023 (0.988)
SSAfrica Dummy	0.200** (0.101)	0.037 (0.100)	0.256** (0.100)	0.115 (0.099)
Ln(Exports/Area)		0.033*** (0.006)		0.028*** (0.006)
Constant	0.300 (0.205)	0.296* (0.160)	0.320** (0.226)	0.490** (0.206)
N	127	127	121	121
R <sup>2</sup>	0.410	0.527	0.428	0.509

\*  $p \leq 0.1$ , \*\*  $p \leq 0.05$ ; \*\*\*  $p \leq 0.01$ ; robust standard errors in parentheses. The analysis excludes the Americas and Oceania inasmuch as the slavery variable is zero for all countries in both regions.

**Table 4: Percent Urban by Area**  
(Source: United Nations (2008))

<b>Region</b>	<b>1950</b>	<b>1975</b>	<b>2000</b>	<b>2030 (projected)</b>
Sub-Saharan Africa	11.1	21.7	32.8	48.2
Asia	16.8	24.0	37.1	54.1
Latin America and Caribbean	42.0	61.2	75.4	84.3
Europe	50.5	65.6	71.7	78.3
Oceania	62.0	71.5	70.5	73.8
North America	63.9	73.8	79.1	86.7



**Table 5: Urbanization and Ethnic Change in Africa**  
(Dependent Variable:  $\Delta$  in ELF 1961-1985)

	(1)	(2)	(3)	(4)
$\Delta$ Urbanization 1960-1985	-0.180** (0.081)	-0.122* (0.070)	-0.195** (0.087)	-0.162** (0.076)
ELF, 1961	0.053** (0.025)	0.012 (0.021)	0.057 (0.025)	0.031 (0.033)
Percentage Urban, 1960	0.043 (0.052)	0.096 (0.057)	0.041 (0.047)	0.087 (0.059)
$\Delta$ Immigrants as a Percentage of the Population, 1960-1985	0.125 (0.148)	0.255 (0.169)	0.203 (0.163)	0.006 (0.199)
$\Delta$ GDP 1961-1985		-0.021 (0.019)		
$\Delta$ Agricultural Productivity 1961-1985			-0.0003* (0.0001)	
$\Delta$ Total Fertility Rate, 1960-1985				0.010 (0.008)
$\Delta$ Crude Mortality Rate, 1960-1985			-0.010	(0.262)
$\Delta$ Sex Ratio, 1960-1985				-0.003 (0.003)
Constant	-0.027 (0.019)	-0.004 (0.018)	-0.035* (0.020)	-0.016 (0.023)
N	44	38	42	44
R <sup>2</sup>	0.200	0.240	0.271	0.251

\*  $p \leq 0.1$ , \*\*  $p \leq 0.05$ , \*\*\*  $p \leq 0.01$ ; robust standard errors in parentheses.

**Table 6: Urbanization and National Identity in Africa  
(Dependent Variable: Identification with Nation vs. Ethnic Group)**

		Excluding Tanzania
	(1)	(2)
Urban Dummy	-0.098* (0.056)	-0.144*** (0.053)
Gender	-0.055** (0.026)	-0.067** (0.027)
Age	0.003** (0.001)	0.003* (0.002)
Age Squared	-0.000** (0.000)	-0.000* (0.000)
Household Head Dummy	0.001 (0.028)	-0.006 (0.029)
# of Adults in Household	0.017 (0.019)	0.008 (0.020)
# of Languages Spoken	-0.029 (0.030)	-0.031 (0.030)
Poverty (Self-Described)	-0.020 (0.018)	-0.001 (0.019)
News from Radio	0.051*** (0.017)	0.048*** (0.017)
News from Newspaper	0.051*** (0.016)	0.051*** (0.017)
Voted in Last Election	-0.003 (0.003)	-0.002 (0.002)
Education	0.002 (0.006)	0.004 (0.006)
Country/Region/Religion/ Tribe Fixed Effects	yes	yes
N	25,579	24,443
Clusters	246	220
Countries	20	19

\*  $p \leq 0.1$ , \*\*  $p \leq 0.05$ , \*\*\*  $p \leq 0.01$ ; robust standard errors clustered at the regional level in parentheses. Source: Afrobarometer.

## Appendix 1: Data Sources for Tables 2, 3 and 5

### Dependent Variables

*Alesina*: ELF as measured by Alesina et al. (2003).

*Fearon*: ELF as measured by Fearon (2003).

*Roeder*: Change in ELF, 1961-1985 as measured by Roeder (2001).

### Independent Variables

$\Delta$  *Agricultural Productivity 1961-1985*: Change in net agricultural production per capita per country, 1961-1985. Source: FAO.

*Area*: Natural log of a country's area in square kilometers. Source: Center of International Development, Harvard University.

*British Colony*: Dummy equals 1 if a state was ever colonized by the UK and 0 otherwise.

$\Delta$  *Crude Mortality Rate, 1960-1985*: Change in crude death rates per country, 1960-1985. Source: United Nations (2010).

*Elevation*: Natural log of a country's mean latitude. Source: Center of International Development, Harvard University.

*Fractal*: Natural log of a measure computing the degree by which a given country's non-coastal borders are straight lines, with the measure decreasing as the border approaches a straight line. Source: Alesina et al. (2011).

*French Colony*: Dummy equals 1 if a state was ever colonized by France and 0 otherwise.

$\Delta$  *GDP 1961-1985*: Change in GDP/capita growth per country, 1961-1985. Source: Heston et al. (2009).

$\Delta$  *Immigrants as a Percentage of the Population, 1960-1985*: Change in immigrants as a percentage of the total population per country, 1960-1985. Source: United Nations (2010).

*Latitude*: Natural log of a country's absolute central latitude. Source: Center of International Development, Harvard University.

*Origin Time*: Length of time a given country has been populated by humans, measured in millions of years. Source: Ahlerup and Olsson (2011).

*Polity IV*: Net change over time in a country's Polity IV rating. Source: Polity IV.

*Portuguese Colony*: Dummy equals 1 if a state was initially colonized by Portugal and 0 otherwise.

$\Delta$  *Sex Ratio, 1960-1985*: Change in sex ratios per country, 1960-1985. Source: United Nations (2010).

*Slave Exports*: Natural log of the total number of slaves exported from a given state between 1400 and 1900, divided by the area of the state and multiplied by 1000. Source: Nunn (2008).

*Temperature*: The natural log of the average mean temperature per country, 1961-1990. Source: Mitchell et al. (2004).

$\Delta$  *Total Fertility Change*: Change in fertility rates per country, 1960-1985. Source: United Nations (2010).

*Percentage Urban, 1960*: Percentage of a country's population living in urban areas in 1960. Source: United Nations (2008).

$\Delta$  *Urbanization 1960-1985*: Change in the percentage of residents living in cities, 1960-1985. Source: United Nations (2008).

## Appendix 2: Estimates of Slave Exports Outside Africa

Region	Slave Exports	Ln(Exports/Area)	Source
Afghanistan	150,000	5.445	Levi (2002: 284)
Iran	350,000	5.358	Hopkins (2008: 645)
South-East Asia	300,000		Seijas (2008: 23) Warren (2007: xxxv)
Indonesia		4.820	
Malaysia		4.820	
Philippines		4.820	
Eastern Europe/Russia	2,000,000		Kizilov (2007: 7)
Belarus		4.684	
Estonia		4.684	
Georgia		4.684	
Latvia		4.684	
Lithuania		4.684	
Moldova		4.684	
Poland		4.684	
Romania		4.684	
Russia		4.684	
Ukraine		4.684	
India and Pakistan	330,000	4.392	
Central Asia	300,000		Levi (2002: 283-284)
Indian Ocean	30,000		Carter (2006: 805)
Bangladesh and Myanmar	5726	1.938	Vink (2003: 141)

---

<sup>1</sup> Here and henceforth we refer to Sub-Saharan Africa merely as Africa.

<sup>2</sup> The next highest region is Asia.

<sup>3</sup> The ELF index is computed by summing the squares of all ethnic groups larger than 1% of the population and subtracting them from 1. While these two datasets actually compute ethnic diversity rather than ethno-linguistic diversity we prefer to use the ELF acronym as it is commonly used in the literature.

<sup>4</sup> See for instance Easterly and Levine (1997) #136 and Miguel and Gugerty (2005), among many others.

<sup>7</sup> However, Green (2012) shows that, far from arbitrary, the artificial shape of many colonial borders were instead a result of low pre-colonial population densities.

<sup>9</sup> More specifically, countries such as Tanzania and Togo do not have their ethnic identity labels listed in the dataset, a problem which does not occur with Fearon (2003)'s data. As a possible result of these coding problems, Alesina et al. (2003) list Tanzania's ethnic diversity score as only 0.74, far below both Fearon (2003)'s score of 0.95 and the 1967 Tanzanian census's score of 0.96 (Government of Tanzania, 1969). (The 1967 census was the last to measure ethnicity in Tanzania.)

<sup>10</sup> For all data sources see Appendix 1.

<sup>11</sup> Interestingly, regressing both ELF measures on indirect rule results in a negative relationship (which is not statistically significant using Fearon (2003)'s data but significant at  $p = 0.081$  with Alesina et al. (2003)'s data), suggesting that, if anything, indirect rule promoted ethnic homogenization rather than fractionalization.

<sup>13</sup> The results below are robust to a variety of higher and lower estimates for Afghanistan, India and Pakistan.

<sup>14</sup> This exercise is thus similar to Englebert (2000)'s efforts in eliminating the statistical significance of the Africa dummy in a series of growth regressions. Interestingly Englebert (2000)'s newly created "legit" variable – which is correlated with the Africa dummy at  $r = -0.53$  – does not eliminate the effect of the Africa dummy when regressing ethnic diversity instead of growth.

<sup>17</sup> Current thresholds for the classification of urban areas in Africa range from agglomerations of 1500 people in Equatorial Guinea to 10,000 in Senegal. Changes in definitions of urbanization do take place over time within countries but these changes are still dwarfed by the cross-country differences.

<sup>19</sup> I also run some Granger causality tests which fail to show any signs of reverse causality (i.e., changes in ELF leading to changes in urbanization); results are available from author.