

Pre-Analysys Plan
for
“Tribe *and* Nation:
Minimum Winning Coalitions and Founding Leader’s
Ethnicity as Causes for Nation-Building Policies”

Constantine Manda *

November 10, 2016

Abstract

[Miguel \(2004\)](#) uses Kenya and Tanzania as natural experiments in nation-building finds that public good provision suffers in contexts where ethnicity rather than the nation is salient among ethnically-heterogeneous citizens and he attributes successful nation-building efforts by Tanzania’s inaugural President Julius K. Nyerere over his Kenyan counterpart, Jomo Kenyatta. I argue that Tanzania’s Nyerere simply faced ethnic diversity that compelled a nation-building strategy over politicizing ethnicity. Specifically, because no Tanzanian ethnicity commands greater than one-tenth of the population, building a *minimum winning coalition* can only occur if national identity emerges rather than local ethnic identity. Moreover, Nyerere’s ethnicity, in such an ethnically-heterogeneous context that existed at the time in Tanzania exacerbated incentives away from ethnic politics. I use these *minimum winning coalition* and *founding leader’s ethnicity* explanations to extend beyond the Kenya-Tanzania comparison to other African countries at independence. I compile a dataset on the ethnicity of African leaders for each 54 African country since their independence, their ethnic fractionalization building from [Fearon et al. \(2007\)](#)’s dataset and data on whether each leader emphasized the tribe or the nation in their political platforms. I estimate OLS regressions to estimate the impact of the leader’s ethnicity and ethnic fractionalization on nationalism. This study is important because it informs literature that relates Africa’s relative economic development to its ethnic diversity.

JEL Codes: N17, O10, O55

*Department of Political Science, 115 Prospect St. #208301, Yale University, New Haven, CT 06511.
E-mail: constantine.manda@yale.edu. Telephone: +13474451779.

1 Introduction

1.1 Motivation

Why are politics in Tanzania not contested along ethnic lines while Kenyan politics are characterized by ethnic conflicts? Is the answer the divergent policies that their respective inaugural leaders took in either building the nation or emphasizing ethnicity? Or are the underlying ethnic fractionalization the relevant constraints on their inaugural leaders' divergent policy-making? Miguel (2004) argues for the former and that specifically, Tanzania's Julius Nyerere was more gifted than Kenya's Jomo Kenyatta because the former muted ethnic identity over national identity while the latter accentuated ethnic rather than national identity.

I argue that Tanzania's Nyerere simply faced ethnic diversity that compelled a nation-building strategy over politicizing ethnicity. Specifically, because no Tanzanian ethnicity commands greater than one-tenth of the population, building a *minimum winning coalition* can only occur if national identity emerges rather than local ethnic identity. Moreover, Nyerere's ethnicity, in such an ethnically-heterogeneous context that existed at the time in Tanzania exacerbated incentives away from ethnic politics.

This is not to say that leadership does not matter. Jones et al. (2005), for instance, find that leadership matters for economic growth while Levitsky et al. (2015) argue that leadership matters for party-building and Magaloni (2006) sees leaders as focal points.¹ It is important, however, to better understand whether it is leadership or the constraints placed on leaders through ethnic fractionalization determines nation-building outcomes given the negative effects of ethnic diversity on economic growth and income (Alesina et al. 2003; Collier and Gunning 1999; Easterly and Levine 1997).

Ethnic diversity is also associated with inferior government performance (La Porta et al. 1999); biased toward provision of excludable public goods (Kimenyi 2006); greater incidence of conflict and coups (Blattman and Miguel 2010; Collier and Hoeffler 2004; Easterly 2001; Fearon and Laitin 2003; Geddes et al. 2016); salient during political com-

¹For more please see Fajardo (2016) who attempts to build a model for the source of leadership reconciling genetic, learning-by-doing, and learning-before-doing explanations for leadership.

petition (Eifert et al. 2010); less public good provision (Alesina et al. 1999; Miguel 2004);² less likely to cooperate (Habyarimana et al. 2007, 2009); less social capital (Collier and Gunning 1999); among other negative economic outcomes (Alesina and La Ferrara 2005).

The African context is an important one to study particularly because of its high ethnic diversity and its relative economic underdevelopment. Kenya in particular is an interesting case study and serves as a motivating example for this study. Burgess et al. (2015), for instance, find that Kenyan districts that share the ethnicity of the president receive twice as much expenditure on roads and have five times the length of paved roads built and although this effect is mitigated by periods of democratic political competition they still find that Kenya's Presidents and Vice Presidents throughout the post-independence period until 2011 do populate their cabinets with co-ethnics. In so far as public goods as such as roads are allocated based on ethnicity rather than economic efficiency Kenya's ethnic politics are costly.

1.2 Research Questions

The main research questions are as follows:

- Is leadership the reason citizens of some African countries today have relatively more of a national identity rather than an ethnic identity?
 - If not, then what other factor explains why some founding African leaders chose to emphasize national identity and mute ethnic identity while others did not?
 - * Can we empirically test whether the founding leader's ethnicity *and* the country's ethnic fractionalization are the binding constraints on the decision to select nation-building over ethnic politics?

To situate ideas imagine a country whose total population is $n = 4$ with each person being from a distinct ethnicity so that there exists no majority ethnicity—each ethnic

²Wimmer (2015), on the other hand, that the relationship between ethnic diversity and low public good provision is spurious because both contemporary ethnic heterogeneity and low public goods provision represent legacies of a weakly developed state capacity inherited from the past.

group has an equal share of the population which is $\frac{1}{4}$. Assume that an ethnicity's political power is a direct function of their relative share of the population.³ Let the founding leader be assigned by *nature* from the population and assume retaining power is a function of having a simple majority so that the leader must convince 2 other citizens to form a minimum winning coalition, otherwise the country degenerates into civil war. Let citizens' ideal points be vertices of an equilateral triangle and assume the leader can only propose ideal points on the perimeter of the triangle. It is trivial that there exists no feasible ideal point that the leader can propose that would lead to a minimum winning coalition.⁴ A unique equilibrium outcome then is civil war regardless of the proposed ideal point.

Alternatively, imagine a different country with similar conditions except there exists at least one ethnicity whose share of the population is larger than either of the remaining ones.⁵ In such a scenario, alliances between this majority ethnicity and any one of the other ethnicity sufficient to form a minimum winning coalition would be possible for multiple equilibria, depending on which ethnicity is assigned as the majority group but each time with a non-civil war equilibrium outcome.

The former country, I argue, is Tanzania while the latter is Kenya, at the time of independence. In the first scenario, the only way for the founding leader to retain power and avert civil war is to diminish ethnic identity by accentuating a national identity given the distribution of shares of the population of the country's ethnic groups. In the second scenario, the founding leader can optimally create minimum winning coalitions by organizing alliances with the *least* possible number of ethnicities sufficient for such a coalition. This is Kenyatta's Kenya at independence. Please also note that in order to capture my second hypothesis—**founding leader's ethnicity**— it matters whether nature assigns the founding leader's ethnicity as the majority ethnicity or not. Extend-

³This assumption allows for abstracting away any difference between political power and ethnic fractionalization given the literature has focused not just on fractionalization but on politically-relevant *configuration*.

⁴Label each vertex A, B, C. If the leader proposes A, both B and C will initiate a civil war; if she proposes B, both A and C will initiate a civil war; and likewise if she proposes C, both A and B will initiate a civil war.

⁵Assume this ethnicity is selected randomly by the state of nature.

ing the framework in that dimension will capture this second incentive behind founding leaders' choice whether to pursue nation-building or not.

2 Research Strategy

2.1 Sampling

2.1.1 Sampling Frame

Because this study is about explaining how leaders, in a context of ethnic diversity, come to choose nation-building over ethnic politics the sample is restricted to all 54 African countries. I do not include countries from the whole world because of the following reasons:

- for North America and Australia (and environs), the founding leader was not necessarily a *native* of the land and was herself a colonist as is the case of the United States that may complicate the theoretical framework
- for Central and South America, in addition to the reason above they also gained independence much earlier than Africa making data availability difficult and also had varying degrees of pre-nation states that may complicate the theoretical framework
- for Europe, in addition to the second reason above there was a lot of interstate wars and other factors that are different in the way nation-building occurred and there was also not necessarily any nations that emerged from colonialism
- for Asia, the last two reasons above apply

so that I focus only on Africa.

2.1.2 Statistical Power

I have not yet looked at the dataset except only to read the paper by [Fearon et al. \(2007\)](#) which reports the sample size of 1,593 leader-years for their African sample. I know that this is a lower bound because a singular look at the observations for Tanzania

shows missing observations during that founding leader Mwalimu Julius Nyerere's entire tenure, which I intend to code. In any case, I take their outcome variable—civil war onset— as indicative of the mean for nation-building or lack thereof so that the mean I assume once my data is complete is 97.42%. I do the following power calculations with further assumptions on the effect size and standard deviations:

- Standard Deviation (40% of Mean)
 - 1% effect size gives me 9.328% power
 - 2% effect size gives me 22.91% power

- Standard Deviation (20% of Mean)
 - 1% effect size gives me 22.91% power
 - 2% effect size gives me 68.13% power

- Standard Deviation (20% of Mean)
 - 1% effect size gives me 68.13% power
 - 2% effect size gives me 99.81% power

which provides some information about how underpowered the study may be with various assumptions. I do not expect to be underpowered as low as the 9.328% but may not be as powered as high as the 99.81% suggests either. The final statistically power may lie between the 68.13% and the 99.81% with expectations of it being close to the latter than the former.

2.1.3 Assignment to Treatment

The source of exogenous variation and the key identifying assumption is that, after controlling for a set of covariates which I will detail in Section 3, a leader's ethnicity and the ethnic fractionalization are exogenously determined and uncorrelated with anything in the error term of the empirical model.

So the treatment variables in this study are not randomly assigned, however, I argue can be viewed as-if randomly assigned after the conditioning set is included, but at the very least provides strongly suggestive correlations about the interaction between a leader's minority ethnicity status and ethnic fractionalization and the decision to choose nation-building over ethnic politics.

2.1.4 Attrition from the Sample

I do not anticipate any attrition from my sample in the classical sense, however, it is important to think about the role that missing data may differentially bias estimates of minority ethnicity status of a leader on nation-building. For instance, because [Posner \(2005\)](#) finds that identifying with one particular ethnicity or another is endogenous to political institutions one may expect to observe leaders identify *less* with their *true* but minority ethnic group for political expediency that affects their ability to choose nation-building or ethnic politics. This will cause the data to attenuate an effect of minority ethnic leaders if they nevertheless act as minority leaders as the present theory suggests and be more likely to implement nation-building policies despite identifying (and accepted by her society) as coming from a non-minority group. The current (2015- present) President of Tanzania, John Pombe Magufuli, for instance identifies, and is accepted, as a *Sukuma*, which is the largest single ethnicity but he is in fact from a smaller ethnicity whose language and customs are similar to the latter are a distinct ethnicity. President Magufuli does this for the political benefits that accrue to politicians by identifying with the larger group. This problem, however, should create a lower bound of the estimates of the effect of the minority ethnicity status of a leader on nation-building.

2.1.5 Data Collection

I expect to complete data collection by November 10th, 2016. The data collection relies on supplementing the [Fearon et al. \(2007\)](#) data where there are missing observations particularly for the variables of interest which include the following:

- a dummy outcome variable equal to 1 if a leader selected nation-building policies

and 0 otherwise

- a dummy explanatory variable equal to 1 if a leader is not from the largest ethnic group and 0 otherwise
- a continuous explanatory variable that is the ethno-linguistic fractionalization measure
- an alternative explanatory variable that is the minimum number of ethnicities that constitute a majority

2.1.6 Data Precession

I hope to finish data processing by November 11th, 2016 and I will host the data on my perosnal website as soon as it is complete.

3 Empirical Analysis

3.1 Variables

Let the names of the variables as explained above be as follows:

- call the dummy outcome variable equal to 1 if a leader selected nation-building policies and 0 otherwise $nation_{ct}$
- call the dummy explanatory variable equal to 1 if a leader is not from the largest ethnic group and 0 otherwise $minority_{ct}$
- call the continuous explanatory variable that is the ethno-linguistic fractionalization measure $frac_{ct}$
- call the alternative explanatory variable that is the minimum number of ethnicities that constitute a majority min_{ct}

where c , l , t indexes countries, leaders, and time, respectively.

The conditioning set includes the following variables:

- a dummy variable equal to 1 if the country experienced prior civil war before and 0 otherwise called $prior_{ct}$
- a continuous variable equal to the country's GDP called $income_{ct}$
- a continuous variable equal to the log of the country's population called $lpop_{ct}$
- a continuous variable equal to country's percentage of mountains called $lmount_{ct}$
- a dummy variable equal to 1 if the country is separated from its neighbors by a water obstacle and 0 otherwise called $noncont_{ct}$
- a dummy variable equal to 1 if the country exports oil and 0 otherwise called oil_{ct}
- a dummy variable equal to 1 for change of POLITY regime in prior 3 years and 0 otherwise called $insta_{ct}$
- a dummy variable equal to 1 if the lagged Polity2 variable is greater than -6 or less than 6 and 0 otherwise called $anoc_{ct}$
- a dummy variable equal to 1 if the country has an Atlantic Ocean coastal border and 0 otherwise called atl_{ct}
- a dummy variable equal to 1 if the country has an Indian Ocean coastal border and 0 otherwise called ind_{ct}
- a continuous variable that equals the total years of education called $educ_{ct}$

which are also in the conditioning set from [Fearon et al. \(2007\)](#), except for the last three, but specifically are included because of the following reasons:

- $prior_{ct}$ to control for a country's prior ethnic conflicts which may be correlated across time to effect the probability of nation-building or lack there of for $t + 1$
- $income_{ct}$ to control for the fact that poorer countries also tend to be more ethnically diverse à la [Fearon et al. \(2007\)](#)

- $lpop_{ct}$ to control for any population threshold effects that may exist for instance for a country that is small but with high salience of ethnicity such as Rwanda that affects nation-building as well
- $lmount_{ct}$ to account for any effect that may include, for instance, that minority ethnicities are more likely to reside in mountaineous regions and therefore a country with a lot of mountains may exhibit more fractionalization in ways that may also affect nation-building choices of its leaders
- $noncont_{ct}$ to account for the effects of being shielded by ethnic conflicts from your neighbors because of the water body boundary
- oil_{ct} to account for the distotinary effects that oil resources may have on political conflict (see [Dunning \(2008\)](#) for a thorough analysis of this)
- $insta_{ct}$ once again to account for prior instability that may be correlated with future instability
- $anoc_{ct}$ similarly as above
- atl_{ct} and ind_{ct} to account for the possible effects of mistrust that stem from the experience of the slavery enterprise as documented by [Nunn and Wantchekon \(2011\)](#)
- edu_{ct} to account for leader quality proxied by highest education attained

3.2 Specification

The specification is a fixed effects OLS regression as follows:

$$\begin{aligned}
 nation_{ct} = & \alpha + \beta minority_{ct} + \xi frac_{ct} + \delta(minority \times frac)_{ct} + \\
 & \sum_{s=1}^S \psi^s \gamma_{ct}^s + \theta_c + \tau_t + \epsilon_{ct}
 \end{aligned} \tag{1}$$

where the variables are as described above with the addition of θ_c and τ_t which are respectively country and time fixed effects. The country fixed effect will pick up any

country-specific variation and unobservable country heterogeneity, such as differences across countries of colonial institutions, while the time fixed effect will pick up time-varying effects on nation-building that are constant across countries such as the Cold War, for instance.

Note that the γ 's are the variables in the conditioning set as described above.

The parameter of interest of course is δ which seeks to identify the impact of having a leader from an ethnic minority in a country that is highly fractionalized on her choice whether to implement nation-building policies.

3.3 Heterogeneous Effects

I expect heterogeneous effects by the following:

- leader quality as proxied by the highest education attained prior to assuming leadership role because although I do not necessarily agree that Nyerere was necessarily more gifted than Kenya's Jomo Kenyatta, the former may certainly be of higher quality than Uganda's Idi Amin, for instance
- the extent to which the country was affected by both the trans-Atlantic slave trade as well as the slave trade that occurred across the Indian ocean in line with [Nunn and Wantchekon \(2011\)](#)'s finding that slavery affected Africans' trust in each other which may plausibly affect contemporaneous ethnic conflicts

3.4 Standard Error Adjustments

Given that the data is presented in a country's leader-years, the least conservative would be to cluster at that level and that would afford me the full 1,593 leader-years (or more given how much missing data I manage to supplement with my own data collection), however the most conservative would be to cluster at the country level which would be at most the 54 African countries. A middle ground would be clustering at the leader level and that is the level I intend to present as the main clustering level. This clustering level is also the most theoretically intuitive given the theory rests on *leaders* choosing

nation-building policies. Nevertheless, for robustness I will still report clustering at the country level in an appendix with the caveat that the estimates will naturally become noisier given the substantial reduction in sample size.

Given that I am only going to look at a singular outcome I do not need to do any adjustment for multiple hypotheses testing. I will, however, do such an adjustment if I do include more than one outcome variable.

4 Research Team

This is a sole-authored study.

5 Deliverables

The main deliverable is a manuscript for publication in an academic journal.

6 Calendar

Data collection and processing to conclude no later than November 11th and data analysis no later than November 12th. A first draft of the manuscript to be delivered no later than November 15th.

Bibliography

- Alesina, A., R. Baqir, and W. Easterly (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics* 114(4).
- Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg (2003). Fractionalization. *Journal of Economic Growth* 8(2), 155–194.
- Alesina, A. and E. La Ferrara (2005). Ethnic diversity economic performance. *Journal of economic literature* 43(3), 762–800.
- Blattman, C. and E. Miguel (2010). Civil War. *Journal of Economic Literature*, 3–57.
- Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miquel (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review* 105(6), 1817–1851.
- Collier, P. and J. W. Gunning (1999). Why has africa grown slowly? *The Journal of Economic Perspectives* 13(3), 3–22.
- Collier, P. and A. Hoeffler (2004). Greed and grievance in civil war. *Oxford economic papers* 56(4), 563–595.
- Dunning, T. (2008). *Crude democracy: Natural resource wealth and political regimes*, Volume 7. Cambridge University Press.
- Easterly, W. (2001). Can institutions resolve ethnic conflict? *Economic Development and Cultural Change* 49(4), 687–706.
- Easterly, W. and R. Levine (1997). Africa’s growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 1203–1250.
- Eifert, B., E. Miguel, and D. N. Posner (2010). Political Competition and Ethnic Identification in Africa. *American Journal of Political Science* 54(2), 494–510.
- Fajardo, A. (2016). Individuals and the Forces of History. *Unpublished Manuscript*.
- Fearon, J. D., K. Kasara, and D. D. Laitin (2007). Ethnic minority rule and civil war onset. *American Political Science Review* 101(01), 187–193.
- Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review* 97(01), 75–90.
- Geddes, B., J. Wright, and E. Frantz (2016). *How Dictatorships Work*. unpublished.
- Habyarimana, J., M. Humphreys, D. Posner, and J. Weinstein (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review* 101(4), 709–725.
- Habyarimana, J., M. Humphreys, D. N. Posner, and J. M. Weinstein (2009). *Coethnicity: Diversity and the Dilemmas of Collective Action: Diversity and the Dilemmas of Collective Action*. Russell Sage Foundation.

- Jones, B. F., B. A. Olken, et al. (2005). Do leaders matter? national leadership and growth since world war ii. *The Quarterly Journal of Economics* 120(3), 835–864.
- Kimenyi, M. S. (2006). Ethnicity, governance and the provision of public goods. *Journal of African Economies* 15(suppl 1), 62–99.
- La Porta, R., F. Lopez-de Silanes, A. Shleifer, and R. Vishny (1999). The quality of government. *Journal of Law, Economics, and Organization* 15(1), 222–279.
- Levitsky, S., J. Loxton, and B. Van Dyck (2015). Introduction: Challenges of party-building in latin america. *Party-Building in Latin America*.
- Magaloni, B. (2006). *Voting for autocracy: Hegemonic party survival and its demise in Mexico*. Cambridge University Press.
- Miguel, E. (2004). Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania. *World Politics* 56(03), 328–362.
- Nunn, N. and L. Wantchekon (2011). The slave trade and the origins of mistrust in africa. *The American Economic Review* 101(7), 3221–3252.
- Posner, D. N. (2005). *Institutions and ethnic politics in Africa*. Cambridge University Press.
- Wimmer, A. (2015). Is Diversity Detrimental? Ethnic Fractionalization, Public Goods Provision, and the Historical Legacies of Stateness. *Comparative Political Studies*, 1–39.