Political Science and Political Economy Working Paper

Department of Government

London School of Economics

Endogenous Ethnicity

Elliott Green (LSE)

THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE ■

Endogenous Ethnicity

Elliott Green[1]
Department of International Development, LSE
e.d.green@lse.ac.uk

This version 23 March 2011

Abstract:

Ethnicity is usually considered an exogenous variable in the field of political economy. Yet there is a growing amount of concern that ethnicity is actually endogenous and not as fixed as is often assumed. Drawing upon a long tradition of social theory leading back to Deutsch, Gellner, Hobsbawm and Marx, among others, here I show for the first time that urbanization contributes to lower levels of ethnic diversity. My results are robust to the use of four different cross-national datasets of ethnic diversity, various control variables, sub-samples, and newly-constructed datasets for Africa and Turkey. I also show how urbanization's effect on ethnic diversity only matters for countries that have yet to "fully urbanize," after which international migration has a much stronger effect.

1. Introduction

Ethnicity has become one of the most important subjects in the field of political economy since the 1990s, with scholars investigating the links between ethnicity and economic growth, civil wars and public goods provision, among other subjects. One of the more consistent findings among these numerous studies has been that ethnic diversity has negative effects on economic and political development.

Yet almost all of these studies contain a very significant flaw, namely the assumption that ethnicity is an exogenous variable unrelated to the dependent variable or other independent variables of interest. This assumption has continued to prevail in the literature despite a long-standing agreement among the numerous "constructivists" located across the social sciences that ethnicity is endogenous to social, political and economic phenomena. Indeed, recent scholarship has begun to endogenize ethnic diversity in explicit ways: (Ahlerup & Olsson, 2009; Michalopoulos, 2008), for instance, find significant correlations between ethnic diversity and state history, land diversity, modern state strength and colonial rule, while within Africa (Nunn, 2008) suggests that ethnic diversity may be a result of the pre-colonial slave trade. However, in none of these or other examples have scholars tested any theories about changes in ethnic diversity over time either in a cross-national or sub-national context.[2]

As a result I use a variety of evidence to show here for the first time that ethnicity is endogenous to the process of urbanization. More specifically, I draw upon a long tradition of social theory and qualitative social science analysis that argues for a strong causal link of increasing levels of urbanization on decreasing levels of ethnic diversity. I first demonstrate a strong correlation between urbanization levels and cross-country estimates of ethnic diversity as measured by (Alesina, Devleeshauwer, Easterly, Kurlat, & Wacziarg, 2003; Annett, 2001; Fearon, 2003). Secondly, I use data on changes in ethnic diversity over time from (Roeder, 2001) to show that a lagged measure of urbanization is strongly correlated with ethnic homogenization, a result

---

[2] (Eifert, Miguel, & Posner, 2010) suggest that Africans may be more likely to identify with their ethnic group (over other identities based on gender, profession or religion) they closer they are surveyed to a presidential election. However, this result still implies a fixed underlying level of ethnic diversity with varying levels of political salience.

which is robust to various control variables, sub-samples and an alternative dataset for Africa. Third, I show that the relationship between urbanization and ethnic diversity changes after countries become "fully urbanized," whereupon it is instead migration that plays a major role in the dynamics of ethnic diversity. Fourth and finally, I use province-level data from mid-20th century Turkey to show that urbanization and ethnic homogenization are correlated at the sub-national level as well.

The paper is organized as follows. Next in section 2 I discuss the use of ethnicity as both an exogenous and endogenous variable, with attention to social theorists like Deutsch, Gellner, Hobsbawm, Marx and others who have argued for a causal effect of urbanization on ethnic homogenization. In section 3 I present a great deal of qualitative material from various parts of the world which demonstrate a link between urbanization and ethnic homogenization. Section 4 presents cross-national quantitative analysis, including the use of three distinct cross-national datasets, various controls, sub-samples and newly constructed datasets for Africa, while Section 5 examines sub-national data from Turkey. Finally, section 6 concludes.

2. Exogenous and Endogenous Ethnicity

"Primordialist" understandings of ethnicity rest upon the assumption that, to take one example, ethnic diversity is an "exogenously determined social state" (Ordeshook & Shvetsova, 1994, p. 108).[3] The amount of literature in the social sciences that makes similar assumptions is far too large to cite here; I merely list some of the more notable examples. For instance, recent studies that assume exogeneity have examined the effects of ethnic diversity on economic growth and development (Easterly & Levine, 1997; Montalvo & Reynal-Querol, 2005a; Spolaore & Wacziarg, 2009), civil wars (Cederman & Girardin, 2007; Collier & Hoeffler, 2004; Fearon & Laitin, 2003; Hegre & Sambanis, 2006; Miguel, Satyanath, & Sergenti, 2004; Montalvo & Reynal-Querol, 2005b; Walter, 2006), public policy and public goods provision (Alesina, Baqir, & Easterly, 1999; Habyarimana, Humphreys, Posner, & Weinstein, 2007; Lieberman, 2007; Miguel & Gugerty, 2005),

---

[3] Cf. (Mauro, 1995, p. 692), who similarly writes that "I assume that the extent to which countries are fractionalized along ethnolinguistic lines is exogenous and unrelated to economic variables."

and political party formation (Clark & Golder, 2006; Harbers, 2010; Ordeshook & Shvetsova, 1994), among other phenomena. Indicative of this literature are (Fearon & Laitin, 2003; Miguel et al., 2004), who record different annual measurements for democracy, GDP and population for all countries in their dataset over a 44 and 20 year period, respectively, but consider ethnic diversity to be fixed over the same time span. In almost all of these studies ethnicity is implicitly assumed to be exogenous to the other variables of interest, with the sole exception of migration,[4] while in the others the authors acknowledge the potential endogeneity of ethnicity but leave its causes and consequences for further investigation.[5]

In contrast, however, a focus on the endogenous nature of ethnicity has a long legacy within the social sciences, in particular in relation to the integrative and assimilationist effects of urbanization on ethnic diversity. For instance, in the *Communist Manifesto* Marx and Engels note how the bourgeoisie "create enormous cities" and "agglomerate population," thereby removing people from the "idiocy of rural life" to a society dominated by "one nation, with one government, one code of laws, one national class interest and one customs-tariff" (Tucker, 1978, p. 477). (Lerner, 1958) similarly proposed that urbanization would lead to higher levels of literacy, communication and political participation, albeit in a nonlinear fashion. And (Hechter, 1978, p. 297) suggested that spatially dispersed populations, such as "peasants in a region of isolated farmsteads," are less likely to develop strong group bonds that those which are more concentrated.

Many scholars have also directly implicated urbanization in the rise of modern nationalism. For instance, Benedict Anderson's claim that "print capitalism" led citizens to create "imagined communities" is contingent upon the existence of a literate urban intelligentsia that could read newspapers and books (which themselves were printed in towns and cities).[6] Eric Hobsbawm also links the rise of nationalism in late 19th-century Europe to urbanization; in particular he argues that urbanization introduced "novel and quite non-traditional classes and strata" that were receptive to nationalist ideologies (Hobsbawm, 1992, p. 109). Thus in places like the Basque country peasants

---

[4] For instance, (Miguel & Gugerty, 2005, p. 2337) assume that levels of ethnic diversity in western Kenya "are largely the product of [pre-colonial] historical accident rather than recent migration," thereby foreclosing other explanations for ethnic diversity.

[5] Cf. (Cederman & Girardin, 2007, p. 176), who note that their use of ethnicity as exogenous in their analysis of civil wars "does not mean that we believe that identities are primordially given."

[6] Indeed, (Anderson, 1991, pp. 79, 133) explicitly links the rise of both early Greek nationalism and the Indonesian nationalist language *bahasa Indonesia* to the presence of a literate urban intelligentsia.

showed much less interest in nationalism than members of the rapidly growing urban lower middle class, where their languages were on the verge of extinction due in part to migration movements tied up with industrialization (Hobsbawm, 1992, p. 119). Finally, John Breuilly similarly links urbanization to nationalism, claiming that urban migrants "need to redefine themselves" to find success in their new environments, and "nationalism can play an important part in this process" (Breuilly, 1993, p. 22).

But perhaps most notable in this regard is Ernest Gellner's theory of nationalism, where industrialization is responsible for the rise of modern national identities. More specifically, inasmuch as industrialization requires an educated and culturally homogenous workforce, migrants to the newly industrialized areas which can assimilate to this new homogenous culture – namely the "Megalomanians" in Gellner's famous story – do so while those that cannot due to cultural differences – a group of people who speak related but distinct "Ruritanian" dialects – react by ignoring their linguistic differences and uniting under a common ideology of nationalism (Gellner, 2006 [1983]).[7] While Gellner does not explicitly conflate industrialization and urbanization, his model nonetheless implicitly ascribes the rise of pan-ethnic identities among both the Megalomanians and Ruritanians to the shift from agrarian society to urban, industrial civilization.

In fact, Gellner's model of assimilation among the Megalomanians is not very different from the role of urbanization in the varieties of modernization theory that were so prominent in the mid-20th century. For instance, within political science Karl Deutsch and others proposed in the 1960s that the creation of new, broad national identities came from the process of modernization, specifically from such phenomena as urbanization, exposure to mass media, greater literacy and economic growth (Deutsch, 1961; Francis, 1968). Since then, however, a good deal of literature has added further evidence that many of these other phenomena are themselves driven by urbanization (Dyson, 2001; McCrone & Cnudde, 1967; Winham, 1970), thereby adding even more importance to urbanization as a causal variable. Indeed, where states impose education in national languages assimilation is more likely to take place, and, at least in the USSR, higher

---

[7] More specifically, it is the Ruritanian urban elite who initiate the growth of a Ruritanian nationalism. For an overview of the role of urban elites in nationalism see (Whitmeyer, 2002).

urban literacy rates historically meant that assimilation was more likely to happen in cities (Harris, 1945).

In addition to modernization theory there are two alternative mechanisms by which urbanization can promote ethnic homogeneity. First, by agglomerating people together urbanization removes the isolation of rural areas and encourages assimilation into larger communities. As such it is the opposite of the isolating effects of high temperature, low latitude, high rainfall and high elevation that have been previously proposed as explaining global levels of ethnic diversity. More specifically, many scholars have suggested that latitude and/or temperature have an inverse relationship with ethnic diversity, in that warm tropical environments are ideal for growing food and thereby create few incentives for inhabitants to migrate elsewhere or trade extensively with other human populations (Ahlerup & Olsson, 2009; Cashdan, 2001; Collard & Foley, 2002; Nichols, 1992; Sutherland, 2003). Other researchers have similarly found a significant positive correlation between ethnic diversity and elevation (Nichols, 1992; Sutherland, 2003) and differential land endowments (Michalopoulos, 2008), both of which also work along the same mechanisms of encouraging isolation.

A second additional theoretical mechanism tying urbanization to ethnic homogenization is through the concept of the "minimum winning coalition," as originally developed by (Riker, 1962) and applied to the theory of ethnic politics more recently by (Chai, 2005; Fearon, 1999; Posner, 2005). More specifically, if ethnic or tribal groups are territorially concentrated and roughly coincide with a set of local government divisions, then most of these local governments will have a majority ethnic group whose members can utilize ethnic identities to take power. In such a scenario there will be no incentive to form broader ethnic identities as any coalition representing a larger ethnic base would therefore have fewer resources *per capita* available for redistribution. On the other hand, due to greater levels of ethnic diversity in cities urban migrants from these territorial ethnic groups would not be able to employ their original ethnic identities in forming majority coalitions. It is thus in their interest to form broader ethnic identities that can capture power in the urban environment, leading to a process of ethnic homogenization.

3. Qualitative Evidence

There is a great deal of qualitative and historical evidence from around the world about the effect of urbanization on ethnic homogenization. To start with Europe, (Weber, 1976, p. 22) famously argued that France became culturally and linguistically homogenous in the late 19th century through urbanization, industrialization and education and the resultant "spread of urban values." In particular urban migrants not only spread urban values back to the countryside and turned "peasants into Frenchmen," but also homogenized cities themselves by "shattering the hold of local speech and lore in the urban centers" (Weber, 1976, p. 290). (Roshwald, 2001) similarly describes the rise of nationalism among the urban intelligentsia of the Austro-Hungarian, Russian and Ottoman Empires in the early 20th century. Finally, in the USSR urban migrants assimilated into Russian society despite Soviet efforts at promoting non-Russian languages and cultures (Harris, 1945), while in Greece ethnic minorities such as the Sarakatsans "tend to drop their identity as Sarakatsans when they leave herding" and migrate to urban areas (Schein, 1975, p. 87).

In the Asia-Pacific region evidence on the integrative effect of urbanization also dates back decades. In Thailand migrants to Bangkok from the northeast found that their local village ties and identities were unhelpful in making sense of the urban environment; in fact, "local differences between the migrant and other North-Easterners became relatively insignificant as he realized his common interests with all who shared the same background" in the city (Keyes, 1966, p. 365). In China there is strong evidence that members of the Bai ethnic minority in Yunan province have retained their ethnic identity in the countryside but shed it in the cities (Wu, 1991). In Papua New Guinea (Levine, 1999) similarly describes how migrants to urban areas abandon their local tribal identities in favor of regional identities or even a "highland" vs. "coastal" dichotomy. Finally, (Chai, 1996) presents evidence of the role of the urban environment in consolidating pan-Malay identities in Malaysia.

In South and South-West Asia, (Khoury, 1983; Özoğlu, 2001) describe the rise of Arab and Kurdish nationalism among the urban intelligentsia of the Middle East in the decades before World War I, respectively, with similar evidence for interwar period as well (Gershoni, 1997). In India there is evidence that attempts by Hindu nationalists to emphasize larger communal identities over caste or ethno-linguistic identities have found much of their support in towns and cities, in particular

through the urban-based Bharatiya Janata Party and the Rashtriya Swayamsevak Sangh (Hansen, 1999; Jaffrelot, 1996). Finally, (Hyman, 2002) argues that one of the reasons nationalism in Afghanistan has had such a minimal political impact is because of its inability to reach beyond the urban elite.

To turn to Africa, (Coleman, 1954, p. 411) argues that urbanization in the mid-20th-century "loosened kinship ties, accelerated social communication between 'detribalized' ethnic groups, and, in general, contributed to 'national' integration." Indeed, as noted by (Freund, 2007, pp. 90-91; Kasfir, 1979, pp. 370-371; Wallerstein, 1960, p. 133), Africans have often assimilated into larger ethnic groups in order to find security and prestige in the difficult urban environment. The Ibo of Nigeria, Jola of Senegal, Duala of Cameroon, Luyia and Mijikenda of Kenya and Bangala of the DRC can thus be seen as classic examples of previously different ethnic groups amalgamating into larger ethnic identities as urban migrants found commonalties among each other and transferred these new identities back to their rural brethren as well (Eckert, 1999; Nugent, 2008; Southall, 1970; Willis, 1993; Young, 1976). The same logic applies to urban migrants who have assimilated into already-existing groups such as northern Nigerian migrants who became Hausa in the southern city of Ibadan or the ethnic minorities who became Batswana in late 20th century Botswana, the country with the highest rate of urbanization of any country in the world since 1950 (Cohen, 1969; Solway, 2004). Finally, (Bryceson, 2010) argues that the cosmopolitan nature of Dar es Salaam promoted the development of a strong Tanzanian national identity in the post-colonial context.

In Latin America, (Blanksten, 1960, p. 470) long ago noted that "the Indian who migrates to the big city becomes a *mestizo* almost in so doing; he has not lived there very long before he is counted as a creole or a 'white'". Indeed, urbanization "contributes to the development of common sets of political attitudes and experiences on the part of the people who live in the growing cities of Latin America… The overall effect of the movement to the cities is integrative" (Blanksten, 1960, p. 475). Evidence for this integrative effect dates back to the colonial period, when ethnic differences by the Spanish upon the indigenous population of Mexico broke down over time in the city of Oaxaca (Chance, 1978). More recently activism among urban intellectuals in Guatemala has led to an attempt to unite the speakers of all twenty-some Maya language groups under a pan-Mayan

identity (Warren, 1998), while in Ecuador Indians continue to assimilate into *mestizo* or *cholo* communities in urban settings (Kyle, 2000).

Finally I turn to North America, where decades of contradictory evidence from the US seems to show that urbanization actually leads to the persistence of ethnic divisions rather than assimilation into larger identities (Dahl, 1961; Wolfinger, 1965). Yet one should remember that this literature has little relevance for my analysis here, for three reasons. First, much of this literature did not disaggregate ethnic persistence according to rural vs. urban areas; in fact, as suggested by (Parenti, 1967), there was very little evidence at the time that ethnic assimilation was any more pronounced in the suburbs or rural areas than in cities. Second, the integrative effects of cities described above all discuss *internal* rather than *international* migrants, with urbanization increasing largely due to people moving from rural to urban areas rather than from one country to another. Third and finally, the US literature focused only on assimilation into a dominant national identity rather than assimilation from smaller ethnic groups to larger ones.

If I instead turn again to evidence from the US and Canada with these points in mind there is a good deal of supporting evidence. Indeed, urbanization among Native Americans, a phenomenon much more comparable to the processes of urbanization and ethnic change described above for the rest of the world, has led to "supratribalism" or the assimilation of disparate Native American tribes into a common Native American, pan-Indian or "Red Power" identity in the late 20th century (Cornell, 1988; Jarvenpa, 1985). Moreover, there is also evidence of the construction of "pan-Indian" alliances and "intertribal bonds" among urban "First Nation" migrants in Canada (Wilson & Peters, 2005, p. 407).

4. Cross-National Quantitative Evidence

If urbanization leads to ethnic homogenization then cross-national data should exhibit this relationship. Indeed, the causal effect of urbanization on nationalism has already been established by (Robinson, 2009) with third-round Afrobarometer survey data from 21,937 respondents across 16 African countries, whereby logistic regressions which include age, gender, education and wealth as controls demonstrate a strong correlation between urbanization and the likelihood of

feeling more strongly attached to one's national identity than one's ethnic identity. I have carried out similar logistic regressions with fourth-round Afrobarometer data from 25,705 respondents across 19 countries and obtained the same robust effect of urbanization on nationalism with a variety of controls (results available from author).

The Afrobarometer survey data are, however, crude inasmuch as urbanization is a dummy variable and the dependent variable only allow for self-identification with either the nation or a single ethnic group. A more profitable exercise would be to regress a measure of ethnic diversity on levels of urbanization in a cross-sectional database in order to see if there is a link between lower levels of diversity and higher levels of urbanization. To do so I employ as my dependent variables three recent datasets that have used the Herfindahl index to calculate ethno-linguistic fractionalization (henceforth ELF) (Alesina et al., 2003; Annett, 2001; Fearon, 2003), as in

$$ELF = 1 - \sum_{i=1}^{i=n}(p_i)^2$$

where p is the percentage of any ethnic group i larger than one percent of the population. Thus, depending upon the dataset, countries like Comoros or South Korea have scores close to 0.0 while countries like Liberia and Uganda have scores above 0.9. I measure urbanization levels in 2000 for (Alesina et al., 2003; Fearon, 2003) and in 1980 for (Annett, 2001) in order to match the date of the datasets while also controlling for variables that previous studies have shown to correlate with ethnic diversity, namely mean elevation, state size and state antiquity (Ahlerup & Olsson, 2009; Green, 2010; Michalopoulos, 2008; Nichols, 1992; Sutherland, 2003). The results, which are presented in Table 1, indicate that urbanization levels are negatively and significantly correlated with levels of ethnic fractionalization across both datasets despite the addition of numerous control variables.[8]

[Insert Table 1 here]

---

[8] The results are robust to the use of additional variables controlling for artificial borders and the slave trade as well as a sub-sample of only former colonies.

This correlation between ethnic diversity and urbanization does not, of course, tell us anything about causality. Moreover, inasmuch as different countries measure urbanization according to vastly different definitions any such static correlation across countries is bound to be inaccurate.[9] Thus using a long-difference regression – where I regress change over two points in time in levels of ethnic diversity on urbanization growth in a cross-sectional dataset – would allow us better insight into the effect of urbanization on ethnic diversity, and would also allow us to control for different definitions for urbanization and other effects like international migration.

As a result I next turn to a dataset from (Roeder, 2001) which estimates country-level ELF scores for 1961 and 1985 based on Soviet data. This data has been criticized previously for overly relying upon a linguistic definition of ethnicity (Laitin & Posner, 2001), but it has been widely cited as an authoritative source on ethnic diversity and remains the only extant dataset that measures changes in ethnicity over time.[10] To obtain a measure of ethnic change over time I thus subtract each country's 1961 ELF score from its 1985 ELF score. I measure urbanization as the annual rate of change in the proportion of people living in urban areas and lag my measure by a decade to account for the time it takes for people to assimilate and alter their identities.[11] I do not measure urbanization as the 1985 level minus the 1961 level both because the annual rate of change in percentage urban is the standard definition of how the urbanization rate is measured and because using mere differences in levels would not take into account initial levels of urbanization. More specifically, while Cameroon and Greece both saw their urbanization levels increase by 18% between 1950 and 1975, the proportion of urban citizens tripled in Cameroon from 9.3% to 27.3% for an average annual growth rate of 4.3%, while in Greece the proportion of urban citizens only grew from 37.3% to 55.3% for a growth rate of 1.6% per year.

---

[9] Current thresholds for the classification of urban areas range from agglomerations of 200 people in Iceland and Norway to 50,000 in Japan and South Korea. Changes in definitions of urbanization do take place over time within countries, such as a reclassification in China for its 2000 census, but these changes are still dwarfed by the cross-country differences.

[10] Examples of authors who have used (Roeder, 2001)'s data include (Harbers, 2010; Lieberman, 2007), while those who have used the 1961 Soviet data as it appeared in (Taylor & Hudson, 1972) include (Easterly & Levine, 1997; Hegre & Sambanis, 2006; Mauro, 1995; Montalvo & Reynal-Querol, 2005b).

[11] I cannot lag the urbanization variable by more than ten years as cross-country data on urbanization only goes back to 1950. All of the results presented here are robust to the use of a 5-year lag as well as well as the separation of the urbanization variable into two variables measuring absolute growth in urban and total population; the signs of the two coefficients are negative and positive, respectively, and both are statistically significant in a variety of specifications.

The dataset is slightly problematic for two reasons. First, a number of countries are missing from the dataset which were included in the (Alesina et al., 2003; Annett, 2001; Fearon, 2003) datasets. The Soviet data does not include, for instance, ELF scores for small countries like Andorra, Belize, Nauru and the Seychelles. More importantly, the dataset also excludes four countries either whose borders changed during the period in question, namely Bangladesh/Pakistan and Vietnam, or which united after 1985, namely Germany and Yemen. However, for four additional countries which split up after 1985, namely Czechoslovakia, Ethiopia, the USSR and Yugoslavia, I was able to compute urbanization rates and other demographic variables based on UN data of their constituent parts.

Second, the data set includes two countries with negative urbanization rates between 1950 and 1975. The first, Trinidad and Tobago, can be explained by its government's decision to use a highly restrictive definition of urbanization that significantly undercounts its actual urban population and leaves it as a notable outlier in the relationship between urbanization and GDP/capita (Satterthwaite, 2007, pp. 33-34). The second, Cambodia, notoriously experienced mass de-urbanization under the Khmer Rouge during the 1970s such that the level of urbanization dropped from 16.0% in 1970 to only 4.4% in 1975. Thus to avoid bias I thus remove these two cases from my analysis.

Having noted these caveats, I control for two other factors that might cause ethnic fractionalization to change over time. If ethnicity is primordial then the only way it will change is due to the entry and exit of people from the population group in question, namely via international migration, which I measure here as net international migration 1960 to 1985 divided by population in 1960. (Inasmuch as variance decreases with population size the migration data suffers from heteroscedasticity, leading me to employ robust standard errors.) I also control for annual change in fertility rates between 1960 and 1985 for two reasons. First, fertility change should affect ELF measurements if changes in these rates are not evenly spread throughout the population (which they never are).[12] Second, there is strong evidence for the causal effects of urbanization on fertility

---

[12] There is a large amount of evidence suggesting differential fertility rates across ethnic groups for a variety of countries. For examples from opposite ends of the global income distribution see (Hout & Goldstein, 1994; Romaniuk, 1980) on the US and the Belgian Congo/DRC, respectively.

decline (Dyson, 2001), and thus by controlling for fertility change I can help to isolate the effects of urbanization here. The descriptive statistics for these variables can be found in Table 2.

[Insert Table 2 here]

My basic regression model is as follows:

$$\Delta ELF_i = \acute{\alpha} + \beta_1 \Delta U_i + \beta_2 \Delta F_i + \beta_3 \Delta M_i + \Delta \acute{\epsilon}_i$$

where ΔELF is change in ELF between 1961 and 1985, ΔU is the annual rate of urbanization growth 1950-1975, ΔF is the annual rate of fertility change 1960-1985 and ΔM is net international migration 1960-1985.[13] My results are presented in Table 3. In column 1 I regress change in ELF on urbanization while controlling for fertility and migration. In column 2 I add additional political economy controls which have been suggested to correlate with ethnic change over time. First, noting evidence that poverty and low economic growth might halt assimilation (as argued by (Gans, 1962) in the context of the United States), I control for annual growth in GDP/capita. (Chandra, 2010) suggests other mechanisms by which ethnic identities change, including violence and democratization; I operationalize the former through the number of years of civil war per country over the time period and the latter through the change in each country's Polity IV score over the period.[14] In column 3 I control for initial levels of ELF in 1961 alongside the other aforementioned variables which have been hypothesized to correlate with levels of ethnic diversity, namely elevation, latitude, and British and Spanish colonialism (Ahlerup & Olsson, 2009; Green, 2010; Michalopoulos, 2008; Nichols, 1992; Sutherland, 2003). In column 4 I control for variables which have been argued to correlate with contemporary urbanization, including annual changes in agricultural production per capita (1961-1985), illiteracy (which I lag 1950-1975 due to the

---

[13] The results do not differ if I include a quadratic term or if I substitute my fertility and migration variables with a more-encompassing variable measuring annual population growth between 1961 and 1985. Controlling for initial levels of urban population or total population (logged) also do not affect the results.

[14] If I instead lag the Polity IV variable by five or ten years I lose 14 and 15 observations, respectively, but urbanization does not lose its significance. I cannot lag GDP growth as I am missing GDP data for most countries in my global sample before 1960. The civil war results do not change if I use a dummy variable for countries which had any years of civil war during the period in question.

unavailability of data from 1960) and crude mortality (1960-1985) (Dyson, 2001; Lerner, 1958; Yuki, 2007). Finally, in column 5 I add a full set of dummy variables for Africa, the Americas, Asia and Europe and former British, French, German, Portuguese and Spanish colonial rule.

[Insert Table 3 here]

In none of these five specifications does urbanization lose its significance despite greatly reduced sample sizes in column 2 and 4, and none of the other variables reaches any notable level of significance, with the exception of migration in column 2. (The weak significance of the GDP/capita growth variable in column 2 disappears if we eliminate the other control variables.) Moreover, the urbanization variable coefficient also does not vary widely, with a median coefficient of -1.4.[15] The data thus suggests that urbanization rates of 2.0%, 4.0% and 8.0% between 1950 and 1975 should have roughly led to drops of 0.03, 0.06 and 0.11 in a given country's ELF score, which is indeed what I find in the cases of Albania, Mongolia and Swaziland, respectively.

To test for robustness I first examine a variety of sub-samples in Table 4. I again control for fertility and migration for sub-samples excluding Africa, the Americas, Asia and Europe in columns 1-4, and in columns 5-6 I exclude former British and French colonies. In none of these regressions does urbanization drop below the 5% level of significance and, while the coefficient now varies more widely, the median is still centered around -1.3.[16]

[Insert Table 4 here]

---

[15] The R-squares across the five specifications are low but this result is by no means unusual with long-differenced or first-differenced data due to the increase in the effects of noise. Indeed, in many cases authors do not even report the R-squares of their long difference models; cf. Tables 2 and 3 in (Acemoglu & Johnson, 2007).

[16] I also used urbanization levels in 1950 as an instrument for subsequent urbanization rates; the resultant 2SLS results (available from the author) are even stronger than the OLS results shown here. A refutability test shows that levels of urbanization do not have a direct effect on subsequent changes in ELF for "fully urbanized" countries above given thresholds (55%, 60% and 65%), corresponding to the result presented below in Table 7. However, there is a question as to whether level of urbanization might have other avenues by which it could affect ethnic change such as through violence, democratization and other factors raised by (Chandra, 2010), which raises question about the validity of the exclusion restriction here.

In Table 5 I examine the effects of urbanization in Sub-Saharan Africa, not only because Africa is the least urbanized and most ethnically diverse continent but also because of a number of studies which show modern processes of economic growth, state formation and war to have very different mechanisms in Africa than elsewhere (Green, 2012; E. A. Henderson, 2008; Nunn & Puga, 2011). In column 1 I regress changes in (Roeder, 2001)'s ELF on urbanization and the two control variables and then control for GDP growth and change in PolityIV scores in column 2.[17] In column 3 I employ an alternative measure of ELF from the two editions of *Black Africa: A Comparative Handbook* (Morrison, Mitchell, & Paden, 1989; Morrison, Mitchell, Paden, & Stevenson, 1972). While the two *Handbooks* list data on ethnicity for all countries in Africa for the years 1967 and 1980, in only fifteen cases was there evidence that the authors of the second edition *Handbook* had updated their data from its first edition.[18] After computing the change in ELF for each of the fifteen countries over this thirteen-year period, I regressed change in ELF on urbanization from 1955 to 1970 alongside the two control variables. (Urbanization data from the UN is only available in 5-year increments so we use the 1955 data as a proxy for 1957 data.)[19] Even in such a small sample urbanization is again negative and significant and does not vary widely in the size of its coefficient.[20]

[Insert Table 5 here]

Finally, I examine whether the effects of urbanization rates on ethnic diversity are contingent upon pre-existing levels of urbanization. More specifically, the data indicates that countries achieve "full urbanization" at a certain point in their development after which urbanization will no longer increase substantially, which (V. Henderson, 2003, p. 65) suggests lies in the range of 60-85% for most countries. Indeed, UN data from 2008 reveals that countries that have reached stable levels of urbanization well below 100% include Austria (at a level of 66-68% urbanized since

---

[17] I use PolityIV scores from 1960 to 1985 as lagging it would leave us with only seven observations.

[18] See this data in more detail in Appendix 1. Note that I did not merely use data that demonstrated a change in a country's ELF score between 1967 and 1980, as ethnic diversity did not change in Togo during this period.

[19] Regressions controlling for GDP growth and PolityIV change leave only five observations each so I do not include them here.

[20] The fifteen countries included in the regressions from Panel B encompass states from all major regions within Africa and from four different colonial empires; a full list along with the data is listed in Appendix 1.

1985), Bahrain (87-89% since 1985), Italy (67-68% urban since 1980), Luxembourg (81-84% since 1985), Oman (72% since 1995), Spain (76-77% since 1995), Switzerland (73-74% since 1990), the United Arab Emirates (78-80% since 1985) and the UK (88-90% since 1980). Moreover, some post-Communist transition countries have reached maximum levels of urbanization at even lower levels: for instance, Azerbaijan peaked at a level of 53.5% urbanized in 1985 while Georgia peaked at 55.1% in 1990 and Slovenia at 50.8% in 2000.

One should therefore not expect urbanization rates to have any effect on ethnic diversity in highly urbanized countries, where it is more probable that ethnic diversity is more endogenous to international migration patterns instead. To test this hypothesis I regress changes in ELF on urbanization in Table 6, first by excluding countries that were more than 55% urban in 1950 in column 1, and then by including only countries above 55% urban in column 2. I then repeat the exercise in columns 3 and 4 with a cut-off of 60% and in columns 5 and 6 with a cut-off of 65%.[21] The results confirm my hypothesis: urbanization is significantly correlated with ethnic homogenization only for countries that have not fully urbanized, whereas for highly urbanized countries migration is instead the sole variable significantly associated with ethnic change. Thus, for instance, Qatar, which was 79.2% urban in 1950, saw a large increase of 0.30 in its level of ethnic diversity between 1960 and 1985. One could surmise that this growth was due to the large influx of migrants from the Middle East and Asia over the same period at a ratio of 1.02 to Qatar's population in 1960, thereby corresponding very well to the median coefficient for migration of 0.35 in the even columns of Table 6. These results thereby confirm the primordialist hypothesis that migration is positively correlated with increasing levels of ethnic diversity, but only in countries where urbanization is no longer a factor.

[Insert Table 6 here]

5. Sub-National Data from Turkey

---

[21] The results are even stronger with a cut-off level of 70% urban (with 130 observations above and 11 observations below the cut-off level). The fourteen countries in the dataset which were 65%+ urbanized in 1950 include Argentina, Australia, Belgium, Denmark, Iceland, Israel, Luxembourg, Malta, New Zealand, Qatar, Singapore, Sweden, United Kingdom and Uruguay.

Ideally one should also be able to observe a relationship between urbanization and ethnic change at a sub-national level, using individual country highest-level administrative units (henceforth provinces) as observations with censuses as data sources. However, such a task is difficult for a number of reasons. First, inasmuch as one would not expect urbanization rates to influence ethnic diversity above a certain level of urbanization, I cannot consider countries with high pre-existing levels of urbanization. Second, a relatively large number of provinces is required to make the statistical exercise meaningful, thereby eliminating most countries in the world which have a small number of provinces.[22] Third, tracking changes over time in ethnic diversity requires the existence of two censuses that disaggregate ethnic diversity by province, while lagging the urbanization variable would require three such censuses. Finally, these censuses should contain provincial-level data on migration to help clarify the role of urbanization.

These requirements eliminate the vast majority of possible case studies from analysis here. First, it is very rare to find provincial ELF data recorded across two censuses: in Latin America, for instance, Mexico is the only country to have data across two censuses (but is unusable here due to very high pre-existing levels of urbanization). Second, even among those countries with a sufficient number of provinces like the Philippines (80 regions since 2006) ELF data is often collected at an even higher regional level (for which there are no local governments). Finally, boundary changes between censuses mean that many provinces are not comparable across time, especially when provinces are consolidated or new provinces are created out of parts of more than one older province.

The only country which I found that can satisfy all of these requirements is Turkey. Inasmuch as the Turkish government recorded data on ethnic groups across its 1935, 1950 and 1965 censuses, I was able to construct an ELF score for each province and for each census.[23] This period corresponds to a shift from one-party rule to a multi-party democracy and Turkey's first democratic change of government in 1950; the 1965 census is also the last Turkish census to

---

[22] As of 2011 only nine countries have more than 40 provinces while just sixteen have more than 30 provinces.
[23] As noted above I exclude groups which comprise less than 1% of the population. Thus, for instance, Adyghe/Circassians are only included in the ELF scores for eleven provinces in 1935, eight in 1950 and three in 1965. Other ethnic groups counted in the 1935 census include Arabs (across seven provinces), Georgians (six provinces) and Pomaks (four provinces), among others.

record ethno-linguistic minorities. While the number of provinces increased over this time span from 57 to 67, the borders between provinces in 1935 did not subsequently alter and thus I was able to reconstruct data on the older provinces for the later censuses.[24] With data on population growth across these censuses I am thus able to indirectly control for the demographic components of population growth, namely fertility, mortality and migration. Moreover, in each census the government recorded the number of citizens per provinces who were born in that province, thereby allowing us to track the non-native population growth as a means to control for migration. Finally, I need only eliminate one province due to high initial levels of urbanization, namely Istanbul (with a urbanization level of 86% in 1935).[25]

Thus in regressions 1-4 of Table 7 I regress change in ELF between 1950 and 1965 on urbanization between 1935 and 1950, thereby allowing for a 15-year lag and avoiding the issue of overlap encountered with the data from (Roeder, 2001) above. In column 1 I control for initial levels of ELF in 1950 and province size, while in column 2 I introduce annual growth in non-native population between 1950 and 1965 and growth in native population between 1950 and 1965 in column 3. In column 4 I control for total population growth between 1950 and 1965. Finally, in column 5 I change the dependent variable to the growth in the percentage of people identified in the censuses as Turkish, while using the same variables as in column 3.

[Insert Table 7 here]

The results only add further evidence in support of my argument: urbanization not only has the right sign but is the only variable to be consistently significantly correlated with change in ELF across columns 1-4. (There is some evidence that population growth, whether in total or just for migrants, has some impact in increasing ELF, while initial levels of ELF are negatively correlated with subsequent change in ELF, but both of these relationships are only weakly significant.) Moreover, in column 5 urbanization is positively and significantly associated with an increase in the

---

[24] The notable exception is the province of Hatay, which joined Turkey in 1939 and is thus not part of the data here. Other new provinces such as Bitlis, Çapakçur, Nevşehir and Tunceli were created between 1940 and 1960 from parts of more than one province but in each case one older province contributed overwhelmingly to the new province, which allowed us to assume that these new provinces were basically split from a single province. In any case the results are robust to the exclusion of these provinces.
[25] The next most-urbanized province was İzmir (48% urban).

percentage of residents identified as Turkish, while initial levels of ELF are positively associated with subsequent change and immigration is negatively associated, both at statistically significant levels.

6. Conclusion

In this paper I have shown that, contrary to large amounts of scholarship, ethnicity is endogenous to the process of urbanization. Based upon a long tradition of social theory and history reaching back to Deutsch, Gellner, Hobsbwam and Marx, among others, I first showed that higher levels of urbanization are correlated with lower levels of ethnic diversity across three recent datasets of global ethnic diversity. Moreover, I also showed that greater levels of urbanization leads to lower ethnic diversity across time with data from a fourth cross-national dataset, and that this result is robust to various control variables, sub-samples and a different cross-national dataset for Africa. I then demonstrated how urbanization's effect on ethnic change only mattered for countries that had yet to "fully urbanize," after which migration had a much stronger effect. Finally, I examined sub-national data from mid-20th century Turkey and found that the same relationship between urbanization and ethnic homogenization hold there as well.

Future work could investigate in more detail the mechanisms through which urbanization leads to ethnic homogenization. For instance, the behind the effects of modernization on ethnic homogenization seem to be mixed: on the one hand, literacy and economic growth are not robustly associated with ethnic change in Table 3, while on the other hand including a variable measuring access to TV (which is correlated with urbanization at a level of 0.454, p=0.000) in regressions with the Afrobarometer data completely removes the effect of urbanization on national vs. ethnic self-identification (results available from author). Secondly, the mechanisms of decreasing isolation would suggest that higher population densities should lead to ethnic homogenization; yet regressions that include population growth as an independent variable (and exclude correlated variables like fertility and migration) fail to add support to this mechanism (results available from author). Finally, the minimal winning coalition mechanism does find some support inasmuch as we can hypothesize that such coalitions would be less likely to form in autocracies where coalitions

could capture power independently of the government, and regressions that restrict the sample to autocracies (with a Polity IV score in 1950 between -6 and -10) do not demonstrate a significant relationship between urbanization and ethnic change (results available from author). However, all of these results need to be supplemented with other empirical data, perhaps including survey data, in order for us to get a better grasp on which mechanism is the most important.

Moreover, there are at least three broader conclusions that I can make from this study. First and most obviously, on an empirical level future political economy work on ethnicity cannot continue to assume that ethnicity is exogenous, especially in regression models that include urbanization as an independent variable. In this sense this paper adds to other recent scholarship that emphasizes the endogeneity of political and economic phenomena previously considered exogenous (see (Przeworski, 2004, pp. 20-21) for an overview). Indeed, at the bare minimum one would hope that the days of using ELF as an instrument for various types of political institutions and phenomena – as in (Mauro, 1995; Miguel & Gugerty, 2005), among others – would come to an end.

Second, this study suggests that the focus within modernization theory on the modernizing effects of demographic change has some substantial plausibility. Modernization theory lost its hold in the social sciences in the 1970s, and more recent studies have attacked its assumptions that modernization necessarily produces democracy (Przeworski & Limongi, 1997). As noted, it is also clear that the mechanisms proposed by modernization theory – namely the influence of growing literacy and economic growth – do not hold for the data.

However, it is possible that modernization theory's emphasis on the causal role of urbanization itself – separate from the influence of increasing literacy or economic growth – may be correct. Indeed, it is often forgotten that modernization theorists such as (Lipset, 1959) not only posited a relationship between economic development and democratization but also suggested that social transformation was dependent on modernization in general, including industrialization and, of course, urbanization. More recent scholarship has suggested that modernization as defined by the demographic transition does influence democratization inasmuch as there is a strong correlation between increasing median age and higher levels of democracy across time and

space (Dyson, 2010). Thus it is plausible that modernization theory has more empirical validity than many have given it credit.

Third and finally, future analysis may wish to address in more detail the policy implications of my analysis. More specifically, if urbanization contributes to lower levels of ethnic diversity then governments should encourage more urbanization if they want to bypass the negative effects of ethnic diversity on economic and political development. This is a markedly different policy prescription than previous attempts to manage the negative effects of ethnic diversity: for instance, (Laitin, 2007) argues that governments should instead create sub-national ethnically distinct or homogenous jurisdictions. He also argues that, "whatever the costs of cultural heterogeneity, the costs of eliminating heterogeneity (at least in the short term) are surely higher," a conclusion that according to my results only holds if the costs of urbanization are higher than the costs of ethnic diversity. Yet a previous concern with the costs of "overurbanization" in the developing world (Timberlake & Kentor, 1983) has been replaced in recent economic scholarship by a much more nuanced approach. In particular, economists now suggest that there is at worst an inconclusive relationship between urbanization and economic growth (V. Henderson, 2003), while others argue for net positive economic effects of urbanization, especially in the developing world (Bertinelli & Black, 2004; Brülhart & Sbergami, 2009). As with other conclusions, however, this proposal remains a topic for further investigation.

Bibliography

Acemoglu, D., & Johnson, S. (2007). Disease and Development: The Effect of Life Expectancy on Economic Growth. *Journal of Political Economy, 115*(6), 925-985.

Ahlerup, P., & Olsson, O. (2009). The Roots of Ethnic Diversity. Department of Economics, Göteborg University.

Alesina, A., Baqir, R., & Easterly, W. (1999). Public Goods and Ethnic Divisions. *Quarterly Journal of Economics, 114*(4), 1243-1284.

Alesina, A., Devleeshauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth, 8*(2), 155-194.

Anderson, B. (1991). *Imagined Communities: Reflections on the Origins and Spread of Nationalism*. London: Verso.

Annett, A. (2001). *Social Fractionalization, Political Instability and the Size of Government*. Washington DC: International Monetary Fund.

Bertinelli, L., & Black, D. (2004). Urbanization and Growth. *Journal of Urban Economics, 56*(1), 80-96.

Blanksten, G. I. (1960). The Politics of Latin America. In G. A. Almond & J. S. Coleman (Eds.), *The Politics of the Developing Areas* (pp. 455-531). Princeton, NJ: Princeton University Press.

Breuilly, J. (1993). *Nationalism and the State*. Manchester: Manchester University Press.

Brülhart, M., & Sbergami, F. (2009). Agglomeration and Growth: Cross-Country Evidence. *Journal of Urban Economics, 65*(1), 48-63.

Bryceson, D. F. (2010). Dar es Salaam as a 'Harbour of Peace' in East Africa: Tracing the Role of Creolized Urban Ethnicity in Nation-State Formation. In J. Beall, B. Guha-Khasnbobis & R. Kanbur (Eds.), *Urbanization and Development: Multidisciplinary Perspectives* (pp. 219-234). Oxford: Oxford University Press.

Cashdan, E. (2001). Ethnic Diversity and Its Environmental Determinants: Effects of Climate, Pathogens and Habitat Diversity. *American Anthropologist, 103*(4), 968-991.

Cederman, L.-E., & Girardin, L. (2007). Beyond Fractionalization: Mapping Ethnicity onto Nationalist Insurgencies. *American Political Science Review, 101*(1), 173-186.

Chai, S.-K. (1996). A Theory of Ethnic Group Boundaries. *Nations and Nationalism, 2*(2), 281-307.

Chai, S.-K. (2005). Predicting Ethnic Boundaries. *European Sociological Review, 21*(4), 375-391.

Chance, J. K. (1978). *Race and Class in Colonial Oaxaca.* Stanford: Stanford University Press.

Chandra, K. (2010). How Ethnic Identities Change. In K. Chandra (Ed.), *Constructivist Theories of Ethnic Politics*: Mimeo, New York University.

Clark, W. R., & Golder, M. (2006). Rehabilitating Duverger's Theory: Testing the Mechanical and Strategic Modifying Effects of Electoral Laws. *Comparative Political Studies, 36*(6), 679-708.

Cohen, A. (1969). *Custom and Politics in Urban Africa: A Study of Hausa Migrants in Yoruba Towns.* London: Routledge.

Coleman, J. S. (1954). Nationalism in Tropical Africa. *American Political Science Review, 48*(2), 404-426.

Collard, I. F., & Foley, R. A. (2002). Latitudinal Patterns and Environmental Determinants of Recent Human Cultural Diversity: Do Humans Follow Biogeographic Rules? *Evolutionary Ecology Research, 4*(3), 371-383.

Collier, P., & Hoeffler, A. (2004). Greed and Grievance in Civil War. *Oxford Economic Papers, 56*(4), 563-595.

Cornell, S. (1988). *The Return of the Native: American Indian Political Resurgence.* Oxford: Oxford University Press.

Dahl, R. (1961). *Who Governs? Democracy and Power in an American City.* New Haven, CT: Yale University Press.

Deutsch, K. (1961). Social Mobilization and Political Development. *American Political Science Review, 55*(3), 493-514.

Dyson, T. (2001). A Partial Theory of World Development: The Neglected Role of the Demographic Transition in the Shaping of Modern Society. *International Journal of Population Geography, 7*(2), 67-90.

Dyson, T. (2010). On the Democratic and Demographic Transitions. Department of International Development, London School of Economics.

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

Eckert, A. (1999). African Rural Entrepreneurs and Labor in the Cameroon Littoral. *Journal of African History, 40*(1), 109-126.

Eifert, B., Miguel, E., & Posner, D. N. (2010). Political Competition and Ethnic Identification in Africa. *American Journal of Political Science, 54*(2), 494-510.

Fearon, J. D. (1999). Why Ethnic Politics and "Pork" Tend to Go Together. Department of Political Science, Stanford University.

Fearon, J. D. (2003). Ethnic and Cultural Diversity by Country. *Journal of Economic Growth, 8*(2), 195-222.

Fearon, J. D., & Laitin, D. D. (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review, 97*(1), 75-90.

Francis, E. K. (1968). The Ethnic Factor in Nation-Building. *Social Forces, 46*(3), 338-346.

Freund, B. (2007). *The African City: A History*. Cambridge: Cambridge University Press.

Gans, H. (1962). *The Urban Villagers*. New York: Free Press.

Gellner, E. (2006 [1983]). *Nations and Nationalism*. Oxford: Blackwell.

Gershoni, I. (1997). Rethinking the Formation of Arab Nationalism in the Middle East, 1920-1945: Old and New Narratives. In I. Gershoni & J. Jankowski (Eds.), *Rethinking Nationalism in the Arab Middle East* (pp. 3-25). New York: Columbia University Press.

Government of Turkey. (1936). *Genel nüfus sayimi, 20 ilktesrin 1935 [Population Census of October 20, 1935]*. Ankara: Office of the Census.

Government of Turkey. (1961). *Genel nüfus sayimi 22 Ekim 1950 [Population Census of October 22, 1950]*. Ankara: Office of the Census.

Government of Turkey. (1969). *Genel Nüfus Sayimi, Nüfusun Sosyal ve Ekonomik Nitelikleri 24.10. 1965 [Social and Economic Characteristics of the Population 1965]*. Ankara: Office of the Census.

Green, E. D. (2010). Explaining African Ethnic Diversity. Mimeo, Development Studies Institute, London School of Economics.

Green, E. D. (2012). On the Size and Shape of African States. *International Studies Quarterly, 56*(forthcoming).

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review, 101*(4), 709-726.

Hansen, T. B. (1999). *The Saffron Wave: Democracy and Hindu Nationalism in Modern India*. Princeton, NJ: Princeton University Press.

Harbers, I. (2010). Decentralization and the Development of Nationalized Party Systems in New Democracies: Evidence from Latin America. *Comparative Political Studies, 43*(5), 606-627.

Harris, C. D. (1945). Ethnic Groups in Cities of the Soviet Union. *Geographical Review, 35*(3), 466-473.

Hechter, M. (1978). Group Formation and the Cultural Division of Labor. *American Journal of Sociology, 84*(2), 293-318.

Hegre, H., & Sambanis, N. (2006). Sensitivity Analysis of Empirical Results on Civil War Onset. *Journal of Conflict Resolution, 50*(4), 508-535.

Henderson, E. A. (2008). Disturbing the Peace: African Warfare, Political Inversion and the Universality of the Democratic Peace Thesis. *British Journal of Political Science, 39*(1), 25-58.

Henderson, V. (2003). The Urbanization Process and Economic Growth: The So-What Question. *Journal of Economic Growth, 8*(1), 47-71.

Heston, A., Summers, R., & Aten, B. (2009). Penn World Table, Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hobsbawm, E. J. (1992). *Nations and Nationalism since 1780: Programme, Myth, Reality* (2nd ed.). Cambridge: Cambridge University Press.

Hout, M., & Goldstein, J. R. (1994). How 4.5 Million Irish Immigrants Became 40 Million Irish Americans: Demographic and Subjective Aspects of the Ethnic Composition of White Americans. *American Sociological Review, 59*(1), 64-82.

Hyman, A. (2002). Nationalism in Afghanistan. *International Journal of Middle East Studies, 34*(2), 299-315.

Jaffrelot, C. (1996). *The Hindu Nationalist Movement and Indian Politics, 1925 to the 1990s*. London: Hurst.

Jarvenpa, R. (1985). The Political Economy and Political Ethnicity of American Indian Adaptations and Identities. *Ethnic and Racial Studies, 8*(1), 29-48.

Kasfir, N. (1979). Explaining Ethnic Political Participation. *World Politics, 31*(3), 365-388.

Keyes, C. (1966). Ethnic Identity and Loyalty of Villagers in Northeastern Thailand. *Asian Survey, 6*(7), 362-369.

Khoury, P. S. (1983). *Urban Notables and Arab Nationalism: The Politics of Damascus 1860-1920*. Cambridge: Cambridge University Press.

Kyle, D. (2000). *Transnational Peasants: Migrations, Networks and Ethnicity in Andean Ecuador*. Baltimore, MD: Johns Hopkins University Press.

Laitin, D. D. (2007). *Nations, States and Violence*. Oxford: Oxford University Press.

Laitin, D. D., & Posner, D. N. (2001). The Implications of Constructivism for Constructing Ethnic Fractionalization Indices. *APSA-CP, 12*, 13-17.

Lerner, D. (1958). *The Passing of Traditional Society: Modernizing the Middle East*. New York: Free Press.

Levine, H. B. (1999). Reconstructing Ethnicity. *Journal of the Royal Anthropological Institute, 5*(2), 165-180.

Lieberman, E. S. (2007). Ethnic Politics, Risk, and Policy-Making: A Cross-National Statistical Analysis of Government Responses to HIV/AIDS. *Comparative Political Studies, 40*(12), 1407-1432.

Lipset, S. M. (1959). Some Social Requisites of Democracy: Economic Development and Political Legitimacy. *American Political Science Review, 53*(1), 69-105.

Mauro, P. (1995). Corruption and Growth. *Quarterly Journal of Economics, 110*(3), 681-712.

McCrone, D. J., & Cnudde, C. F. (1967). Towards a Communications Theory of Democratic Political Development: A Causal Model. *American Political Science Review, 61*(1), 72-79.

Michalopoulos, S. (2008). The Origins of Ethnolinguistic Diversity: Theory and Evidence. Department of Economics, Tufts University.

Miguel, E., & Gugerty, M. K. (2005). Ethnic Diversity, Social Sanctions, and Public Goods in Kenya. *Journal of Public Economics, 89*(11-12), 2325-2368.

Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy, 112*(4), 725-753.

Montalvo, J. G., & Reynal-Querol, M. (2005a). Ethnic Diversity and Economic Development. *Journal of Development Economics, 76*(2), 293-323.

Montalvo, J. G., & Reynal-Querol, M. (2005b). Ethnic Polarization, Potential Conflict, and Civil Wars. *American Economic Review, 95*, 796-816.

Morrison, D. G., Mitchell, R. C., & Paden, J. N. (1989). *Black Africa: A Comparative Handbook* (2nd ed.). London: Macmillan Reference.

Morrison, D. G., Mitchell, R. C., Paden, J. N., & Stevenson, H. M. (1972). *Black Africa: A Comparative Handbook* (1 ed.). New York: Free Press.

Nichols, J. (1992). *Linguistic Diversity in Space and Time.* Chicago: University of Chicago Press.

Nugent, P. (2008). Putting the History Back in Ethnicity: Enslavement, Religion and Cultural Brokerage in the Construction of Mandika/Jola and Ewe/Agotime Identities in West Africa, c. 1650-1930. *Comparative Studies in Society and History, 50*(4), 920-948.

Nunn, N. (2008). The Long Term Effects of Africa's Slave Trade. *Quarterly Journal of Economics, 123*(1), 139-176.

Nunn, N., & Puga, D. (2011). Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics, forthcoming.*

Ordeshook, P. C., & Shvetsova, O. V. (1994). Ethnic Heterogeneity, District Magnitude and the Number of Parties. *American Journal of Political Science, 38*(1), 100-123.

Özoğlu, H. (2001). "Nationalism" and Kurdish Notables in the Late Ottoman-Early Republican Era. *International Journal of Middle East Studies, 33*(3), 383-409.

Parenti, M. (1967). Ethnic Politics and the Persistence of Ethnic Identification. *American Political Science Review, 61*(3), 717-726.

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

Przeworski, A. (2004). Geography vs. Institutions: Were Fortunes Reversed? Department of Politics, New York University.

Przeworski, A., & Limongi, F. (1997). Modernization: Theory and Facts. *World Politics, 49*(2), 155-183.

Riker, W. (1962). *The Theory of Political Coalitions*. New Haven, CT: Yale University Press.

Robinson, A. L. (2009). National versus Ethnic Identity in Africa: State, Group and Individual Level Correlates of National Identification. Afrobarometer Working Paper #112, Michigan State University.

Roeder, P. (2001). Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985. http://weber.ucsd.edu/~proeder/elf.htm.

Romaniuk, A. (1980). Increase in Natural Fertility During the Early Stages of Modernization: Evidence from an African Case Study, Zaire. *Population Studies, 34*(2), 293-310.

Roshwald, A. (2001). *Ethnic Nationalism and the Fall of Empires: Central Europe, Russia and the Middle East, 1914-1923*. London: Routledge.

Satterthwaite, D. (2007). The Transition to a Predominantly Urban World and its Underpinnings. International Institute for Environment and Development, Human Settlements Discussion Paper Series.

Schein, M. D. (1975). When Is an Ethnic Group? Ecology and Class Structure in Northern Greece. *Ethnology, 14*(1), 83-97.

Solway, J. (2004). Reaching the Limits of Universal Citizenship: 'Minority' Struggles in Botswana. In B. Berman, D. Eyoh & W. Kymlicka (Eds.), *Ethnicity and Democracy in Africa* (pp. 128-147). Oxford: James Currey.

Southall, A. (1970). The Illusion of Tribe. In P. C. W. Gutkind (Ed.), *The Passing of Tribal Man* (pp. 28-50). Leiden: Brill.

Spolaore, E., & Wacziarg, R. (2009). The Diffusion of Development. *Quarterly Journal of Economics, 124*(2), 469-529.

Sutherland, W. J. (2003). Parallel Extinction Risk and Global Distribution of Languages and Species. *Nature, 423*(6937), 276-279.

Taylor, C. L., & Hudson, M. C. (1972). *World Handbook of Political and Social Indicators*. Ann Arbor, MI: ICSPR.

Timberlake, M., & Kentor, J. (1983). Economic Dependence, Overurbanization and Economic Growth: A Study of Less Developed Countries. *Sociological Quarterly, 24*(4), 489-507.

Tucker, R. C. (Ed.). (1978). *The Marx-Engels Reader* (2 ed.). New York: Norton.

Wallerstein, I. (1960). Ethnicity and National Integration in West Africa. *Cahiers d'Études Africaines, 1*(3), 129-139.

Walter, B. F. (2006). Information, Uncertainty and the Decision to Secede. *International Organization, 60*(1), 105-135.

Warren, K. B. (1998). *Indigenous Movements and Their Critics: Pan-Maya Activism in Guatemala*. Princeton, NJ: Princeton University Press.

Weber, E. J. (1976). *Peasants into Frenchmen: The Modernization of Rural France, 1870-1914*. Stanford, CA: Stanford University Press.

Whitmeyer, J. M. (2002). Elites and Popular Nationalism. *British Journal of Sociology, 53*(3), 321-341.

Willis, J. (1993). *Mombasa, the Swahili and the Making of the Mijikenda*. Oxford: Clarendon.

Wilson, K., & Peters, E. J. (2005). "You Can Make a Place For It": Remapping Urban First Nations Spaces of Identity. *Environment and Planning D: Society and Space, 23*(3), 395-413.

Winham, G. R. (1970). Political Development and Lerner's Theory: Further Test of a Causal Model. *American Political Science Review, 64*(3), 810-818.

Wolfinger, R. E. (1965). The Development and Persistence of Ethnic Voting. *American Political Science Review, 59*(4), 896-908.

Wu, D. Y.-h. (1991). The Construction of Chinese and Non-Chinese Identities. *Daedalus, 120*(2), 159-179.

Young, M. C. (1976). *The Politics of Cultural Pluralism*. Madison, WI: University of Wisconsin Press.

Yuki, K. (2007). Urbanization, Informal Sector and Development. *Journal of Development Economics, 84*(1), 76-103.

Table 1: Urbanization and Ethnic Diversity

Dependent Variable	Alesina	Fearon	Annett
	(1)	(2)	(3)
Urbanization	-0.310***	-0.366***	-0.449***
	(0.109)	(0.117)	(0.105)
State Antiquity	-0.352***	-0.305***	-0.301***
	(0.097)	(0.103)	(0.095)
Km2 (log)	0.044***	0.055***	0.074***
	(0.016)	(0.017)	(0.016)
Elevation	0.012	0.010	-0.016
	(0.027)	(0.027)	(0.026)
Constant	0.134	0.049	-0.012
	(0.206)	(0.214)	(0.200)
N	96	93	96
Adjusted R^2	0.270	0.292	0.369

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

Table 2: Descriptive Statistics

	N	Mean	St. Deviation	Minimum	Maximum
ELF, 1961	141	0.465	0.278	0	0.909
ELF, 1985	141	0.471	0.278	0	0.922
Change in ELF 1961-1985	141	0.006	0.086	-0.218	0.599
Urbanization, 1950	141	0.296	0.229	0.014	0.994
Urbanization Rate 1950-1975	141	0.022	0.016	0.0002	0.082
Fertility Rate 1960-1985	141	-0.013	0.012	-0.050	0.008
Net Migration 1960-1985	141	0.025	0.212	0.109	2.2

Table 3: Urbanization and Ethnic Change
(Dependent Variable: Change in ELF, 1961-1985 (Roeder, 2001))

	(1)	(2)	(3)	(4)	(5)
Urbanization 1950-1975	-1.339***	-1.053***	-1.393***	-1.440**	-1.472***
	(0.402)	(0.359)	(0.445)	(0.558)	(0.459)
Fertility 1960-1985	-0.235	-0.410	0.018	0.467	-0.518
	(0.608)	(0.469)	(0.662)	(0.842)	(0.670)
Net Migration 1960-1985	0.086	0.599**	0.072	0.061	0.088
	(0.084)	(0.283)	(0.080)	(0.082)	(0.085)
GDP 1960-1985		-0.571*			
		(0.294)			
PolityIV 1960-1985		0.0001			
		(0.001)			
Civil War Years		-0.001			
		(0.001)			
ELF 1961			-0.013		
			(0.029)		
Elevation			-0.008		
			(0.008)		
Absolute Latitude			-0.0001		
			(0.005)		
Km2 (log)			-0.004		
			(0.007)		
British Colony			0.005		
			(0.018)		
Spanish Colony			-0.006		
			(0.016)		
Agricultural Productivity				-0.157	
				(0.792)	
Illiteracy 1950-1975				0.074	
				(0.401)	
Mortality 1960-1985				-0.905	
				(1.098)	
Colonial and Continent Dummies	no	no	no	no	yes
Constant	0.031**	0.036***	0.137*	0.018	-0.016
	(0.014)	(0.013)	(0.082)	(0.027)	(0.037)
N	141	74	125	91	141
R^2	0.127	0.226	0.156	0.132	0.169

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

Table 4: Urbanization and Ethnic Change, Robustness Checks
(Dependent Variable: Change in ELF, 1961-1985 (Roeder, 2001))

	Excluding Africa	Excluding Americas	Excluding Asia	Excluding Europe	Excluding Ex-British Colonies	Excluding Ex-French Colonies
	(1)	(2)	(3)	(4)	(5)	(6)
Urbanization 1950-1975	-2.357***	-1.343***	-1.232***	-1.270***	-1.063**	-1.590***
	(0.862)	(0.442)	(0.380)	(0.410)	(0.435)	(0.515)
Fertility 1960-1985	-0.581	-0.340	-0.133	0.034	-0.120	-0.283
	(0.788)	(0.806)	(0.525)	(0.636)	(0.452)	(0.689)
Net Migration 1960-1985	0.094	0.083	0.049	0.086	0.107	0.087
	(0.090)	(0.083)	(0.139)	(0.083)	(0.202)	(0.087)
Constant	0.038	0.030	0.030	0.030	0.026*	0.038
	(0.021)	(0.016)	(0.014)	(0.015)	(0.014)	(0.017)
N	96	115	107	114	96	117
R^2	0.136	0.135	0.131	0.122	0.074	0.137

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

Table 5: Urbanization and Ethnic Change in Africa

Dependent Variable	Roeder	Roeder	Morrison et al.
	(1)	(2)	(3)
Urbanization Lagged	-1.016**	-1.079**	-1.205**
	(0.408)	(0.442)	(0.433)
Fertility	0.619	0.662	1.265
	(0.634)	(1.192)	(3.872)
Net Migration	-0.117***	0.172	0.281
	(0.027)	(0.159)	(0.600)
PolityIV 1960-1985		0.003	
		(0.002)	
GDP 1960-1985		0.230	
		(0.437)	
Constant	0.029	0.039	0.047
	(0.014)	(0.023)	(0.021)
N	45	18	15
R^2	0.244	0.370	0.262

* $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; robust standard errors in parentheses. The dependent variable in columns 1-2 is change in ELF from 1961 to 1985 (Roeder, 2001); in column 3 it is change in ELF from 1967 to 1980 (Morrison et al., 1989; Morrison et al., 1972). Urbanization is measured from 1950 to 1975 in columns 1-2 and from 1955 to 1970 in column 3. Fertility and Migration are measured from 1960 to 1985 in columns 1-2 and from 1965 to 1980 in column 3.

Table 6: Urbanization and Ethnic Change in Highly Urbanized Countries
(Dependent Variable: Change in ELF, 1961-1985 (Roeder, 2001))

	Excluding 55%+ Urbanized Countries	Only 55%+ Urbanized Countries	Excluding 60%+ Urbanized Countries	Only 60%+ Urbanized Countries	Excluding 65%+ Urbanized Countries	Only 65%+ Urbanized Countries
	(1)	(2)	(3)	(4)	(5)	(6)
Urbanization 1950-1975	-0.870**	12.998	-0.940**	17.068	-1.025**	2.105
	(0.387)	(8.319)	(0.402)	(8.673)	(0.405)	(5.441)
Fertility	0.306	-1.177	0.096	-1.657	-0.414	0.323
	(0.500)	(2.079)	(0.506)	(2.446)	(0.644)	(1.914)
Net Migration	0.001	0.367***	0.001	0.345***	0.050	0.258***
	(0.009)	(0.137)	(0.009)	(0.111)	(0.066)	(0.023)
Constant	0.015	-0.037	0.017	-0.063	0.018	0.044
	(0.014)	(0.078)	(0.014)	(0.081)	(0.014)	(0.042)
N	120	21	122	19	127	14
R^2	0.053	0.508	0.060	0.607	0.075	0.484

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

Table 7: Urbanization and Ethnic Change in Turkey

Dependent Variable	Change in ELF 1950-65	Change in ELF 1950-65	Change in ELF 1950-65	Change in ELF 1950-65	Change in % Turkish 1950-65
	(1)	(2)	(3)	(4)	(5)
Urbanization, 1935-50	-1.934**	-1.953**	-2.013**	-2.132**	0.960**
	(0.960)	(0.929)	(0.978)	(0.995)	(0.469)
ELF 1950	-0.121	-0.135*	-0.144*	-0.137*	0.214***
	(0.074)	(0.074)	(0.077)	(0.076)	(0.047)
Km2 (log)	0.030	0.028	0.027	0.026	-0.011
	(0.018)	(0.018)	(0.017)	(0.017)	(0.008)
Migrant Pop. Growth, 1950-65		0.311*	0.174		-0.273**
		(0.180)	(0.229)		(0.118)
Native Pop. Growth, 1950-65			1.103		0.229
			(0.947)		(0.363)
Population Growth, 1950-65				1.283*	
				(0.686)	
Constant	-0.289*	-0.281*	-0.288*	-0.283*	0.109
	(0.163)	(0.164)	(0.166)	(0.159)	(0.071)
N	56	56	56	56	56
R^2	0.183	0.204	0.222	0.218	0.657

* $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$; robust standard errors in parentheses. All data is from (Government of Turkey, 1936, 1961, 1969).

Appendix 1: African Countries included in Table 5 (Morrison et al.)

Country	ELF (1967)	ELF (1980)
Botswana	0.1900	0.0591
Burkina Faso/Upper Volta	0.7048	0.7049
Burundi	0.2748	0.3276
DRC/Zaire	0.8278	0.8309
Ivory Coast	0.8451	0.8408
Kenya	0.7854	0.7634
Lesotho	0.0975	0.0199
Liberia	0.6622	0.6631
Mali	0.7390	0.7565
Niger	0.7229	0.6521
Rwanda	0.0959	0.1979
Senegal	0.7669	0.7613
Sierra Leone	0.6614	0.6579
Sudan	0.6759	0.7623
Togo	0.7387	0.7387
Mean	0.5859	0.5824
Standard Deviation	0.2712	0.2823

Appendix 2: Data Sources

<u>Dependent Variables</u>

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

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

Morrison: Change in African ELF between 1967 and 1980 as measured by (Morrison et al., 1989; Morrison et al., 1972).

Roeder: Change in ELF between 1961 and 1985 as measured by (Roeder, 2001).

<u>Independent Variables</u>

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

Agricultural Productivity: Annual growth rate in net agricultural production per capita per country, 1961-1985. Source: FAO.

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

ELF 1961: ELF per country as measured in 1961. Source: (Roeder, 2001).

Fertility: Annual change in fertility rates per country. Source: United Nations Population Division, World Population Prospects.

GDP and GDP/capita Growth: GDP and GDP/capita growth per country. Source: (Heston, Summers, & Aten, 2009).

Illiteracy: Annual change in illiteracy rate per country. Source: UNESCO.

Mortality: Annual change in crude death rates per country. Source: United Nations Population Division, World Population Prospects.

Net Migration: Net migration per country over a given period divided by the country's population in 1960. Source: United Nations Population Division, World Population Prospects.

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

Urbanization: Annual growth rate in the percentage of a country's population living in urban areas. Source: United Nations, 'World Urbanization Prospects: The 2007 Revision', Department of Economic and Social Affairs/Population Division (2008).