National versus Ethnic Identity in Africa:

State, Group, and Individual Level Correlates of National Identification

Amanda Lea Robinson[*]

prepared for

Working Group in African Political Economy
University of California, Berkeley
December 11, 2009

Abstract:

While sub-Saharan African states are not generally considered to be true nation-states, there is still considerable variation across countries in the level of nationalism expressed by their citizens. This paper explores the relative importance of national and ethnic identities in sixteen sub-Saharan African countries, using individual-level survey data, and tries to determine how much of that variation is explained by existing theories of nationalism and ethnic politics. Largely descriptive, the paper presents the correlates of national identification at the state, ethnic group, and individual level, offering a first step towards understanding the micro-level processes of nation-building. Four broad conclusions are offered. First, "classic" modernization theories of nationalism, based primarily on the rise of European nations, explain a considerable amount of the variation observed in Africa. Second, the arguments for why African states are less nationalist than other regions of the world, such as their high levels of ethnic diversity and artificial borders, do less well at explaining variation within-Africa. In particular, and contrary to expectations, the degree of ethnic partition is positively related to national identification. Third, different colonial experiences, such as the particular colonial power and having fought an anti-colonial war, influence present day levels of nationalism. Finally, Tanzania is an outlier at the state-level, suggesting that factors other than those considered here were important in the relatively successful nation-building process there.

The nation-state is the primary mode of political organization in the international system today. Based on the success of the nation-state in Europe, many thought it natural for post-colonial African leaders to attempt to culturally homogenize and build nations from the states that they inherited (Rothchild, 1983). However, with the possible exception of Tanzania, it is generally noted that post-colonial African states are not nation-states. Why not? A number of theories have been given for this failure, but, to my knowledge, this is the first cross-sectional empirical investigation of the plethora of conditions purported to thwart African nationalism. To the extent that some African states have partially succeeded in nationalizing their citizens to a greater degree than others, how well do classic theories of nationalism account for that success? The goal of this project is to take a first step towards answering these questions. Four broad findings result: classic theories of nationalism explain a significant amount of variation across Africa, theories that focus on qualities that are more common to African states than states from other regions (e.g., high ethnic diversity) do less well at capturing variation within the continent, variation in colonial experience is related to present day levels of nationalism, and Tanzanian nationalism is not well explained by the theories tested here.

This project makes three important contributions. First, it uses theories developed in both the study of state nationalism and the literature on ethnic politics, to try to understand the choice between two potential group identities: ethnic and national. Second, it focuses the study of nationalism on the African continent, where the states are relatively young and nationalism is not ubiquitous. This yields interesting cross-country variation, and the possibility of studying the processes of nation-building as they are still occurring. Third, it uses individual level survey data from a representative sample of citizens across sixteen African countries. This is important, as nationalism – the widespread identification with the state – is fundamentally a mass phenomenon. In most previous studies, especially historical accounts of the rise of nationalism in Europe, non-elite perspectives have been very rare (Conner, 1990). Thus, analyses of these individual level survey data from "not yet nation-states" make a significant contribution towards understanding the causes of national identification.

The main goal of this project is to determine how much variation in African nationalism is explained by existing theories of nationalism. Building on theories of nationalism and ethnic identity developed across multiple disciplines, I identify potential correlates of national identification at three different levels. At the state level, I evaluate the impact of income, diversity, colonial power, and anti-colonial war experience on the percentage of a state's citizens

that identify with the state over their ethnic identity. At the ethnic group level, the group's size, degree of partition, and relative economic condition are considered. Finally, I look at the effect of individual level characteristics, including age, gender, employment status, education level, urban/rural location, and wealth.

The paper proceeds as follows. First, I provide motivation for the analysis, by suggesting several reasons, both positive and normative, that we should care about understanding variation in African nationalism. In the next section, theories from the study of nationalism and ethnic identification are used to generate nine hypotheses. Then, the measurement strategy and the sources of the data are laid out. Finally, the empirical results are reported at the state, ethnic group, and individual level, and are followed by a general discussion of the findings.

Motivation

The main goal of this project is to empirically establish correlates of nationalism in African states. One might reasonably ask why we care. There are two main reasons. First, it is useful to determine if there are *general* causes of national attachment, such that the same conditions apply in Africa as did in the rest of the world. For example, it is useful to know whether temporal trends in the rise of nationalism are due to changes in norms and conventions over time, or if such trends are due to specific conditions that happened to be more prevalent in some time periods than others. More specifically, from the observation that a rise in nationalism corresponded in time to the rise of industrialization in Europe, we cannot say anything about the causal impact of industrialization. If, however, we see that industrialization in another place and time also leads to greater nationalism, we can be more confident in a true causal effect.

Beyond this desire to understand nationalism as a social phenomenon, there are normative reasons that we should care about nationalism. In fact, it may be that nationalism itself explains variation in political and economic outcomes of interest, including cooperation, economic development, democracy, and civil conflict.

A long tradition in social psychology has studied the ways in which social identities form, and the impact they have on individual and group behavior. Most theories start with the assumption that individuals are constantly categorizing other individuals into one of two categories: in-group or out-group. Experiments have shown that such categorization, even based on artificial and temporary distinctions, can greatly affect individuals' perception and evaluation of others, and, importantly, cooperative behavior (Billig and Tajfel, 1973). This is important for

political forms of cooperation because social identity theorists also argue that the benefits enjoyed by in-group members can be extended to out-group members by re-categorization. If we think of building a national identity that supercedes sub-national identities (like ethnicity) as a stable form of identity re-categorization, we should observe greater cooperation in states with higher levels of nationalism. Transue (2007) has shown that by increasing the salience of the American identity experimentally, one can induce greater support for minority-favoring policies. For a real world example, Miguel (2004) shows that inter-ethnic cooperation is higher in Tanzania than in Kenya, and he argues that this is because a superceding national identity is more salient in Tanzania than in Kenya.

While higher levels of cooperation may produce a multitude of positive outcomes, economic development should be particularly impacted. Many studies have found a negative relationship between ethnic diversity and the under provision of public goods, or other 'bad' policies (Alesina and Ferrara, 2004; Easterly and Levine, 1997). The purported mechanism at work, however, is not based on the mere presence of ethnic difference, but on the salience of ethnic differences. If national identification can reduce the salience of ethnicity, the negative impact of diversity should be ameliorated. Additionally, identifying with the state may be an impetus to the "capitalist spirit", such that individuals work beyond the subsistence level towards the national good (Greenfield, 2001). For either or both of these reasons, nationalism should have a positive impact on economic growth.

Democracy should also operate more effectively when there is a strong sense of national identity. Neuberger (2000) provides four reasons why strong ethnic identification is detrimental to democracy. First, with no overarching identity, there is very little consensus on what is in the state's best interest. Second, ethnic identification can lead to ethnic bloc voting; but, for democracy to be successful, individuals need to vote on policy and not along ethnic (or other sub-national) lines (Horowitz, 1982).[1] Third, ethnic bloc voting means that there is no chance for election losers to win next time, deligitimizing peaceful pursuits of power. Finally, ethnic minorities have no incentive to support the success or persistence of democratization, suggesting that support for democracy will be low among ethnic minorities in states with low levels of nationalism. The impact of nationalism on democracy is vitally important in Africa, as a majority of African states have become at least nominal democracies in the last twenty years.

---

[1] However, see Chandra (2005) for a more nuanced argument, showing that the negative impact of ethnic politics is mediated by institutional design.

Finally, the salience of a national identity may be important for understanding the risk of intra-state conflict. In particular, it may be that the same conditions that fail to foster a national identity – thus, making ethnic identity relatively more salient – also increase the chance of civil war. For example, low levels of economic growth may lead to low levels of nationalism (and high levels of ethnic identification), as well as increase the likelihood of civil war. If so, it could resolve the paradox of why so many civil conflicts are fought along ethnic lines, yet measures of a state's ethnic make-up are poor predictors of conflict risk (Fearon and Laitin, 2003). In short, if the same factors lead to *both* a higher risk of insurgency *and* increased (decreased) ethnic salience (national salience), it may explain why insurgencies often take on ethnic organization.

In sum, there are potentially very important reasons to care about what factors are related to national identification in African states. A goal of future work is to determine the extent to which nationalism does, in fact, influence cooperation, development, democracy, and conflict.

Nationalism and Ethnic Identification in Africa

Nationalism, as a subject of study, is generally defined in one of two ways. In the first, nationalism is defined as the doctrine that state and cultural boundaries should be congruent. This often entails studying the causes of succession and civil war. A second definition of nationalism refers to the feelings of affection, loyalty, and identification with a politically defined group of people. Within the context of this project, I primarily use the term in the latter sense. More specifically, for the African cases, I will take nationalism to mean *identification with the state* (rather than a sub- or supra-state group). "Nationalist" attachments to an ethnic group, by contrast, will be called "ethnic identification"[2] and the measure of nationalism is relative to ethnic identification. Because I consider ethnicity to be the primary alternative to national identity,[3] theories of ethnic identification are considered in addition to those that focus on processes of state-level identification. In contrast to most of the literature on nationalism, recent work on ethnic attachment has generally operated at the micro-level, by exploring the instrumental value of ethnic identities and predicting when they should be salient. Though these studies are rarely connected to the literature on nationalism, they are highly relevant, because as

---

[2] Others have distinguished these two phenomenon as ethnonationalism and civic or patriotic nationalism (Hutchinson, 1994).

[3] Of all possible sub-national identity groups – class, occupation, etc. – ethnic identities pose the greatest alternative to national identification, as the components of the attachment to each group are very similar (Geertz, 1963). In fact, Conner (1990) argues that if you want to study nationalism, you have to study the ways in which the state is able to break nationalistic ties with sub-national ethnic groups.

national attachment increases ethnic attachment should decrease. Not because those identities are mutually exclusive – as constructivists have cogently argued (Laitin and Posner, 2001; Chandra, 2001), they are not – but because in multicultural states, the increased salience of a national identity must produce a concurrent reduction of the *relative* salience of ethnic difference. In fact, decreasing ethnic salience may be one of the most important mechanisms of turning a citizenry into a nation. Consequently, correlates of ethnic salience should be negative predictors of national attachment, as I have defined nationalism here.

Most scholars agree that nationalism is a modern phenomenon, preceded by small-scale local identification. While nationalist ideologies may have existed before the eighteenth or nineteenth centuries, mass identification with large political units did not. The processes proposed to account for this relatively recent rise of national attachment include war (Tilly, 1990), industrialization (Gellner, 1983), print capitalism (Anderson, 1983), and strategic state policy (Weber, 1976). All of these accounts are studied at the macro-historical level, and typically explain the rise of nationalism in eighteenth century Europe, despite motivation from post-colonial cases (Anderson, 1983; Gellner, 1983). A strong prediction that emerges from this literature is that "modernization" – increased education, industrialization, and urbanization – is positively correlated with national identification. Gellner (1983) argues that this relationship exists because new social structures arise in response to industrialization, changing the way individuals identify politically. One mechanism by which this operates is the need, induced by industrialization, for mass, state-sponsored education, which leads to country-wide homogenization and breaks ties to communal identities. Similarly, jobs in the capital or other urban centers lead to greater intra-state migration. Weber (1976) similarly argues that mass education, increased government employment, and military drafting produced Frenchmen out of peasants. This line of reasoning suggests that more modern states should have higher levels of nationalism, all else equal. In the African context, the degree of modernization can be proxied by a state's per capita income[4], generating the first hypothesis:

H1: Richer countries will have higher levels of nationalism than poorer countries.

---

[4] Within this dataset, a country's per capita income is significantly correlated with aggregate levels of urbanization (r=0.7), education (r=0.4), and formal employment (r=0.7).

Note that this hypothesis is based on the observation that modernization and nationalism arose concurrently in Europe, but does not make any claims or predictions about causation. As indicated in the discussion of the motivations for understanding nationalism, if such a relationship between income and nationalism exists, causation could be operating in either or both directions.

In addition to a macro-level correlation between modernization and the degree of national sentiment, the mechanism by which this relationship is purported to operate should hold at the individual level, as well. This suggests that, all else equal, nationalism should be strongest in those individuals who are employed in the formal sector, educated, and living in urban areas. Hyden (1983) argues, in line with this reasoning, that the low levels of nationalism in Africa are due, in part, to the peasant mode of production, which is dominant in many parts of Africa. If true, this suggests a micro-level modernization hypothesis:

> *H2: Formal sector employment, education, and urbanization should be positively related to an individual's propensity to identify with the state.*

There is an important qualification to the modernist argument. Modernization and industrialization will not necessarily lead to increased nationalism for *all* members of the state. If the benefits of modernization are uneven, and there are ethnic groups that are excluded from those benefits, we will see increased *ethnic identification* among members of that group in response to increased modernization (Horowitz, 1989, 1990; Gellner, 1983; Nairin, 1977). Horowitz, for example, argues that ethnic identification should be highest in ethnic groups that are "backward," having less education, being primarily agriculturalists, and holding a lower socio-economic status than other groups. Similarly, Gellner attributes nationalist failures (sub-national successions) to the uneven distribution of the gains of modernization. Thus, we should expect ethnicity to be most salient among disadvantaged groups, and, thus, that nationalism should be lowest among such groups.

> *H3: Members of poorer ethnic groups will be less likely to identify with the state than members of relatively richer ethnic groups.*

In contrast to the first two hypotheses, and, perhaps, as an extension of the negative impact of modernization stipulated in H3, Bates (1983) has argued that *ethnic* salience (rather than *national* salience) increased in post-independence Africa in response to modernization. According to Bates, this was because most of the benefits of modernization flowed through state governments, and ethnic identities provided viable bases for competition over such spoils of modernization. As a consequence, individuals involved in such competition should be more likely to identify ethnically than less modernized individuals. At a first approximation, then, we might expect 'modern' individuals to identify primarily as ethnic rather than national. However, as noted above, this contradicts the logic given in H1 and H2. It is important to note, though, that Bates expects individuals not integrated into the modern sector to identify with a *more local* identity than ethnicity, like the village.[5] Thus, it is unclear what predictions would follow from Bates when individuals are given the choice between an ethnic and a national identity. From the two literatures, we get the predictions that both national identification and ethnic identification should increase with modernization, but neither directly addresses the *relative* importance of these two identities. If increased ethnic salience trumps gains in nationalism, we should see the opposite relationship to the one proposed in H2, namely that modernization has a net negative impact on nationalism.

These first three hypotheses take national attachment to be a byproduct of modernization. Nationalism could be related to modernization in a number of ways, but one potential way may be that the benefits of modernization coming from the state increase individuals' feelings of identification with the state. For example, Snyder (1993) has argued that within the former Soviet states, individuals identify with the state to the extent that it is able to provide public services to its citizens. Azam (2001) makes a similar argument for the failure of state identification (via the strength of sub-national ethnic identification) in Africa due to states' under provision of such goods. If this mechanism is in operation, we should expect:

> *H4: The provision of basic public services/goods by the state will be positively correlated with nationalism.*

---

[5] Eifert, Miguel, and Posner (2008) find support for Bates' theory across African states using earlier Afrobarometer data. In particular, they show that individuals employed in the formal sector are more likely to choose their ethnic identity over a religious or occupational identity. Importantly, their measure of identification **explicitly excludes national identity**, asking respondents only to differentiate among potential sub-national identities.

Hypotheses H1-H4 have all assumed that African states will follow more or less the same trajectory towards nationhood as states in other regions of the world. However, there are several reasons to expect that the process may look different on the African continent. First, the processes of state and nation building have not been concurrent. In Europe, national identities arose within state borders that, while every bit as "arbitrary" as those found in the post-colonial world, were the product of endogenous processes within those borders – generating what Engelbert (2000) calls "vertical legitimacy." In contrast, the borders in Africa were mostly determined by colonial negotiation at the 1885 Berlin Conference (Herbst, 1998; Jackson and Rosberg, 1982), resulting in arbitrary borders with respect to pre-colonial cultural boundaries. This partition, in the terms borrowed from Geertz by Engelbert, Tarango, and Carter (2000), led to "suffocation" – the amalgamation of multiple ethnic groups within a single state – and "dismemberment" – the partition of single ethnic groups into two ore more states – of ethnic groups in Africa.

First, "suffocation" resulted in African states being among the most ethnically diverse in the world. Ethnic diversity, in turn, makes nation building more difficult than it would be with a culturally homogenous group. Horowitz (1985), for example, laments the difficulty of quelling sub-national disputes within "plural societies." Gellner (1983), too, argues that multiple "potential nations" within state boundaries makes engendering a national identity more difficult. Thus:

> H5: Culturally diverse states will have lower levels of nationalism than more homogeneous states.

Scholars of ethnic politics have shown how, in such multi-ethnic states, ethnic demographics influence how individuals identify. For example, Daniel Posner (2004, 2005) has forcefully argued that ethnic identity will be most salient when ethnic cleavages create groups big enough to produce a collective outcome (e.g., win an election). Thus, we would expect larger ethnic groups (relative to the rest of the country) to have higher levels of ethnic attachment, and thus, lower relative levels of nationalism. However, at the extreme, when an ethnic group makes up almost the entire population, such as in Lesotho or Botswana, the citizens may see the state and the ethnic group as synonymous. In such situations, supra-majority groups may espouse the national identification, producing a non-monotonic relationship between ethnic group size, and national attachment.

H6: The larger an ethnic group is as a share of a country's total population, the less likely members of that ethnic group should be to identify with the state, until they make up a large majority of the population.

In addition to creating culturally diverse states, the 1885 partitioning of Africa also divided many cultural groups into more than one colony, and subsequently, into multiple states – the "dismemberment" of ethnic groups. This partitioning of Africans is thought to have had profound effects on the likelihood of successful nation building (Bienen, 1983; Asiwaju, 1985). The partition of cultural groups, Asiwaju argues, creates incentives for the redrawing of state borders in order to rejoin ethnic kin. Such redrawing of boundaries poses a direct threat to a state's ability to engender a statewide common identity.[6] As a result, we should expect:

H7a: Members of partitioned ethnic groups will be less nationalist than members of groups residing entirely within state borders.

Of those that are partitioned, the severity of their split may be related to the strength of their national attachment. If the rejection of a national identity among members of a partitioned group is predicated upon the illegitimacy of a state that does not contain all members of the group (Engelbert, 2000), then the effect should be stronger in groups that are more severely partitioned. The severity of the split, from the point of view of members on one side of the border, can be measured as the share of the total group living on the other side.

H7b: If partitioned, the more co-ethnics residing outside state boundaries, the less likely members of that group will be to identify with the state.

---

[6] While most of the literature on ethnic group partition anticipates that it will have a negative impact on nation-building, there are dissenting opinions. For example, Miles (1994) finds that Hausa on both sides of the Niger-Nigeria border identify more with their respective states than as Hausa. He argues that the differences that states do make in the lives of their citizens are most evident to unitary cultural groups that still interact and can observe differences in education, language, currency, etc. In other words, being a single ethnic group may make any differences between people on either side of the border most attributable to differing national identities. Thus, Miles would predict an opposite relationship to that suggested in H7a.

Beyond the effects of partition, many scholars have argued that colonial administration often exacerbated the saliency of any ethnic differences present, through the processes of indirect rule and ethnic favoritism (Horowitz, 1982; Laitin, 1986; Hechter, 2000). The use and degree of such policies varied among the colonial powers, however, with the British being the most likely to use, and most extreme in their use of, ethnic categorization (Young, 1985; Mazrui, 1983). In addition, unlike the French, the British did not attempt to homogenize the colonial population, and, instead, had a policy of respecting local languages and customs (Mazrui, 1983). If the resulting increased salience of cultural differences persisted past the colonial period, we should expect:

> *H8: Former British colonies will have lower levels of nationalism than non-British former colonies.*

Finally, there is one consequence of colonialism that is thought to have led to higher levels of nationalism in African states – anti-colonial struggles. Anti-colonial campaigns often took on national rhetoric and unified individuals from different cultural groups against a common enemy (Mazrui, 1983; Neuberger, 2000). The most extreme form of anti-colonialism is represented by those colonies that fought anti-colonial wars. As a result:

> *H9: States that fought anti-colonial wars will have higher levels of nationalism than states that did not fight an anti-colonial war.*

The nine hypotheses above, which were derived from theories of nationalism, ethnic identification, and African politics, make predictions about the types of relationships we should expect to see. However, current tests of these hypotheses are not able to establish the existence or direction of causation, which must be left to future work. Tests of these hypotheses do provide a significant first step towards that end, though, by elucidating the strength of relationships between variables of interest and nationalism in Africa.

Data

To test these hypotheses, I begin with individual level survey data from the third round of Afrobarometer (Afrobarometer, 2009), which are available for sixteen Sub-Saharan African countries: Benin (2005), Botswana (2005), Ghana (2005), Kenya (2005), Lesotho (2005),

Madagascar (2005), Malawi (2005), Mali (2005), Mozambique (2005), Namibia (2005), Nigeria (2005), Senegal (2005), South Africa (2006), Tanzania (2005), Uganda (2005), and Zambia (2005).[7] Within each of the 16 countries, 1200 to 2400 individuals were interviewed, face-to-face, on a range of topics including democracy, governance, economics, social capital, conflict and crime, and political identity. Samples were designed to reflect the voting age population of each country, and were stratified in order to capture the opinions of all major segments of society. Afrobarometer provides within- and across-country weights in the cases where there was over- or under-sampling of a segment of the population. Random selection was used at every stage of the sampling, and the response rate was greater than sixty percent in all countries.

By construction, the sample of countries included by Afrobarometer is biased towards liberalizing regimes of Sub-Saharan Africa. To the extent that countries represented are different from Sub-Saharan African countries not included, the findings of this study will be limited in their generality. Table 1 lists a few descriptive statistics for countries in the dataset, as well as the same information for the remaining Sub-Saharan African countries.

	Afrobarometer Countries N=16	Other SS African Countries N=27
Democracy - Polity Score, 1999	3.1	-2.7 *
GPD/capita (thousands), 1999	1.22	1.30
Ethnic Fractionalization (ELF)	0.67	0.63
Share of Largest Ethnic Group	0.39	0.43
Low Level Civil War Occurrence	0.63	0.92 *
Full-Scale Civil War Occurrence	0.25	0.20
Population (millions)	19987	9552 *
Illiteracy Rate	0.34	0.45 *
Former British Colonies	0.63	0.26 *

Table 1. Characteristics of countries included in the Afrobarometer, and other Sub-Saharan African countries not included in the sample. Stars denote significant differences at the 0.1 level.

As expected, polity scores are significantly higher in the sample than in countries not included. Included countries have similar levels of income and diversity, but countries in the sample are more populous[8] and have lower illiteracy rates than countries not in the sample. British colonies are over represented in the sample. Finally, while there is no significant difference in the

---

[7] Survey data were also collected in Cape Verde and Zimbabwe, but Cape Verde is not included due to its low population size and lack of pre-colonial inhabitation, and Zimbabwe is not included because the question from which the dependent variable is derived was not asked there.

[8] The difference in population size is largely driven by Nigeria. When Nigeria is excluded, the average population size of countries in the Afrobarometer sample drops to 13901 million, and the difference is no longer statistically significant.

occurrence of full-scale civil wars since independence, countries in the sample are much less likely to have experienced small-scale conflicts.[9] Thus, to the degree that democracy, small-scale conflict, population, illiteracy, and British colonialism are related to nationalism, the results reported in this paper may not necessarily generalize to other Sub-Saharan African countries.

Measuring Nationalism

As a measure of nationalism, I use the following question from the third round of Afrobarometer:

Let us suppose that you had to choose between being a [Ghanaian/Kenyan/etc.] and being a [respondent's ethnic group]. Which of these two groups do you feel most strongly attached to?

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.]

The respondent's ethnic group was determined prior to this question when he or she was asked:

What is your tribe? You know, your ethnic or cultural group.[10]

Figure 1 shows the distribution of responses to the identity question. Over 70% of respondents answered that they were either 'only national', or 'equally national and ethnic,' with the latter category as the modal answer. Because these two categories seem to be the most normatively appealing ways for individuals to answer the question on identity, I dichotomize the measure in order to capture the strongest signal in the data. Thus, I operationalize 'nationalism' as

---

[9] Measured as 25 battle deaths per year, by the UCDP/PRIO Armed Conflict Dataset.

[10] The relevant ethnic groups for each country were predetermined by Afrobarometer, and over 95% of the respondents identified with one of the given groups, with only 4.5% specifying that they belonged to an "other" ethnic group. However, over 60% of those who chose "other" were from Tanzania: almost half of all Tanzanians chose "other" when asked which tribe they belonged to. For those individuals, I was unable to determine the characteristics of their identity group (such as size, average wealth, etc.), resulting in those respondents being dropped from the ethnic group level of analysis. Note that while the data do not provide the name of the ethnic group to which such respondents belong, these individuals were still asked to compare their strength of national attachment to that of ethnic attachment. Thus, the 46% of Tanzanian respondents that were coded as an "other" ethnic group still provide data on dependent variable of interest, and are included in the state and individual level analyses. Further, they have, in the aggregate, virtually identical levels of nationalism as those whose ethnic group was on the Afroabarometer list (89% vs. 87%).

identifying with the nation-state *more* than one's ethnic group (categories 4 and 5, "national over ethnic" and "national only"), making my measure of nationalism binary.
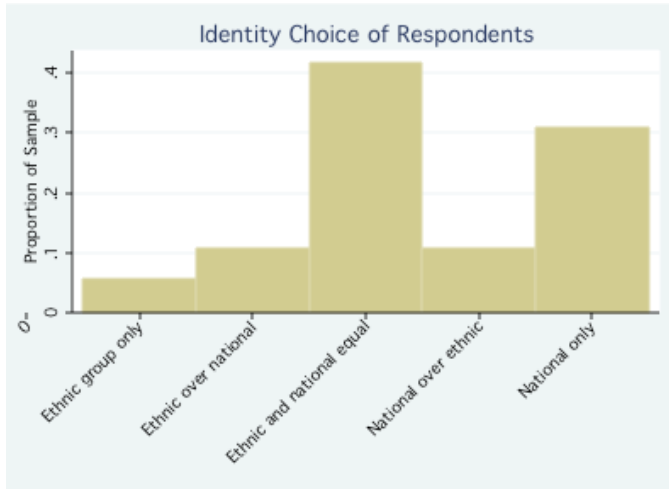

Figure 1. Histogram of Respondent Identification

This dichotomization preserves aggregate national identification across countries. Figure 2 shows the average level of nationalism (% nationals in each country using my dichotomous measure) by the average identity rank (1-5 scale, from above). Further, all results presented below are very similar when using the original ranked dependent variable, but this dichotomization makes the substantive interpretation of results more clear.[11]
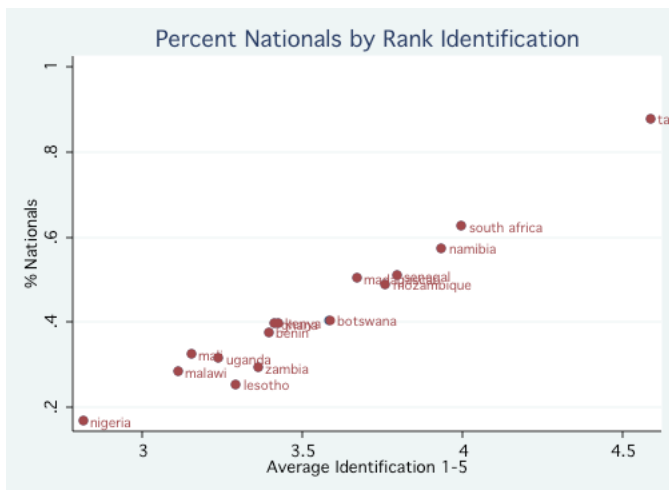

Figure 2. Average nationalism by average identity rank.

---

[11] Analyses using this alternative dependent variable are available from the author.

Across all sixteen countries, forty-two percent (42%) of the 22,765 respondents are coded as nationals. Figure 3 shows the percentage of respondents that are coded as nationals within each country. Tanzania is an outlier, with 88% of respondents identifying as nationals. The outlier status of Tanzania is not due to such a large percentage of respondents choosing "other" as their ethnic group – 87% of those whose ethnic group *was* on the list still identify as nationals.[7] Due to its outlier status, all analyses will be reported with and without Tanzania.
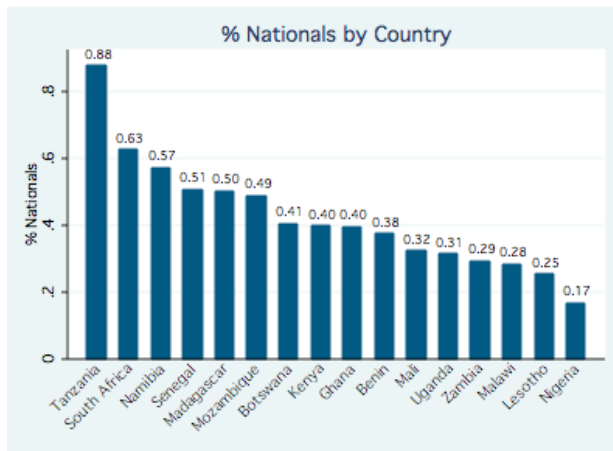

Figure 3. Nationalism by Country.

Is this a valid measure of nationalism? There are several reasons to be skeptical. First, we know that context is an important determinant of how individuals identify, and the Afrobarometer survey question does not account for or control for variation in situational context. Kuo and Margalit (2009) provide some evidence from the developed world that individuals do not necessarily answer questions about primary identification consistently over time, suggested either that identities are unstable, or that survey questions are not tapping into stable identities. For a specific example of how context matters for identification, Eifert, Miguel, and Posner (2008) find that the proximity to an election increases the salience of ethnic identities in Africa, *relative to other sub-national identities*. While they do not measure the impact of elections on ethnic versus national identification, it is possible that elections influence the relative importance of the two identities in a systematic way. A second concern is that individuals may not answer identity questions honestly for reasons of social desirability. In particular we might be worried that the taboo of "tribalism" would lead to an overestimation of the level of nationalism in Africa, and that this bias would be systematic across differ segments

of a population. This particular problem will be addressed again in the discussion of results. Finally, even if respondents answer honestly and the question provides an appropriate context, we can still not be sure that measured attitudes substantively influence behavior (Zaller and Feldmann, 1992).

Still, in spite of these reasons for caution, the cross-country pattern above seems to reflect common wisdom. First, Tanzania is much more nationalist than any other country in the sample – a fact which has been noted by many scholars (see Miguel, 2004 for a nice summary). Second, in analyses not shown here, regions associated with past separatist movements, such as the predominately Igbo regions of Nigeria, have very low levels of nationalism. An important next step in this research project will be to determine if this measure of nationalism is capturing identities that are stable over time. An initial way to do this will be to compare the results presented here with the subsequent round of Afrobarometer surveys due to be released later this year.[12] At the moment, there is no clear cross-national alternative measure of nationalism.

Measuring the Independent and Control Variables
State level

Hypotheses H1, H5, H8, and H9 will be tested at the state-level. Measures of income, ethnic composition, colonial power, and anti-colonial war experience all come from the dataset compiled by Fearon and Laitin (2003). *Income* is measured as the per capita gross domestic product in 1999. For *Diversity*, two measures are used. The first, the commonly employed Ethno-linguistic Fractionalization (ELF) index, represents the probability that two randomly selected individuals from the country are of a different ethnicity. The second measure is the percentage of the population made up by the largest ethnic group. Thus, while the first measure captures diversity, the second is more akin to the degree of ethnic dominance, and is negatively related to diversity (r=-0.55). The following ten countries are coded as *British colonies*: Botswana, Ghana, Kenya, Lesotho, Malawi, Nigeria, South Africa, Tanzania, Uganda, and Zambia.[13] Countries that fought an *Anti-Colonial War* were identified using the Fearon and Laitin (2003) dataset, and

---

[12] Though the first two Afrobarometer rounds did ask questions about political identity, the questions were worded such that cross-round comparisons are very difficult. However, the wording on round four surveys is identical to the question used here, allowing for such comparison across rounds.

[13] Uganda and Tanzania were not technically British colonies. Uganda was a protectorate. Tanganyika (the former name of mainland Tanzania) was a German colony, and it was under German rule that Swahili became the language of administration. It only came under British authority as a mandate under the League of Nations. However, using the Fearon and Laitin (2003) dataset, both countries are coded as British colonies.

include only Kenya, Madagascar, and Mozambique.[14] Descriptive statistics for all these variables can be found in Table 1 of the supplementary materials.

Ethnic Group Level

　　Hypotheses H3, H6, and H7 are evaluated using ethnic group level data. Two hundred and eighty-seven ethnic groups are represented in the Afrobarometer sample, with 176 groups represented by at least one percent of the sample from their respective countries. The number of such main groups per country ranges from two in Lesotho to seventeen in Uganda. Figure 4 shows the percentage of nationals within each of the main ethnic groups by country, which gives the reader some idea of the amount of variation in the nationalism across ethnic groups within a single country.
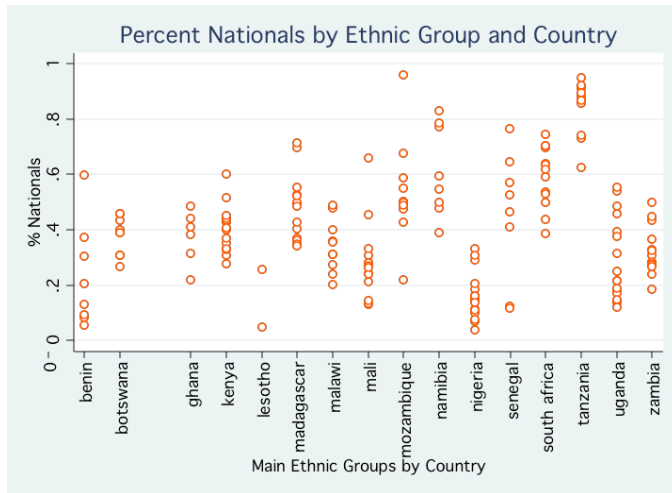

Figure 4. Nationalism by ethnic group and country. Each dot represents a single ethnic group.

Note that ethnic groups that are found in more than one country in the sample (such as the Chewa in Malawi, Mozambique, and Zambia) are coded as separate groups. This is because the dependent variable is the choice between ethnic group and national group, and individuals within a single group may have different choices depending on which side of the border they are on.

　　Measuring ethnic group level variables proved to be more difficult than the other levels First, many of the ethnic group names listed in Afrobarometer did not match the group names in other datasets, as the same groups are often called by different names when they straddle

---

[14] South Africa and Namibia could arguably be coded as having fought anti-colonial wars. When they are coded as such, the state-level analyses are similar, though the effect of income per capita is weakened.

international borders (Asiwaju, 1985) and local spellings can be quite different than English spellings. Thus, I put together a list of alternative names for each ethnic group using a variety of sources, including Ethnologue (Gordon, 2005) and the Joshua Project (2009) – an online database of ethnic group demographics based on country census data. I then used Joshua's Project to determine the relative size and the degree of partition of each ethnic group in each country. *Ethnic Group Size* is simply the share of the country's population made up of members of that ethnic group (e.g., Kikuyu of Kenya=20%, Wolof of Senegal=36%, Sesotho of Lesotho=79%, Tumbuka of Malawi=8%, Tumbuka of Zambia=4%).[15] Thus, it captures the political power of an ethnic group based on an ethnic head count. To account for the anticipated non-linear effect of group size, the square of ethnic group size is also included. *Not Partitioned* indicates that all members of the ethnic group reside within the state. The *Degree of Partition* is measured as the percentage of the total ethnic group that resides outside the state boundaries (e.g., Kikuyu of Kenya=2%, Wolof of Senegal=6%, Sesotho of Lesotho=69%, Tumbuka of Malawi=32%, Tumbuka of Zambia=78%). Thus, the higher the degree of partition, the more that international boundaries separate a member of that group from the majority of his or her ethnic group members. Note that when *Not Partitioned* is one, *Degree of Partition* is zero. When *Not Partitioned* is zero, I use the continuous variable *Degree of Partition* to capture the difference in a group that is split 30-70 across a border and one that is split 1-99 across a border.[16]

Next, *Group Poverty* is measured as an ethnic group's average response to the following question:

> *Think about the condition of [respondent's ethnic group]. Are their economic conditions worse, the same, or better than other groups in this country?*
>
> *1=Much Better*
> *2=Better*
> *3=Same*
> *4=Worse*
> *5=Much Worse*

---

[15] This measure of ethnic group share of the population is highly correlated (r=0.91) with the ethnic group share in the Afrobarometer sample, suggesting that my coding is fairly accurate.

[16] Perhaps contrary to what one might expect, an ethnic group's size and degree of partition are not highly correlated (r=0.03). While partitioned groups tend to be slightly larger, on average, than non-partitioned groups (t=2.6, df=186, p<0.01), when each variable is split into two groups by the median, the degree of partition and the ethnic group size are statistically independent ($X^2(1, 260) = 0.02$, p>0.05).

Thus, higher values indicate more economically disadvantaged groups. This is obviously not an ideal measure, as it is based on a subjective perception of economic status, and is averaged over members of the group. However, it does provide a better measure than simply averaging the individual level wealth estimate, as it asks for the status of the group as a whole, not just the individuals interviewed.[17]

Finally, as a control, I coded an ethnic group as having a *Co-Ethnic Head of State* if the head of state at the time of the survey was a member of the same ethnic group as the respondent. This was determined using data from Fearon, Kasara, and Laitin (2007), supplemented by my own research when a leader was not included in their dataset.[18]

Individual level

Hypotheses H2 and H4 require individual level data. All individual level data came from questions within the survey. *Formal Sector Employment* was determined by combining two questions: the first asked whether the individual received a wage, and the second asked the specific occupation. The occupational categories were split into formal (military/police, clerical worker, business person, professional worker, civil servant, teacher, etc.) and informal (subsistence farmer, informal manual labor, herder, housewife, etc.). Individuals receiving a wage and employed in the formal sector were coded as formal sector employees. *Education* level was measured on a scale from no formal schooling (0) to post-graduate education (9). Whether the individual lived in an *Urban* or rural location was recorded by the enumerator, and was part of the sample stratification. Enumerators also recorded whether a number of services, including post office, school, police station, electricity grid, piped water, sewage system, health clinic, and paved roads were in the immediate area. As a measure of public services, I use *Piped Water*, under the assumption that it is primarily provided by the state. *Age, Gender* (Male=1, Female=0), and *Wealth* were recorded for each respondent, as control variables, with the first two coming directly from the survey. Socio-economic status was determined by three questions asking whether the individual owned a radio, a television, and a car. *Wealth*, then, is a four-point scale including not owning a radio (1), owning a radio (2), owning a television (3), and owning a car

---

[17] The average measure of subjective ethnic group poverty is negatively related to the average measure of individual wealth across ethnic groups (r=-0.38, p<0.01), suggesting that this is a valid measure of group poverty.

[18] The Fearon, Kasara, and Laitin (2007) dataset only goes through 1999. Thus, it was used in cases where the head of state in the 2005-2006 survey period was already serving in 1999 or had previously held office sometime before 1999. When this was not the case, I read basic sources on the leader to determine his ethnicity.

(4). This measure was chosen because it is objective, and it seems to represent a natural gradient (96% of those with a TV also have a radio and 85% of those owning a car also have a television).[19] Finally, to control for enumerator effects, I include a control variable indicating whether the enumerator and the respondent were of the *Same Ethnicity*. This indicator was constructed by comparing the home language of the enumerator and the home language of the respondent. Again, descriptive statistics for each of these variables are reported in Table 1 of the supplementary materials.

Results and Discussion

To determine if and how state, ethnic group, and individual level variables are related to identification with the state, models at three separate levels were estimated.

[In the next version of this paper, I plan to combine these three models into one multi-level, random intercepts model. Please see Table 5 at the end of this document for a first cut at that estimation. Most of the results are consistent with the results modeled separately at each level, in terms of statistical significance.]

State Level

To determine the effect of the four state level variables, nationalism (percent of respondents in a country that chose national over ethnic identity) is regressed on log of income per capita, ethnic diversity, a British colonialism indicator, and an anti-colonial war indicator using ordinary least squares estimation. The model is presented in Table 2 with each of the measures of ethnic composition, and with and without Tanzania.[20]

With the full sample (Model 1), the model produces a very poor fit: the negative adjusted R-squared means that the fit is worse than what we would expect by chance, given the number of observations and covariates. However, once Tanzania is excluded, Model 2 is able to account for 65% of the variation in nationalism. The results are likely to be strongly driven by per capita income. In Model 2, we can see that a 1% increase in $1000 per capita income translates into to an additional 11% of a state's population identifying with the nation over an ethnic group. Figure 5 shows the bivariate relationship between income and nationalism.

---

[19] An alternative measure of socio-economic status, based on a subjective question about the respondent's living conditions relative to others, is highly correlated with this objective measure, and produces very similar results in regression analyses.

[20] The models were also estimated with the exclusion of Nigeria, the least nationalist country in the sample, in addition to Tanzania, but the results were similar to those estimated with the sole exclusion of Tanzania.

Table 2.
State Level OLS: Percent National over Ethnic Identity

	(1) Full Sample	(2) Excluding Tanzania	(3) Full Sample	(4) Excluding Tanzania
Ln(Income per capita)	0.07	0.11***	0.09	0.12***
	(0.06)	(0.02)	(0.05)	(0.02)
Ethnic Fractionalization	0.19	0.04		
	(0.21)	(0.09)		
Share of Largest Ethnic Group			-0.42*	-0.19*
			(0.22)	(0.09)
British Colonialism	-0.07	-0.10**	-0.08	-0.11**
	(0.10)	(0.04)	(0.09)	(0.04)
Anti-colonial War	0.09	0.12*	0.02	0.09*
	(0.13)	(0.05)	(0.11)	(0.05)
Constant	0.24	0.27***	0.52***	0.37***
	(0.18)	(0.07)	(0.13)	(0.06)
Observations	16	15	16	15
Adjusted R-squared	-0.09	0.65	0.11	0.75

*** p<0.01, ** p<0.05, * p<0.1
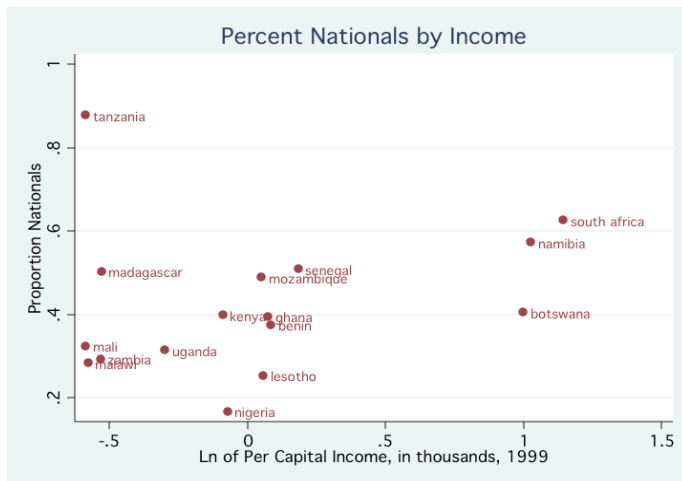Standard errors in parentheses

Figure 5. Percent nationals by income.

One interpretation of this relationship is that higher income leads to greater modernization, which, in turn, increases nationalism via education, industrialization, and urbanization.[21] This would be consistent with the literature showing that the rise of nationalism in Europe happened in conjunction with increased modernization (H1). An alternative explanation for the strong positive relationship between income and nationalism is that the two are endogenous. In fact, as I argued to be a motivation of the project, a strong degree of nationalism may lead to good economic outcomes by reducing the saliency of sub-national differences, and, in turn, increasing trust and cooperation among citizens. The results reported here are only able to establish a relationship. It must be left to future work to determine the mechanism by which income and nationalism are related.

The effect of ethnic diversity, as measured by ELF, is not significantly different from zero. However, contrary to H5, the coefficients are positive. The second measure, population share of the largest ethnic group, *is* a significant negative predictor of nationalism (Models 3 and 4), following the same trend – more diverse states are more nationalist. The larger that the largest ethnic group is, the less nationalist a country is on average – again, contrary to H5. Figure 6 shows this bivariate relationship. This seems slightly counterintuitive. We should expect, as articulated in H5, that the more homogenous a population, the easier it would be to build a national culture. However, given that a state contains multiple cultural groups – which all the states in the sample do – then having a large group may make the possibility of incorporating members of other groups into a single national culture less likely. This may explain part of Tanzania's success – it has no dominant ethnic group, with its largest group comprising only 12% of the population. We will return to questions about the relative sizes of ethnic groups in the next section.

Finally, as predicted by H7 and H8, British colonialism and anti-colonial war experience are significantly related to nationalism. Former British colonies have, on average, around 10% less of their population identifying with the state. Having fought an anti-colonial war, on the other hand, produces about 9-12% more nationals, holding all other variables constant. However, only three of the sixteen cases fought such wars, suggesting caution in interpretation.[22] Further,

---

[21] Within this dataset, a country's per capita income is significantly correlated with urbanization (r=0.7), education (r=0.4), and formal employment (r=0.7).

[22] The significance of anti-colonial wars is even more surprising considering that such wars did not always serve as a unifying force. In fact, Jeremy Weinstein points out that in Mozambique the favored ethnic groups under colonialism fought on the side of Portugal, rather than with the nationalists. With so few cases, the finding remains tentative.

the relationship may be endogenous if the level of nationalism present under colonial rule determined which colonies went on to violently oppose foreign rule, and which did not. Still, the results lend further empirical support to qualitative arguments proposing such relationships.
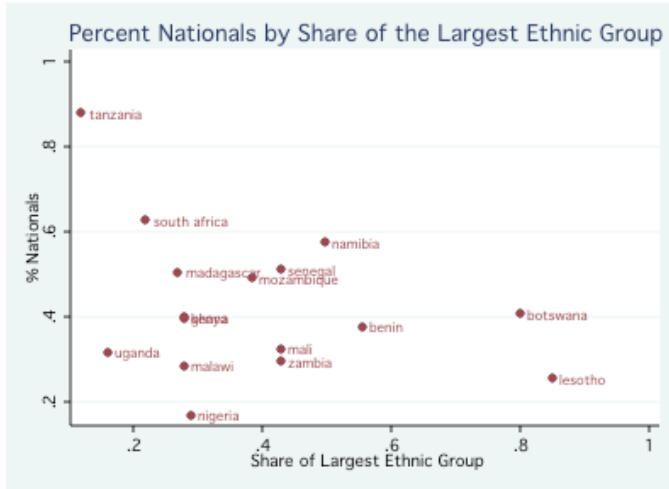

Figure 6. Percent nationals by population share of the largest ethnic group.

Ethnic Group Level

Because of the small sample size for many ethnic groups, the ethnic group level analysis is modeled as an individual level logistic regression, with ethnic group variables, country fixed-effects, and standard errors clustered by ethnic group. Table 3 reports the estimated effect of the ethnic group's size (ethnic group size and ethnic group size squared), degree of partition (an indicator for partition and the proportion of the total ethnic group living outside the state borders), relative economic conditions, and having a co-ethnic head of state. For brevity, the coefficients on state fixed effects are not reported. As a quick comparison of Models 1 and 2 shows, the exclusion of Tanzanian ethnic groups does not significantly affect the estimated coefficients. As a result, only Model 1 is discussed.

Table 3.
Ethnic Group Level Logistic Regression: National over Ethnic Identity

| | (1) | (2) |
| | | Excluding |
	Full Sample	Tanzania
Group Size (Share of State Population)	4.97***	4.90***
	(1.30)	(1.35)
Group Size Squared	-5.88***	-5.82***
	(1.67)	(1.74)
Not Partitioned	0.34**	0.36**
	(0.15)	(0.16)
Degree of Partition	0.56***	0.57***
	(0.19)	(0.20)
Group Poverty	-0.08***	-0.08**
	(0.03)	(0.03)
Co-Ethnic Head of State	-0.45***	-0.43***
	(0.13)	(0.14)
Constant	-0.83**	-0.84**
	(0.37)	(0.37)
Observations	18666	18057
Clusters	252	235

Clustered standard errors in parentheses; state fixed effects included.
*** p<0.01, ** p<0.05, * p<0.1

As expected, the state population share that a group makes up is significantly related to national attachment. However, contrary to the prediction put forth in H6, the non-linear relationship between ethnic group size and degree of nationalism is the following: increasing group size leads to increasing attachment to the state up to a certain threshold, at which point increased size leads to decreased nationalism. More simply, the least national groups, on average, are the very small and the very large. All else equal, national identification is maximized at an ethnic group making up 41% of the population. This means that for any groups smaller 41%, average national identification increases with size. For those larger than 41% of the state population – which, in this sample, includes only the Sotho of Lesotho, the Tswana of Botswana, and the Chewa of Malawi – national identification decreases with size. Posner (2004, 2005) and others have argued that *ethnic* salience should increase with ethnic group size. While this result does not rule out that claim, it may qualify it. If ethnic salience does increase with ethnic group share of the population, this increased salience is trumped by an increased salience in national identity for most groups.

What about the effect of ethnic group partition? The coefficient on *Not Partitioned* suggests that groups that fall completely within the state boundaries (i.e., groups which are not partitioned) are about 8% more likely to identify with the state, consistent with H7a. However, if a group is partitioned, the degree of the split (measured as the share of the group outside one's state) is *positively* related to identifying with the state. This means that if your group is partitioned, the more members of your group that live over the border, the more likely you are to identify with your own state. There are several potential explanations for this relationship. First, highly partitioned groups may be more likely to identify themselves as nationals due to social desirability. In other words, members of these highly partitioned groups may feel they have to "prove" their nationalism more. Second, to the degree that ethnic identification is maintained through the provision of services or benefits to members of that ethnic group, members of groups that have been split off from the majority of their co-ethnics are also more likely to be cut off from the institutions or networks that operate within that group, and, thus, accrue fewer benefits by identifying ethnically. In the absence of a strong ethnic identity, national identification may be more likely. The third and fourth potential explanations stem from the observations that members of highly portioned groups often have their co-ethnics make-up a substantial proportion of the population in the neighboring country. For example, the Yoruba of Benin, the Yao of Mozambique, the Bambara of Senegal, and the Tswana and Sotho of South Africa are all among the most nationalist groups in their respective countries, and each are co-ethnics with a group that is of a significant size in a neighboring country (Nigeria, Malawi, Mali, Botswana, and Lesotho, respectively). Thus, the third potential explanation is that in such highly partitioned cases, the ethnic identity becomes so strongly affiliated with the neighboring state, that the few individuals living in other states do not seem themselves as similar any longer. Finally, there may be a political story, if the high levels of nationalism in highly dismembered groups is a reflection of the central government's efforts to buy them off and prevent irredentism. More work is needed in order to determine the mechanism that leads highly partitioned groups to self-identify as nationals to a greater degree than less partitioned groups.

In support of H3, poorer ethnic groups – or, more specifically here, ethnic groups that see themselves as relatively poorer than other ethnic groups – have lower levels of state identification. The substantive effect is, at most, about an 8% lower probability of identifying with the state within a group with "much worse" economic conditions, as compared to members of a group with "much better" economic conditions, relative to other groups.

Finally, having a co-ethnic as head of state is negatively related to national identification. This may be because individuals identify with their ethnic group when there are material benefits to doing so (Padro i Miquel, 2007), and patronage from co-ethnic politicians is one source of such benefits. Though, empirically, having a co-ethnic head of state in Africa can actually leads to higher levels of taxation rather than group patronage (Kasara, 2007). Alternatively, this may be a case of reverse causality, where an ethnic group is most likely to have a co-ethnic elected *because* ethnicity is very salient among members of that group.[23]

Individual Level

An individual-level logistic regression with state fixed-effects is used to estimate the effect of individual characteristics on the propensity to identify with the state. Table 4 reports the average effects across the sixteen states.

These individual level results are consistent with the state level results reported above. In particular, note that modernization (education, urbanization, formal employment) and income are positively related to national identification. This finding is consistent with H2, and suggests that increased development may lead to greater nationalism. An alternative explanation for this relationship would be that more "modern" individuals are more likely to think that national identification is normatively better than identifying with one's ethnic group. Thus, the relationship may be due to individuals giving what they believe to be the socially desirable answer, and the strength of social desirability varying by degree of individual modernization. While I cannot rule this alternative out, it is promising that these results hold even when I restrict the sample to those respondents that the enumerator judged to be "honest" rather than "in between" or "misleading."

Contrary to H4, the provision of piped water has no significant effect on national identification. In unreported analyses, the result is the same for indicators of access to an electricity grid and paved roads, suggesting that development infrastructure is unrelated to nationalism at the individual level. This is consistent with the finding that such infrastructure has

---

[23] As a robustness check, the same model was estimated with the inclusion of individual level variables of interest, and the results are presented in Table 2 of the supplemental materials. The fact that the group level estimates are very similar even with the inclusion of individual level variables gives confidence that the group level effects are not driven by differences at the individual level, with different types of groups (big/small, partitioned/not partitioned) being composed of different kinds of individuals.

little impact on democratization (Bratton and Chang, 2006). Thus, the data do not support the mechanism that modernization impacts nationalism via increased state-driven development.

Table 4.
Individual Level Logistic Regression: National over Ethnic Identity

	(1) Full Sample	(2) Excluding Tanzania
Age	0.01**	0.01**
	(0.001)	(0.001)
Age Squared	-0.001**	-0.001**
	(0.00)	(0.00)
Male	0.13***	0.13***
	(0.03)	(0.03)
Formal Employment	0.11***	0.11**
	(0.04)	(0.04)
Level of Education	0.05***	0.05***
	(0.01)	(0.01)
Urban	0.09**	0.09**
	(0.04)	(0.04)
Same Ethnic Group	0.03	0.04
	(0.03)	(0.03)
Piped Water	0.02	0.05
	(0.04)	(0.04)
Wealth	0.06***	0.05**
	(0.02)	(0.02)
Constant	-1.13***	-1.10***
	(0.12)	(0.12)
Observations	21937	20753

Standard errors in parentheses; state fixed effects included.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Finally, while entered as controls, it is interesting to note that age and gender are significant predictors of national identification, even when controlling for characteristics likely to be correlated with age and gender, such as education level and employment status. As for gender, males are significantly more likely to identity with the state. This pattern is consistent with the fact that African women are much less likely than men to speak the *lingua franca* of a state (Laitin, 1992), suggesting that investment in a state-level language and the importance of one's national identity may be related, and differ across genders. Interestingly, on average, age has a non-linear relationship to nationalism, as is shown in Figure 7. From this graph, and the signs on the age and age-squared coefficients, it is clear that younger and older people are less nationalist

than middle-aged individuals. It would be interesting to compare this result with trends on other continents to determine if the trend reflects life-cycle patterns, or whether it is specific to African states and the timing of their independence. If it is the latter, it could be that individuals who lived under colonialism (those over 60) and individuals who came of age under democracy (those under 30) are more likely to identify in ethnic terms than are middle aged individuals who came of age during the independence era. Consistent with this interpretation, in Namibia, the newest state in the sample, age and age squared have no significant effect on national identification. Alternatively, the pattern of greater ethnic attachment in young and old age may reflect a more general life-cycle, where the young and the old are likely to be the most dependent on the support of ethnic kin (Jeremy Weinstein, personal communication).
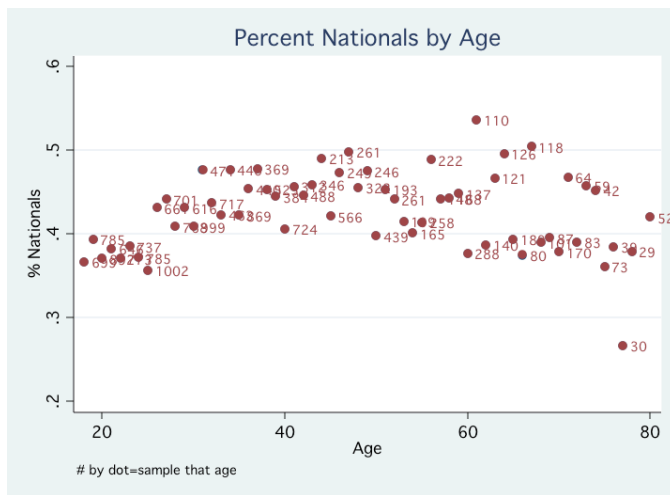

Figure 7. Percent nationals by age (aggregated across all countries).

Table 4 reports the *average* effects of individual characteristics across states. These estimates, however, vary across countries. Table 3 of the supplementary materials lists the estimated effect of the individual level variables for each of the sixteen states separately. For brevity I will not discuss those results here, but they suggest many interesting questions for future research. As just one example, note the differing effect of being interviewed by a co-ethnic on an individual's choice of national or ethnic identity. In Botswana, Lesotho, Madagascar, Senegal, and Zambia co-ethnics influence one towards identifying with the state. In contrast, co-ethnic enumeration is negatively correlated with national identification in Nigeria, South Africa, Tanzania, and Uganda.

Conclusion

We can draw four broad conclusions from the preliminary results reported above. First, "classic" theories of nationalism, mostly based on the rise of nations in 18th century Europe, explain a considerable amount of variation in nationalism across the sixteen African states included in this study. This suggests that nationalism – identification with the state – follows general patterns across time and space. In particular, hypotheses derived from the modernization school of thought were mostly supported with the African data. Rich states have higher levels of nationalism than poor states. Across all states, the richer, more educated, urban, and formally employed individuals are more likely to identify with the state than poorer, less educated, rural agriculturalists. However, consistent with the predictions about the adverse effects of uneven development along ethnic lines, poorer ethnic groups are less nationalist than richer ones. Thus, we should expect nationalism to continue to increase as African states continue to industrialize and modernize their peripheries. Contrary to theories of state identification based on the provision of goods and services, I find no relationship between running water (or other indicators of infrastructure) and state identification.

Second, the qualitative arguments for why African states are less nationalist than other regions of the world – high levels of ethnic diversity and ethnic partition due to artificial boundaries – do less well at explaining variation *within*-Africa. First, ethnic diversity seems to be *positively* related to nationalism. This may be an artifact of the fact that all the states in the sample are heterogeneous, and, in multi-cultural states, a supra-ethnic identity is most easily fostered in the absence of ethnic dominance. In other words, in the absence of a completely homogeneous citizenry, creating a national identity may be easiest in highly heterogeneous populations. In addition, ethnic share of the population shows a pattern contrary to the expectation that ethnic salience, relative to national salience, should increase with size. The results suggest, instead, that nationalism generally increases with size. One possibility, related to the above discussion of ethnic dominance, is that ethnic group share of the population may only matter relative to the size of the largest group. When the largest group has a low share of the population, the size of other groups may matter a lot. When the largest group is much larger than any other group, group share of the population may explain much less. Finally, the effect of group partition, consistent with literature on partitioned groups, seems to be that groups that fall completely within a state's borders are more nationalist than members of groups that are partitioned. However, contrary to expectations, the results show that, among partitioned groups,

the larger the proportion of the total ethnic group that resides across the border, the *more* nationalist a group will be. There are several potential explanations for this finding: more work is required to determine if the pattern is robust, and, if so, why.

Third, the characteristics that have been suggested to account for within-Africa variation, such as British colonialism and anti-colonial wars, are tentatively supported. Once Tanzania is excluded, both variables influence nationalism in the hypothesized direction. British colonies are less nationalist, on average, than the French, German, and Portuguese colonies in the sample. Countries that used violence in the struggle for independence, on the other hand, have higher average levels of nationalism.

The final conclusion is that even after all of the characteristics considered here, Tanzania remains an outlier at the state level. As the most nationalist state in the sample, Tanzania runs counter to all predictions – it is very poor, highly ethnically diverse, a former British colony, and did not fight an anti-colonial war. However, these differences do not stem from differences at the ethnic group or individual levels, as the exclusion of Tanzanian data at those two levels did not influence the results. This suggests that there are state-level conditions that are not considered here that are very important for nationalism. Miguel (2004) gives four such conditions that have contributed to Tanzanian nationalism: the widespread use of Kiswahili as a second language, the nationalist content of primary school education and the use of Kiswahili in schools, equitable regional distribution of state resources in the early post-independence era, and the personal attributes of the first political leader, Julius Nyerere. An understanding of the causes of Tanzania's success, and measures of those attributes for all African states, would undoubtedly explain much more of the cross-country variation that remains after this analysis.

This project has made an important first step towards understanding the correlates – and potential causes – of national identification in Africa. Significant work remains to be done in order to establish the robustness of these findings, as well as the mechanisms that underly them.

Table 5.
Multi-Level Logistic Model: National over Ethnic Identity w/ Country and Group Random Intercepts

		(1)	(3)	(2)	(4)
		Full Sample	*Excl. Tanzania*	*Full Sample*	*Excl. Tanzania*
Country-Level IVs	Ln(Income per capita)	0.120	0.517***	0.446	0.639***
		(0.369)	(0.177)	(0.355)	(0.169)
	Ethnic Fractionalization	1.274	0.103		
		(1.050)	(0.552)		
	Share of Largest Ethnic Group			-2.883**	-1.412**
				(1.162)	(0.642)
	British Colonialism	-0.218	-0.289	-0.333	-0.418**
		(0.468)	(0.228)	(0.415)	(0.203)
	Anti-colonial War	0.634	0.687**	0.173	0.545**
		(0.584)	(0.271)	(0.502)	(0.238)
Group-Level IVs	Group Size (Share of Population)	2.608*	3.023**	2.333	2.697*
		(1.433)	(1.462)	(1.437)	(1.460)
	Group Size^2	-2.166	-3.574*	-1.425	-2.508
		(2.269)	(2.152)	(2.279)	(2.144)
	Not Partitioned	0.045	0.061	0.037	0.061
		(0.151)	(0.157)	(0.151)	(0.154)
	Degree of Partition	0.489**	0.556**	0.500**	0.598***
		(0.214)	(0.217)	(0.215)	(0.218)
	Group Poverty	-0.035*	-0.032*	-0.035*	-0.032*
		(0.019)	(0.019)	(0.019)	(0.019)
	Co-Ethnic Head of State	-0.514**	-0.425*	-0.518**	-0.437*
		(0.233)	(0.243)	(0.233)	(0.242)
Individual-Level IVs	Age	0.020***	0.020***	0.020***	0.021***
		(0.006)	(0.006)	(0.006)	(0.006)
	Age Squared	-0.001***	-0.001***	-0.001***	-0.001***
		(0.00)	(0.00)	(0.00)	(0.00)
	Male	0.149***	0.145***	0.149***	0.145***
		(0.034)	(0.034)	(0.034)	(0.034)
	Formal Employment	0.085*	0.085*	0.085*	0.084*
		(0.048)	(0.048)	(0.048)	(0.048)
	Level of Education	0.056***	0.057***	0.056***	0.056***
		(0.011)	(0.011)	(0.011)	(0.011)
	Urban	0.135***	0.136***	0.135***	0.134***
		(0.042)	(0.043)	(0.042)	(0.043)
	Same Ethnic Group	0.189***	0.191***	0.190***	0.193***
		(0.041)	(0.041)	(0.041)	(0.041)
	Piped Water	0.004	0.014	0.004	0.017
		(0.043)	(0.043)	(0.043)	(0.043)
	Wealth	0.045**	0.043*	0.045**	0.042*
		(0.022)	(0.022)	(0.022)	(0.022)
	Constant	-2.302***	-1.713***	-0.173	-1.008**
		(0.797)	(0.443)	(0.663)	(0.396)
	SD of country-level random intercept	0.791	0.321	0.697	0.278
		(0.150)	(0.084)	(0.133)	(0.075)
	SD of group-level random intercept	0.555	0.546	0.556	0.543
		(0.043)	(0.044)	(0.043)	(0.044)
	Individual-Level Observations	18076	17489	18076	17489
	Ethnic Group-Level Observations	251	234	251	234
	State-Level Observations	16	15	16	15

Standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1

Supplemental materials are available at:
http://www.stanford.edu/~alrobins/Amanda_Lea_Robinson/Research.html.

References

Afrobarometer. 2008. AfrobarometerSurveys, Round 3. Retrieved on 2008-09-10.
     <http://www.afrobarometer.org/countries.newmap.html>.

Alesina, A., W. Easterly, and J. Matuszeski. 2006. "Artificial States." NBER Working Paper No. 12328.

Alesina, A. and E. La Ferrara. 2004. "Ethnic Diversity and Economic Performance." National Bureau of Economic Research WP 10313.

Anderson, B. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism.* London: Verso Press.

Asiwaju, A. I. 1985. *Partitioned Africans : ethnic relations across Africa's international boundaries, 1884-1984.* New York: St. Martin's Press.

Azam, J. 2001. "The redistributive state and conflicts in Africa." *Journal of Peace Research,* 38(4):429.

Bates, R. 1983. "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In D. Rothchild and V.A. Olorunsola, eds. *State versus Ethnic Claims: Africa Policy Dilemma.* Boulder, Colorado: Westview Press. pp. 152-171.

Bienen, H. 1983. Sthe state and ethnicity: Integrative formulas in Africa. In D. Rothchild and V. A. Olorunsola, eds. *State versus Ethnic Claims: African Policy Dilemma.* pp. 100-126. Boulder, Colorado: Westview Press.

Billig, M. and H. Tajfel. 1973. "Social categorization and similarity in intergroup behavior." *European Journal of Social Psychology,* 3(1).

Bratton, M. and E.C.C. Chang. 2006. "State Building and Democratization in Sub-Saharan Africa." *Comparative Political Studies,* 39 (9): 1059-1083.

Chandra, K. 2001. "Cumulative Findings in the Study of Ethnic Politics." *APSA-CP,* 12 (1): 7-11.

Chandra, K. 2005. "Ethnic Parties and Democratic Stability." *Perspectives on Politics* 3(2): 235-252.

Conner, W. 1990. "When is a Nation?" *Ethnic and Racial Studies* 13 (1): 92-100.

Easterly, W. and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *The Quarterly Journal of Economics* 112 (4): 1203-1250.

Eifert, B., E. Miguel, and D. Posner. 2008. "Political Competition and Ethnic Identification in Africa." Unpublished manuscript.

Engelbert, P. 2000. *State Legitimacy and Development in Africa.* Boulder, Colorado: Lynne Rienner Publishers.

Engelbert, P., S. Tarango, and M. Carter. 2002. "Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries." *Comparative Political Studies,* 35 (10): 1093-1118.

Fearon, J. and D. Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review,* 97: 75-90.

Fearon, J., K. Kasara, and D. Laitin. 2007. "Ethnic Minority Rule and Civil War Onset." *American Political Science Review,* 101: 187-193.

Geertz, C. 1963. "The Integrative Revolution: Primoridal Sentiments and Civil Politics in the New States." In C. Geertz, ed. *Old Societies and New States: The Quest for Modernity in Asia and Africa.* Free Press: New York.

Gellner, E. 1964. "Nationalism." In *Thought and Change.* London: Weidenfeld and Nicholson. 147-178.

Gellner, E. 1983. *Nations and Nationalism.* Ithaca, NY: Cornell University Press.

Padro i Miquel, G. 2007. "The Control of Politicians in Divided Societies: The Politics of Fear." *Review of Economic Studies* 74(4): 1259-1274.

Gordon, Raymond G., Jr. (ed.), 2005. *Ethnologue: Languages of the World, Fifteenth edition.* Dallas, Tex.: SIL International. Online version: http://www.ethnologue.com/

Greenfeld, L. 2001. *The Spirit of Capitalism: Nationalism and Economic Growth.* Cambridge: Harvard University Press.

Hechter, M. 2000. *Containing Nationalism.* Oxford: Oxford University Press.

Herbst, J. 1998. "The creation and maintenance of national boundaries in Africa." *International Organization,* 43 (4): 673-692.

Herbst, J. 2001. *States and Power in Africa: Comparative Lessons in Authority and Control.* Princeton: Princeton University Press.

Horowitz, D. L. 1985. *Ethnic Groups in Conflict.* Berkeley: University of California Press.

Jackson, R.H., C.G. Rosberg. 1982. "Why Africa's weak states persist: the empirical and the juridical in statehood*." World Politics: A Quarterly Journal of International Relations,* 1-24.

Hutchinson, J. 1994. *Modern nationalism.* London: Fontana Press.

Hyden, G. 1983. Problems and Prospects of State Coherence. In D. Rothchild and V. A. Olorunsola, editors. *State versus Ethnic Claims: African Policy Dilemma.* pages 67-84. Westview Press: London.

Jackson, R. and C. Rosberg. 1982. "Why Africa's Weak States Persist: The Empirical and the Juridical in Statehood." *World Politics,* 35 (1): 1-24.

Joshua Project. 2009. www.joshuaproject.net.

Kasara, K. 2007. "Tax me if you can: Ethnic geography, democracy, and the taxation of agricultural Africa." *American Political Science Review,* 101 (1): 159-172.

Kuo, A. and Y. Margalit. 2009. "Measuring Individual Identity: Experimental Evidence." Unpublished manuscript, Stanford University.

Laitin, D. 1986. *Hegemony and Culture: Politics and Religious Change among the Yoruba.* Chicago: Chicago University Press.

Laitin, D. 1992. *Language Repertoires and State Construction in Africa.* Cambrige: Cambridge University Press.

Laitin, D. and Posner, D. 2001. 'The Implications of Constructivism for Constructing Ethnic Fractionalization Indices' APSA_CP 12: 1-14.

Mazrui, A.A. 1983. "Francophone nations and english-speaking states: imperial ethnicity and African political formations." In D. Rothchild and V. A. Olorunsola, editors. *State versus Ethnic Claims: African Policy Dilemma.* Boulder, Colorado:Westview Press. pp. 25-43.

Miguel, E. 2004. "Tribe or Nation? Nation-building and Public Goods in Kenya versus Tanzania." *World Politics,* 56 (3): 327-362.

Miles, W. 1994. *Hausaland Divided.* Ithaca, NY: Cornell University Press.

Nairn, T. 1977. "The twilight of the British state." *New Left Review*, 101(2): 3–61.

Neuberger, Benyamin. 2000. "Ethnic Groups and the State in Africa." In *Ethnic Challenges to the Modern Nation State*, eds. Sholomo Ben-Ami, Yoav Peled and Alberto Spektorowski. pp. 294-308.

Posner, D. 2004. "The Political Salience of Cultural Difference: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi," *American Political Science Review*, 98, 4: 529-545.

Posner, D. 2005. *Institutions and Ethnic Politics in Africa.* New York: Cambridge University Press.

Rothchild, D. 1983. "Collective demands for improved distributions." In D. Rothchild and V. A. Olorunsola, editors. *State versus Ethnic Claims: African Policy Dilemma.* Westview Press: London. pp. 172–198.

Snyder, J. 1993. "Nationalism and the Crisis of the Post-Soviet State." In *Ethnic Conflict and International Security.* Princeton: Princeton University Press. pp. 79-102.

Tilly, C. 1990. *Coercion, Capital, and European States, AD 990-1992*. Oxford: Basil Blackwell.

Transue, J.E. 2007. "Identity salience, identity acceptance, and racial policy attitudes: American national identity as a uniting force." *American Journal of Political Science*, 78–91.

Weber, E. 1976. *Peasants Into Frenchmen: The Modernization of Rural France, 1880–1914.* Stanford: Stanford University Press.

Young, C. 1985. "Ethnicity and the colonial and post-colonial state in Africa." In P. Brass (ed) *Ethnic Groups and the State.* London: Croom Heim Press. pp. 57–93.

Zaller, J. and S. Feldmann. 1992. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science,* 36: 579-616.