

**ETHNIC AND STATE HISTORY AS DETERMINANTS OF  
ETHNIC FRACTIONALIZATION<sup>1</sup>**

**ERIC KAUFMANN**

*Department of Politics, Birkbeck College, University of London*

*Malet Street, London, United Kingdom WC1E 7HX*

*e.kaufmann@bbk.ac.uk*

---

<sup>1</sup> The author acknowledges the support of British Academy grant # SG091163 which made this research possible and is grateful for the assistance of Anthony D. Smith, who helped scrutinize the dates obtained in the research. I am also grateful for feedback on an earlier version of the paper presented at the Intrastate Conflict Seminar at the Kennedy School of Government, Harvard, 2008-9.

## **Ethnic and State History as Determinants of Ethnic Fractionalization**

This paper presents a new dataset which codes for the approximate founding date of the largest ethnic group in the world's states as well as the wave of state creation in which the particular country emerged.

All contemporary questions 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 in a state. Ethnic fractionalization is produced by the overlay of political fragmentation atop ethnic fragmentation. There are currently close to 200 states overlaid on top of some 6000 ethnic groups producing a variegated ethnic grid that varies from state to state. Fractionalization will tend to decrease as ethnicity and politics move toward congruence. This can occur for three major reasons: a) the number of political units increases; b) the number of ethnic groups decreases or certain large ethnic nuclei attract members away from smaller groups; and/or c) the fit between ethnicity and politics improves.

Unfortunately, with the partial exception of Philip Roeder's 1961 and 1985 dataset (Roeder 2001), only cross-sectional ethnic fractionalization data is available. This means that we cannot adequately test proposition a). However, we can derive indicators for the arguably more important processes b) and c). (Bockstette et al. 2002) have already demonstrated a significant negative association between the date of state formation (and degree of indigenous control of the state at the time) and the ethnolinguistic fractionalization index (ELF). Older indigenous political histories tend to be associated with less fractionalized states today. However this was not the primary purpose of their work, so the association was simply noted as an interesting univariate relationship. This paper attempts to build on this state history finding, but goes beyond it by parsing out ethnic and state formation variables.

The foundation date of the largest ethnic group in a state serves as a measure of b) above, because we assume that older state-dominant ethnic groups have had more time to assimilate subaltern groups than newer state-dominant groups. According to ethnosymbolist theory, modern nations typically form around premodern ethnic 'core' groups which often engaged in social and ethnic assimilation prior to the modern epoch (Smith 1986). For modernists, modern state creation begins the process of nation-building and diversity reduction. Regardless of which approach one favours, it is worth noting that ethnic groups and states serve as assimilationist actors - large fractionalization-reducing ethnic nuclei - within multi-ethnic states. One question this paper seeks to adjudicate is whether ethnic pluralities or states are more important actors in this process.

At one time, dominant minorities were common. With the decline of empire, dominant minorities give way to dominant pluralities or majorities as popular sovereignty and democratization spread (Kaufmann 2004). (Connor 1994a: 39) notes that while fewer than 10 percent of states are ethnically homogeneous, a substantial majority of them have an ethnic majority. All but a handful of states have a significant plurality group. How to measure the founding date of the largest ethnic group? This date is defined as the first imagining of the

group by a putative member of the group and is determined through consultation with historians and social scientists who possess expertise on the appropriate country.

To tap c) above, the paper distinguishes five waves of state formation: a) pre-1815; b) 1815-1917; c) 1918-45; d) 1946-89; e) 1990-present. Here we posit a connection between the date of state formation and the degree of ethnic-nationalist self-determination. This is not a continuous or ordinal variable as there have been discrete waves of state formation punctuating long periods of relative calm. The relative balance between ethnic self-determination, desire for territorial aggrandizement and other-determination of boundaries governs the congruence of borders. We assume that states which largely emerge through secession from a modern state (1990-present) are most homogeneous. Those springing from the transformation of an early modern state (pre-1815) would be predicted to be next in ethno-national congruence. These should be followed by those which emerged in the wake of Versailles (1918-45) and which were animated at least in part by Wilsonian principles of ethno-national congruence, though largely by Great Power motives and the domestic desire for territory. Early secessions from land empires (many of 1815-1917) should follow, and, finally, decolonizations, where ethnic self-determination is least evident (most of 1946-89 state creations). We also test other popular determinants of ELF such as climate, terrain, colonizer, state size and various indicators of modernization and economic development from several leading datasets.

All contemporary questions 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 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. The greater the number of ethnic groups and the more even their relative size, the more fractionalized the population<sup>1</sup>. More recent formulations attempt to modify the original 1964 ELF measure by considering the size of the largest and second largest ethnic groups (Fearon and Laitin 2003: 84). Hence the very high correlation ( $r^2=.96$ ) between Fearon and Laitin's fractionalization measure and the size of the largest ethnic group in a country. This paper pursues this line of reasoning, seeking to evaluate *which variables best explain the proportion of a state's population made up of the largest ethnic group and a state's level of ethnic fractionalization.*

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 a series of political and economic outcome variables such as economic growth, public goods provision and violent conflict. The list - far from exhaustive - of work which tests the link between ethnic fractionalization and violence is impressive, viz. (Urdal 2008; Montalvo and Reynal-Querol 2005; Ellingsen 2000; Vanhanen 1999a; Collier and Hoeffler 1998; Lujala et al. 2005; Ostby et al. 2009 ; Buhaug et al. 2008; Schneider and Wiesehomeier 2009; Collier and Hoeffler 2004; Sambanis 2001; Cederman et al. 2010).

However, fractionalization is frequently taken as given. Literature inquiring into the root causes of ethnic fractionalization - i.e. ethnic diversity as a dependent rather than independent variable - is a good deal more sparse. Many of those with this particular (I dare

not say peculiar!) interest are assembled here. Working papers by Michalopoulos (2011), on geographic factors, and Green (2011), on the role of urbanization - to be presented in more recent versions on this panel - are part of this new development. To this we may add the work of (Nunn 2008) on the role of the slave trade in increasing ethnolinguistic diversity.

## Theories of Ethnicity and Nationalism

Any attempt to make sense of the ethnicity-nation link demands that we define our concepts with greater precision. This is a perilous but necessary task for a field of study which has been impaired by its arrival onto the academic scene during an era in which conceptual contestation has been more greatly esteemed than simplification (Connor 1994b). Even so, since 1980, great strides have been made toward differentiating the concepts of state, nation, and ethnic group, and sketching the linkages between such phenomena.<sup>2</sup> The term ethnic group, or *ethnie*, should be reserved for communities which possess a belief in shared genealogical descent and meet a threshold requirement that distinguishes them from smaller-scale *gemeinde* like clans and tribes or larger ones like pan-ethnies.<sup>3</sup> There are two major forms of ethnicity, according to the sadly forgotten schema of Francis (1976): *primary* ethnicity, in which members of the group occupy their ancestral 'homeland' territory; and *secondary* ethnicity, whereby groups acknowledge that they are not native, and thus their homeland lies elsewhere.<sup>4</sup> It is primary ethnic fractionalization which is generally captured by ELF indices and is by far the most form of fractionalization when it comes to economic development and conflict.

The power of ethnicity is explained in different ways by the three main families of nationalism and ethnicity theory, primordialism, ethnosymbolism and modernism.

Primordialism and ethnosymbolism can be grouped under the rubric of perennialist theories of ethnicity as they deem ethnicity to be a long-standing feature of humanity which pre-dates the modern, post-1789 period. Primordialism locates ethnicity in our genes or in universal aspects of human psychology. For primordialists like Pierre Van den berghe, our evolutionary psychology represents the successful adaptation of our genes to conditions obtaining during prehistoric time. The evolutionary fitness of genes was maximized by groups that cooperated along nepotistic lines, i.e. in such a way as to favour those with the most similar genes.

Primordialism argues that individual sacrifice in battle could often be a successful strategy in evolutionary terms if one's genetic relatives gained a long-term advantage over competitors. Individuals that failed to cooperate in this way (by free-riding, for instance) had lower long-term genetic success. (Salter 2001; Van den Berghe 2002)

Though genetic variation between humans is very slight, even a marginal difference is enough to drive selection. These inherited drives continue to shape individual behaviour today, argue primordialists. Thus 'type I' appeals to shared ethnicity tend to trump 'type II' appeals along the lines of interest or ideology, and are much more emotionally fervent. (Van den Berghe 1979) This also explains the high degree of relatedness – equivalent to that between grandparent and grandchild - between random members of the world's major ethnic groups. (Salter 2003) Primordialism's power is also manifested in human psychological universals, notably our attachments to territory and shared culture. (Geertz 1963; Shils 1995)

Critics of primordialism suggest that beyond the family, evolutionary psychological impulses may be deflected toward constructed forms of community. Moreover, primordialists admit that their theory cannot account for the drawing of group boundaries. For these reasons,

primordialist explanations remain less favoured than those which invoke the social creations of mankind.

A more popular alternative within perennialism is ethnosymbolism. For the ethnosymbolist school, ethnic groups predate the modern era, but arise only during the era of written civilization after 6000 B.C. They arose in the late neolithic period as 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 Englishness of Bede and King Alfred (Hastings 1997: 35-9). Some ethnic groups (i.e. Jews, Amhara, Armenians, Persians) have ancient origins, while many more emerged in the medieval and early modern periods through tribal confederation (i.e. Arabs, Kurds), conquest agglomeration (i.e. Germanic war bands) or dynastic competition (i.e. Scots, Catalans). In all cases, territorial identities extending beyond the locale came to be established (Smith 1986; Armstrong 1982).

Most of these premodern entities were primary ethnic groups, even if they sometimes spawned secondary offshoots like the Jewish, Parsee and Armenian diasporas. The appearance of translations of religious texts like the bible into vernacular languages, and the numerous premodern written references to nations and peoples 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). Ethnosymbolist theory predicts that ethnic identities, once formed, are highly path-dependent and durable. Ethnic sentiments are typically reproduced by popular institutions in civil society and are therefore capable of inspiring collective action and resisting 'artificial' political constructs which may be imposed by subsequent conquerors.

The *modernist* account goes beyond the ethnosymbolist 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; 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, more intensive transport networks and secularization orient all identities toward a common, this-worldly community.

Shared ethnic identities are only created in modern times as a product of state socialization (which integrates localities into a shared nationality), state internal boundary demarcation (which institutionalizes ethnic diversity, as with Soviet Republics or colonial administrative departments) or anti-state mobilization by political entrepreneurs (Trevor-Roper 1983; Brass 1991; Brubaker 1996). Such entrepreneurs may have experienced blocked upward mobility within central state structures (Gellner 1983), or merely see an opportunity to acquire more power or wealth than they can through established central political channels (Breuilly 1993).

Modernists therefore maintain that elites’ invention of ethnicity is typically instrumental, for their own private gain. The response of the populace to ethnic appeals is sometimes explained as the outcome of coercion (rendered easier by a shared language and sunk costs which tie individuals to groups) or attributed to the idea that people believe their interests to be bound up with the fate of their group (Hechter 2000; Laitin 2007).



## Theorizing Ethnic Fractionalization

Now consider the sources of ethnic diversity. A popular explanation in the immigrant-rich urban west is that immigrant societies are more diverse than emigrant societies. But only 2.7 percent of the world's population are immigrants, and not all are ethnically different from their hosts and permanently settled in the hostland (Demeny and McNicoll 2006, ch. 1). Moreover, diasporas are heavily concentrated in cities. Diasporic, or secondary ethnicity (Francis 1976) is therefore really only important in a small number of countries. To explain ethnic diversity and its relationship to conflict, we really need to explain variation in primary ethnic diversity, captured in the now-established measure of ethnic fractionalization (ELF). ELF measures the likelihood that any two random individuals in a state's population are members of the same ethnic group. The greater the number of ethnic groups and the more even their relative size, the more fractionalized the population. This paper seeks to evaluate which theories best explain why certain countries are more ethnically fractionalized than others.

Particular theories of ethnicity make specific predictions with regard to ELF. The primordialist perspective maintains that rougher terrain, by isolating populations, leads to cultural and genetic drift, hence greater numbers of languages, religions and ethnic identities. The ethnosymbolist contention is that rougher terrain makes it more difficult for premodern tribal confederations, conquest agglomerations and kingdoms (which operate at a larger scale than face-to-face interaction) to integrate peripheral populations into their community of shared ancestry. The spread of a common consciousness of myths, symbols, culture and memory is thereby impeded. Extreme isolation produces mere localism, impeding ethnogenesis. However, above this threshold, more difficult terrain would be expected to

increase the number of competing premodern polities. Rival ethnic identities can thereby take root so long as there is a literate clerisy that can spread myths, symbols and memories beyond the local. Given some level of extra-local mobilization, rougher terrain would be expected to lead to greater ethnic heterogeneity, but less so than in the primordialist case.

For modernists, rough terrain makes it more difficult for the state to homogenize populations and easier for secessionists to escape to peripheral redoubts from which they can invent their interest-reinforcing ethnic projects. Creating ethnic identity is functional for social communication between group members, and serves the practical purpose of enabling members to distinguish insiders from outsiders. A softer variant of this argument is to allow that group members are attached to their ethnic identities, but to claim that this attachment is manipulated by instrumental ethnic elites to attain wealth, power and prestige. Though rough terrain 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 the importance of spatial variation by taking the simple relationship between the unevenness of a country's topography, which does not vary over time and cannot be reduced to a prior social cause, and ethnic diversity. Thus our first hypothesis:

H<sub>1</sub>: Rougher terrain is associated with higher levels of ethnic fractionalization. The stronger the relationship, the greater the support for primordialist theories.

A strong relationship between these two variables can only be consistent with rough terrain shaping primary ethnic diversity unless we can explain how ethnically-diverse ecologies

cause rougher terrain. Another possibility is that immigrants prefer mountainous lands, an unlikely prospect.

We should also note that the most developed literature on the sources of ethnic fractionalization seems to be that which examines the effects of geography and climate on diversity. Generally, the verdict is that variations of rainfall, soil type and elevation matter. Areas of low rainfall require larger social networks, and hence produce less diversity over a given area than wetter regions. Zones where soil type and terrain varies tend to produce greater ecological variation, which is associated with more ethnolinguistic variety (Michalopoulos 2010; Sutherland 2003; Collard and Foley 2002; Moore et al. 2002; Nettle 1996; Nichols 1992).

#### Political and Economic Factors

Elliott Green and Nathan Nunn usefully introduce economic forces into the analysis of ethnic fractionalization. For Nunn, the slave trade inhibited ethnic fusion by setting groups against each other, hence a correlation between slave exports and contemporary ethnic fractionalization in Africa (Nunn 2008: 164). Green makes the case that the growth of cities acts as an agent of ethnic agglomeration, reducing fractionalization (Green 2010). This echoes qualitative work which noted the presence of ethnic fusion in modernizing locations such as the colonial Zambian Copperbelt settlements of the 1930s where miners from formerly distinct kin groups amalgamated into larger ethnic entities based on cultural relatedness. They even invented new rituals to mark themselves off from other ethnic groups (Eriksen : 20-21). The case has also been made for heavily regionally-minded European

groups like the Ukrainians and Italians, who often developed an awareness of themselves only after settling together in ethnic neighbourhoods of American cities in the late nineteenth and early twentieth centuries (i.e. Thomas and Znaniecki 1918).

Political factors must also be considered. Ethnic fractionalization is produced by the overlay of political fragmentation atop ethnolinguistic topography. There are currently close to 200 states superimposed on some 6000 ethnic groups producing a variegated ethnic grid that varies from state to state. Fractionalization will tend to decrease as ethnicity and politics move toward congruence. This can occur for three major reasons: a) the number of political units increases; b) certain large ethnic nuclei absorb smaller groups or attract members away from them ; and/or c) ethnicity and politics come into closer alignment through secession and partition. Fractionalization will be increased to the extent that states unite or conquer others, or ethnic fission takes place.

Clearly a fine mesh of 6000 squares overlaid onto a 6000-group map will result in fewer groups per square than a coarser 200-square grid. Unfortunately, with the partial exception of Philip Roeder's 1961 and 1985 datasets (Roeder 2001), only cross-sectional ethnic fractionalization data is available. This means that we cannot adequately test the power of new state formation in reducing ELF. If we had annualized data on ELF, we would find an important temporal connection between the secession/creation of new states and declines in ELF. Think of the shift from the Ottoman Empire to rump Turkey after 1918, the USSR to Russia, or Yugoslavia to a lesser Serbia after the secessions of the 1990s. In such cases, ELF scores would nosedive in regions where the implosion of multiethnic states took place. Political and economic variables which change over time are thus poorly captured by ELF data, which is time invariant. Given data limitations, the best we can do is to record the number of successful secessions or partitions that have taken place in a particular state over the course of its history.

We would expect that, all else being equal:

H<sub>2</sub>: States which have experienced more secessions or partitions will have lower levels of ethnic fractionalization than others

The proliferation of states is not especially relevant for the theoretical debate between modernists, primordialists and ethnosymbolists. However, b) above is pregnant with theoretical implications because ethnic fusion or assimilation tends to be viewed by ethnosymbolists as perennial and continuous with premodernity, while modernists treat ethnic accretion as a post-1789 development arising out of the nation-building project (Wimmer 2002; Birch 1989)

A critical test of these differing predictions is the relationship between the date at which an ethnic group or state is established, and ELF. Extant research (Bockstette et al. 2002; Chanda and Putterman 2007) finds a significant negative association between the date of state formation (and degree of indigenous control of the state at the time) and a state's ELF. The logic is that older states, and those where the indigenous population was more often in control, are less diverse than newer states (or those where foreign rulers predominated over much of the territory's history). However the connection between state history and ethnic fractionalization is a bivariate finding as this relationship was not the primary focus of the authors. This paper builds on Bockstette, Putterman et al.'s state history finding, but moves beyond it by building in an ethnic continuity dimension and parsing out ethnic and state formation variables.

The founding date of the largest ethnic group in a state serves as a measure of ethnic absorption, i.e. b) above, because an ethnosymbolist would assume that older plurality ethnic

groups have had more time in which to assimilate subaltern groups than newer ones.

According to ethnosymbolist theory, modern nations typically form around premodern ethnic 'core' groups (Smith 1986). Popular sovereignty has emerged as a powerful norm in the modern era, rendering dominant minorities rare. This means that in many cases, the largest ethnic group is also state-dominant (Kaufmann and Haklai 2008).

Dominant ethnic groups frequently emerge as assimilationist actors - fractionalization-reducing nuclei - within multi-ethnic states. Connor (1994a) notes that homogeneous nation-states occur in less than 10 percent of the world, but that a substantial majority of states in the world have an ethnic majority and all but five (of 156 countries in the Vanhanen dataset) have 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 explain this as a result of nation-building and the ethnic exclusions practiced by modern states. Older states would be expected to contain less ethnic diversity, but premodern ethnomorphology should not affect contemporary ELF.

How to measure the founding date of the largest ethnic group? This is defined as the first imagining of the group by a putative member of the group. The Ethnic Group Foundation Dates dataset will be introduced later. It has been developed through consultation with historians and social scientists with expertise on particular country(s). Accepting the 'reality' of ethnic founding dates need not entail an ethnosymbolist position. Premodern ethnic imaginings could be explained by modernists as the mere musings of individuals - with no consequences whatsoever for social and political behaviour. Primordialists, too, would consider these visions to be subsidiary to spontaneous mass nepotism in the process of ethnic fusion. Therefore, an ethnosymbolist would expect that:

H<sub>3</sub>: States with plurality groups that are comparatively old will have lower degrees of ethnic fractionalization

Modernists would qualify this as follows:

H<sub>3a</sub>: Older states (defined as post-1789 phenomena) will have lower degrees of ethnic fractionalization, but pre-1789 ethnic founding dates should not affect ELF

Primordialists are agnostic on boundary construction. They hold few views about social and political boundaries beyond the idea that multi-ethnic states will tend to break apart (Vanhanen 1999b). Ethnic and state founding dates would be relevant for ELF insofar as older groups and states would have had more time for spontaneous ethnic fusion (of genetically-related groups) and political secession to take place. This yields:

H<sub>3b</sub>: Older states and ethnic groups will have lower degrees of ethnic fractionalization

Moving to consider political 'fit', c) above, it can be observed that the degree of ethnic determination of state boundaries issues from either 'top-down' variation in the ethno-cartographic accuracy of Great Power boundary demarcation, or a 'bottom-up' variable: the nature and extent of ethno-national self-determination. Green alludes to the 'top-down' aspect, citing variation in colonial practices in Africa, suggesting that the large size of states and the arbitrariness of political boundaries produced high levels of fractionalization:

In particular (Englebert, Tarango, & Carter, 2002, p. 1096) have noted that some 44% of colonial borders were straight lines, leading to as many as 177 ethnic groups split across two and sometimes three colonial borders; when added together these partitioned ethnic groups represented 43% of the average African state's population (Green 2010).

What of bottom-up processes? Clearly, for an ethnosymbolist, the more a movement for national self-determination defines itself in ethnic terms (i.e. 'ethnic' Irish nationalism v. 'civic' Eritrean nationalism) and the greater a movement's degree of efficacy in securing its ethnic territory, the lower its ELF.

But modernists advance their own 'bottom-up' processes which are also important: the worse an ethnonational movement fares in grabbing others' territory (Romanians too successful, Kurds too weak, Poles and Hungarians optimal), the lower will be its level of ethnic fractionalization. An ethnic group must be strong *vis á vis* the powers-that-be, but not too strong, to achieve low ELF. Ulster Protestants and Lebanese Christians, in consultation with their Great Power guarantors, aimed at about a two-thirds majority in their new states (or provinces). This satisfied these dominant groups' desire to be in a comfortable majority but also slaked their ambition to control significant swathes of minority-dominated territory. In other cases, 'winning' nationalizing states, like Romania after WWI, blatantly sought out the most territory they could hold (Brubaker 1996; Zimmer 2003). Most irridentisms, if successful, would increase ethnic diversity (Saideman and Ayres 2008). Conversely, ex-empires often emerge from collapse as more homogeneous: the partitioning of Hungary and



Ireland, for instance, or the dismemberment of the Turkish, Soviet, Austrian and Japanese empires, greatly increased these states' ethnic homogeneity (Brubaker 1998).

The relative balance between ethnonationalists' degree of self-determination, their success at territorial aggrandizement and the ethno-political accuracy of others' determination of boundaries governs the congruence of borders. To tap ethnonational 'fit', i.e. c) above, this paper offers two indicators. First, it distinguishes five waves of state formation, given that there have been discrete waves of state formation punctuating long periods of relative calm: 1) pre-1815; 2) 1816-1918; 3) 1919-45; 4) 1946-90; 5) 1991-present. Roughly speaking, these correspond to 1) state nations of the western core; 2) secessions from European land empires; 3) self-governing white dominions; 4) newly independent overseas ex-colonies; and 5) post-Cold War secessions.

This posits a connection between the date of state formation and the degree of ethno-national self-determination. This is not a continuous or ordinal variable. States which emerge through secession from a modern state (1990-present) are predicted to be most homogeneous. Those springing from the transformation of an early modern state (pre-1815) are expected to be next in terms of ethno-national congruence. These should be followed by those which emerged in the wake of Versailles (1918-45) and which were animated at least in part by Wilsonian principles of self-determination, though distorted by Great Power imperatives and winners' desire for territory. Early secessions from land empires (1815-1917) should follow, and, finally, decolonizations, where ethnic self-determination is least evident (most of new 1946-89 states). Hence:

H<sub>4</sub>: The wave in which a state was created should be associated with ethnic fractionalization, though not in an ordinal way; and, as a corollary:

H<sub>4a</sub>: Ex-colonies should have higher levels of ethnic fractionalization than societies which were never colonized

It is difficult to mark either H<sub>4</sub> or H<sub>4a</sub> as reflecting any particular theoretical orientation. All three major theories would make similar predictions.

Where there is more dispute is over the importance of the content of ethnic myths. Ethnosymbolists would claim that states founded on the basis of national self-determination movements which defined themselves in explicitly ethnic terms (i.e. Zionism) should be more homogeneous than those which did not (i.e. Eritrea). This generates:

H<sub>5</sub>: States formed on the basis of nationalist movements which define themselves in ethnic terms should have lower levels of ethnic fractionalization than other states

Modernists would argue that independence movements are a route to state power, which, once achieved, is followed by cultural homogenization (Breuilly 1993; Birch 1989; Weber 1976). Ethnic exclusion is constitutive of modern statehood, with culture following, rather than leading, politics (Wimmer 2002). For modernists then, the content of nationalist myths at independence is of little importance and H<sub>5</sub> should not hold.

Finally, the valence of assimilation tends to favour prestigious or powerful groups (Deutsch 1966 [1953]: 156-62). Therefore, states with historically dominant minorities in which the largest ethnic group did not control the state - even if these groups have recently lost power - are expected to be more fractionalized. This is because the powerlessness of the

plurality group retards its progress in assimilating its neighbours.<sup>5</sup> Therefore we would expect:

H<sub>6</sub>: States with a historic dominant minority should register higher levels of ethnic fractionalization

This holds for all major theoretical orientations, though modernists would limit their purview to post-1789 dominant groups.

Global Region

The primordialist account conceives of the ‘ethny’ as a unit that can be as small as several hundred people. Genetic cooperation begins at local level and radiates outward, thus it is difficult to see how macro-level continental factors could make much difference to the degree of ELF obtaining in a country. This is especially true where we control for geography and religion. Hence if the predictions of primordialism are correct, we should not expect to see world regions associated with different levels of ELF. The modernist account makes a similar prediction, but arrives at it differently. For modernists, it is the political and economic environment, notably income per capita, democracy, colonialism, instability or resource wealth which count most. Controlling for these properties, modernists would view world regions as relatively unimportant determinants of contemporary ELF.

Ethnosymbolists, by contrast, would predict that the historical aspects of major world regions are important. For them, the collision between expansive world religions and empires in the premodern period helped to lay the groundwork for modern ethnic diversity. Two key factors in this process are the integrating power of premodern polities and the size of the empires which they commanded. When empires collapse, fracturing is rarely culturally perfect, so homogeneity is better achieved when culture and power expand in dialectical fashion, as in western Europe. Islamic world regions involved a peripatetic elite governing nomadic populations. Muslim polities united temporal and cultural authority in the figure of the caliph, thus ethnicity and politics were less strongly aligned (Armstrong 1982).

Ethnosymbolists also maintain that certain world regions were embedded in circuits of 'archaic globalization' and exchange (Eurasia) prior to the modern age, while others (sub-Saharan Africa, Oceania, the pre-Columbian Americas) were largely excluded for biogeographical reasons (Abu-Lughod 1989; Diamond 1997; Bayly 2004). The resulting political, ethnic and kinship structure - and their congruence - thereby differs from region to region.

In Western Europe, the more rapid development of military technology began to favour medium-sized, integrationist states over loosely sovereign empires by the absolutist period, limiting the degree to which large states ruled over diverse populations, i.e. ELF (McNeill 1991). In Eastern Europe, by contrast, ethnosymbolists would aver that empires hung on longer, leading to more ELF. Finally, Northeast Asian premodern kingdoms (Chinese, Korean, Japanese), despite periods of expansion and contraction, may have been better able than their European counterparts to align ideology with the state. They were thus highly successful in assimilating outsiders into the ethnic core, reducing ELF. Overall, we would predict that:

H<sub>7</sub>: If primordialist theories are correct, there should be no association between world region and ethnic fractionalization beyond what can be accounted for by ecological differences. The same should be true for modernist theories when economic and political factors are held constant. Finally, if ethnosymbolist theories are accurate, levels of ethnic fractionalization should vary significantly by world region.

We must treat the foregoing with as indicative rather than definitive, since regions may be serving as a proxy for unobserved material factors (unspecified ecological aspects in the case of primordialism, or material features for modernism).

## Methodology

The Ethnic Group Foundation Dates dataset was compiled by PhD student researchers with the assistance of a grant from the British Academy. Work was completed by eight PhD students who are working on issues of nationalism and ethnic conflict. All are members of, and have access to, the expertise base of the Association for the Study of Ethnicity and Nationalism (ASEN) at LSE in London. This is a 20-year old student-run organisation which publishes *Nations and Nationalism*, the leading nationalism journal, and *Studies in Ethnicity and Nationalism*, and has several hundred members at any one time. Directories of members and their research interests are one resource, alongside more conventional forms of bibliographic inquiry.

Experts were contacted and asked to reply to the statement in Appendix IV. Experts were directed to address the question of 'first imagining' by an in-group member rather than the formation of fully-fledged ethnicity, as this involves making determinations about the degree to which a set of myths of origin must penetrate down the social scale before a group can be said to be in existence. Though the label 'ethnic group foundation date' makes an ethnosymbolist presumption, it is perfectly possible to interpret this date in modernist fashion as an ideational curiosity whose significance and continuity has been retrospectively constructed by political entrepreneurs. This is precisely how Eric Hobsbawm, for instance, treats the 'cultural nationalist' activities of romantic intellectuals, i.e. philology, folk anthropology, festivals, archaeology (Hobsbawm 1990).

The dating of a subjective phenomenon like the first imagining of a group confounds attempts to use standard techniques of inter-coder reliability. In the event, the initial 'scientific' aim of achieving inter-coder reliability had to be replaced by a more historical approach in which the accounts of scholars were triangulated because there was no simple process by which coders could 'code' data. This is the method, for instance, used in the State Antiquity Data project dataset (Bockstette et al. 2002; Chanda and Putterman 2007). All dates were reviewed for consistency by the author and by Anthony D. Smith, to whom the author is especially grateful.

The majority of those contacted did not respond (see Appendix VI). Even among those who did, some found the question difficult to answer while others thought the matter controversial given the theoretical debate over the question of the modernity of ethnic groups. We had, however, anticipated such controversies and were able to offer guidance (see Appendix IV) regarding which changes we felt represented continuity of historical consciousness and which signified rupture. We also advised upon when consciousness was deemed to exist and when the evidence was insufficient to support such a claim. Based on

this criteria, Sunni Arab Egyptians and Iraqis are deemed to have lost contact with the narrative created by ancient progenitors while Greeks, Persians and the Jews are seen to have maintained it, however minimally.

The great shifts in consciousness, i.e. from ancient Hellenes to Byzantines to modern Greeks, are viewed as insufficient to constitute a full rupture. By contrast, the collapse of Sumer, Babylon or Assyria clearly sundered any continuity of Iraqi consciousness. It is a difficult exercise to erect a pure standard for borderline cases to decide when a historic consciousness has been broken or not. This has been attempted here, however, and distinguishes this dataset from the State Antiquity dataset which only seeks to establish whether an 'indigenous' dynasty was in control in a portion of the geographic area of the present-day state in a given half-century. This is an important difference, but one which may account for the stronger connection between ethnic group founding date and current group size as compared with the State Antiquity measure.

We were able to compile a set of dates for 129 of the 156 countries in the Vanhanen dataset. Crucial to bear in mind is that there is a considerable distance to travel between the first imagining of a group by a putative insider and the fully-fledged emergence of mass ethnic consciousness. Walker Connor (1994) argues that a majority of members must be conscious of their ethnic status in order for a group to be considered to be in existence. By this stringent criteria, there would be few, if any, groups in existence prior to 1800. This may be the reality - but the astounding relationship between ethnic first imagining and modern-day ethnic group morphology suggests that the date of origin is far from meaningless.

Naturally the path from initial imagining to active ethnic behaviour is episodic. In some cases, the time between first imagining and collective action is brief. In others, it is protracted. For example, kingdoms which formed the ultimate origin of the largest ethnic

groups in the Democratic Republic of Congo, Nigeria, Sierra Leone and several West African states date to the 14th and 15th centuries. Yet it is unclear how much this gives these groups an advantage over their twentieth century equivalents in neighbouring Liberia or Cameroon. In conjunction with ethnic date, a measure of both continuity and social (vertical) penetration would most likely improve the predictive power of how well an ethnic group is able to expand geographically (horizontally) into neighbouring populations. These offer avenues for future research.

Another possibility is that even if their early kingdoms achieved low social penetration, older ethnic groups (like the Wolof of Senegal or Kikuyu of Kenya, both successful assimilators) can point to these early antecedents as a form of cultural capital which endows them with a patina of prestige (Johnson 2004). This in turn attracts power and adherents from other groups, increasing the group's share of the population.

In many cases, historians proved to be the best sources, though scholars from a wide range of fields contributed (see Appendix V and VI). In most cases, scholars generally agreed on a ballpark founding date. However, it must be emphasised that in some cases there was significant disagreement on dates between scholars, and in these cases we were forced to compile an 'early' and 'late' founding year for the group. Arriving at an intermediate estimate in these cases was a detailed exercise which involved a great deal of trial and error, with discussions between researchers, scholars, outside experts and myself. The triangulation of dates received from scholars was not always unproblematic. In general, the older dates were selected unless there was reason to suspect these were mythical, post hoc inventions, or described a kingdom with no connection to the present ethnic group. We used textual sources to make estimates of the remaining 27 countries. Tests show that these did not differ significantly from other dates in the data.<sup>6</sup> The dates were also subject to consistency across



the dataset and checked by the author and by Anthony Smith, one of the world's leading scholars of comparative ethnic history.<sup>7</sup>

Three variables require further explication. Namely, 1) the number of secessions per state; 2) states which sprang from successful secessionist movements which defined their national boundaries in ethnic terms; and 3) states which contain a current or historic dominant minority. Secessions are based on the COW dataset: a tally per state of secessions that took place in lands contiguous to the state. 2) and 3) above have been coded by inspection in the dataset but await inter-coder reliability tests. Findings from these measures should therefore be treated with some caution. Other variables have been drawn for the year 1999 from Fearon and Laitin (2003), Vanhanen (1999), the EPR (Wimmer et al. 2009) and the Correlates of War datasets. Data on the area of states is derived from the CIA World Factbook.

## Results

Results of a series of OLS regressions on several leading measures of ethnic fractionalization appear in table 1. Related models regressing measures of the largest ethnic group in a country appear in table 2. What is missing from the tables is arguably as important as what is in them. Modern economic indicators - economic growth, urban share and growth, infant mortality and income per capita - have no predictive power in any model. Modern political factors - the presence of oil, democracy and anocracy, political instability - also fail to enter any of the models. Turning to demographic factors, population density, which is a function of total population and area, as well as economic development, is significant in some models.

Surprisingly, neither area nor population (logged) seem to predict the ethnic fractionalization of a state.

Table 1. Models of Ethnic Fractionalization, 1999 Data

	ELF (Sov. 1964)	EF (Fearon and Laitin 2003)	ELF-PREG (Wimmer et al 2009)†
(Std.) Mountainous (highest v lowest points)	.08*** (.02)	.07*** (.02)	n.s.
Early Medieval Ethnic Founding Date	-.15** (.05)	-.17** (.05)	-.17** (.06)
State Founding Date (COW)	.001** (.0003)	.001** (.0003)	n.s.
Ethnic Nationalism	n.s.	-.12* (.05)	n.s.
(Std.) Population Density	n.s.	n.s.	-.27*** (.07)
Sub-Saharan Africa	.27*** (.05)	.26*** (.05)	.35*** (.05)
North Africa / Middle East	n.s.	.14* (.06)	.36*** (.07)
Asia	n.s.	n.s.	.29*** (.06)
constant	-1.88*** (0.65)	-1.49** (0.59)	0.40*** (0.03)
R <sup>2</sup>	0.420	0.494	0.467
N	138	138	112

\*p<.05; \*\*p<.01; \*\*\*p<.001. Unstarred coefficients are significant at p<.1 level. Source: Author's analysis in Stata 10.0 using data from sources above. † Data from Wimmer et. al excludes countries where ethnicity is not held to be politically relevant.

Table 2. Models of Plurality Group Population Share, 1999 Data

	Largest Group % (Vanhanen 2000)	Largest Group % (Fearon & Laitin 2003)	Largest Politically-Relevant Group % (Wimmer et. al 2009)‡	Largest Power-Holding Politically-Relevant Group % (Wimmer <>)‡
(Std.) Mountainous	-3.62* (1.53)	-.06** (.02)	-.06** (.02)	-.06** (.02)
Early Medieval Ethnic Founding Date	16.62*** (3.93)	.14** (.04)	.10 (.06)	.11 (.06)
Ethnic Nationalism	12.80* (3.93)	.08 (.04)	n.s.	n.s.
Dominant Minority	n.s.	n.s.	n.s.	-.15 (.06)
Ex-colony	n.s.	-.09* (.04)	-.17** (.05)	-.20** (.05)
(Std.) Population Density	n.s.	n.s.	.16* (.08)	.17* (.07)
State Founding Date (COW)	-.06* (.03)	n.s.	n.s.	n.s.
Sub-Saharan Africa	-11.88** (3.91)	-.22*** (.05)	-.24*** (.06)	-.26*** (.06)
North Africa / Middle East	n.s.	n.s.	-.24*** (.06)	-.29*** (.07)
(Std.) Muslim %	n.s.	-.04* (.02)	n.s.	n.s.
constant	178.46*** (50.44)	.75*** (.07)	.83*** (.04)	.83*** (.05)
R <sup>2</sup>	.386	.482	.522	.593
N	138	138	109	109

\*p<.05; \*\*p<.01; \*\*\*p<.001. Unstarred coefficients are significant at p<.1 level. Source: Author's analysis in Stata 10.0 using data from sources above. ‡ Data from Wimmer et. al excludes countries where ethnicity is not held to be politically relevant.

The only important political factors - state founding date and ex-colony status - are historical-institutional and, by definition, time invariant. These results intimate that quicker-moving politico-economic modernization variables have a limited effect on cross-sectional

ethnic fractionalization scores. This which would appear to cast doubt on modernist theory. Yet a full test of modernism would demand a more time-sensitive dataset which could register how changes in political and economic indicators affect fractionalization over time as well as space. As noted earlier, the paucity of more than 'snapshot' ethnic census data renders such a task impossible at present.<sup>8</sup>

We find strong confirmation of the primordialist hypothesis  $H_1$ , since rougher terrain is associated with higher levels of ethnic fractionalization in all but one of the seven models (even here its sign is in the expected direction). This terrain measure is based on the difference between the highest and lowest points in a state. By contrast, an alternate measure of rough terrain based on 'proportion mountainous' used by Fearon and Laitin proved insignificant in all models.

Why might this be? The two measures align at the upper end of the elevation difference scale which suggests that it may be the differences between relatively flat and moderately mountainous countries that best predicts ethnic fractionalization. Data has yet to be collected on other aspects of rough terrain, such as forest, deserts and swamps (Fearon and Laitin 2003; 2003b). These almost certainly play a part in ethnic diversity. New Guinea is arguably the paradigm case, with its 1000 languages and highly-forested, mountainous landscape (Diamond 1997). The Caucasus, Alps and the Himalayan regions likewise come to mind as diverse and rugged, though these zones were also the sites of important trade routes and imperial frontiers which, as in the case of the Levant, also play a part in the genesis of diversity (Smith 1986: 84, 94).

Our second hypothesis,  $H^2$ , tested the effect of secession on fractionalization. This hypothesis was not confirmed in any of the models but  $H^3$  proved a very different story. The founding date of an ethnic group was discovered to be a robust predictor of ethnic

fractionalization in many models. However, its impact is not linear. In particular, states in which the largest groups were founded in the early medieval period (400 to 1100 A.D.) seem to be the least fractionalized. This may be because this was a period in which some of the earliest continuous ethnic states were formed. These include many states which emerged in the wake of the Arab Muslim conquests of the 7th-13th centuries. In addition, a number of lasting East Asian kingdoms arose at this time and many west European states emerged from the Germanic barbarian kingdoms which devastated the western Roman Empire. Those in which plurality groups formed in Antiquity (pre-450 A.D.) appear relatively fractionalized, however. This may reflect a more interrupted continuity with modern states, as in Indonesia or Iran.

The Medieval Ethnic Founding Date indicator is significant (at least  $p < .01$  level) in five of seven models. It comes close to significance ( $p < .1$ ) in the other two. The models where it is insignificant all pertain to the EPR (Ethnic Power Relations) dataset, which omits countries where ethnicity is deemed irrelevant for politics. Most of these 25 omitted cases are very homogeneous in ethnic terms. When values from the Vanhanen index are substituted for these omitted countries, this indicator regains its significance at the  $p < .05$  level. Furthermore, there is a degree of amalgamation of groups in the EPR dataset as higher levels of ethnic identity achieve greater political salience, notably in Africa. We may surmise that while older ethnic groups have had more time to assimilate surrounding groups, lowering ELF, as ethnolinguistic diversity accretes into politically-relevant ethnicity in the modern context, some of the power of ethnic history to determine fractionalization dissipates. In this case, modern differences between core and periphery in mobilizing capacity or the spread of the ideology of nationalism play a larger role (Marshall and Gurr 2003).

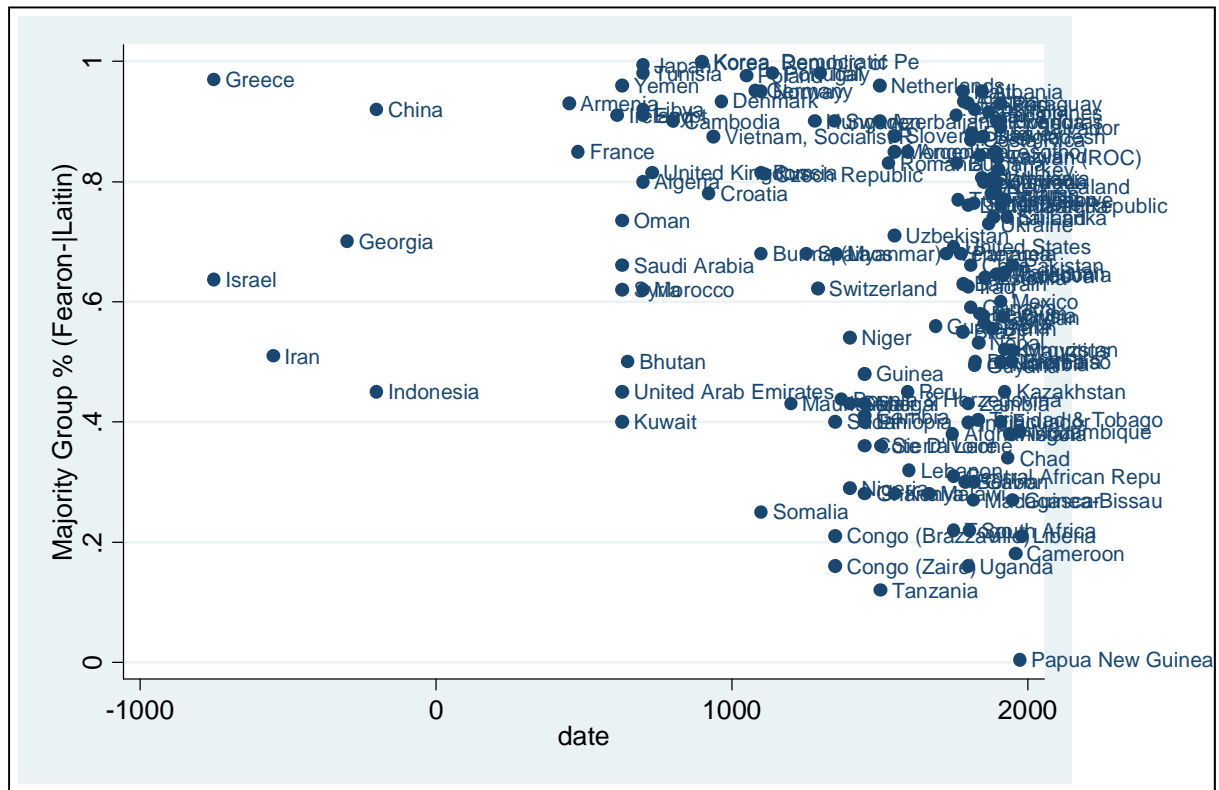
$H_{3a}$  and  $H_{3b}$  provide something of a theoretical test: is the modern state the only agent of assimilation or can this function also attach to (sometimes) premodern ethnic groups? In

fact, both are important. The zero-order correlation between the age of a state and measures of fractionalization is higher than that for the age of the largest ethnic group in the state on all but one measure. However, a portion of this is due to the correlation between state founding date and decolonization. Once a term for ex-colonies enters the analysis, the state date measure weakens and is less consistent than ethnic founding date, and, especially, early medieval ethnic founding date. When it comes to predicting the size of the largest group, the age of the state falls out of most models.

These results broadly suggest that the age of a state is important, and contributes to the progressive erosion of diversity. It is generally as strong an influence as ethnic assimilation processes which predate the state, but not on all measures. Of course, from an ethnosymbolist point of view, many modern states are continuous with premodern antecedent states which are viewed as merely 'ethnic' in this dataset, thus the importance of state founding date for the size of the largest group also makes sense. Here we find confirmation of  $H_3$  and  $H_{3b}$  but slightly weaker support for  $H_{3a}$ . This again lends somewhat more weight to ethnosymbolist and primordialist premises than modernist ones.

What is crucial is that the relationship between the age of an ethnic group and its share of a state's population is nonlinear unlike the connection between state age and the same dependent variable is linear. Figure 1 plots the relationship. Here we compare Fearon and Laitin's measure of the largest group's share of state population with the ethnic group founding date.

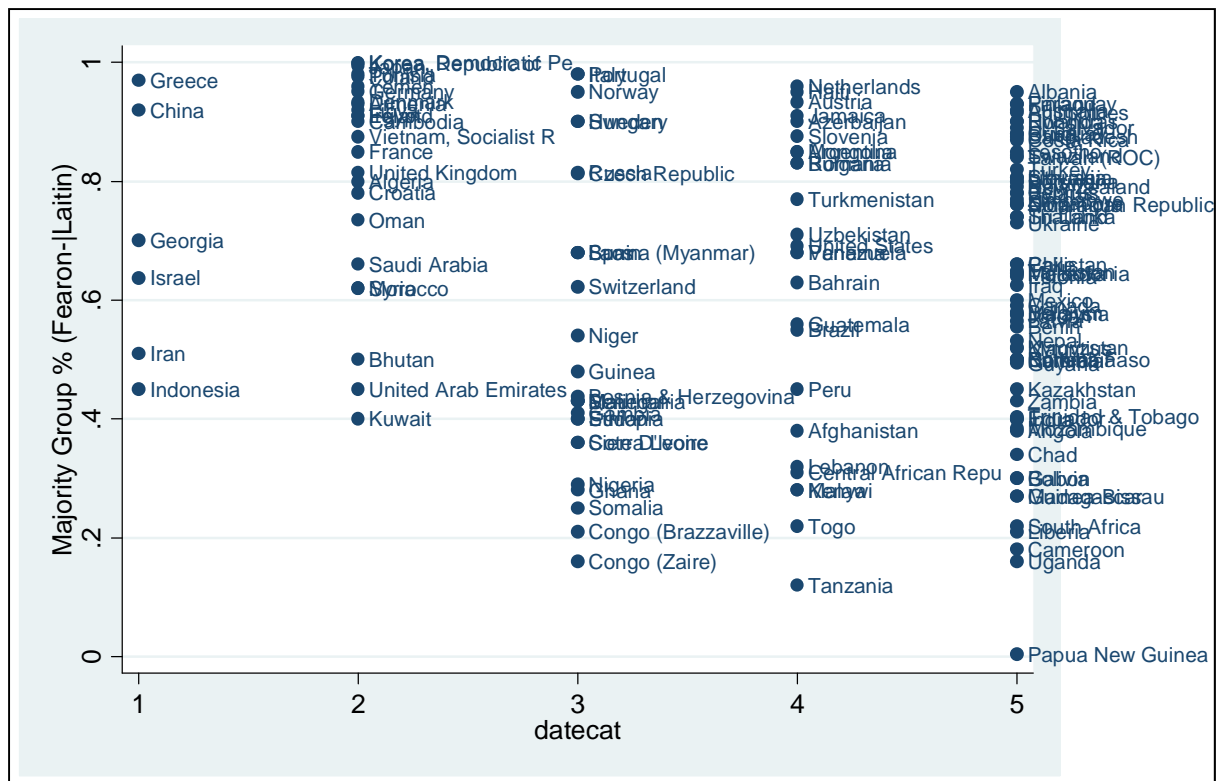
Figure 1. Proportion of Largest Ethnic Group in a State and Ethnic Group Founding Date



Further, the strength of the connection between ethnic group founding date and a group's share of the population is especially strong for the early medieval Period (category 2 in figure 2). This category contains many west European, Sunni Arab and East Asian states which are relatively homogeneous today.

Figure 2. Proportion of Largest Ethnic Group in a State and Ethnic Group Founding Date

Wave



Ethnic group founding date is a significant predictor in three of seven models in its own right, attaining greater power when cases with ancient (pre-400 A.D.) founding dates are dropped. Models with this specification appear in tables 3 and 4 in Appendix III. As Ernest Gellner noted, not all nations must have premodern ethnic navels, but equally, as Anthony Smith surmised, those which do tend to be more ethnically homogeneous (Gellner 1996; Smith 1995).

H<sub>4</sub>, which surmises that the historical wave in which a state was created affects its degree of fractionalization, is partly supported, but only in its form H<sub>4a</sub>. In other words, only the decolonization wave (1945-89) is associated with greater fractionalization. The premise that there is a relationship between other waves of state creation, such as those following the



Napoleonic War and Treaty of Versailles, and degree of ethno-political congruence, is not confirmed.

There is, however, moderate support offered to a related hypothesis, H<sub>5</sub>, that states formed on the basis of nationalist movements which defined their membership and *raison d'être* in ethnic terms should have lower levels of ethnic fractionalization than 'civic' states. The 25 states coded as ethno-national in their rationale for state formation (i.e. Israel, Hungary) were significantly more likely to be homogeneous (on two measures) than others. This lends some credence to the ethnosymbolist account as against both modernist and primordialist alternatives.

What of hypothesis 6: that countries which contain minorities of less than 25 percent of the population who hold, or have historically held, power are more fractionalized? Despite glaring examples like the Sunnis of Iraq, Kikuyu of Kenya or Americo-Liberians, there is little here to support this claim. Only when it comes to one measure in the EPR dataset - the proportion of the population comprised of members of the largest ethnic group in power - is there a relationship that approaches significance. Moreover, this is to brush aside a potential endogeneity problem in that a dominant minority will, *ipso facto*, hold power and constitute a relatively small portion of the population.

What appears most important across all models is geographic region, specifically sub-Saharan Africa. Sub-Saharan Africa is strongly associated with higher levels of fractionalization and smaller plurality groups, even with controls for rough terrain, land area, income, oil, colonialism and other political and economic factors. To a somewhat lesser extent, the Middle East/North Africa, Asia and more Muslim countries are more diverse than other parts of the globe, but only for politically-relevant ethnic diversity. This hints at the fact that in these regions, dominant ethnic groups had less of an advantage over subaltern

competitors, perhaps because of late modernization in the core or early integration efforts by subaltern ethnic leaders.

The importance of region supports H<sup>7</sup>, which lends some support to ethnosymbolist accounts which stress the role of historical particularity over primordialist and modernist arguments. That said, region could be standing in for many varieties of unspecified variable. Primordial factors like forest cover, soil variety or distance to the coast and rivers may explain sub-Saharan African exceptionalism.<sup>9</sup> Economic patterns of regions - such as reduced differences between core and periphery modernization rates, or variation in primary commodities or forms of agriculture and land use - might account for some of the power of the regional dummy variables. Inter-regional transportation networks are related to economic development indicators, but not fully so. Acknowledging this, the imprint of culture and history, whether in the form of religion, kinship systems, civilization and/or history, must be considered a major source of regional particularity.

## Conclusion

This paper attempts to explain variation in ethnic fractionalization between countries, as of 1999, assessing explanations associated with competing theories of nationalism. Established work on the importance of rugged terrain is borne out. States with a greater difference between their highest and lowest points are significantly more ethnically fractionalized and have a smaller plurality ethnic group than other states. This speaks to the importance of primordialism in laying the basis for ethnolinguistic diversity, as exemplified by the case of Papua New Guinea and its linguistic diversity.

But other factors are as or more important. Even in the presence of numerous control variables, sub-Saharan Africa remains a highly fractionalized region, and is the strongest predictor of ethnic fractionalization in the dataset. North Africa and the Middle East are not more fractionalized than other regions in ethnolinguistic terms, but are significantly more so in ethnopolitical terms. These regional effects were important across several specifications.

Most political and economic variables were not associated with ethnic fractionalization. Differences in urbanization, education, income, health, political instability and democratization between countries did not explain fractionalization or plurality size. A state's land area and population were not significant, but population density was important in several models, and predicted lower levels of diversity. These tests furnish weak evidence for modernist theories. However, the time invariant nature of the dependent variable biases the data against these predictors, so these results do not decisively reject the importance of modernizing processes in reducing diversity. Furthermore, the age of the state was significant in several specifications, suggesting that state integration is an important solvent of ethnic bonds, a cornerstone of modernist theory.

Fractionalization will tend to decrease as ethnicity and politics move toward congruence, and vice-versa. Congruence may occur for three major reasons: a) the number of political units increases; b) certain large ethnic nuclei absorb smaller groups or attract members away from them ; and/or c) ethnicity and politics come into closer alignment through secession and partition. Tests for the effects of new state creation are limited by the lack of a time dimension in global ethnic census and fractionalization data. That said, states which had experienced secessions were not noticeably more homogeneous than those which did not.

Turning to b), the evidence from the Ethnic Group Foundation Date data is quite striking. States in which the largest group has an earlier ethnic founding date are significantly more homogeneous than those with a later founding date, though the relationship only holds for plurality groups originating after the birth of Christ. This is assumed to be because groups which have existed for longer are better able to ethnically assimilate or absorb neighbours than younger groups. The pattern is especially true for states whose largest ethnic group was founded in the Early Middle Ages, between 400 and 1100 A.D. This may be because these groups were better able to maintain a continuity of consciousness into the modern era than those dating from the Ancient period.

Finally, with respect to the alignment between ethnicity and politics, we find evidence for the proposition that states formed at least partly through ethnically-defined national self-determination movements are more homogeneous than more 'civic' states originating more squarely on the basis of ideology or realpolitik. However, this finding does not explain fractionalization levels when data is restricted to politically-relevant ethnic groups.

Broadly speaking, this paper highlights the importance of ethnic, and to a lesser extent, state history, in predicting the degree of ethnic fractionalization of the world's states. While replicating findings regarding the importance of primordial ecological factors, the paper goes well beyond this. In particular, it flags the importance of ethnosymbolist theories of nationalism and group formation and offers only partial support for the importance of modernizing factors. As noted, however, the definitive test of modernist theories awaits a more time-sensitive base of evidence which can accurately track shifts in levels of ethnic fractionalization across time.

## Appendix I Descriptive Statistics for Major Variables

	<b>Mean</b>	<b>Sd</b>	<b>N</b>
<b>Ethnic date</b>	1467.551	574.2849	156
<b>State date</b>	1924.858	56.39161	148
<b>Early medieval</b>	0.1602564	0.3680252	156
<b>Elevation difference</b>	3082.667	2037.411	153
<b>Dominant minority</b>	0.0705128	0.2568338	156
<b>State wave</b>	3.243243	1.22125	153
<b>Ethnic Nationalism</b>	1.160256	0.3680252	156
<b>Secessions</b>	0.25	1.231967	156
<b>Population density</b>	162.9403	591.4514	156
<b>Sub-Saharan Africa</b>	0.2745098	0.4477325	153
<b>N. Africa/ Mideast</b>	0.124183	0.3308733	153
<b>Asia</b>	0.1437908	0.3520298	153
<b>Muslim percentage</b>	27.34059	37.61453	153
<b>Plurality (Fearon-Laitin)</b>	0.6456993	0.2434683	153
<b>Plurality (Vanhanen)</b>	69.27564	21.37336	156
<b>Plurality (EPR)</b>	0.5205075	0.3289768	147
<b>Largest Group in Power (EPR)</b>	0.4881673	0.3422546	147
<b>ELF (1964)</b>	0.4080659	0.2767858	153
<b>ELF (Fearon-Laitin)</b>	0.4741106	0.261591	153
<b>ELF (EPR)</b>	0.4135845	0.3064493	147

## Appendix II List of States for Ethnic Nationalism and Dominant Minority Variables

<i>Ethnic Basis for National Self-Determination</i>	<i>Historic Dominant Minority</i>
Albania	Bahrain
Armenia	Burundi
Azerbaijan	Ethiopia
Bangladesh	Iraq
Bulgaria	Kenya
Croatia	Liberia
Estonia	Lebanon
Finland	Niger
Georgia	Rwanda
Germany	Syria
Greece	United Arab Emirates
Hungary	
Ireland	
Israel	
Italy	
Latvia	
Lithuania	
Mongolia	
Norway	
Poland	
Romania	
Serbia	
Singapore	
Slovakia	
Slovenia	

### Appendix III Tables for Models Using Ethnic Group Founding Date Specification

Table 3. Models of Ethnic Fractionalization, 1999 Data. Ethnic Founding Date Specification.

	ELF (Sov. 1964)	EF (Fearon and Laitin 2003)	PREG (Wimmer et al <>)#
Mountainous (highest v lowest points)	39.60*** (9.61)	37.75*** (9.16)	18.50 (10.20)
(post-Antiquity) Ethnic Founding Date	15.38** (5.11)	13.96** (5.25)	n.s.
State Founding Date (COW)	10.08** (3.43)	8.66** (3.21)	n.s.
Ethnic Nationalism	n.s.	11.59* (5.04)	n.s.
Population Density	n.s.	n.s.	-38.33** (13.50)
Sub-Saharan Africa	28.29*** (4.63)	28.56*** (4.62)	39.41*** (4.71)
North Africa / Middle East	n.s.	11.30 (6.14)	25.44*** (6.19)
Asia	n.s.	n.s.	24.80*** (6.49)
constant	-2.86*** (0.73)	-2.28** (0.68)	0.31*** (0.05)
R <sup>2</sup>	0.434	0.466	0.434
N	132	132	121

\*p<.05; \*\*p<.01; \*\*\*p<.001. Certain coefficients and error terms inflated for ease of interpretation.  
Source: Author's analysis in Stata 10.0 using data from sources above. † Data from Wimmer et. al  
excludes countries where ethnicity is not politically relevant.

Table 4. Models of Plurality Group Population Share, 1999 Data. Ethnic Group Founding Date Specification.

	Largest Group % (Vanhanen 2000)	Largest Group % (Fearon & Laitin 2003)	Largest Politically-Relevant Group % (Wimmer <>)#	Largest Power-Holding Politically-Relevant Group % (Wimmer <>)#
Mountainous (highest v lowest points)	-16.64* (7.44)	-36.30*** (8.35)	-30.50** (8.61)	-31.60*** (8.46)
(post-Antiquity) Ethnic Founding Date	-17.33*** (3.86)	-13.09** (4.67)	n.s.	n.s.
Ethnic Nationalism	9.73* (3.94)	n.s.	n.s.	n.s.
Dominant Minority	n.s.	n.s.	n.s.	-16.25* (7.56)
Ex-colony	n.s.	-9.23* (3.94)	-12.68* (5.00)	-15.21** (4.86)
Polity <sup>2</sup> IV Democracy Score	16.01*** (4.40)	n.s.	n.s.	n.s.
Population Density	n.s.	n.s.	25.27* (12.32)	22.49 (12.13)
Sub-Saharan Africa	-11.84** (3.83)	-25.15*** (4.58)	-29.22*** (5.61)	-31.59*** (5.61)
North Africa / Middle East	n.s.	n.s.	-21.80*** (5.87)	-27.54*** (6.04)
Muslim %	n.s.	-11.32* (4.55)	n.s.	n.s.
constant	185.41*** (28.40)	1.85*** (.34)	.85*** (.04)	.86*** (.04)
R <sup>2</sup>	.450	.472	.506	.602
N	130	132	117	115

\*p<.05; \*\*p<.01; \*\*\*p<.001. Certain coefficients and error terms inflated for ease of interpretation. Source: Author's analysis in Stata 10.0 using data from sources above. † Data from Wimmer et. al excludes cases where largest ethnic group share of the population is coded zero.



## Appendix IV Ethnic Group Foundation Date Questionnaire

Subject: the Imagining of Ethnic Groups

Dear [expert],

I am involved in a British Academy research project which seeks to determine the approximate year in which all of the world's majority ethnic communities were first 'imagined' by a person who considered themselves a member of the group. Ethnic groups are treated as communities which share a belief in common ancestry (i.e. Smith 1991), but are larger than face-to-face tribal entities, or *gemeinschaften*. In other words, their existence must be imagined rather than derived from first-hand relationships.

As an expert on [], we thought you could suggest an approximate founding date - as generally accepted by scholars - for the following ethnic group[s]:

[ethnic group] in [state]

[ethnic group] in [state]

....

Note that this ethnic consciousness must spring from a ruler or thinker who considers themselves an 'insider'. Thus we are not referring to the perceptions of outsiders (i.e. the Roman historian Tacitus writing about 'Germans' and English, or colonial missionaries and explorers writing about African and Amerindian groups ) even if these subsequently shaped the imaginings of insiders. Nor are we concerned with the group's own claims, which often stretch back much further than what is warranted by the historical record.

Finally, we are not trying to determine the date upon which the ethnic group emerged as a social actor because this involves both the hotly-debated question of when the mass of the population became conscious and what degree of mass consciousness is necessary before we can speak of ethnicity. Thus we focus instead on the more clearly-delineated problem of founding intellectuals or elites.

Where no clear record of a founding intellectual exists, it should be drawn from the date of tribal confederation, sect formation or the creation of the kingdom from which the current group springs. In some cases, the group may undergo a change of name but with a continuity of consciousness. Our conception also allows for interruptions of consciousness and changes in the shape of homeland and form of culture, as in the Greek, Turkish or Armenian cases. However, where there is a radical shift and change of name and the current group subsequently claims to originate from an ancient kingdom, the founding year should date from the more recent imagining. Thus the Greek, Jewish and Persian founding dates would lie in the ancient period, but the Iraqi Arab consciousness should not date from Assyria or Sumer, the Shona from Great Zimbabwe, nor the Dutch from the ancient Batavians, the Portuguese from the Lusitanians or the English from the Celtic Britons.

There is undoubtedly a debate about the first stirrings of ethnic consciousness, and what individuals really meant when they first used the group's name, so if you could indicate a set of earlier and more recent dates, as well as the date you would endorse, this would be most welcome.

If you feel you cannot provide an estimate, please let me know of those whom you feel can do so.

Your help in this matter is greatly appreciated.

Sincerely yours,

[ ]

### Appendix V Sample Answers to Ethnic Group Foundation Date Survey

State	Expert	Institutional Affiliation	Answer from Expert
Albania	Prof. Nicholas Pano	Western Illinois University	1844/1845
Belarus	Prof. Andrew Savchenko	University of Rhode Island	1880-1906
Bosnia & Herzegovina	Dr. John Paul Newman	University College Dublin	1353-1391
Bulgaria	Vassilis Maragos	Bulgarian European Community Studies Association	1762
Congo (Brazzaville)	Professor Wyatt MacGaffey	Haverford College	1910
	Professor Crawford Young	Professor Emeritus of Political Science at the University of Wisconsin–Madison	14th Century
Congo (Zaire)	Professor Wyatt MacGaffey	Haverford College	1910
	Professor Crawford Young	Professor Emeritus of Political Science at the University of Wisconsin–Madison	14th Century
Croatia	Marcus Tanner	London-based writer and journalist on the Balkans	923
Czech Republic	Dr. Florian Soporan	Researcher at Centre for Transylvanien Studies	1100-1120, Kosmas of Prague, in this work "Chronica Bohemorum
Estonia	Dr. Meike Wulf	University of Maastricht	1857
Ethiopia	Prof. Asafa Jalata	University of Tennessee	15th Century

Georgia	Prof. Donald Rayfield	Queen Mary, University of London	299 BC
Germany	Dr. Len Scales	Durham University	1080
Hungary	Dr. Florian Soporan	Researcher at Centre for Transylvanien Studies	1270-1290, Simon of Keza, the bishop of Oradea, "Gesta Hunorum et Hungarorum"
Latvia	Prof. Aldis Purs	Eastern Michigan Univeristy	mid 1850s
Lithuania	Dr. Tomas Balkelis	Unversity College Dublin	1883
Macedonia	Dr. John Paul Newman	Unversity College Dublin	1896
Moldova	Professor Wim van Meurs	Radboud University Nijmegen	1917-1918
	Professor Charles King	Georgetown University	1924, the date of the founding of the Moldovan Autonomous Soviet Socialist republic, inside Soviet Ukraine.
Poland	Dr George Lukowski	University of Birmingham	1000 - 1100
Romania	Dr. Florian Soporan	Centre for Transylvanien Studies	1520-1540, Nicolaus Olachus, a Roumanian Scholar and clergyman
Russia	Dr. Tiago Ferreira Lopes	University of Lisbon	First 1702, but more commonly 1832
	A.D. Smith	LSE	1100. Russian Primary Chronicle
Rwanda	Prof. Jan Vansina	University of Wisconsin	1895
	Prof. David Newbury	Smith College	959, though contested

Serbia	Dr. John Paul Newman	University College Dublin	1113-1199
Slovakia	Dr. Erika Harris	Liverpool University	8th Century, which was decimated by the absorption of Slovaks in the year 1000 into Hungary
	Prof. Tibor Pichler	Slovak Academy of Sciences	1846
Slovenia	Dr. Erika Harris	Liverpool University	16th Century; it is argued that Slovene language has not changed since 16 century which makes it one of the oldest Slav languages, because unlike other Slav tribes, Slovenes have never moved since their early settlements when incorporated into the Roman Empire.
	Prof. Mitja Zagar	University of Ljubljana	16th Century; Very long answer with more details
Somalia	Prof. Peter Little	Emory University	11 <sup>th</sup> century (1100s)
Sudan	Prof. Peter Woodward	University of Reading	14th Century
Ukraine	Prof. Paul J. Kubicek	Oakland University	Some might say the 1648 revolts, but he thinks rather the 1870s

### Appendix VI Sample of Response Rate for Ethnic Group Founding Date Survey

State	Expert	Date	Indication of who gave the definitive answer
Albania	Dr. Anna DiLellio	08.06.2010	
	Dr. Douglas Saltmarshe	05.07.2010	
	Prof. Robert Austin	05.07.2010	
	<b>Prof. Nicholas Pano</b>		x
Belarus	Prof. Timothy Snyder	08.06.2010	
	<b>Dr. Andrew Savchenko</b>	05.07.2010	x
Bosnia & Herzegovina	Noel Malcolm	08.06.2010	-
	Prof. Brendan Simms	08.06.2010	
	Dr. Karl Bahm	21.06.2010	
	Prof. Francine Friedman	05.07.2010	
	Prof. Edith Klein	05.07.2010	
	Prof. Ralph Bogert	11.08.2010	
	<b>Dr. John Paul Newman</b>	23.08.2010	x
	Dr. Uğur Ümit Üngör	14.09.2010	
Bulgaria	<b>Dr. Vassilis Maragos</b>	21.06.2010	x
Congo (Brazzaville)	Prof. Georges Nzongola-Ntalaja	08.06.2010	
	Prof. Michael Schatzberg	08.06.2010	
	<b>Prof. Crawford Young</b>	09.06.2010	x
	<b>Prof. Wyatt MacGaffey</b>	09.06.2010	x
Congo (Zaire)	Prof. Georges Nzongola-Ntalaja	08.06.2010	
	Prof. Michael Schatzberg	08.06.2010	
	<b>Prof. Crawford Young</b>	09.06.2010	x
	<b>Prof. Wyatt MacGaffey</b>	09.06.2010	x
Croatia	Noel Malcolm	08.06.2010	
	Dr. Nina Caspersen	18.06.2010	

	Prof. Ivo Goldstein	21.06.2010	
	<b>Mr. Marcus Tanner</b>	21.06.2010	x
	Prof. Edith Klein	05.07.2010	
Czech Republic	Dr. Karl Bahm	21.06.2010	
	Dr. Stefan Auer	28.06.2010	
	Dr. Florian Soporan	25.07.2010	
Estonia	Prof. Antonio Ramos dos Santos	21.06.2010	
	<b>Dr. Meike Wulf</b>	21.06.2010	x
Ethiopia	<b>Prof. Asafa Jalata</b>	08.06.2010	x
	Dr. Ruth Iyob	08.06.2010	
	Dr. Sarah Vaughan	21.06.2010	
	Prof. Donald Crummey	05.07.2010	
Georgia	Prof. Ronald Grigor Suny	08.06.2010	
	Prof. David Braund	08.06.2010	
	<b>Prof. Donald Rayfield</b>	23.08.2010	x
Germany	Dr. James Bjork	08.06.2010	
	Prof. David Blackbourn	05.07.2010	
	Prof. Simon Kitchen	05.07.2010	
	Prof. Steven Ozment	11.07.2010	
	Prof. John Breuilly	23.08.2010	
	<b>Dr. Len Scales</b>	23.08.2010	x
	Prof. Patrick Geary	23.08.2010	
Hungary	Prof. Martyn Rady	08.06.2010	
	Dr. Stefan Auer	28.06.2010	
	Dr. Florian Soporan	25.07.2010	
Latvia	Prof. Antonio Ramos dos Santos	21.06.2010	
	Dr. Meike Wulf	21.06.2010	
	<b>Prof. Aldis Purs</b>	22.06.2010	x
Lithuania	Prof. Timothy Snyder	08.06.2010	
	Prof. Antonio Ramos dos Santos	21.06.2010	
	Dr. Meike Wulf	21.06.2010	
	Dr. Egle Rindzeviciute		
	<b>Dr. Tomas Balkelis</b>	23.08.2010	x
Macedonia	Prof. Daniel Ogden	08.06.2010	

	Prof. Roderick Beaton	08.06.2010	
	Dr. Philip Carabott	09.06.2010	
	Prof. Andrew Rossos	11.08.2010	
	<b>Dr. John Paul Newman</b>	23.08.2010	x
Moldova	Prof. Nicholas Pappas	08.06.2010	
	Prof. Don Dyer	08.06.2010	
	<b>Prof. Charles King</b>	09.06.2010	x
	Prof. Valentina Iepuri	09.06.2010	
	<b>Prof. Wim van Meurs</b>	21.06.2010	x
Poland	Prof. Timothy Snyder	08.06.2010	
	<b>Dr. George Lukowski</b>	05.07.2010	x
	Dr. Anita Prazmowska	05.07.2010	
Romania	Prof. Klaus Bochmann	08.06.2010	
	Prof. Dennis Deletant	08.06.2010	
	Dr. Radu Cinpoes	05.07.2010	
	Prof. Ioan-Aurel Pop	05.07.2010	
Russia	Prof. Henry Hale	08.06.2010	
	Prof. Geoffrey Hosking	08.06.2010	
	<b>Dr. Tiago Ferreira Lopes</b>	21.06.2010	x
	Prof. Roman Szporluk	21.06.2010	
	Dr. Andreas Umland	28.06.2010	
	Dr. John Westwood	05.07.2010	
	Prof. Robert Legvold	05.07.2010	
Rwanda	Prof. Patrick Chabal	28.06.2010	
	Dr. Devon Curtis	04.08.2010	
	Prof. René Lemarchand	11.08.2010	
	Dr. Liisa Malkki	11.08.2010	
	Prof. Thomas Spear	23.08.2010	
	Prof. Margee Ensign	23.08.2010	
	Dr. Redie Bereketeab	02.09.2010	
	Prof. Gregory Mann	10.09.2010	
	Dr. Gerard Prunier	10.09.2010	
	Ms. Danielle De Lame	14.09.2010	
	Prof. David Maxwell	21.09.2010	
	Prof. Helena Pohlandt-McCormick	21.09.2010	



	Prof. Wim van Binsbergen	21.09.2010	
	Prof. Michael Dietler	21.09.2010	
	Prof. R. Brian Ferguson	21.09.2010	
	Prof. Mahmood Mamdani	23.09.2010	
	Prof. Johann Pottier	24.09.2010	
	Prof. Scott Straus	24.09.2010	
	Prof. Lee Ann Fujii	24.09.2010	
	Prof. Aimable Twagilimana	24.09.2010	
	Prof. Timothy Longman	24.09.2010	
	Prof. Learthen Dorsey	24.09.2010	
	Prof. Catherine Newbury	26.09.2010	
	Dr. Phil Clark	27.09.2010	
	Dr. Jude Murison	27.09.2010	
	Prof. David Newbury	23.10.2010	
	Prof. Eugenia Herbert	23.10.2010	
	Dr. David Henige	23.10.2010	
	Prof. Scott MacEachern	23.10.2010	
	<b>Prof. Jan Vansina</b>	<b>25.10.2010</b>	<b>x</b>
	Dr. Helen Hintjens	25.10.2010	
	Prof. Richard Roberts	25.10.2010	
	Prof. Kenneth Harrow	25.10.2010	
Serbia	Noel Malcolm	08.06.2010	
	Dr. Eric Gordy	08.06.2010	
	Prof. Francine Friedman	05.07.2010	
	Prof. Edith Klein	05.07.2010	
	Prof. Ralph Bogert	11.08.2010	
	Prof. Stevan Pavlowitch	23.08.2010	
	Mr. Ioannis Armakolas	02.09.2010	
	<b>Dr. John Paul Newman</b>		<b>x</b>
	Dr. Uğur Ümit Üngör	14.09.2010	
Slovakia	Dr. Erika Harris	08.06.2010	<b>x</b>
	Dr. Karl Bahm	21.06.2010	

	<b>Prof. Tibor Pichler</b>	21.06.2010	x
Slovenia	<b>Dr. Erika Harris</b>	08.06.2010	x
	<b>Prof. Mitja Zagar</b>	21.06.2010	x
Somalia	Prof. Patrick Chabal	28.06.2010	
	Prof. Jane Lewis	05.07.2010	
	<b>Prof. Peter Little</b>	04.08.2010 and 25.08.2010	x
	Prof. James Mayall	04.08.2010	
	Prof. Thomas Spear	23.08.2010	
Sudan	Dr. Stephanie Beswick	08.06.2010	
	Prof. Gaby Warburg	08.06.2010	
	Dr. Ahmed Al-Shahi	09.06.2010	
	Prof. Richard Lobban	21.06.2010	
	Ms. Mareike Schomerus	05.07.2010	
	Prof. Donald Crummey	05.07.2010	
	Dr. Joost Fontein	04.08.2010	
	Prof. James Mayall	04.08.2010	
	Prof. Bruce Hall	23.08.2010	
	Prof. Jay Spaulding	23.08.2010	
	Prof. Rex Sean O'Fahey	23.08.2010	
	Dr. Redie Bereketeab	02.09.2010	
	Dr. John Cooper	02.09.2010	
	Dr Mark Leopold	10.09.2010	
	Prof. Justin Willis	10.09.2010	
	<b>Prof. Peter Woodward</b>	10.09.2010	x
	Dr. Gerard Prunier	10.09.2010	
	Prof. Louis Brenner	14.09.2010	
Ukraine	Prof. Philipp Ther	08.06.2010	
	Prof. Henry Hale	08.06.2010	
	Prof. Roman Szporluk	21.06.2010	
	Dr. Alexander Motyl	21.06.2010	
	Prof. Orest Subtelny	05.07.2010	
	Prof. David Marples	05.07.2010	
	Dr. Frank Sysyn	25.07.2010 and 11.08.2010	
	<b>Prof. Paul J. Kubicek</b>	23.08.2010	x

## References

- Abu-Lughod, Janet L. 1989. *Before European hegemony : the world system A.D. 1250-1350*. New York: Oxford University Press.
- Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quarterly Journal of Economics* 114 (November):1243–84.
- Anderson, Benedict. 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Armstrong, John. 1982. *Nations Before Nationalism*. Hill: University of North Carolina Press.
- Bayly, C. 2004. *The Birth of the Modern World, 1780-1914: Global Connections and Comparisons*. Oxford: Blackwell.
- Birch, Anthony H. 1989. *Nationalism and National Integration*. London: Unwin Hyman.
- 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.
- Brubaker, R. 1998. "Migrations of ethnic unmixing in the "New Europe"." *International Migration Review* 32 (4):1047-65.
- Brubaker, Rogers. 1996. *Nationalism Reframed: Nationhood and the National Question in the New Europe*. Cambridge: Cambridge University Press.
- Buhaug, H., L. E. Cederman, and J. K. Rod. 2008. "Disaggregating ethno-nationalist civil wars: A dyadic test of exclusion theory." *International Organization* 62 (3):531-51.
- Cederman, L. E., A. Wimmer, and B. Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62 (1):87-+.
- Chanda, A., and L. Putterman. 2007. "Early starts, reversals and catch-up in the process of economic development." *Scandinavian Journal of Economics* 109 (2):387-413.
- Collard, I., and R. Foley. 2002. "Latitudinal Patterns and Environmental Determinants of Recent Human Cultural Diversity: Do Humans Follow Biogeographic Rules?" *Evolutionary Ecology Research* 4 (3):371-83.
- Collier, P., and A. Hoeffler. 1998. "On economic causes of civil war." *Oxford Economic Papers-New Series* 50 (4):563-73.
- . 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.
- Connor, Walker. 1994b. "A Nation is a Nation, is a State, is an Ethnic Group, is a ..." In *Nationalism*, ed. J. Hutchinson and A. D. Smith. Oxford: Oxford University Press.
- Demeny, P., and G. McNicoll, eds. 2006. *The Political Economy of Global Population Change, 1950-2050*. New York: Population Council.
- Deutsch, Karl. 1966 (1953). *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality, 2nd Ed*. Cambridge, MA: M.I.T. Press.
- Diamond, J. 1997. *Guns, Germs, and Steel: the Fates of Human Societies*. New York; London: W.W. Norton & Co.
- Easterly, W., and R. Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 111 (4):1203-50.

- 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 David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1):75-90.
- Fearon, James, and David Laitin. 2003b. "'Additional tables for "ethnicity, insurgency, and civil war" '." Stanford University.
- Francis, E.K. 1976. *Interethnic Relations: An Essay in Sociological Theory*. New York, NY: Elsevier Scientific.
- Geertz, Clifford. 1963. "The integrative revolution. Primordial sentiments and civil politics in the new states." In *Old Societies and New States. The Quest for Modernity in Asia and Africa.*, ed. C. Geertz. New York, NY: Free Press.
- Gellner, E. 1983. *Nations and Nationalism*. Oxford: Blackwell.
- . 1996. "Do Nations Have Navels?" *Nations and Nationalism* 2 (3):366-71.
- Giddens, Anthony. [1985] 1996. *The Nation-State and Violence: Volume Two of A Contemporary Critique of Historical Materialism*. Cambridge: Polity Press.
- Hastings, Adrian. 1997. *The Construction of Nationhood: Ethnicity, Religion and Nationalism*. Cambridge & New York: Cambridge University Press.
- Hechter, M. 2000. *Containing Nationalism*. Oxford: Oxford University Press.
- Hobsbawm, E. J. 1990. *Nations and Nationalism since 1780. Programme, Myth, Reality*. Cambridge: Cambridge University Press.
- Johnson, Nancy Kwang. 2004. "Senegalese into Frenchman or Peasants into Senegalese?: The Politics of Language, Culture and Assimilation: A Colonial and Post-Independence Critique (Senegal)", *International Studies Association*. Montreal.
- Kaufmann, E. P. 2004. "Dominant ethnicity: from background to foreground." In *Rethinking Ethnicity: Majority Groups and Dominant Minorities*, ed. E. Kaufmann. London: Routledge.
- Laitin, David D. 2007. *Nations, states, and violence*. Oxford ; New York: Oxford University Press.
- Lujala, P., N. P. Gleditsch, and E. Gilmore. 2005. "A diamond curse? Civil war and a lootable resource." *Journal of Conflict Resolution* 49 (4):538-62.
- Marshall, Monty G., and Ted Robert Gurr. 2003. *Peace and Conflict 2003: A Global Survey of Armed Conflicts, Self-Determination Movements, and Democracy*. College Park, MD: Integrated Network for Societal Conflict Research.
- McNeill, William Hardy. 1991. *The rise of the West : a history of the human community : with a retrospective essay*. Chicago: University of Chicago Press.
- Michalopoulos, Stelios. 2010. "The Origins of Ethnolinguistic Diversity: Theory and Evidence." Tufts University.
- Montalvo, J. G., and M. Reynal-Querol. 2005. "Ethnic polarization, potential conflict, and civil wars." *American Economic Review* 95 (3):796-816.
- Moore, J. L., L. Manne, T. Brooks, N. D. Burgess, R. Davies, and C. Rahbek. 2002. "The Distribution of Cultural and Biological Diversity in Africa." *Proceedings of the Royal Society of London* 296 (1501):1645-53.
- Nairn, Tom. 1977. *The Break-Up of Britain*. London: NLB.
- Nettle, D. . 1996. "Language Diversity in West Africa." *Journal of Anthropological Archaeology* 15 (4):403-38.
- Nichols, J. . 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

- Nunn, N. 2008. "The long-term effects of Africa's slave trades." *Quarterly Journal of Economics* 123 (1):139-76.
- Ostby, Gudrun, Ragnhild Nordas, and Jan Ketil Rod. 2009 "Regional Inequalities and Civil Conflict in Sub-Saharan Africa " *International Studies Quarterly* 53 (2):301-24.
- Roeder, P. 2010. *Ethnolinguistic Fractionalization (ELF) Indices, 1961 and 1985* 2001 [cited 2010]. Available from <http://weber.ucsd.edu/~proeder/elf.htm>.
- Saideman, Steve, and R. W Ayres. 2008. *For Kin or Country: Xenophobia, Nationalism, and War* New York, N.Y.: Columbia University Press.
- Salter , Frank K. 2001. "A Defense and an Extension of Pierre van den Berghe's Theory of Ethnic Nepotism." In *Evolutionary Theory and Ethnic Conflict*, ed. P. James and D. Goetze. Westport, CT and London: Praeger.
- . 2003. *On Genetic Interests: Family , Ethny , and Humanity in an Age of Mass Migration*. Berlin: Peter Lang.
- 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*.
- Shils, Edward. 1995. "Nation, Nationality, Nationalism and Civil Society." *Nations and Nationalism* 1 (1):93-118.
- Smith, Anthony. 1986. *The Ethnic Origins of Nations*. Oxford: Blackwell.
- . 1991. *National Identity*. London: Penguin.
- Smith, Anthony D. 1995. "Gastronomy or Geology? The Role of Nationalism in the Reconstruction of Nations." *Nations and Nationalism* 1 (1):3-23.
- Sutherland, W.J. 2003. "Parallel Extinction Risk and Global Distribution of Languages and Species." *Nature* 423 (6937):276-9.
- Thomas, William Isaac, and Florian Znaniecki. 1918. *The Polish peasant in Europe and America; monograph of an immigrant group*. Chicago, Ill.,: The University of Chicago press.
- 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. "Common Wealth: Economics for a Crowded Planet." *Journal of Peace Research* 45 (6):856-7.
- Van den Berghe, P.L. 2002. "Multicultural democracy: can it work?" *Nations and Nationalism* 8 (4):433-49.
- Van Dyke, Vernon. 1995. "The Individual, the State, and Ethnic Communities in Political Theory." In *The Rights of Minority Cultures*, ed. W. Kymlicka. Oxford: Oxford University Press.
- Vanhanen, T. 1999a. "Domestic Ethnic Conflict and Ethnic Nepotism: A Comparative Analysis." *Journal of Peace Research* 36 (1):55-73.
- . 1999b. *Ethnic Conflicts Explained by Ethnic Nepotism*. Stamford, CT: Jai Press.
- Weber, Eugen Joseph. 1976. *Peasants into Frenchmen: the Modernization of Rural France, 1870-1914*. Stanford, California: Stanford University Press.
- Weber, Max. 1996. "The Origins of Ethnic Groups." In *Ethnicity*, ed. A. D. Smith and J. Hutchinson. Oxford: Oxford University Press.
- Wimmer, A., L. E. Cederman, and B. Min. 2009. "Ethnic Politics and Armed Conflict: A Configurational Analysis of a New Global Data Set." *American Sociological Review* 74 (2):316-37.
- Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflict*. Cambridge: Cambridge University Press.

Zimmer, O., and Len Scales, eds. 2005. *Power and the Nation in History*. Cambridge: CUP.  
Zimmer, Oliver. 2003. *Nationalism in Europe: 1890-1940*. London: Palgrave/Macmillan.

## Endnotes

---

<sup>1</sup> Initial research on ELF used data culled 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.)

<sup>2</sup> Leading works in this area include (1986; 1991), Gellner (1983), Anderson (1983); Hobsbawm [1990] 1993, and Connor (1994).

<sup>3</sup> See (Weber 1996): 35; (Smith 1991): 40; (Francis 1976): 6; (Van Dyke 1995)

<sup>4</sup> (Eriksen 1993): 12; Francis 1976: 6.

<sup>5</sup> For discussion of this in relation to ELF, see Wimmer, Cederman and Min 2009; and Fearon and Laitin 2006.

<sup>6</sup> A dummy variable for estimated dates proved insignificant, and results without the estimated dates closely matched those for the full dataset.

<sup>7</sup> The dataset will be released upon publication of this article.

<sup>8</sup> Though Green (2011), using Roeder's unique two time-points dataset, finds a relationship between urbanization and decreases in ethnic fractionalization between 1961 and 1987.

<sup>9</sup> Landlocked countries were more homogeneous, but the relationship only approached significance in one model.