European Sociological Review Advance Access published November 4, 2014
European Sociological Review, 2014, 1–18
doi: 10.1093/esr/jcu078
Original article

Nation Building. A Long-Term Perspective and Global Analysis

Andreas Wimmer*

Princeton University, 147 Wallace Hall, Princeton, NJ 08544, USA

*Corresponding author. Email: awimmer@princeton.edu

Submitted 22 February 2014; revised 17 July 2014; accepted 22 September 2014

Abstract

Why have some states been captured by specific ethnic elites and their clienteles, excluding all others from access to government power? Conversely, what explains political inclusion across ethnic divides or, in other words, successful 'nation building'? Assuming a relational theoretical perspective, I argue that high state capacity to deliver public goods, well developed voluntary organizations, and low levels of linguistic diversity enhance nation building because they make it easier to extend networks of political alliances across an entire territory. Contemporary state capacity and linguistic diversity are in turn related to levels of state formation achieved during the late 19th century. On average, such long-term factors of political development are more important for explaining contemporary nation building than political institutions (including democracy) or the legacies of imperial rule. This is demonstrated on the basis of a cross-national data set covering all countries of the world since 1945.

Previous research has shown that ethnic inequality and exclusion increase the likelihood of civil war (Wimmer *et al.*, 2009) and hamper economic growth (Birnir and Waguespack, 2011) as well as public good provision (Baldwin and Huber, 2011). It remains unclear, however, how to understand variation in levels of ethno-political inclusion. Why has nation building in Switzerland, for example, led to the successful integration of French-, German-, and Italian-speaking political elites and populations into an informal power sharing arrangement, while countries such as Iraq, Liberia, Bolivia, or South Africa have been ruled ethnocratically by Sunni, Americo-Liberian, Creole, and White elites during most of their history?

This article tests various competing hypotheses using a global data set with information on all countries of the world since 1945. It shows that long-term, slow-moving processes of political development—highlighted by an earlier generation of scholarship—are indeed crucial for explaining ethno-political inclusion. This political development argument is based on a relational approach that focuses on the conditions under which alliance and support relationships between state elites and the population at large are more likely to stretch across a country's entire territory—and thus across ethnic divides. A political economy, an organizational, and a communication aspect of these alliance networks are distinguished and three corresponding conditions for nation building identified.

First, state elites capable of providing basic public goods—including security, infrastructure, and rule of law—across a territory can establish far-reaching networks of support across different ethnic groups, rather than restricting such networks to their own ethnic clientele. Second, dense networks of voluntary organizations will make it less likely that political elites build ethnic patronage pyramids to mobilize followers because they

© The Author 2014. Published by Oxford University Press. All rights reserved.
For permissions, please e-mail: journals.permissions@oup.com

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

can instead rely on these organizational networks to build multi-ethnic coalitions. Third, networks of alliances can proliferate more easily across a territory the more homogenous the population in linguistic terms and the lower therefore communicative transaction costs of establishing political relationships.

State capacity to provide public goods and linguistic diversity are path-dependent phenomena and change slowly over generations, not years. In a second step of analysis that reaches deeper into the past, I will show that both variables are in turn related to late 19th century levels of state formation because centralized states homogenized the population in linguistic terms and left a legacy of bureaucratic capacity that facilitated public goods provision in the post-war era.

This political development argument contrasts with major alternative explanations of nation building, including those centred on democratic regimes or particular democratic institutions, most famously proportional representation and parliamentarianism (Lijphart, 1977), colonial legacies such as minority over-representation in government or the army (Horowitz, 1985, Chapters 11–13), or globalization processes such as the increasing hegemony of minority rights regimes (Kymlicka, 2007).

Using a pooled time-series cross-sectional data set with 155 countries since 1945 or independence, I evaluate these competing claims as well as possible, given data limitations. I do not find consistent support for the imperial legacy, globalization, and democracy hypotheses, but confirm the political development argument, using data on pre-colonial levels of state centralization assembled by anthropologists. These results are mostly robust to a country-fixed effect design that takes care of unobserved heterogeneity, to subsample analysis of more and less homogenous countries, to a coding of independent and dependent variables (where available) with an 80-year lag between them, thus minimizing reverse causation problems, as well as further endogeneity tests using 'placebo' regressions with switched dependent and independent variables and long time lags (reported in the Supplementary Appendix).

Defining Nation Building

Among the large majority of countries where ethnicity is politically salient, we find substantial variation in the distribution of power between representatives of different ethnic communities. In some countries, large ethnic groups remain outside the alliance and support networks stretching from the seats of government down to the villages of the hinterland. In Syria, for example, the Asad clan and their fellow Alawite held a firm grip on all high-level government and military positions over the past decades. In other countries, more inclusionary ethno-political configurations have emerged and most of the population is integrated into the web of alliances and support centred on a national government—the definition of 'political inclusion' or 'nation building' that I adopt for the present discussion (for similar understandings of nation building, see Bendix, 1964: pp. 18–19; Lemarchand, 1972: p. 68). Examples are Switzerland, Malaysia, or Burkina Faso—all ethnically heterogeneous countries.

Figure 1 (inspired by the well-known polity model of Tilly, 1975) shows an inclusionary and an exclusionary ethno-political configuration of power. Nodes represent political actors (organizations or individuals), lines describe exchange relationships (or alliance), and actors higher up in the drawing wield more political power, those at the top representing national government. As will be discussed in greater detail below, we can operationalize this concept of political inclusion by measuring the percentage of the population that is excluded from these webs of alliances (the white nodes in the right hand panel), i.e. is not directly or indirectly linked to the highest level of government.

Two clarifications follow from that. First, both countries represented in Figure 1 are characterized by the same ethno-demographic structure. Thus, ethnic diversity and ethnic inclusion are conceptually and empirically distinct. Second, political inclusion and nation building also need to be distinguished from democratization (in contrast e.g. to Dobbins, 2003/2004). Access to state power can also be organized through ethnic patronage networks within a non-democratic one-party regime (Rothchild, 1986; as in Burkina Faso).

A Theory of State Formation, Political Development, and Nation Building

How do we explain the variation in degrees of ethno-political inclusiveness illustrated by Figure 1? The following gives an overview of the major arguments that have been proposed in the literature in sociology and political science and describes the variables and data sources used to test them. I first outline a relational theory of state formation and nation building.

Nation Building as a Consequence of Political Development

It seeks to understand how the political arena transformed over generations, rather than years or electoral cycles—in line with the now rather unfashionable political development theories introduced by the likes of

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

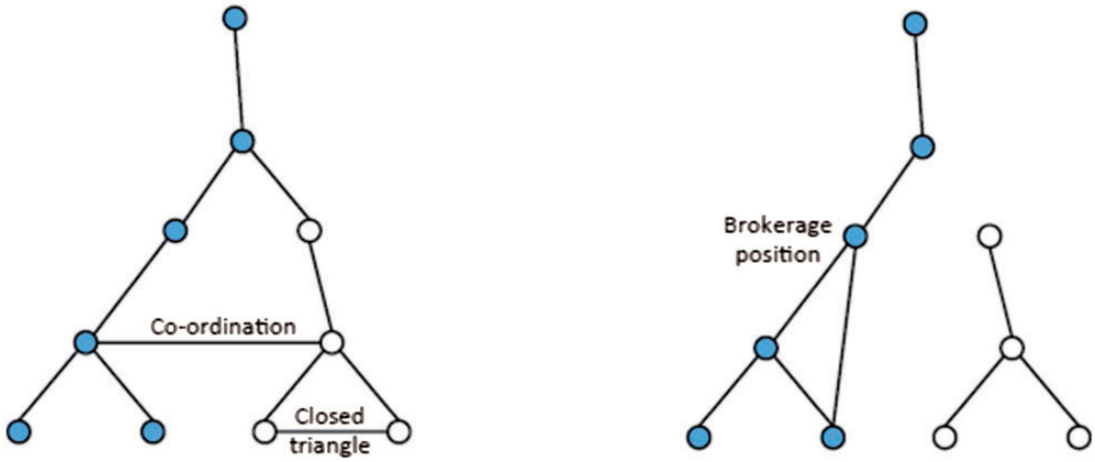

Figure 1. An inclusionary (left panel) and exclusionary (right panel) configuration of power

Deutsch (1953), Geertz (1963), Reinhard Bendix (1964), and others. Adopting a relational perspective on politics (Tilly, 2006, among others), I focus on the long-term evolution of the networks of political alliances—both formal and informal—that bind state elites and the population at large together (see again Figure 1).

In the following, I briefly elaborate the micro foundations of the argument, relying on exchange theoretic principles (cf. Blau, 1964/1986). We can distinguish between three basic and irreducible aspects of an exchange relationship: the political economy aspect refers to the resources that actors trade; the organizational aspect describes whether and how the relationship is institutionalized; the exchange of meaning and information within a dyad represent the communication aspect. For each aspect, we can formulate a hypothesis regarding the conditions under which an exchange relationship is more likely to cross existing ethnic divides, thus leading to nation building in the aggregate.

Seen from the resource point of view, government individuals and organizations offer public goods and influence over political decisions, while non-elite individuals and organizations offer political support (including votes), military loyalty, and taxes. Depending on how control over and demand for these various resources is distributed over actors, they will establish a mutually beneficial exchange relationship with each other (the micro foundations of this model have been elaborated in Kroneberg and Wimmer, 2012).

In states with a large capacity to provide public goods, government elites represent more attractive exchange partners and more non-elite individuals and organizations (from whatever ethnic background) will want to establish ties with them, offering military and political support and reducing their resistance to taxation. In other words, nation building is easier in states capable of providing public goods (ibid.; for a related analysis, see Levi, 1988). Conversely, the rulers of weaker states will have to limit the circle of recipients of public goods. Because in modern nation-states government elites are supposed to care for 'their own people', they will choose an ethnically defined circle of beneficiaries. Public goods then become ethnic pork (cf. Fearon, 1999; or 'excludable club goods' in Congleton, 1995), alliance networks compartmentalize along ethnic divides, and parts of the population will remain disconnected from the exchange networks centred on governing elites (Wimmer, 2002).

From the organizational point of view, resources are exchanged through different organizational channels, which can be of an informal nature (such as in patronage networks) or formalized into relationships between organizational units (such as in inter-organizational alliances). Well-developed networks of voluntary associations (clubs, trade unions, party youth organizations, etc.) will increase the chances that exchange relationships will cut across ethnic divides and thus prevent the politicization of ethnicity and the emergence of exclusionary power structures. Why should this be the case?

Voluntary organizations have a 'built-in' tendency to cross ethnic divisions—obviously a probabilistic tendency, rather than an iron law (for examples, see Varshney, 2003). Associational networks are characterized by a high number of horizontal 'co-ordination' positions (Hillmann, 2008; see Figure 1) and many closed

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

triangles (Baldassari and Diani, 2007). Patronage structures, by contrast, are associated with vertical brokerage structures (Gould and Fernandez, 1989) and open triangles (the *locus classicus* is Scott, 1972). Political alliances therefore tend to spread horizontally across an ethnic divide when there is a high density of voluntary organizations, while they tend to spread vertically in patronage systems and thus have a higher likelihood of forming along ethnic divisions (for counter examples, see Scott, 1972; Dunning and Nilekani, 2013).[1] Cross-ethnic exchange relationships will thus be more rare and ethno-political exclusion more likely.

The communicative aspect describes how actors exchange information about the resources they offer and demand. Communication can be more or less costly in terms of the efforts needed to minimize errors, involve more or less complete information, and can be based on more or less trust. Linguistic heterogeneity tends to slow-down the proliferation of network ties across a territory because initiating, coordinating, and stabilizing exchange relationship are more costly and difficult due to misinformation about others' resources, demands, and intentions, which in turn decreases generalized trust in strangers (for evidence, see Knack and Keefer, 1997: p. 1281). This argument formed part of Deutsch's (1966) theory of nationalism, which posited that successful nation building depends on an even social mobilization of the population, which might be hampered by communication barriers.

In sum, high infrastructural capacity to deliver public goods, well-developed organizational networks, and low communicative barriers increase the likelihood that encompassing networks of political alliances will eventually emerge. But couldn't the causal arrow point in the other direction as well? In the empirical analysis below, I will use a variety of temporal lags between dependent and independent variables as well as inverse 'placebo' regressions (reported in the Supplementary Appendix) in which independent variables are turned into dependent variables with a time lag between them, to test whether such reverse causation drives a possible association between political development and ethno-political inclusion.

Political Development as Legacy of Previous State Formation

How can we explain why different countries have achieved different levels of linguistic homogeneity, organizational development, and the infrastructural capacity to provide public goods? Pushing the analysis further down the road of macro-political history I hypothesize that linguistic homogeneity and infrastructural capacity are influenced by past levels of state building—in the case of the post-colonial world, those achieved before colonization and the emergence of the colonial state (in line with Englebert, 2000; Bockstette *et al.*, 2002; Gennaioli and Rainer, 2007). Historic levels of state building affect these two political development variables in the following ways.

First, centralized states with strong bureaucratic capacity offered incentives to adopt the culture and language of the dominant *Staatsvolk* to be able to communicate with state officials in their language or become a civil servant oneself. Linguistic heterogeneity should thus decrease over time (see Weber, 1979 for the French case; for other explanations of linguistic homogeneity, see Nunn, 2008; Ahlerup and Olsson, 2012; Michalopoulos, 2012). Second, centralized states also provided institutional capacity on which later administrations (including colonial ones) could build and which allowed for the infrastructural integration of the territory and an effective provision of public goods in the post-war era (for African evidence, see Gennaioli and Rainer, 2007). Historical levels of state centralization are not linked in the same way to the development of voluntary organizations. On the one hand, voluntary organizations often did historically emerge in opposition to a centralizing state (Mann, 1993). On the other hand, however, strongly centralized states in some other cases policed or even suppressed the emergence of such organizations (as was the case in Japan), thus producing, in the aggregate and on average, no clear pattern of relationships.

Figure 2 summarizes the argument developed so far and indicates which of the various variables can be directly observed and which one will remain unobserved (in italics). Overall, the theory developed here represents a 'pre-colonial legacy' argument of sort (similar to Herbst, 2000), to be distinguished from the various colonial legacy hypotheses that I will discuss below. The next section introduces the measurements used to test the political development argument.

Measurements

Data on the density of voluntary organizations is provided by Schofer and Longhofer's (2011) count of non-governmental organizations, based on an encyclopedia of NGOs. It covers many countries and all years from 1970 to 2005 (for descriptive statistics, see Supplementary Table S1). Unfortunately, these data are not including government-controlled voluntary organizations (prominent in many communist countries and beyond), but only classical NGOs.

There is also no ideal measurement of state capacity (see discussion in Hendrix, 2010). I rely on two different

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

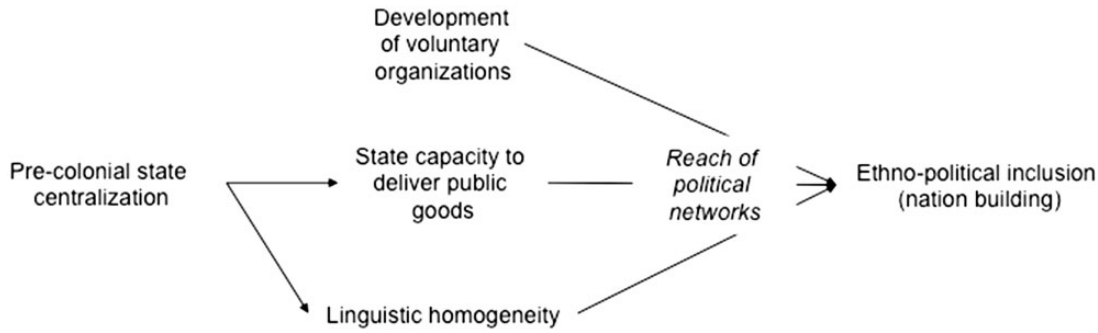

Figure 2. A theory of state formation, political development, and nation building. Note: Unmeasured intervening variable in italics

measurement of public good provision. The more commonly used is adult literacy rates (e.g. La Porta *et al.*, 1999; Gennaioli and Rainer, 2007), which are strongly influenced by public school systems as well as state-led alphabetization campaigns.[2] The data were assembled from various sources (see Wimmer and Feinstein, 2010). They refer to the percentage of alphabetized adults in the overall population and are provided for most countries of the world since the early 19th century.

The second, less often used measurement is the length of railroad tracks per square kilometre. This measurement is also available since the early 19th century (data again from Wimmer and Feinstein, 2010), which will allow addressing endogeneity issues by introducing long time lags. Railroads often represent a public good in themselves, provided and maintained by the state—though some railroads also served military purposes or to transport natural resources to the coast or ran (at least initially) without any state subsidies. Still, railway length comes closer to a measurement of public goods provision than tax rates or government share of GDP—commonly used indicators for overall state capacity that do not distinguish between different types of government expenditures, most importantly, between military and other expenditures. Other interpretations of the railway length variable remain certainly possible and plausible (see ibid.). Because I will use literacy rates and railway track length as two alternative measurements of state capacity to provide public goods, they will be integrated into two separate models.

To measure linguistic diversity, several possible indices are available, and all produce substantially similar results. In the following analysis, I use the earliest available data, which was assembled by Soviet ethnographers in the 50s and 60s (data adopted from Fearon and Laitin, 2003). Linguistic data are ideal for the specific purpose at hand because we need an 'objective' description of the communicative landscape that disregards the political relevance of ethnic ties.

Finally, how could we measure pre-colonial levels of state centralization? The Human Relations Area Files were assembled by anthropologists on the basis of thousands of ethnographies referring to pre-colonial economic, social, political, and cultural features. These rich data have been aggregated to the country-level by Müller *et al.* (1999; for other recent use of these data for Africa, see Gennaioli and Rainer, 2007). They mapped the pre-colonial ethnic groups onto the grid of today's states and estimated contemporary population shares of these groups. Using their data, I calculate the percentage of today's population that was ruled in pre-colonial times by some form of state with a minimal degree of bureaucratic development and hierarchical political differentiation.[3] Unfortunately, these data are available for only half of the countries of the world, excluding the Americas and Europe. I thus conducted some additional and preliminary tests, using government expenditure per capita as a proxy variable for state centralization, for these Western countries as well (reported in the Supplementary Appendix).

Competing Arguments: Political Institutions, Imperial Legacies, Global Pressures

The political development argument outlined above contrasts with a series of other approaches in comparative politics and political sociology. I discuss the most prominent ones briefly here.

Political Institutions: Democracy, Proportionalism, and Parliamentarianism

Many comparativists and political theorists believe that democracy represents, at least in the long-term (Huntington, 1996), the most effective institution to

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

achieve ethnic inclusion and nation building—so much so that many popular authors and policymakers use the terms nation building and democratization synonymously (Dobbins, 2003/2004). After all, democracy provides political leaders with incentives to reach beyond the circles of co-ethnics and seek votes across ethnic divides, one could argue. Ethnic minority elites, on the other hand, can organize freely in democracies (Diamond, 1994), pressure for participation, and offer themselves as partners to form minimum winning coalitions. Competitive elections, finally, can bring about a shift in government and thus minimize the risk that ethnic minorities remain permanently excluded from representation in the executive. Democracies should therefore be less exclusionary than non-democratic regimes.

To test this hypothesis, I use the Polity IV data set and the standard cut-off points of +6 on the combined autocracy and democracy scale (which ranges from −10 to +10) to identify democracies. 'Anocracies' are defined as regimes that display both autocratic and democratic features (from −6 to +6 on the combined scale), and autocracies contain no democratic elements (−6 to −10 on the combined scale).

Beyond the general democracy-breeds-inclusion argument, students of 'ethnically divided' democracies have debated whether parliamentarian (Lijphart, 1977; Linz, 1990) or presidential systems (Horowitz, 2002; Saideman *et al.*, 2002; Roeder, 2005; Reilly, 2006) are more conducive to political inclusion and nation building. Another, closely related debate is whether proportional systems of electing members of parliament enhance political power sharing when compared with majoritarian rules (Lijphart, 1994, 1999).[4]

To evaluate these competing hypothesis, I rely on three data sets on political institutions: Gerring and Thacker (2008) have coded whether a system is presidential, mixed, or parliamentarian, as well as proportional, mixed, or majoritarian. These data are available for all non-authoritarian systems. I also use a more fine-grained coding of political institutions after 1975 provided by a team of researchers from the World Bank (Thorsten Beck *et al.*, 2001), which includes autocratic regimes as well. Finally, the equally granular Institutions and Elections Project data set (Regan and Clark, 2011) covers all countries from 1972 onward. The results for the WB and IAEP data will be displayed in the Supplementary Appendix.

Imperial Legacies

A second group of approaches highlights the historical legacies of imperialism, thus putting the finger on the long-term consequences of Europe's conquest and century long domination of the world. According to a first argument, conquerors often recruited members of minority ethnic groups into the imperial administration (such as Tamils in British Sri Lanka) or into the colonial army (such as Berbers in French Morocco). This colonial divide-and-rule strategy laid the ground for the domination of the post-colonial state by these historically privileged minority groups (Horowitz, 1985, Chapters 11–13).

There are two versions of this argument, a weaker and a stronger one. According to the weak version (Chandra and Wilkinson, 2008), early post-colonial power configurations are highly influenced by late-colonial constellations. This weak version thus folds into a general path-dependency argument: power configurations tend to change slowly except when turned upside down by revolution or war. The stronger version argues that countries colonized in the past should display starker ethno-political inequality today because it is *only* in colonial polities that minority rule is promoted systematically. To test this stronger version of the argument, I count the percentage of years since 1816 that a territory has been controlled by an imperial or colonial power (most data are from Wimmer and Min, 2006).

Other authors focus on the consequences of different styles of imperial rule, rather than emphasizing the effects of colonialism in general. Mahoney (2010; see also Olsson, 2007) argues that mercantilist forms of colonial domination, based on natural resource extraction, controlled trade, and a sharply drawn boundary between conquerors and conquered, leave behind a legacy of internal colonialism. By contrast, 'liberal' forms of colonial domination, which combine free trade with a laxer political control of the native population, result in a less exclusionary post-colonial political system. Former Spanish colonies, most of which were conquered during the mercantilist period, should therefore display higher contemporary levels of ethno-political exclusion compared with other colonies, especially compared with British colonies in which 'liberal' regimes were prevalent most of the time. Others have argued that French rule relied on more direct forms of control through a colonial bureaucracy staffed with assimilated locals from various ethnic and regional backgrounds, leading to sharper post-colonial ethno-political inequality than indirect rule more common among other, especially British colonial powers (Blanton *et al.*, 2001). All data on former colonial and imperial masters were adopted from Wimmer and Feinstein (2010). They coded the maximum percentage of territory that was ever controlled by an empire. While a measurement of the degree of directness of rule would obviously be preferable, this is currently only available for a subset of British colonies (Lange, 2005).

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

World Polity: The Global Diffusion of Multi-Culturalist Inclusion

While the institutionalist and colonial legacy arguments refer to domestic political factors, world polity theory points at global, trans-continental forces that shape ethno-political configurations around the world. According to John Meyer and collaborators, a rational world culture based on the normative principles of the enlightenment has emerged over the past 200 years and eventually become hegemonic (Meyer *et al.*, 1997; see also constructivist scholarship in international relations). Since the American civil rights movement, the political empowerment of minorities has increasingly become part of this canon and corresponding policies—such as reserved parliamentary seats, electoral district engineering, quotas at the cabinet or senior government level, etc.—diffused widely across the world (Kymlicka, 2007).

World culture theory offers a cross-sectional as well as a longitudinal argument. First, the more linkages a country maintains to the centres of global culture and power, the more its elites are exposed to world cultural models and the more likely they will pursue a policy of multi-ethnic inclusion and open the ranks of government to hitherto excluded groups. To test this hypothesis, I use data on the number of memberships in international governmental organizations provided by the Correlates of War project (Pevehouse *et al.*, 2004). Note that this is a more appropriate measurement of world cultural exposure than the often-used number of international NGOs because governmental elites, rather than civil society actors, need to be networked into global institutions to eventually open their palace doors to ethnic minorities. To be on the safe side, however, I also used a count of INGOs to test the world polity argument (results will be provided in the Supplementary Appendix).

Focusing on the longitudinal dimension, we can hypothesize that levels of ethnic exclusion should decrease in all countries of the world over time, a process that should accelerate from the 1970s onward after the global hegemon, the United States, had finally overcome racial restrictions to voting rights. This longitudinal argument can be tested statistically with natural cubic splines coded on calendar time. They allow tracing non-linear trends (accelerations, peaks, and lows, etc.) in how the passing of chronological time affects levels of ethno-political inclusion in the world. These trends will be graphed out to ease interpretation.

Dependent Variable and Modeling Approach

Levels of ethno-political inclusion (or nation building) will be measured with data from the Ethnic Power Relations data set (Wimmer *et al.*, 2009). EPR includes 733 politically relevant categories, defined as ethnic groups that are either discriminated against in the political domain or for which at least one significant actor claims representation in the national political arena. These categories can fuse or fission over time, become politically irrelevant, and so on, in line with constructivist theories of ethnicity. For each country and each year, EPR codes in how far actors who claim to represent these various communities have access to central-level government power.

Access to government ranges from monopoly power (total control of executive government by representatives of a particular group) to discrimination (i.e. targeted exclusion from any level of government), through an ordinal scale that includes dominance, senior partner in a power sharing arrangement, junior partner, regional autonomy, and powerless. For the purposes of this study, I calculate the share of the population that is either discriminated, powerless, or is represented in regional governments only, all of which are thus excluded from executive power at the level of the central state.

A few points about model specification are in order. First, since most variables vary over time (with the exception of the linguistic fractionalization index), a pooled-time series cross-sectional research design is appropriate. Units of observation are country-years. Included are all years between 1946 or the year of independence and 2005 for the 156 countries that have a surface of more than 50,000 square kilometres and a population of at least 1 million individuals. Robust standard errors will be clustered by country, following standard model specifications.

Second, since the dependent variable is a proportion and not over-dispersed, the appropriate statistical model is a general linear regression with a logistic link function and the specification of a binomial distribution of the dependent variable. This takes into account that possible values are bound by 0 and 1. Because many countries do not exclude any citizen on the basis of ethnicity, there is a problem of excess zeros (about 30% of observations), which the above specification of the link function is supposed to handle effectively (McCullagh and Nelder, 1989). To be on the safe side in terms of interpretation, I also ran models without countries where ethnicity was never relevant and where, therefore, ethno-political exclusion cannot possibly exceed 0. This is to make sure that the models do not depend on these 0s, which is not the case.

This approach only makes sense, however, if we can be confident that the political processes that lead to full inclusion (the 0s) are of a similar nature as those that

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

produce exclusion (the non-0s). But maybe homogenous countries, which by definition cannot exclude anybody on ethnic grounds, are the result of previous ethnic cleansing or the secession from heterogeneous countries, rather than the forces of trans-ethnic tie formation and gradual assimilation, as the political development model assumes? This possibility of causal heterogeneity will be explored in the robustness section by analyzing subsamples of more homogenous and more heterogeneous countries in separate models.

Third, all models need to control for the ethno-demographic composition of the population. Imagine a country with only two politically relevant ethnic groups, each 50% of the population. The size of the excluded population is either 0 or 50%. In a country with 10 politically relevant ethnic groups with a 10% share of the population each, the size of the excluded population can be 0, 10, 20, 30, etc. up to 90%. All models thus include controls for the number of ethno-politically relevant groups as well as one for the population share of the largest politically relevant ethnic group. The main models were also tested using a polarization and fractionalization index based on EPR's group list, which generated substantially similar results.

Findings

To ensure transparency, I proceed in a stepwise fashion, separately evaluating the effects of groups of variables that speak to the same theory. As Table 1 shows, there are no collinearity problem except in models that include democracy, or International Governmental Organizations (IGOs), or linguistic heterogeneity and literacy, or associational density and democracy. None of these correlations exceed 0.55, however. I conducted collinearity tests on all models and removed the time trends from problematic models, which will be noted in footnotes and the tables.

Main Findings

Model 1 in Table 2 incorporates a number of terms to test the various colonial legacies arguments outlined above. It shows that having spent many years under imperial rule since 1816 is not associated with exclusion, contrary to what the strong variant of the colonial legacy argument predicts. Regarding membership in specific empires, it turns out that former French dependencies are indeed characterized by higher levels of exclusion (compared to countries without any history of imperial domination), in line with expectations, but so are British dependencies, thus raising doubts about whether directness of rule affects post-imperial nation-building. Former Spanish colonies are not excluding larger shares of their populations today, as maintained by the legacy of mercantilism argument. We will see further below that the effects of having been a British or French colony disappear once the political development variables are introduced in Table 3.

Model 2 reveals a highly significant influence of the number of memberships in international governmental organizations, indicating that countries that are more integrated into the world polity and thus more exposed to hegemonic notions of minority rights and multi-cultural justice exclude a smaller proportion of their population from access to central government power. This association, however, will again disappear as soon as we introduce variables associated with the political development argument (see Table 3). A count of international NGOs per capita—the most often used variable to measure world polity exposure—would miss standard levels of significance in Model 2 of Table 2 (not shown, see Supplementary Table S2).

How about the longitudinal aspect of the world polity argument? One of the coefficients for the natural cubic splines on chronological time is significant in Model 2. When charting out the predicted levels of inclusion per calendar year, we do not see the expected downward trend in Figure 3, however, but an inverted U shape. To disentangle panel composition effects from world polity influence, this figure only includes 'new' countries that became independent after 1945. A graph with all countries included would be even less encouraging.

Model 3 moves on to the political institutions variables and uncovers a highly significant and substantially important association between democracy and ethno-political inclusion, which is also highly significant in a fixed effects model, which takes care of other, non-measured differences between countries that might influence their ethno-political configuration of power (see Supplementary Table S3). Model 4 reveals, however, that neither proportional representation nor parliamentarism is associated with more inclusion—a finding that runs against important strands of the literature that had emphasized the inclusionary power of consociationalism. These results are based on Gerring and Thacker's (2008) data (which is not coded for autocracies), but also hold if we use the World Bank data set on political institutions to code parliamentarism and proportionalism (see Supplementary Table S4), or when we code both variables with IAEP data (see Supplementary Table S5), or when observations are restricted to democracies only (see Supplementary Table S6).

Note that democracy might be associated with inclusion, as Model 3 demonstrates, through reverse

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Table 1. Correlation matrix (variables for Tables 2 and 3 only; correlations above 0.4 in bold)

	Proportion excluded	Prop. years imperial	Spanish	Habsburg	Ottoman	Russian	French	British	Portuguese	Other empires	Observations
Proportion of excluded population	1										7,116
Proportion years under imperial rule since 1816	0.0343	1									7,116
Fomer Spanish dependency	−0.0234	**−0.4453**	1								7,116
Former Hapsburg dependency	−0.0615	0.0967	−0.0743	1							7,116
Fomer Ottoman dependency	0.0784	0.2569	−0.1709	−0.038	1						7,116
Former Russian dependency	−0.0308	0.2082	−0.0941	−0.0312	−0.0507	1					7,116
Former French dependency	0.08	0.1101	−0.1946	−0.0758	0.0768	−0.096	1				7,116
Former British dependency	0.0397	0.3215	−0.2597	−0.1011	0.0832	−0.1281	−0.2547	1			7,116
Former Portuguese dependency	0.0373	0.031	−0.0743	−0.0289	−0.0665	−0.0366	−0.0758	−0.0105	1		7,116
Former dependency of other empires	0.0399	0.1464	−0.007	−0.0753	−0.0217	−0.0954	0.0513	−0.0997	−0.0753	1	7,116
Number of memberships in IGOs	−0.1352	−0.2872	0.0495	−0.0019	−0.0933	−0.0485	−0.0468	−0.0814	−0.0059	−0.1093	7,116
Democracy, lagged	−0.2346	−0.1122	0.0552	0.0255	−0.1023	0.0657	−0.2746	0.035	−0.0668	−0.1218	7,116
Percentage literates among adult population	−0.2981	−0.1072	0.145	0.1686	−0.0671	0.2389	−0.416	−0.1318	−0.1053	−0.0518	7,116
Length of railway tracks (kilometres) per 1000 km^2	−0.2658	−0.1043	−0.1354	0.3055	−0.0802	0.0462	−0.2229	−0.2263	−0.0942	−0.0518	7,116
Linguistic fractionalization	0.3449	0.0639	−0.1498	−0.0325	−0.231	−0.0107	0.1864	0.196	0.093	0.0008	7,116
Number of politically relevant ethnic groups	0.2105	0.2396	−0.1182	0.0013	−0.0574	−0.002	−0.0227	−0.0024	−0.0241	−0.0433	7,116
Largest share of politically relevant ethnic group	0.0273	−0.1516	0.2234	0.097	0.0093	0.1904	−0.0533	−0.2639	−0.0675	−0.0507	7,116
Year	0.0029	0.0127	−0.0927	0.0279	−0.0175	0.1602	0.0425	0.0482	0.05	−0.0117	7,116
Number of associations per capita	−0.1925	−0.0611	−0.1001	−0.0188	−0.0821	0.0202	−0.1933	0.177	−0.0711	−0.0623	4,611
Proportion of population governed by states before colonization	−0.2022	0.0507	0.129		0.0956		−0.1132	−0.1716	−0.211	0.0205	2,507

	IGOs	Democracy	Literacy	Railways	Ling. heter.	No of grps	Max. grp size	Year	No of ass.	Precol. State	
Number of memberships in IGOs	1										7,116
Democracy, lagged	**0.4839**	1									7,116
Percentage literates among adult population	**0.4621**	**0.5113**	1								7,116
Length of railway tracks (kilometres) per 1000 square kilometres	0.3055	0.3579	**0.5227**	1							7,116
Linguistic fractionalization	−0.0747	−0.2061	**−0.4014**	−0.3011	1						7,116
Number of politically relevant ethnic groups	−0.0525	−0.1183	−0.016	−0.1168	0.19	1					7,116
Largest share of politically relevant ethnic group	0.022	0.0723	0.2032	0.0638	−0.1639	0.07	1				7,116
Year	**0.5443**	0.1689	0.2852	−0.0894	0.0949	0.0071	−0.0513	1			7,116
Number of associations per capita	0.3451	**0.4414**	0.3957	0.3213	−0.1253	−0.1299	−0.0038	0.0911	1		4,611
Proportion of population governed by states before colonization	−0.0659	0.1416	0.3298	0.3529	**−0.5033**	0.0792	0.3547	−0.0211	−0.0549	1	2,507

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Table 2. Testing imperial legacy, world polity, and political institutions arguments (Generalized Linear Models of proportion of population excluded from executive government)

	1	2	3	4
Imperial legacy variables				
Proportion years under imperial rule since 1816	−0.6510 (0.502)			
Former Spanish dependency	0.1718 (0.470)			
Former Hapsburg dependency	−0.3882 (0.411)			
Former Ottoman dependency	0.5572 (0.381)			
Former Russian dependency	0.3761 (0.496)			
Former French dependency	0.6347* (0.346)			
Former British dependency	0.5802* (0.347)			
Former Portuguese dependency	0.7743 (0.602)			
Former dependency of other empires	0.3723 (0.378)			
World polity variables				
Number of memberships in IGOs		−0.0159*** (0.005)		
Political institutions variables				
Democracy, lagged			−0.9460*** (0.219)	−1.1050*** (0.305)
Fully proportional systems				0.4463 (0.278)
Fully parliamentary systems				−0.0550 (0.346)
Control variables				
Natural cubic spline on calendar year 1	Not shown	0.0178** (0.008)		Not shown
Natural cubic spline on calendar year 2	Not shown	−0.0065 (0.008)		Not shown
Constant and ethno-demographic controls	Not shown	Not shown	Not shown	Not shown
Observations	7,134	6,542[a]	6969[b]	3,387[c]

Notes: [a]Missing data for years after 2001; [b]Missing values for the democracy variable during wars and anarchy; [c] Sample restricted to democracies and anocracies because of missing values on the proportional/parliamentary systems variables for autocracies, missing values for the democracy variable during wars and anarchy.
***$P < 0.01$; **$P < 0.05$; *$P < 0.1$.

causation: minority regimes (such as Iraq under Saddam Hussein or Syria under Assad) can be expected to be much more reluctant to democratize than governments that rest on a broader ethnic power base (see Horowitz, 1993: pp. 21–22; Tilly, 2000: p. 10). I do not pursue this possibility any further here (see Wimmer, 2013, chapter 6) because in any case the association between nation building and democracy will become insignificant once we introduce the political development variables, to which I now turn.

The first panel of Table 3 (Models 1–9) shows results with literacy rates as a proxy for state capacity to deliver public goods. The second panel (Models 10–18 in Table 3) refers to models in which public good provision is measured with railroad density. I discuss both results conjointly. The first models (Models 1 and 10) show full

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Table 3. Testing the political development argument I (Generalized Linear Models of proportion of population excluded from executive government)

			Fixed effects	Cross-section with IV set at 1900	Heterogenous countries only		Homogenous countries only		
Models with literacy	1	2	3	4	5	6	7	8	9
Political development variables									
Percentage of literates among the adult population	−0.0120** (0.005)	−0.0119** (0.006)	−0.0012*** (0.000)		−0.0078* (0.004)		−0.0327*** (0.009)		−0.0114* (0.007)
Percentage of literates among the adult population in 1900				−0.0245*** (0.006)					
Number of associations per capita	−14.4758* (7.612)	−9.1236 (7.334)	−0.0141*** (0.003)			−17.4000* (10.205)		−37.6134* (20.251)	−24.4112 (53.355)
Linguistic fractionalization	1.3303** (0.555)	1.2941** (0.553)			0.6390 (0.633)	0.3951 (0.677)	2.3644*** (0.915)	1.1607 (1.287)	−0.1399 (0.690)
Proportion of population governed by states before colonization									−1.3235** (0.577)
Variables significant in Table 2									
Former French dependency		−0.1973 (0.440)							
Former British dependency		−0.0053 (0.347)							
Number of memberships in International Government Organizations		−0.0037 (0.006)							
Democracy, lagged		−0.3898 (0.261)							
Constant, time controls, and ethno-demographic controls	Not shown	Not shown	Not shown	Not shown, no time control	Not shown, no time control	Not shown	Not shown	Not shown	Not shown, no time control
Observations	4,611[a]	4584[a]	4611[a]	144	3,858	2,539[a]	3,280	2072[a]	2507[a,b]
Number of countries	147	147	147	144	87	82	68	65	74

(continued)

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Table 3. Continued

			Fixed effects	Cross-section with IV set at 1900	Heterogenous countries only		Homogenous countries only		
Models with railroad density	10	11	12	13	14	15	16	17	18
Political development variables									
Length of railway tracks (kilometres) per 1,000 square kilometres	−0.0226*** (0.006)	−0.0209*** (0.006)	0.0001 (0.001)		−0.0191*** (0.005)		−0.0174** (0.009)		−0.0303* (0.018)
Length of railway tracks in 1900				−0.0283*** (0.007)					
Number of associations per capita	−14.7201* (7.915)	−10.3356 (7.503)	−0.0062** (0.003)			−17.4000* (10.205)		−37.6134* (20.251)	−39.5100 (52.096)
Linguistic fractionalization	1.4137*** (0.528)	1.3840** (0.547)			0.8122 (0.653)	0.3951 (0.677)	1.0476 (1.247)	1.1607 (1.287)	−0.0746 (0.689)
Proportion of population governed by states before colonization									−1.3623** (0.563)
Variables significant in Table 2									
Former French dependency		−0.0718 (0.388)							
Former British dependency		−0.0213 (0.349)							
Number of memberships in International Government Organizations		−0.0012 (0.006)							
Democracy, lagged		−0.3581 (0.241)							
Constant, time controls, and ethno-demographic controls	Not shown	Not shown	Not shown, no time control	Not shown, no time control	Not shown	Not shown	Not shown	Not shown	Not shown
Observations	4,611[a]	4,584[a]	4,611[a]	149	3,858	2,539[a]	3,280	2,072[a]	2,507[a,b]
Number of countries	147	147	147	149	87	82	68	65	74

Notes: [a] Missing values on the number of associations per capita variable before 1970 and for eight countries; [b] without Europe and the Americas.
***$P < 0.01$; **$P < 0.05$; *$P < 0.1$.

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

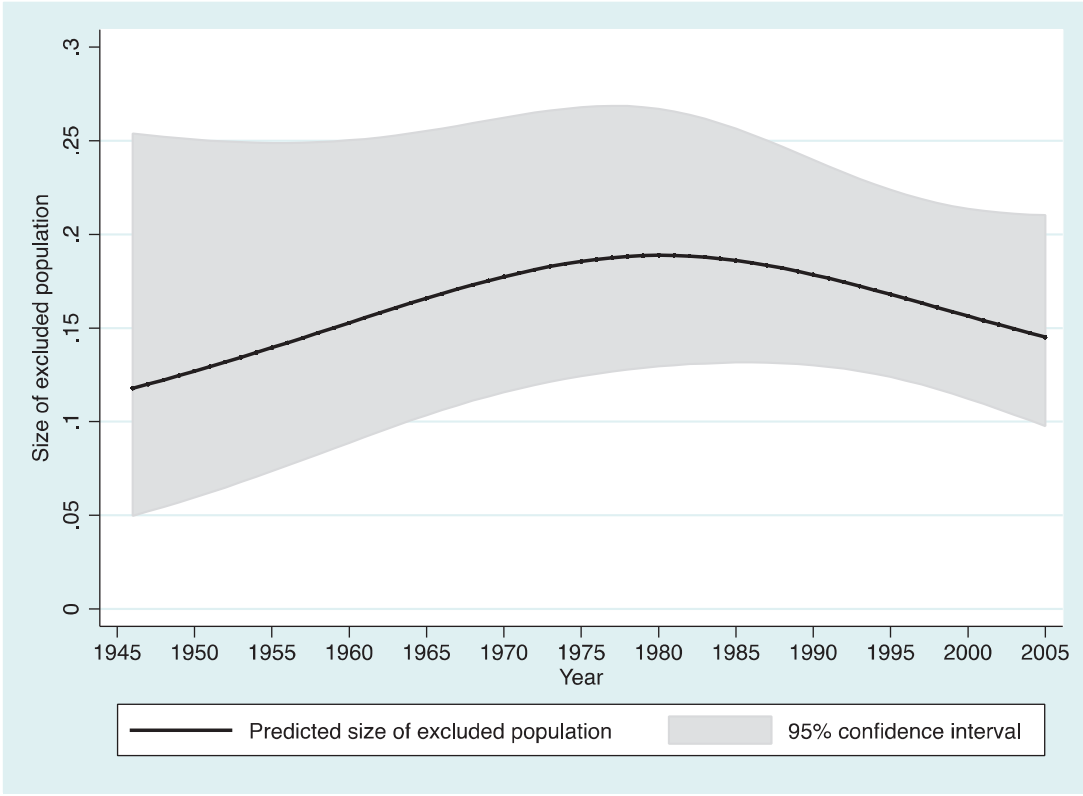

Figure 3. Predicted size of the excluded population over time in 98 countries that became independent after 1945

support for the political development argument. The share of adult literates in the population, the length of railway tracks per square kilometre (both proxying for public goods provision by the state), the number of associations per capita, as well as the fractionalization of a population in linguistic terms are all associated with levels of inclusion. Additional analysis shows that it is unlikely that the association between the two public goods variables and nation-building are generated by other than the proposed mechanisms: literacy could foster nationalism or tolerance toward minorities and thus ethno-political inclusion—but they do not (see Supplementary Table S7); high railroad density could be the consequence of war and ethnic cleansings, which could in turn decrease the share of minority populations excluded from power—but that is not the case either (see Model 2 in Table 4).

Models 2 and 11 in Table 3 combine all variables that were statistically significant in previous models from Table 2, including the democracy variable for which there might be a reverse causation problem. The coefficients and standard errors of the political development variables remain largely the same when these additional variables are included, with the exception of the number of NGOs per capita variable, which is no longer significant, mostly due to collinearity with the democracy variable. The variables indicating that a country was a British or French colony, the number of memberships in international organizations, and the democracy variable become statistically insignificant—lending additional support for the argument that we need to focus on slow moving, domestic processes of political development to understand nation building.

Sensitivity, Endogeneity, and Causal Heterogeneity

The remaining models in Table 3 (Models 3–9, as well as 12–18) check for unobserved heterogeneity across countries, for causal heterogeneity, and possible reverse causation problems. Models 3 and 12[5] in Table 3 report the results of country-fixed-effects specifications, which test whether changes in the values of independent variables over time *within* a country affect the dependent

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Table 4. Testing the political development argument II (models of associational density, railway length, literacy, and linguistic fractionalization)

	1 Dependent variable: Number of associations per capita	2 Dependent variable: Length of railway tracks	3 Dependent variable: % literates among the adult population	4 Dependent variable: Linguistic fractionalization
Legacy of stateness variable				
Proportion of population governed by states before colonization	−0.0016* (0.001)	10.1478*** (3.013)	1.0107*** (0.302)	−1.6201*** (0.327)
Control variables				
Number of years since 1816 with constant borders	0.0000 (0.000)	−0.0173 (0.027)	0.0004 (0.002)	0.0021 (0.003)
GDP per capita in 1000 dollars, lagged	0.0002*** (0.000)	0.1298 (0.268)	0.0343* (0.018)	−0.0637*** (0.016)
Difference between highest and lowest elevation (in kilometers)	−0.0002* (0.000)	−0.1155 (0.472)	−0.0249 (0.062)	0.1507** (0.061)
Number of ethno-national wars between 1816 and first year in data	0.0004 (0.000)	−0.0318 (1.539)	0.1622 (0.158)	0.0707 (0.177)
Democracy, Polity 4	0.0014* (0.001)			
Political instability (pol. regime change during past 3 years)	−0.0004 (0.000)			
Constant and time controls	Not shown	Not shown	Not shown	Not shown
Observations	2,409[a]	3,505	3,505	71[b]

Notes: All models without Europe and the Americas; [a] Missing values on dependent variable before 1970; [b] Cross-sectional model.
$^{***}P < 0.01$; $^{**}P < 0.05$, $^{*}P < 0.1$.

variable in the expected way. In this way, they control for all time-invariant unobserved attributes of countries. The models do not include the term for linguistic fractionalization because there is no time-varying data available. The literacy and organizational density variables are all significant in the expected direction, making it less likely that their association with nation building is due to some other, unobserved country characteristics. Unfortunately, the railways variable fails to achieve standard levels of statistical significance Model 12. This may be due to the fact that some countries increased their capacity to provide public goods in the age of mass automobility from the 1960s onwards all the while developing more inclusionary ethno-political regimes (think of the United States during the Great Society period).

Models 4 and 13 in Table 3 test the possibility of reverse causation by introducing a long time lag between independent and outcome variables, excluding linguistic fractionalization and the count of NGOs because there is no data for these two variables before the 1960–1970s. I measured literacy and the length of railway tracks for the year 1900 and tested whether these variables have an effect on the average level of exclusion between 1945 (or the year of independence for newer countries) and 2005. The results are encouraging: both the literacy and the railways variable are significant in the expected direction. In the 'placebo' regressions shown in Supplementary Tables S8, I further assess possible reverse causation problems by switching independent and dependent variables. Ethno-political exclusion measured in first year of available data does not affect literacy, organizational density, railways, and linguistic heterogeneity later on as soon as we also control for pre-colonial state centralization.

Models 5–8 and 14–17 in Table 3 explore whether we need to distinguish between different causal pathways leading to nation building. Models 5[6] and 6 as well as 14 and 15 in Table 3 refer to the subset of heterogeneous countries with a fractionalization index above the mean. Models 7 and 8 as well as 16 and 17 do the same for the subset of more homogenous countries. This is to see whether the path to integrative power structures is different for the homogenous countries of Europe and East Asia, where wars, ethnic cleansings, separations, and forced assimilation over long centuries have done much to reduce ethno-demographic the obstacles for successful nation-building, compared with the

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

more heterogeneous countries of the global South that have inherited more diversity from the colonial era. To avoid collinearity problems when using a reduced number of observations, the public goods provision (literacy or railways) and organizational development variables are run in separate models. The results suggest that political development influences nation building both in diverse and homogenous countries—though not surprisingly the linguistic fractionalization variable is rarely significant in these models owing to the reduced variation within the subsamples, which were defined on the very basis of the degree of linguistic heterogeneity.

State Building and Political Development

In the final step of analysis, I evaluate whether two of the core political development variables—infrastructural capacity and linguistic heterogeneity—are in turn associated with historically achieved levels of state formation, as argued in the theory section. Table 4 reports the results. The variable that measures pre-colonial state centralization in Africa and Asia strongly influences post-colonial diversity as well as infrastructural capacity: the pre-colonial state-building variable is highly significant and with large coefficients in Ordinary Least Square regressions on railroad density (Model 2), in general linear regressions on the proportion of literate adults (Model 3), and on linguistic fractionalization (Model 4, which is a cross-sectional model because the dependent variable does not vary over time). In line with theoretical expectations, pre-colonial centralization does not affect organizational development (Model 1).

Models 1 to 4 include the following theoretically meaningful controls, many of which have been the focus of past research. The 'artificiality' of contemporary states, measured with the number of years with constant borders since 1816,[7] is believed by some authors to affect contemporary state capacity (Englebert, 2000; Bockstette *et al.*, 2002) and thus perhaps railway density and literacy rates. GDP per capita could obviously be associated with all four dependent variables. Topography, measured as the difference between highest and lowest elevation in a country, could influence railway length, organizational density, and linguistic fractionalization (Michalopoulos, 2012). The number of ethno-nationalist wars fought between 1816 and the first year of data could increase linguistic homogeneity through ethnic cleansings or state partitions along ethnic divides; they would also increase the density of railway tracks if these were mainly built for military purposes. Finally, I replicate Schofer and Longhofer's (2011) model to explain associational density, adding democracy and regime change to the list of independent variables in Model 1.

While post-war state capacity to provide public goods and linguistic homogeneity are strongly influenced by pre-colonial state centralization, as shown in Table 4, pre-colonial stateness has also a direct, unmediated effect on post-war nation building. This is shown in Models 9[8] and 18 of Table 3. Pre-colonial state centralization is significantly and negatively associated with ethno-political exclusion. Because this measurement of state formation refers to pre-colonial polities, reverse causation can—for once—be safely excluded from consideration. Only the coefficients of the two measurements of public goods provision are still statistically significant in Models 9 and 18. This indicates either that pre-colonial centralization affects nation-building through other mechanisms than the ones explored in this article or that measurement problems for the mediator variables prevent a precise identification of these mechanisms. Additional tests using the Sobel-Goodman Mediation test reveal, however, that the relationship between pre-colonial state centralization and post-War exclusion is indeed mediated by capacity to provide public goods, organizational density, and linguistic heterogeneity (results not shown).

But how about the Americas and Europe, which were so far excluded from consideration because no data on the pre-colonial stateness variable are available for these continents? A preliminary test uses government expenditure per capita in 1920 as an *Ersatz* variable instead of late 19th century state centralization. The results are reported in Supplementary Table S9 and show a significant and negative association with post-war ethno-political exclusion—thus suggesting that the long-term political development argument might well hold for these independent states as well.

Conclusion

This study represents a first attempt at exploring the long-term dynamics of nation building and ethno-political inclusion from a systematic comparative point of view, thus revitalizing the neo-Weberian political development school that had emerged in the 1950s and 1960s. Using an exchange theoretic framework, I distinguished between a resource, an organizational, and an informational aspect of political alliances. According to this theory, alliances will stretch across ethnic divides and throughout a territory if state elites have the infrastructural capacity to provide public goods and to make themselves attractive from a resource exchange point of view; if associational networks have developed that will

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

make the building of non-ethnic coalitions easier; and if few linguistic barriers hamper the exchange of information and thus the establishment of wide-reaching relationships of alliance and support. State capacity to deliver public goods and linguistic homogeneity in turn result from high levels of state centralization achieved in previous epochs.

Nation building, in other words, is the outcome of these slow-moving historical forces rather than of particular institutional incentive structures (such as democracy), global pressures to emulate templates of Western state- and nationhood, or the lack of a divisive colonial experience. This does not imply that these competing accounts are irrelevant, since the statistical approach chosen in this article obviously captures average effects only. If one were more ambitious and searched for a complete causal account of all cases, for example, using Qualitative Comparative Analysis as a technique (Ragin, 1989), it may well be that democratization or specific colonial experiences are important factors to understand a particular subset of cases.

The theory of state formation, political development, and nation building advanced here also does not include potentially important other factors and mechanisms that affect levels of inclusiveness on a faster temporal scale. It disregards the policies of neighbouring states or regional hegemons that might affect the decision of whether to grant political participation to minorities (see Mylonas, 2012) as well as changing coalitions and elite bargains that will affect the exact power configuration over the short run (see Acemoglu and Robinson, 2006; for Africa: Roessler, 2011). The argument offered here also cannot explore the historical antecedent conditions that influence whether strongly centralized states have emerged in the 19th century or before.

Obvious data limitations should be noted as well. Future work should aim at a more direct measurement of state capacity to provide public goods (for example by focusing on publicly financed or subsidies schools or clinics and hospitals) and of the development of voluntary organizations that include organizations controlled by government. Measurements of linguistic diversity would greatly improve if they would include a longitudinal dimension—tracing ethno-demographic developments over the past century—as would the organizational data were they available before 1970. Second, there is currently no cross-national and time-varying data on the structure of political alliance networks. It is thus impossible to determine, for example, whether public goods provision and dense networks of voluntary organizations lead to ties that crosscut politically relevant ethnic divides or whether they prevent the politicization of ethnicity in the first place. Creating such data represents a major challenge to be addressed in the future.

Notes

1 For other explanations for the association between ethnicity and patronage, see Fearon (1999) and Chandra (2004).
2 Non-state institutions such as religious schools had historically a large impact on literacy levels as well. This should be less of a problem, however, in post-colonial societies and more generally in the post-war world, after most states had taken over the task of providing elementary education from religious organizations (Meyer *et al.*, 1992).
3 I include the following four political types in the definition of statehood: states in stratified societies; states in non-stratified societies; feudal states; complex states.
4 More fine-grained analyses both of institutional rules and of their interaction with ethno-demographics have been offered by Mozaffar *et al.*, 2003; Birnir, 2007. Such more fine-grained analysis is beyond the reach of this article.
5 Model 5 has no time controls because tests indicated that they create a collinearity problem.
6 This model does not include time controls because tests indicated that they create a collinearity problem.
7 For data sources for these additional variables, see Supplementary Table S1.
8 This model does not contain a time trend because tests indicated that they create a collinearity problem.

Supplementary Data

Supplementary data are available at *ESR* online.

References

Acemoglu, D. and Robinson, J. A. (2006). *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Ahlerup, P. and Olsson, O. (2012). The roots of ethnic diversity. *Journal of Economic Growth*, 17, 17–102.

Baldassari, D. and Diani, M. (2007). The integrative power of civic networks. *American Journal of Sociology*, 113, 735–780.

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Baldwin, K. and Huber, J. D. (2011). Economic versus cultural difference. Forms of ethnic diversity and public goods provision. *American Political Science Review*, 104, 644–662.

Bendix, R. (1964). *Nation-building and Citizenship. Studies in Our Changing Social Order*. New York: John Wiley.

Birnir, J. K. (2007). *Ethnicity and Electoral Politics*. Cambridge: Cambridge University Press.

Birnir, J. K. and Waguespack, D. M. (2011). Ethnic inclusion and economic growth. *Party Politics*, 17, 243–260.

Blanton, R., Mason, D. T. and Athow, B. (2001). Colonial style and post-colonial ethnic conflict in Africa. *Journal of Peace Research*, 38, 473–491.

Blau, P. (1964/1986). *Exchange and Power in Social Life*. New Brunswick: Transactions Publisher.

Bockstette, V., Chanda, A. and Putterman, L. (2002). States and markets: the advantage of an early start. *Journal of Economic Growth*, 7, 347–369.

Chandra, K. (2004). *Why Ethnic Parties Succeed. Patronage and Ethnic Head Counts in India*. Cambridge: Cambridge University Press.

Chandra, K. and Wilkinson, S. (2008). Measuring the effect of "ethnicity". *Comparative Political Studies*, 41, 515–563.

Congleton, R. D. (1995). Ethnic clubs, ethnic conflict, and the rise of ethnic nationalism. In A., Breton *et al*. (Eds.), *Nationalism and Rationality*. Cambridge: Cambridge University Press, pp. 71–97.

Deutsch, K. W. (1953). *Nationalism and Social Communication: An Inquiry into the Foundations of Nationality*. Cambridge: MIT-Press.

Deutsch, K. W. (1966). *Nationalism and Social Communication. An Inquiry into the Foundation of Nationality.* Boston: MIT Press.

Diamond, L. J. (1994). Toward democratic consolidation. *Journal of Democracy*, 5, 4–17.

Dobbins, J. F. (2003/2004). America's role in nation-building: from Germany to Iraq. *Survival*, 45, 87–110.

Dunning, T. and Nilekani, J. (2013). Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils. *American Political Science Review*, 107, 35–56.

Englebert, P. (2000). Pre-colonial institutions, post-colonial states, and economic development in tropical Africa. *Political Research Quarterly*, 53, 7–36.

Fearon, J. D. (1999). *Why Ethnic Politics and "Pork" Tend to go Together*. Standford: Department of Political Science, unpublished manuscript.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97, 1–16.

Geertz, C. (1963). The integrative revolution. Primordial sentiments and civil politics in the new states. In G. Clifford (Ed.), *Old Societies and New States. The Quest for Modernity in Asia and Africa*. New York: The Free Press. 105–157.

Gennaioli, N. and Rainer, I. (2007). The modern impact of precolonial centralization in Africa. *Journal of Economic Growth*, 12, 185–234.

Gerring, J. and Thacker, S. (2008). *A Centripetal Theory of Democratic Governance*. Cambridge: Cambridge University Press.

Gould, R. V. and Fernandez, R. M. (1989). Structures of mediation: a formal approach to brokerage in transaction networks. *Social Methodology*, 19, 89–126.

Hendrix, C. S. (2010). Measuring state capacity: theoretical and empirical implications for the study of civil war. *Journal of Peace Research*, 47, 273–285.

Herbst, J. (2000). *States and Power in Africa. Comparative Lessons in Authority and Control*. Princeton: Princeton University Press.

Hillmann, H. (2008). Localism and the limits of political brokerage: evidence from revolutionary Vermont. *American Journal of Sociology*, 114, 287–331.

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

Horowitz, D. L. (1993). Democracy in divided societies. *Journal of Democracy*, 4, 18–38.

Horowitz, D. L. (2002). Constitutional design: proposals vs. process. In A., Reynolds (Ed.), *The Architecture of Democracy: Constitutional Design, Conflict Management, and Democracy*. Oxford: Oxford University Press.

Huntington, S. (1996). Democracy for the long haul. *Journal of Democracy*, 7, 3–13.

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics*, 112, 1251–1288.

Kroneberg, C. and Wimmer, A. (2012). Struggling over the boundaries of belonging. A formal model of nation building, ethnic closure, and populism. *American Journal of Sociology*, 118, 176–230.

Kymlicka, W. (2007). *Multicultural Odysseys. Navigating the New International Politics of Diversity*. Oxford: Oxford University Press.

La Porta, R. *et al.* (1999). The quality of government. *The Journal of Law, Economics, and Organization*, 15, 222–279.

Lange, M. (2005). British colonial state legacies and development trajectories: A statistical analsis of direct and indirect rule. In M., Lange and D., Rueschemeyer (Eds.), *States and Development: Historical Antecedents of Stagnation and Advance*. Palgrave: Macmillan, pp. 117–140.

Lemarchand, R. (1972). Political clientelism and ethnicity in tropical Africa: competing solidarities in nation-building. *American Political Science Review*, 66, 68–90.

Levi, M. (1988). *Of Rule and Revenue*. Berkeley: University of California Press.

Lijphart, A. (1977). *Democracy in Plural Societies: A Comparative Exploration*. New Haven: Yale University Press.

Lijphart, A. (1994). *Electoral Systems and Party Systems: A Study of Twenty-Seven Democracies 1945-1990.* Oxford: Oxford University Press.

Lijphart, A. (1999). *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven: Yale University Press.

Linz, J. J. (1990). The perils of presidentialism. *Journal of Democracy*, 1, 51–60.

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014

Mahoney, J. (2010). *Colonialism and Postcolonial Development: Spanish America in Comparative Perspective*. Cambridge: Cambridge University Press.

Mann, M. (1993). *The Sources of Social Power, vol. 2. The Rise of Classes and Nation States, 1760-1914*. Cambridge: Cambridge University Press.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman and Hall.

Meyer, J. *et al.* (1997). World society and the nation-state. *American Journal of Sociology*, 103, 144–181.

Meyer, J. W., Ramirez, F. O. and Nuhoglu Soysal, Y. (1992). World expansion of mass education, 1870-1980. *Sociology of Education*, 65, 128–149.

Michalopoulos, S. (2012). The origins of ethnic diversity. *American Economic Review*, 102, 1509–1539.

Mozaffar, S., Scarritt, J. R. and Galaich, G. (2003). Electoral institutions, ethnopolitical cleavages, and party systems in Africa's emerging democracies. *American Political Science Review*, 97, 379–389.

Müller, H. P. (1999). *Atlas vorkolonialer Gesellschaften. Sozialstrukturen und kulturelles Erbe der Staaten Afrikas, Asiens und Melanesiens. Ein ethnologisches Kartenwerk für 95 Länder mit Texten, Datenbanken und Dokumentationen auf CD-ROM*. Berlin: Reimer.

Mylonas, H. (2012). *The Politics of Nation-Building: Making Co-Nationals, Refugees, and Minorities*. Cambridge: Cambridge University Press.

Nunn, N. (2008). The long-term effects of Africa's slave trades. *The Quarterly Journal of Economics*, 123, 139–176.

Olsson, O. (2007). *On the Institutional Legacy of Mercantilist and Imperialist Colonialism*. Working Paper in Economics, University of Göteborg 247.

Pevehouse, J. C., Nordstrom, T. and Warnke, K. (2004). The COW-2 International Organizations dataset version 2.0. *Conflict Management and Peace Science*, 21, 101–119.

Ragin, C. (1989). *The Comparative Method. Moving Beyond Qualitative and Quantiative Strategies*. Berkeley: University of California Press.

Regan, P. and Clark, D. (2011). *Institutions and Elections Project*. Department of Political Science, Binghampton University. Available at http://www.binghamton.edu/political-science/institutions-and-elections-project.htm.

Reilly, B. (2006). *Democracy and Diversity: Political Engineering in the Asia-Pacific*. Oxford: Oxford University Press.

Roeder, P. G. (2005). Power dividing as an alternative to ethnic power sharing. In P. G., Roeder and D., Rothchild (Eds.), *Sustainable Peace. Power and Democracy after Civil War*. Ithaca: Cornell University Press. 51–82.

Roessler, P. G. (2011). The enemy from within. Personal rule, coups, and civil wars in Africa. *World Politics*, 63, 399–346.

Rothchild, D. (1986). Hegemonial exchange: an alternative model for managing conflict in Middle Africa. In D., Thompson and D., Ronen (Eds.), *Ethnicity, Politics, and Development*. Boulder: Lynne Rienner, pp. 65–104.

Saideman, S. M. *et al.* (2002). Democratization, political institutions, and ethnic conflict. A pooled time-series analysis, 1985-1998. *Comparative Political Studies*, 35, 103–129.

Schofer, E. and Longhofer, W. (2011). The structural sources of association. *American Journal of Sociology*, 117, 539–585.

Scott, J. C. (1972). Patron-client politics and politican change in Southeast Asia. *The American Political Science Review*, 66, 91–113.

Thorsten B. *et al.* (2001). New tools in comparative political economy: the database of political institutions. *World Bank Economic Review*, 15, 165–176.

Tilly, C. (1975). Western state-making and theories of political transformation. In C., Tilly (Ed.), *The Formation of National States in Western Europe*. Princeton: Princeton University Press, pp. 601–638.

Tilly, C. (2000). Processes and mechanisms of democratization. *Sociological Theory*, 18, 1–16.

Tilly, C. (2006). *Identities, Boundaries, and Social Ties*. Boulder: Paradigm Press.

Varshney, A. (2003). *Ethnic Conflict and Civil Life*. New Haven: Yale University Press.

Weber, E. (1979). *Peasants into Frenchmen: The Modernisation of Rural France, 1870-1914*. London: Chatto and Windus.

Wimmer, A. (2002). *Nationalist Exclusion and Ethnic Conflicts. Shadows of Modernity*. Cambridge: Cambridge University Press.

Wimmer, A. (2013). *Waves of War. Nationalism, State Formation, and Ethnic Exclusion in the Modern World*. Cambridge: Cambridge University Press.

Wimmer, A., Cederman, L. -E. and Min, B. (2009). Ethnic politics and armed conflict. A configurational analysis of a new global dataset. *American Sociological Review*, 74, 316–337.

Wimmer, A. and Feinstein, Y. (2010). The rise of the nation-state across the world, 1816 to 2001. *American Sociological Review*, 75, 764–790.

Wimmer, A. and Min, B. (2006). From empire to nation-state. Explaining wars in the modern world, 1816-2001. *American Sociological Review*, 71, 867–897.

Downloaded from http://esr.oxfordjournals.org/ at Princeton University on November 5, 2014