LAND, HISTORY OR MODERNIZATION?: EXPLAINING ETHNIC FRACTIONALIZATION

Journal:	*Ethnic and Racial Studies*
Manuscript ID:	RERS-2013-0089.R2
Manuscript Type:	Original Manuscript
Keywords:	geoclimatic, ethnic fractionalization, theories of nationalism, ethnogenesis, dominant ethnicity, ethnic group founding date

Abstract

Ethnic fractionalization is frequently used as an explanatory tool in models of economic development, civil war and public goods provision. However, if ethnic fractionalization is endogenous to political and economic change, its utility for further research diminishes. This turns out not to be the case. This paper provides the first comprehensive model of ethnic fractionalization as a dependent variable. It contributes new data on the founding date of the largest ethnic group in each state. It builds political and international variables into the analysis alongside historical and geoclimactic parameters. It extends previous work by testing models of politically-relevant ethnic fractionalization. In addition, this research interprets model results in light of competing theories of nationalism and political change. Results show that cross-national variation in ethnic fractionalization is largely exogenous to modern politico-economic change. However the data are inconclusive with respect to competing geoclimactic, historical institutional and modernist theories of ethnogenesis.

Keywords: ethnic fractionalization, geoclimatic, theories of nationalism, ethnogenesis, dominant ethnicity, founding date

Why are some countries more ethnically diverse than others? This is not a question that has received much attention in the social sciences. Many intuitively cite immigration as a critical factor, yet the foreign-born comprise under 3 percent of the world's population. Hence it is variation in native, or *primary* ethnic diversity (Francis 1976) that lies at the heart of the question. This paper breaks new ground by testing the relative weight played by geoclimatic variation, historical factors and politico-economic modernization in predicting interstate variation in ethnic fractionalization. In so doing, it offers a large-N evaluation of theories of nationalism. It introduces new data on historical institutional and international predictors of ethnic diversity. Finally, expanding beyond previous work on ethnic fractionalization, it employs linguistic, identity-based and politically-relevant ethnic heterogeneity measures.

All contemporary issues of nationalism and ethnic conflict begin with the 'imperfect' overlap between ethnic/national communities and political units. Ethnic fractionalization indices provide a quantitative measure of one manifestation of this: the degree of ethnic pluralism contained under the political roof of each of the world's states. The ethno-linguistic fractionalization index, or ELF, measures the likelihood that any two random individuals in a

state's population are members of the same ethnic group.[1] The greater the number of ethnic groups and the more even their relative size, the more fractionalized the population[2].

An ELF of 0 describes a very homogeneous state and 1 a highly diverse one. A more recent formulation attempts to modify the original 1964 ELF measure by considering the size of the largest and second largest ethnic groups to create a new metric, EF (Fearon and Laitin 2003: 84). This has been mapped in figure 1.

[Figure 1 here]

Critically, cultural diversity in the form of language, as measured by the Ethnologue dataset, for instance, must be distinguished from ethnic diversity, which is based on self-identity measures. Ethnic identity in turn is not coterminous with politically-relevant ethnicity: in some societies, notably in sub-Saharan Africa, ethnicity has an 'onion'-like character, with several different levels, only the highest of which may be politically important (Posner 2005). Similarly, in North America, 'white' is now politically relevant in a way Croatian or German is not. Jewish and Mormon, however, remain politically relevant groups despite their size. In order to account for the imperfection of any single fractionalization measure, this paper utilizes a wide range of measures, seeking to evaluate which factors are most closely associated with different forms of ethnic fractionalization across the world's states.

Incumbent upon the pathbreaking work of Easterly and Levine (1997), Alesina et al. (1999) and Fearon and Laitin (2003), an extensive literature now exists on the relationship between ethnic fractionalization and political and economic outcomes such as economic

growth (Easterly et al. 2006), public goods provision (Banerjee and Somanathan 2007) and violent conflict. In terms of violent conflict, the literature is divided. Studies which take conflict *onset* as the independent variable tend to find no relationship with ethnic fractionalization (i.e. Sambanis 2001; Fearon and Laitin 2003; Collier and Hoeffler 2004; Schneider and Wiesehomeier 2009). Those that focus on the *incidence* of civil war, by contrast, typically report a significant association (i.e. Ellingsen 2000; Montalvo and Reynal-Querol 2005; Urdal 2008).

Few have turned the question around to ask why some states are more fractionalized than others. Such questions are of more than intrinsic interest. They are also important because it is vital to understand the upstream determinants of the fractionalization which may be producing malign political and economic effects. Moreover, identifying fixed or slow-changing correlates of ethnic fractionalization enables scholars to deploy these measures as bedrock independent or instrumental variables in their analyses. On the other hand, if ethnic fractionalization is at the mercy of faster-moving political and economic modernization factors, then it should properly be viewed as endogenous to modern economic and political change - and therefore less important.

Theories of Nationalism and Ethnic Fractionalization

The three main theories of nationalism - primordialism, ethnosymbolism and modernism - offer competing explanations for the existence of ethnic diversity (Özkirimili 2010). Primordialism locates ethnicity in universal aspects of human psychology. For primordialists, our evolutionary psychology represents the successful adaptation of humanity to conditions obtaining during prehistoric time. Primary among these is geography. Geoclimatic isolation of people produces cultural and genetic drift. Genetic distance creates

divergent genetic interests which are the basis for group boundaries and conflict. Phenotype is used where available as an indicator of genetic difference but because neighbouring groups rarely look different, culture is activated as a marker of genetic boundaries when groups leave their ecologies and come into contact with others (Van den Berghe 2002; Pinker 2011: 353-55).

In effect, ethnic boundaries based on genetic and cultural difference preserve the variations incubated in ecological niches. In this manner, geoclimatic explanations for variations in ethnic fractionalization are most closely related to primordialism. An alternative 'geoculturalist' interpretation would claim that geography shapes ethnogenesis via cultural diversity and its usefulness as a group marker, but that this diversity is subsequently amenable to being eroded and reshaped by social processes (Cavalli-Sforza 2001). This second formulation would predict that while traces of geography's effects may remain in ethnic fractionalization patterns, the two will be more loosely connected than if primordialist assumptions hold.

Critics of primordialism suggest that kin-selection impulses are deflected toward constructed forms of community like teams, religious groups or political nations (Brigandt 2001). They also reject the primordialist position that ethnicity can exist in small groups, averring that by definition, ethnic communities must involve a larger scale of human community. Bonds therefore need to be culturally imagined rather than merely experienced in the form of face-to-face *gemeinschaft* relationships (Anderson 1983). The ethnosymbolist school, for example, concurs with primordialists that ethnic groups predate the modern era, but emphasizes the importance of political and cultural institutions rather than geography. Ethnosymbolists claim that ethnic groups do not arise until the late neolithic period when writing, religion, recorded history and extra-local mobilization allowed for the formation of communities knit together by ‘imagined’ bonds of territory, memory and ancestry (Smith

1986: 44-5). Often ethnic consciousness remained the preserve of a small elite, as with the Anglo-Saxon English consciousness of the Venerable Bede and King Alfred (Hastings 1997: 35-9). Some ethnic groups (i.e. Jews, Amhara, Armenians, Persians) had ancient origins, while many more emerged in the medieval and early modern periods through tribal confederation (i.e. Arabs, Kurds), conquest agglomeration (i.e. Gothic founders of Spain) or dynastic competition (i.e. Scots, Catalans). In all cases, territorial identities extending beyond the locale came to be established (Smith 1986; Armstrong 1982).

There are two major forms of ethnicity, according to Francis' (1976: 6) schema: *primary* ethnicity, in which members of the group occupy their ancestral 'homeland' territory; and *secondary* ethnicity, whereby groups acknowledge that they are diaspora and not native, and thus their homeland lies elsewhere. Since immigrants form just 2.7 percent of the world's population, it is primary ethnic fractionalization which is generally captured by ELF and is by far the most important form of fractionalization when it comes to economic development and conflict (Demeny and McNicoll 2006, ch.1). Most of the premodern entities studied by ethnosymbolists are primary ethnic groups, even if they sometimes spawn secondary offshoots like the Jewish, Parsee and Armenian diasporas.

The appearance of translations of religious texts such as the bible into vernacular languages, and the numerous recorded premodern references to *natio*, *gens* and *ethnos* is cited in favour of the theory (Hastings 1997). This explains why historians of the medieval period tend to be ethnosymbolists (Zimmer and Scales 2005). This argument predicts that ethnic identities, once formed, are highly path-dependent and durable. Ethnic sentiments are reproduced by both state and vernacular institutions. The vernacular rootedness of ethnicity means that it is capable of inspiring collective action and resisting 'official' political and identity constructs imposed by the state. Though more culturalist than materialist in

orientation, ethnosymbolist theory nests most comfortably within historical institutionalism, evincing greater skepticism of explanations tied to geoclimatic or modernizing factors.

The *modernist* account contests the ethnosymbolist and primordialist version of events. It argues that premodern identities were strictly local – for the peasant masses, or imperial-religious for military and religious elites (Anderson [1983] 1991; Gellner 1983; Giddens [1985] 1996). Modernity fractures the horizontal ties between cosmopolitan elites, as with Latin Christendom, which fragmented into nation-states with their own vernacular languages. Beneath them, locals were 'invited into history', and came to be connected to wider, self-conscious territorial communities (Nairn 1977). Print capitalism, mass conscription, mass education, secularization and more intensive transport networks combine to orient local identities toward a common, this-worldly community. Pre-existing cultures are orthogonal to this process, which is driven by political and economic imperatives. 'Nationalism, which sometimes takes preexisting cultures and turns them into nations, sometimes invents them, and often obliterates preexisting cultures: that is a reality,' writes Eric Hobsbawm (1990: 10).

Beyond the majority ethnic identity - created by the new nation-state - shared ethnic identities are forged in modern times through states' internal administrative boundary marking activity. This reflexive demarcation institutionalizes ethnic diversity, as with Soviet Republics or colonial administrative departments. Divide-and-rule policies by imperial rulers and missionaries contribute to the process (Trevor-Roper 1983; Brass 1991; Brubaker 1996; Wimmer 2002). Anti-state mobilization by political entrepreneurs, often driven by the imperative to control important industrial resources such as oil, is another vector of ethnogenesis. Ethnic entrepeneurs may have experienced blocked upward mobility within central state structures (Gellner 1983), or may use ethnic and national movements as a vehicle to acquire more power or wealth than they might through conventional political

channels (Breuilly 1993). Modern processes, not geography and premodern history, are responsible for spawning ethnic diversity.

Existing Work on the Genesis of Ethnic Fractionalization

Work on ethnic diversity as a dependent variable is in its early stages. The best developed line of inquiry concerns geoclimatic predictors. Thomas Sowell remarks that sub-Saharan Africa's lack of navigable rivers, smooth shallow coasts and mesa-like terrain has left a legacy of exceptional linguistic diversity (Sowell 2011: 317-18). Others find that the wider the diversity in land quality and topography in a territory, the greater the ethnic fractionalization (Michalopoulos 2012; Sutherland 2003). Ahlerup and Olsson (2011) add that an early incidence of initial prehistoric human settlement, together with geoclimatic factors, predicts enhanced diversity. States far from mankind's East African origins, such as Sweden, were settled later than equatorial regions, and hence possess less ethnic diversity. Laitin and Robinson (2011) also advance a geoclimatic argument, applying Jared Diamond's continental axis theory to individual states. They uncover some evidence that linguistic diversity is greater in states characterized by a North-South cartographic skew as compared with those which spread in a more East-West direction.

Historical institutional factors feature in work with the State Antiquity dataset (Bockstette et al. 2002). This uncovers a significant negative association between ELF and the date of initial state formation coupled with the degree of indigenous control of the state in the ensuing period. The logic is that older states, and those where the indigenous population had greater political control, could spread their culture and identity and are therefore less diverse than newer states - or those ruled by foreigners. However the connection between state history and ethnic fractionalization is a bivariate finding since this relationship was not

the primary focus of the authors. Fletcher and Iyigun (2009) claim a higher incidence of Muslim-Christian conflict between 1400 and 1900 predicts lower ethnic fractionalization in today's European and Middle Eastern states while those which experienced Protestant-Catholic conflict or anti-Jewish pogroms are more fractionalized. Nunn (2008: 164), also working from a historical institutional perspective, considers the role of four historic slave trades in producing ethnolinguistic diversity in Africa. The internal tribe-on-tribe raiding that characterized African slave economies is linked to weaker precolonial states and, by extension, more ethnic fractionalization (Acemoglu and Robinson 2012: 87).

Modernist approaches to this subject are in their infancy. Green (2013), using Philip Roeder's 1961 and 1985 datasets (Roeder 2001), contends that urbanization in postcolonial states in Africa during 1961-85 is associated with declining levels of ethnic diversity. This echoes qualitative work which notes the presence of ethnic fusion in modernizing locations such as the colonial Zambian Copperbelt settlements of the 1930s where miners from formerly distinct groups amalgamated into larger ethnic entities based on cultural relatedness (Eriksen 1993: 20-21). This paper breaks new ground by adopting a comprehensive approach that incorporates geoclimatic diversity, historical indicators, modernization and international determinants. Finally, expanding beyond previous work, it encompasses linguistic, identity-based and politically-relevant ethnic fractionalization measures.

Geoclimatic Variation and Ethnic Fractionalization

The three major theories of nationalism make different predictions regarding the relationship between geography and ethnic fractionalization. Primordialists would view extreme ecological diversity - as in the New Guinea case - as diversity-enhancing. Yet for ethnosymbolism, extreme isolation produces sub-ethnic localism, impeding imagined community. This localism produces disorganization, reducing resistance to the modern state

when it eventually penetrates the periphery, and renders the task of nation-building easier than might be the case in a situation where larger ethnic groups have mobilized. However, above a certain threshold - perhaps several thousand in population - an ethnosymbolist would grant that difficult terrain acts to increase the number of competing premodern polities. Rival ethnic identities can take root so long as there are literate intellectuals and institutions that help spread myths, symbols and memories beyond the local. Terrain that permits this mobilization while preventing wider integration is optimal in producing ethnic fragmentation. Thus more challenging terrain would be expected to lead to greater ethnic heterogeneity, but less so than in the primordialist case.

For modernists, varied terrain acts as a barrier to the state in its quest to homogenize populations and facilitates secessionists' strategy of escaping to peripheral redoubts from which they can construct their interest-reinforcing ethnic projects. For Horowitz, ethnic fusion tends to occur with political amalgamation; fission with political division (Horowitz 1975: 139-40). Meanwhile, variegated terrain hampers the networks of coordination (Laitin 2007) which incentivize participants to join an imagined community. Though ecological variation is associated with greater ethnic fractionalization in all three theories, this relationship would be expected to be stronger under assumptions of primordialism than for competing theories. We can test for this by examining the relationship between a country's geoclimatic diversity and its ethnic diversity.

Thus our first hypothesis:

H_1: Geoclimatic variation is associated with ethnic fractionalization. The stronger the relationship, the greater the support for primordialist theories

Historical Institutions and Ethnic Fractionalization

According to ethnosymbolist theory, modern nations typically form around premodern dominant ethnic groups (Smith 1986). In most cases, the dominant ethnic group is also largest because popular sovereignty and democratization spread their influence down the social scale and render exclusive dominant minorities like Syria's Alawis rare (Kaufmann and Haklai 2008). Therefore we may approach the question of ethnic absorption through measures of ethnic and state antiquity. The State Antiquity dataset asks when a polity above the tribal level was founded on the territory of an existing state; whether this was indigenously-controlled or foreign; and further, what proportion of the territory of the present-day state was under native rule. This is determined for every 50-year period since 1 A.D. Different rates of discounting past periods are applied by the authors, with the most common measure being a 5 percent discount every 50 years (Putterman 2007).

An alternative approach is to attempt to code the founding date of the largest ethnic group in a state. This serves as a measure of ethnic absorption because older dominant ethnic groups will have had more time in which to assimilate neighbouring or subaltern groups than newer groups. Dominant ethnic groups frequently emerge as assimilationist actors - fractionalization-reducing nuclei - within multi-ethnic states. Connor (1994a: 96) notes that homogeneous nation-states occur in less than 10 percent of the world, but that a substantial majority of states contain an ethnic majority. All but five of 156 countries in Vanhanen's (1999) dataset feature a plurality group of a third or more of the population. In other words, some form of ethnic dominance appears to be nearly universal (Kaufmann 2004).

Modernists would be somewhat more circumspect. They would explain ethnic homogeneity as a result of nation-building and the ethnic exclusions practiced by modern states (Wimmer 2002). States established earlier in the modern period would be expected to contain less ethnic diversity than more recent states, but the age of premodern *ethnic* groups should bear no relation to contemporary fractionalization after controlling for the age of the modern state.

How to measure the founding date of the largest ethnic group? This is nowhere near as straightforward as the founding date of states. In this paper, ethnic founding dates are operationalized as the *first imagining of the group by a putative member of the group*. The Ethnic Plurality Group Founding Dates dataset has been developed through a survey of historians and social scientists with expertise on particular countries. Accepting the 'reality' of these ethnic founding dates does not entail embracing the ethnosymbolist perspective. Premodern ethnic imaginings are explained by modernists as the musings of individuals - with no consequences for mass social and political behaviour. Primordialists, too, would consider these visions to be subsidiary to spontaneous collective nepotism in the process of ethnic fusion.

Therefore, an ethnosymbolist would expect that:

H_2: States with a plurality ethnic group that is comparatively old will have lower degrees of ethnic fractionalization; and

H_3: States that have an older tradition of indigenous control will have lower degrees of ethnic fractionalization

Modernists would qualify this as follows:

H_4: Older modern states (defined as post-1789 phenomena) will have lower degrees of ethnic fractionalization, but pre-1789 ethnic plurality founding dates or state antiquity scores should not affect ethnic fractionalization net of the age of the modern state

Primordialism grounds its claims in the ethnic substratum. It is skeptical of states' ability to fuse ethnic groups together, even in the long run. It treats ethnicity as being of prehistoric provenance. Therefore neither ethnic nor state age should affect fractionalization measures.

Modernists locate the source of ethnic fragmentation in modern political and economic factors. States with an abundance of exportable primary commodities such as oil are more likely to experience rent-seeking and ethnic entrepreneurialism and consequently will be more ethnically fractionalized than others. States experiencing political instability in the form of a transition between autocratic and democratic governance are more likely to offer opportunities to political entrepreneurs. Democracy, however, once attained, should permit dissent to be expressed through the state rather than via extra-statal insurgent movements, lowering fractionalization. Finally, urbanization and income per capita are indicators of the intensity of modernity in a state. Thus modernists would predict that:

H_5: States with higher levels of urbanization, democracy and income per capita should be less ethnically fractionalized; and

$H_{6:}$ States whose economy is based on exportable primary commodities and/or those undergoing political instability should be more ethnically fractionalized

Data

Dependent Variables

The dependent variable consists of seven distinct measures of ethnic fractionalization. These cover a range, from those which tap cultural diversity (but which may not be salient for identity) to those which measure politically-salient constructs (which may consist of supra-ethnic amalgams). NUMBRLANG, the number of languages in a country as derived from Ethnologue (Michalopoulos 2012), or ELF, the ethnolinguistic fractionalization index as derived from the 1964 Soviet ethnographic atlas, lie on the cultural side of the spectrum. ELFPREG lies at the political end, and measures a country's politically-relevant ethnic fractionalization index. Homogeneous countries deemed not to have politicized ethnic divisions are excluded from the analysis. MAXPOP is the share of the state's population made up the largest politically-relevant ethnic group (Wimmer et. al. 2009). In between the culturalist and political measures lie those that focus more squarely on ethnic identity: PLURAL (Fearon and Laitin 2003) and PCTMAJ (Vanhanen 1999) are measures of the share of the population made up of the largest ethnic group. EF, the fractionalization measure used by Fearon and Laitin (2003), combines information on the size of the largest and second largest ethnic groups, with data on the total number of linguistic groups exceeding 1 percent of the population.

Independent Variables

A full list of variables used in the analysis, with data sources and frequencies, appears in appendices 1 and 2.[3] Geoclimatic data covers mean temperature, mean elevation, mean precipitation, distance to the sea, standard deviation of mean elevation, mean agricultural suitability and standard deviation of mean agricultural suitability. Historical institutional variables include state antiquity, origin date of largest ethnic group, date of transition to agriculture, log of population density in 1500, and historic slave exports per capita.[4] Data on the founding date of the largest ethnic group in each state has been collected through our British Academy-funded survey of experts, supplemented with textual sources. Methodology, questionnaire and detailed response data for the Ethnic Plurality Group Founding Dates dataset may be found at: www.sneps.net/ethnic/ethnicdates.htm.

Modernization variables encompass state founding date, political instability, proportion urbanized, democracy (Polity IV) score, GDP per capita, oil output per capita, commodity exports per capita and infant mortality rate. In addition there are a series of parameters which do not easily fit one of the major theories. These include population density, which could stem from geoclimatic, historical or modern influences, land area, a dummy variable for ex-colonies, wave of state formation, world region and number of historic secessions. We also code for states which emerged from successful ethnic national self-determination movements as well as states with longstanding or current dominant minorities (see Appendix 3)[5], refining the data through discussion with Anthony Smith, a leading expert on the history and sociology of ethnonationalist movements.

Results

Results are shown in table 1. We are restricted to cross-sectional models due to the absence of time-series data on ethnic fractionalization. Many independent variables are also unavailable over time. Most datasets ignore smaller island states, reducing the universe of cases. Finally, gaps in the data result in a small degree of listwise deletion. The number of cases (112 to 136 depending on the specification) limits our degrees of freedom, thus the need for a iterative approach beginning with strictly geoclimatic, historical and modernist specifications to reduce the number of parameters before bringing the most promising variables together into a combined Ordinary Least Squares (OLS) analysis using Stata 7.0.[6]

[Table 1 here]

This yields the results in table 1. This model predicts over 60 percent of the variation in the four identity-based ethnic fractionalization measures. It performs less well when the dependent variable is pure linguistic diversity or politically-relevant ethnic fractionalization, but still captures half or more of the variation. Of geoclimatic predictors, higher and drier countries contain less diversity. Those with greater variation in elevation and soil productivity are more fractionalized. As the dependent variable moves from language to identity to politically-organized groups, the power of geoclimatic parameters weakens. In our two models of politically-relevant ethnic fractionalization (ELFPREG and MAXPOP), no geoclimatic predictors remain significant.

One should not overplay the importance of this finding. Of non-geographic predictors, only oil output per capita and the sub-Saharan Africa dummy are significantly associated with politically-relevant ethnic fractionalization. Moreover, variation in elevation and soil

type approach significance and relationships are signed in the expected direction in these models. One must bear in mind, however, that homogeneous countries are deemed not 'relevant' in ethnic terms on the PREG measure: the 22 states excluded in this way tend to be highly homogeneous, such as Denmark, or, in a few instances such as Tanzania, extremely heterogeneous. Since the average excluded state has an 81 percent ethnic majority as against 67 percent for included states, this weakens the predictive power of parameters which distinguish highly homogeneous states from moderately diverse ones. Furthermore, politically-relevant ethnic group measures represent an aggregation of ethnic groups into politically-relevant entities on the national stage which in some cases may be considered pan-ethnic rather than ethnic. Thus it taps processes of supra-ethnic amalgamation as well as ethnic diversity.

The combined model delivers a verdict on H_2 ,H_3 and H_4. First of all, while the antiquity of civilization and indigenous control is correlated with ELF and EF, table 1 shows that this bivariate relationship washes out with the addition of other parameters to the model, disconfirming H_3. H_2, concerning plurality ethnic group antiquity, cannot be dismissed as summarily. While the significance of plurality ethnic group founding date for fractionalization disappears in multivariate analysis, it exhibits nonlinearity which can best be captured by a variable for states whose largest ethnic group was founded between 0 and 1100 A.D.[7] The most likely explanation for this pattern is that the period from 0-1100 A.D. was one in which some of the earliest continuous 'ethnic states' (Smith 1986), with elite myths of descent and cultural codes, were formed. It encompasses many states which occupy lands captured during the Sunni Arab conquests of the 7th-11th centuries. In addition, a number of durable East Asian kingdoms arose at this time and many West European states emerged out of the Germanic barbarian successor dynasties which replaced the western Roman Empire (see figure 2).

[Figure 2 here]

This raises the question posed by H_4, namely, is the founding date of the largest ethnic group or that of the state more central in predicting the four identity-based fractionalization measures (ELF, EF, PLURALITY, PCTMAJ)? The two are close, but state date proved a somewhat more powerful predictor than ethnic date in two of three models, confirming H_4.

Confirming, but with caveats: one reason to be skeptical of the superiority of the modernist explanation based on state founding date rather than the ethnosymbolist (ethnic) date measure is that reverse causation is a stronger possibility for state founding date. Namely, it is highly plausible that a more ethnically fractionalized territory might hinder the formation of a modern state. By contrast, fractionalization prior to state formation is unlikely to affect the founding date of the ethnic group that emerges as the largest in the state.

Modernist theory is more unambiguously supported by the predictive power of oil output per capita, lending support to H_6. This is a significant parameter in four of seven models, backing 'greed'- based modernist arguments based on ethnic entrepreneurialism (Collier and Hoeffler 1994). Structural modernization variables offer a mixed picture: democracy enters just one of seven models and GDP per capita none. Infant mortality rate (not shown) did not approach significance in any specifications. We therefore find little support for H_5.

Modern political and economic dynamics count for more when it comes to reducing ethnic fractionalization over time. The founding date of the state is significant in several specifications, suggesting that national integration is an important solvent of ethnic bonds. It

is also vital to appreciate that this is an analysis of variation between countries rather than time points. The static nature of the dependent variable biases the data against faster-moving predictors, hence these results do not negate the importance of modernizing processes in reducing diversity in time. Predictors of variation between countries at one point in time are often different from those which predict variation within countries over time (Kittel 2001: 233).

As noted, there is qualified support here for historical institutionalist arguments. The sub-Saharan Africa dummy, the only variable to perform well across all models, partitions fairly evenly into effects related to the volume of historic slave exports per capita and regional effects unrelated to the legacy of slavery.[8] Along with ethnic group founding date effects, this thereby lends some credence to ethnosymbolic approaches. All told, the results are inconclusive in arbitrating between ethnosymbolism and modernism, though both underperform primordialism in the sense that geoclimatic variables are more powerfully associated with all but the politically-relevant measures of ethnic fractionalization.

Other variables, not readily assignable to the three major theories of nationalism, figure prominently in the combined model. Population density in 1995, which springs from geoclimatic, historical institutional and modernist sources, is significant in five of seven specifications, though its sign changes when the dependent variable moves from linguistic to identity-based measures of fractionalization. In general, denser populations are associated with less ethnic fractionalization. This effect persists with a control for population density in 1500 indicating that more recent variation in population growth may underpin this correlation.[9] Larger territories are expected to contain more groups, and this is borne out in the data: land area is significant in three models and signed in the expected direction in all seven.

Among other variables unrelated to major theories of nationalism, neither a state's number of historic secessions nor its rule by a historic dominant minority were associated with ethnic fractionalization. This may be because more fractionalized states such as Russia/USSR and Serbia/Yugoslavia are more apt to experience secession such that the two effects negate each other. Finally, the historic era in which states were created is important. Those which formed prior to the Congress of Vienna in 1815 are most homogeneous, while those emerging during the period of decolonization are most diverse. Importantly however, a term capturing whether a country is an ex-European colony did not reach significance in the model in table 1, though it was important in several more restricted specifications.[10] This questions the general wisdom that colonization, by running roughshod over ethnic boundaries, is primarily responsible for sub-Saharan Africa's high ethnic fractionalization. It seems the ethnically fractionalizing inheritance of slavery plays a more important role in this development, as Nunn (2008) and Acemoglu and Robinson (2012) surmise.

Finally, states which emerged on the back of nationalist movements defined in ethnic terms (see appendix 3) are somewhat more homogeneous: ethnic nationalism reaches significance in two of seven models, though it is signed in the expected direction in all. This intimates that successful ethnic nationalists partially attain their homogeneous utopias. Their failure to fully realize their dreams probably stems from the fact that most, i.e. Lebanese Christians, Romanians or Ulster Protestants in the interwar period, bear few qualms about annexing territory populated by other groups (Brubaker 1996).

Discussion

Why are some countries more ethnically diverse than others? This paper weighs explanations based on geoclimatic, historical and politico-economic factors and concludes that geoclimatic

diversity is the most important predictor of international differences in ethnic fractionalization. States with a greater difference between their highest and lowest points are significantly more diverse than others. The same holds for those with a wider range of soil types. Higher and drier countries are more homogeneous than low, wet ones. Overall, geoclimatic variables are more strongly associated with ethnic fractionalization than premodern historical or modern politico-economic predictors. This speaks to the importance of interpretations which hold that ecological diversity lays the basis for linguistic and ethnic diversity, as exemplified by the case of Papua New Guinea.

This primordialist thesis must be qualified, however, by noting that geoclimatic variables are not strongly associated with politically-relevant ethnic diversity. This is an important corrective to the existing literature on geographic determinants. Ethnosymbolic legacies from the premodern period also have an important bearing on ethnic fractionalization, though less so than geography. Ethnic group founding date is a weak (inverse) predictor of ethnic diversity, probably because older groups have had longer to assimilate proximal neighbours. Moreover, their patina of age confers prestige. This variable is not linear, however: plurality ethnic groups formed in the period between 0 and 1100 A.D. are associated with highly homogeneous states while ancient plurality groups and those formed after 1100 are located in more diverse ones. In terms of historical factors, this study corroborates the claims of Nunn (2008) and Acemoglu and Robinson (2012) that African states with a history of slave exports are more fractionalized than other countries. There is modest evidence that a late transition to agriculture is associated with greater ethnic fractionalization.

Most modern economic and political variables are not associated with ethnic fractionalization. Cross-national differences in urbanization, health and income had little or no predictive power. Of all modernization variables, only the founding date of the modern

state proved robustly associated with ethnic fractionalization. Modernist explanations, based on the competition for lootable resources driving ethnogenesis, receive more support from this study. Oil exports per capita are strongly associated with ethnic fractionalization, and one of the few consistently significant predictors of politically-relevant ethnic diversity. Per capita commodity exports is significantly associated with some fractionalization measures in a few models. On the other hand, political instability fails to predict diversity in any model.

'Every 14 days a language dies,' claims the National Geographic's *Enduring Voices Project* team. 'By 2100, more than half of the more than 7,000 languages spoken on Earth—many of them not yet recorded—may disappear.'[11] Qualitative evidence that modernization is driving this decline, or, in a similarly intuitive vein, that secession reduced ethnic fractionalization in Russia/USSR after 1989, can only be uncovered using time-series data. The global coverage of the large-scale Demographic and Health Survey (DHS), while imperfect, will eventually furnish a global database to conduct time-series ethnic fractionalization research. Meanwhile innovative historical approaches, such as those utilizing sub-state census data for one or more countries (Urdal 2008) or providing fractionalization data at two time points (Roeder 2011; Green 2013) offer new pathways toward understanding the temporal aspects of this phenomenon.

There is a less theoretically-classifiable input into fractionalization from international factors. Ethnic diversity tends to decrease when state and ethnic boundaries converge, and vice-versa. This can occur as a) the number of political units increases and/or b) ethnicity and politics come into closer alignment through secession and partition. States which experienced secessions are no more homogeneous than others in the data and those with historic dominant minorities are not more fractionalized. However, the data shows that states formed through ethnically-defined national self-determination movements are more homogeneous than more 'civic' states originating on the basis of ideology or Great Power machinations.

Finally, the historic era in which states were created is important. Those which formed prior to the Congress of Vienna in 1815 are most homogeneous, while those emerging during the period of decolonization are most diverse. Having said this, ex-colonies are not significantly more fractionalized than other countries once geographic, historical institutional and modernist factors are introduced into the model. This questions the conventional wisdom that colonial borders condemned African states to ethnic fractionalization. Instead, these results suggest that variegated and low-lying terrain, a history of slavery and the presence of lootable resources better explains the ethnic diversity of sub-Saharan Africa. Indirect effects of colonization, such as the large size and recent vintage of Africa's states, are contributing factors.

These results question the strong variant of constructionist theory which claims that ethnolinguistic diversity may be created *ex nihilo*. This suggests that ethnic entrepreneurs can politicize pre-existing linguistic divisions or activate previously latent ethnic identities, but will have difficulty creating language and ethnicity anew. The Ijaw movement in Nigeria in the late 1990s, for example, represents the emergence of a new politically-relevant ethnic group in response to the glaring inequalities and pollution generated by local oil resources (Osaghae 2008). However, the Ijaw have been constructed on the basis of subgroups which the Ethnologue dataset classifies as speaking a related language. Linguistic invention for political reasons is not impossible: Bosnian, Croat and Serb variants of Serbo-Croat have only recently been developed. Yet the principal source of ethnolinguistic difference is geographical and historical. Ethnogenesis requires a plausibility structure and degree of popular resonance which limits the scope for invention (Zimmer 2003: 174). Hence nakedly political attempts at ethnogenesis such as the Padanian movement in Northern Italy, the Cruithin interpretation of Ulster Protestant origins or the Arab myth among Trinidadian Muslim Indians have proven conspicuous failures (Kaufmann 2008).

On the whole, the most striking finding of this paper is that cross-national differences in ethnic fractionalization are largely rooted in the geography, climate and historical institutions of a country. Ethnic diversity broadly predates modern political and economic change, and is therefore a useful independent parameter for analyzing contemporary political and economic life.

References

Acemoglu, Daron, and James Robinson. 2012. *Why Nations Fail: the Origins of Power, Prosperity and Poverty*. London: Profile Books.

Ahlerup, Pelle, and Ola Olsson. 2011. "The Roots of Ethnic Diversity." Stockholm Institute of Transition Economics.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114 (November):1243–84.

Anderson, Benedict. [1983] 1991. *Imagined Communities*. London: Verso.

Armstrong, John. 1982. *Nations Before Nationalism*. Chapel Hill: University of North Carolina Press201-3.

Banerjee, A., and R. Somanathan. 2007. "The political economy of public goods: Some evidence from India." *Journal of Development Economics* 82 (2):287-314.

Bockstette, V., A. Chanda, and L. Putterman. 2002. "States and markets: The advantage of an early start." *Journal of Economic Growth* 7 (4):347-69.

Brass, P. 1991. "Ethnic Groups and Ethnic Identity Formation." In *Ethnicity and Nationalism*, ed. Brass. New Delhi; Newbury Park, CA: Sage Publications.

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

Brigandt, I. 2001. "The Homeopathy of Kin Selection: An Evaluation of van den Berghe's Sociobiological Approach to Ethnic Nepotism." *Politics and Life Sciences* 20 (2):203-15.

Brubaker, R. 1996. *Nationalism Reframed: Nationhood and the National Question in the New Europe*. Cambridge: Cambridge University Press.

Cavalli-Sforza, L. L. 2001. *Genes, Peoples and Languages*. Stanford, CA: University of California Press.

CIESIN 2012. Center for International Earth Science Information Network (CIESIN), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw/lecz. (downloaded January 18, 2012).

Collier, P., and A. Hoeffler. 2004. "Greed and grievance in civil war." *Oxford Economic Papers-New Series* 56 (4):563-95.

Connor, W. 1994a. *Ethnonationalism: The Quest for Understanding*. Princeton, N.J: Princeton University Press.

———. 1994b. "A nation is a nation, is a state…." In *Ethno-nationalism: The Quest for Understanding*, ed. Connor. Princeton, N.J.: Princeton University Press.

Demeny, P, and G. McNicoll, eds. 2006. *The Political Economy of Global Population Change, 1950-2050*. New York: Population Council.

Easterly, W., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 111 (4):1203-50.
Easterly, William, Jozef Ritzen, and M. Woolcock. 2006. "Social Cohesion, Institutions, And Growth." *Economics and Politics* 18 (2):103-20.
Ellingsen, T. 2000. "Colorful Community or Ethnic Witches' Brew: multiethnicity and domestic conflict during and after the Cold War." *The Journal of Conflict Resolution* 44 (2):228-49.
Eriksen, Thomas H. 1993. *Ethnicity and Nationalism: Anthropological Perspectives*. London: Pluto Press.
Fearon, James D, and D Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1):75-90.
Fearon, J. D. 2005. "Primary commodity exports and civil war." *Journal of Conflict Resolution* 49 (4):483-507.
Fletcher, E., and M. Iyigun. 2009. "Culture, Clashes and Peace." In *IZA Discussion Paper no. 4116*.
Francis, E.K. 1976. *Interethnic Relations: An Essay in Sociological Theory*. New York, NY: Elsevier Scientific.
Gellner, Ernest. 1983. *Nations and Nationalism*. Oxford: Blackwell.
Giddens, Anthony. [1985] 1996. *The Nation-State and Violence*. Cambridge: Polity Press.
Green, Elliott. 2013. "Explaining African Ethnic Diversity," *International Political Science Review* 34, 3: 235-253
Hastings, Adrian. 1997. *The Construction of Nationhood: Ethnicity, Religion and Nationalism*. Cambridge & New York: Cambridge University Press.
Hobsbawm, E. J. 1990. *Nations and Nationalism since 1780. Programme, Myth, Reality*. Cambridge: Cambridge University Press.
Kaufmann, E. P. 2004. "Dominant ethnicity: from background to foreground." in *Rethinking Ethnicity: Majority Groups and Dominant Minorities*, ed. E. Kaufmann. London: Routledge.
_____ 2008. "The Lenses of Nationhood: An Optical Model of Identity." *Nations & Nationalism* 14 (3):449-77.
Kaufmann, Eric., and Oded Haklai. 2008. "Dominant Ethnicity: From Minority to Majority." *Nations and Nationalism* 14 (4):743-67.
Kittel, Bernhard. 2001. "Sense and Sensitivity in Pooled Analysis of Political Data." *European Journal of Political Research* (35):225-53.
Laitin, David D. 2007. *Nations, states, and violence*. Oxford ; New York: Oxford University Press.
Laitin, David., and Amanda Robinson. 2011. "The Continental Axis Theory Revisited," unpublished paper, *American Political Science Association*. Seattle, WA.
Mansergh, Nicholas. 1978. *The prelude to partition : concepts and aims in Ireland and India*. Cambridge [Eng.] ; New York: Cambridge University Press.
Michalopoulos, Stelios. 2012. "The Origins of Ethnolinguistic Diversity." *American Economic Review*, 102(4): 1508-39
Nairn, Tom. 1977. *The Break-Up of Britain*. London: NLB.
Nunn, Nathan. 2008. "The Long Term Effects of Africa's Slave Trade." *Quarterly Journal of Economics* 123 (1):139-76.
Osaghae, E. E. 2008. "Constructions of belonging: Igbo communities and the Nigerian state in the twentieth century." *African Affairs* 107 (429):666-7.
Özkirimli, Umut. 2010. *Theories of nationalism : a critical introduction*. 2nd ed. ed. Basingstoke: Palgrave Macmillan.

Pinker, Steven. 2011. *The better angels of our nature : why violence has declined*. New York: Viking.
Posner, Daniel. 2005. "Measuring Ethnic Fractionalization in Africa." *American Journal of Political Science* 48 (4):849-63.
Putterman, L. 2007. *The State Antiquity Dataset*. Brown University [cited February 2, 2012.] Available from http://www.econ.brown.edu/fac/Louis_Putterman/antiquity%20index.htm.
Roeder, P. 2010. *Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985* 2001 [cited 2010]. Available from http//:weber.ucsd.edu\~proeder\elf.htm.
Sambanis, N. 2001. "Do ethnic and nonethnic civil wars have the same causes? A theoretical and empirical inquiry (part 1)." *Journal of Conflict Resolution* 45 (3):259-82.
Schneider, Robert A, and N. Wiesehomeier. 2009. "Ethnic Polarization, Potential Conflict, and Civil Wars: Comment." *working paper, University of Konstanz*.
Smith, Anthony. 1986. *The Ethnic Origins of Nations*. Oxford: Blackwell.
Sowell, Thomas. 2011. *The Thomas Sowell Reader* (New York: Basic Books)
Sutherland, W.J. 2003. "Parallel Extinction Risk and Global Distribution of Languages and Species." *Nature* 423 (6937):276-9.
Trevor-Roper, Hugh. 1983. "The Invention of Tradition: The Highland Tradition of Scotland," in *The Invention of Tradition*, ed. E. Hobsbawm and T. Ranger. Cambridge: Cambridge University Press.
Urdal, H. 2008. "Population, resources, and political violence - A subnational study of India, 1956-2002." *Journal of Conflict Resolution* 52 (4):590-617.
Van den Berghe, P.L. 2002. "Multicultural democracy: can it work?" *Nations and Nationalism* 8 (4):433-49.
Vanhanen, T. 1999. *Ethnic Conflicts Explained by Ethnic Nepotism*. Stamford, CT: Jai Press.
Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflict. Shadows of Modernity*. Cambridge: Cambridge University Press.
Zimmer, O. 2003. "Boundary mechanisms and symbolic resources: towards a process-oriented approach to national identity." *Nations and Nationalism* 9 (2):173-93.
Zimmer, O., and Len Scales, eds. 2005. *Power and the Nation in History*. Cambridge: CUP.

Notes

[1] Ethnic group as used here encompasses both ethnic categories and groups. For the distinction, see Eriksen 1993, p. 44.
[2] Initial research on ELF used data from a 1960s Soviet ethnographic atlas (Bruk, S. I. and V. S. Apenchenko, eds. 1964. Atlas narodov mira. Moscow: Glavnoe upravlenie geodezii i kartografii gosudarstvennogo geologicheskogo komiteta SSSR and Institut etnografii im. H. H. Miklukho-Maklaia, Akademiia nauk SSSR.)
[3] All appendices at: www.sneps.net/ethnic/append.htm.
[4] Exports encompasses a total for all four African slave trades - trans-Atlantic, Indian Ocean, Red Sea, and trans-Saharan. For more, see Nunn (2008).
[5] All appendices at: www.sneps.net/ethnic/append.htm.
[6] These partial models can be viewed at: www.sneps.net/ethnic/alternative.htm.
[7] Logged variants of plurality ethnic group founding date also improve performance, but not as dramatically. See www.sneps.net/ethnic/alternative.htm for alternative specification with ethnic group founding date instead of 0-1100 A.D. origin.
[8] Rerunning the four identity-based models (EF, ELF, PLURAL, PCTMAJ) with an interaction term for sub-Saharan African slave exports results in the sub-Saharan Africa dummy and sub-Saharan African slave exports each falling out of two models. Both remain signed in the expected direction across all models. Sub-Saharan African slave exports was not included in the combined model due to restricted degrees of freedom, but this specification is shown in table 5 in www.sneps.net/ethnic/alternative.htm.
[9] See table 2 in www.sneps.net/ethnic/alternative.htm.

[10] See table 4 in www.sneps.net/ethnic/alternative.htm.
[11] See the National Geographic's Enduring Voices project website at: http://travel.nationalgeographic.com/travel/enduring-voices/.

ERIC KAUFMANN is Professor in the Department of Politics at Birkbeck College, University of London.
ADDRESS: Department of Politics, Birkbeck College, WC1E 7HX, UK. Email: e.kaufmann@bbk.ac.uk

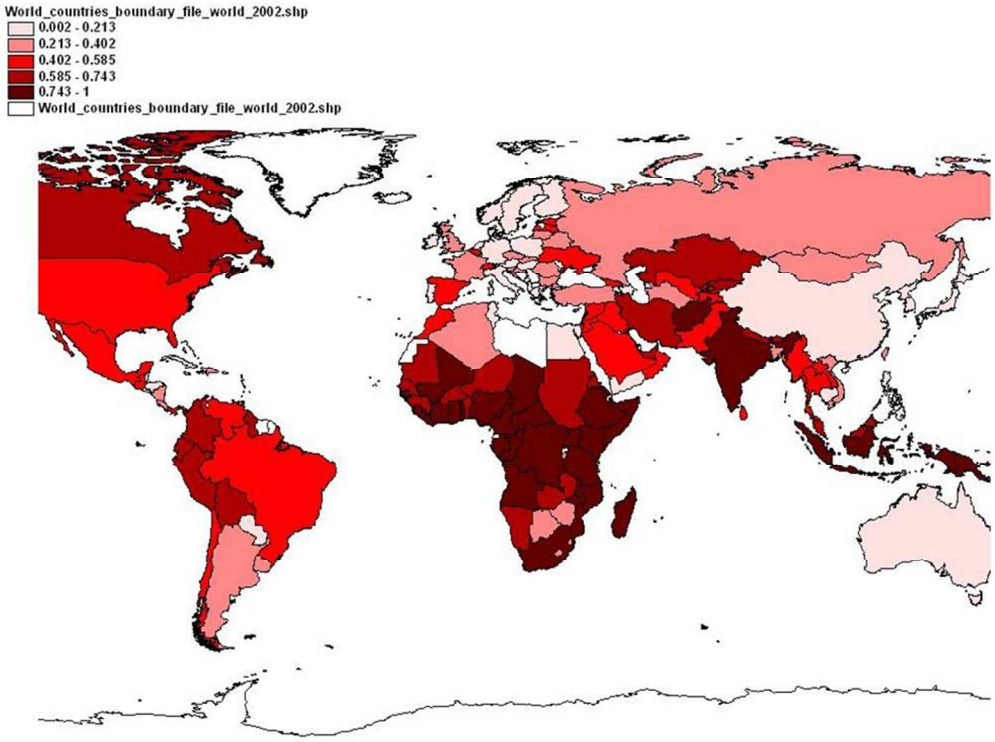

Figure 1. Global Ethnic Fractionalization Index (EF), 1999
245x175mm (104 x 111 DPI)

Table 1. Combined Model of Ethnic Fractionalization

	NMBRLANG	ELF	EF	ELFPREG	PLURALITY	PCTMAJ	MAXPOP
State	0.002	0.001***	0.001**	0.001	-0.001*	-0.046	0.000
Found. Date	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.027)	(0.000)
Land Area	0.555***	0.038**	0.017	0.038*	-0.014	-1.054	-0.024
(Square km)	(0.070)	(0.012)	(0.011)	(0.016)	(0.010)	(0.912)	(0.015)
Democracy	-0.006	-0.001	-0.001	0.000	0.001	0.148**	0.000
(Polity IV)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.049)	(0.001)
GDP per	-0.047	0.006	-0.017	-0.016	0.026	0.893	0.026
capita	(0.156)	(0.026)	(0.025)	(0.033)	(0.023)	(2.046)	(0.030)
Oil Output	0.009	-0.001	0.009**	0.015**	-0.009**	-0.184	-0.015***
Per Capita	(0.016)	(0.003)	(0.003)	(0.004)	(0.002)	(0.215)	(0.004)
Ethnic Origin	-0.287	-0.141**	-0.126**	-0.077	0.085*	12.921**	0.030
0-1100 A.D.	(0.288)	(0.049)	(0.047)	(0.068)	(0.043)	(3.772)	(0.062)
Mean	0.008***	0.001	0.001*	0.000	-0.001**	-0.061**	0.000
Precipitation	(0.002)	(0.000)	(0.000)	(0.000)	(0.000)	(0.023)	(0.000)
S.d. of Mean	2.043	0.518*	0.582**	0.506	-0.507**	-41.721*	-0.344
Agric. Suitability	(1.214)	(0.206)	(0.198)	(0.257)	(0.182)	(15.928)	(0.232)
S.d. of Mean	1.208*	0.171*	0.190*	0.048	-0.186*	-9.561	-0.127
Elevation	(0.491)	(0.083)	(0.080)	(0.102)	(0.074)	(6.437)	(0.093)
Mean	-0.751*	-0.105	-0.116*	0.003	0.114*	7.097	0.019
Elevation	(0.325)	(0.055)	(0.053)	(0.068)	(0.049)	(4.267)	(0.062)
Sub-Saharan	0.695*	0.239***	0.222***	0.242**	-0.234***	-12.801**	-0.315***
Africa	(0.315)	(0.054)	(0.052)	(0.068)	(0.047)	(4.134)	(0.061)
Ethnic	-0.134	-0.091	-0.110*	-0.100	0.076	7.960*	0.031
Nationalism	(0.276)	(0.047)	(0.045)	(0.061)	(0.041)	(3.626)	(0.056)
Population	0.063*	-0.267**	-0.587**	-0.223	0.486*	51.546**	0.182
Density (1995)	(1.235)	(0.210)	(0.202)	(0.253)	(0.185)	(16.204)	(0.230)
constant	-1.384	-2.428**	-1.439	-1.052	1.858**	144.659*	1.067
	(4.546)	(0.772)	(0.743)	(0.980)	(0.681)	(59.625)	(0.893)
N	136	136	136	112	136	136	113
R^2	0.565	0.636	0.618	0.497	0.621	0.617	0.537

* p ≤ 0.05, ** p ≤ 0.01; *** p ≤ 0.001; robust standard errors in parentheses. Dependent variable at top of each column. NMBRLANG, ELF, EF and ELFPREG are measures of ethnic fractionalization, PLURAL, PCTMAJ and MAXPOP of the size of the largest ethnic group. See Appendix 1 for further details.

Figure 2. Plurality Ethnic Group Founding Dates and Ethnic Fractionalization
167x104mm (72 x 72 DPI)