

The Rise of the Nation-State across the World, 1816 to 2001

Andreas Wimmer^a and Yuval Feinstein^a

Abstract

Why did the nation-state proliferate across the world over the past 200 years, replacing empires, kingdoms, city-states, and the like? Using a new dataset with information on 145 of today's states from 1816 to the year they achieved nation-statehood, we test key aspects of modernization, world polity, and historical institutionalist theories. Event history analysis shows that a nation-state is more likely to emerge when a power shift allows nationalists to overthrow or absorb the established regime. Diffusion of the nation-state within an empire or among neighbors also tilts the balance of power in favor of nationalists. We find no evidence for the effects of industrialization, the advent of mass literacy, or increasingly direct rule, which are associated with the modernization theories of Gellner, Anderson, Tilly, and Hechter. Nor is the growing global hegemony of the nation-state model a good predictor of individual instances of nation-state formation, as Meyer's world polity theory would suggest. We conclude that the global rise of the nation-state is driven by proximate and contextual political factors situated at the local and regional levels, in line with historical institutionalist arguments, rather than by domestic or global structural forces that operate over the *long durée*.

Keywords

nationalism, nation-state, diffusion, historical institutionalism

The French and American revolutions of the late-eighteenth century gave birth to the ideal of the modern nation-state—an independent state with a written constitution, ruled in the name of a nation of equal citizens. During those days, all other states were still governed on the basis of other principles of legitimacy. In dynastic states, a prince was entitled to assume the mantle of power upon the death of his father (as in the multi-ethnic Habsburg and Ethiopian empires); in theocracies, religious leaders guided their flocks in worldly matters as well (e.g., in Tibet and Montenegro); Ottoman and Spanish elites spread the true faith across

the globe, British governors brought progress to “backward” peoples in far-away places, and, during the twentieth century, the party cadres of the Soviet Union advanced a revolutionary, transnational project in the name of the world's working classes. Kings, theocrats, and imperial elites attempted to extend their states' boundaries irrespective of the

^aUniversity of California, Los Angeles

Corresponding Author:

Andreas Wimmer, Department of Sociology,
University of California, Los Angeles, 264 Haines
Hall, Box 951551, Los Angeles, CA 90095-1551
E-mail: awimmer@soc.ucla.edu

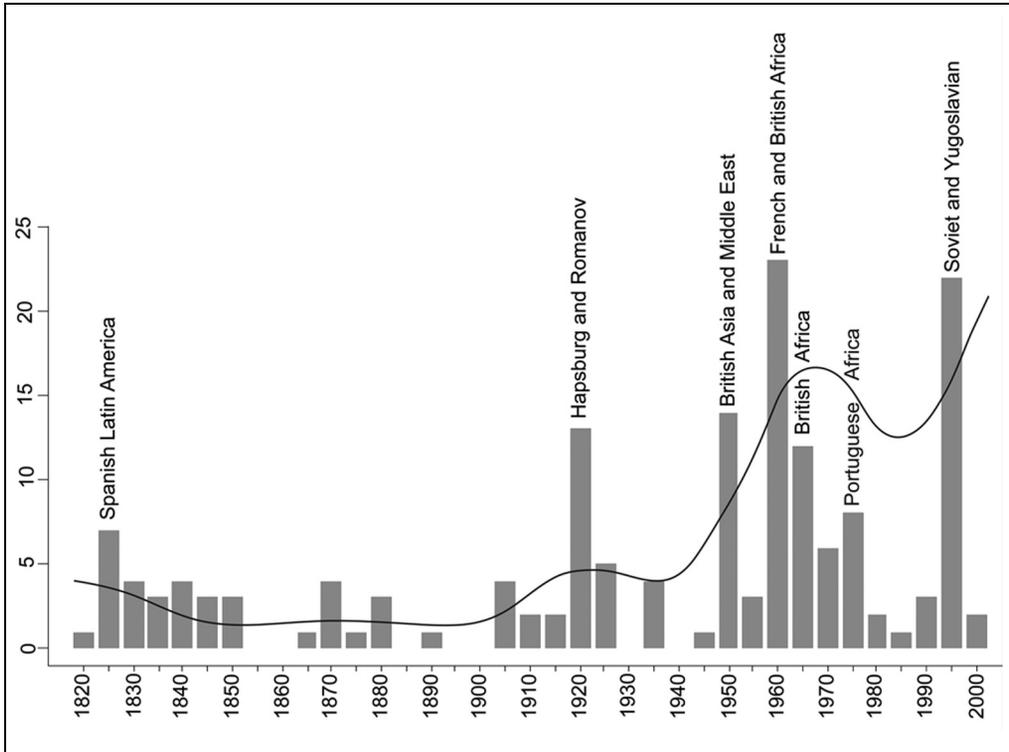


Figure 1. Number of Nation-States Created per Five-Year Period, Smoothed Hazard Rate

ethnic backgrounds of those who came under their rule.

Compare that situation to the world today: empires have dissolved, theocracies have been dethroned, and only a handful of countries, mostly in the Middle East, are still governed as absolutist monarchies comparable to pre-revolutionary France, where the king ruled in the name of God and represented the House of Bourbon, not the French nation. The once revolutionary template of political legitimacy—self-rule in the name of a nation of equal citizens—is now almost universally adopted. This framework is recognized as the essence of modern statehood, so much so that the terms “nations” and “states” are often used interchangeably. Figure 1 shows that the global ascent of the nation-state over the past 200 years was a discontinuous process, unfolding in various waves linked to the break-up of large empires.

Understanding the global rise of the nation-state is one of the most formidable tasks of comparative historical sociology—on par with the analysis of the emergence of sovereign, territorial states in early modern Europe (see Tilly’s [1975] pioneering work). Why did modern states—once they emerged out of the dynamics of war-making, bureaucratic centralization, and increasing taxation—become nation-states? A rich literature has developed to answer this question, including the well-known oeuvres of Gellner, Anderson, Smith, Hechter, and Meyer. This research tradition displays two main weaknesses. First, many general theoretical statements are meant to explore universal processes that could account for the rise of the nation-state in the modern world as a whole, but empirical support for these generalizations is often based on examples picked selectively, sometimes in a merely

illustrative manner (deplored by Breuilly 2005; Wimmer 2008). Second, more empirical research on particular trajectories of nation-state creation tends to be segmented along regional and disciplinary lines. For example, the political science literature on decolonization (Spruyt 2005; Strang 1990) and nation-building (Bendix 1964) in the postcolonial world developed quite independently from debates among historical sociologists about the origins of the nation-state in the West. Yet another strand of scholarship investigates the historical developments that led to the collapse of the land-based Ottoman, Habsburg, or Soviet empires and subsequent waves of nation-state creation (e.g., Barkey and von Hagen 1997; Roshwald 2001; Saideman and Ayres 2008). Given that nation-states cover almost the entire world by now, one wonders whether an integrated view might be within reach.

To overcome some of these limitations, we assembled a new, global dataset that allows identification of those patterns of nation-state formation that recur across continents, empires, and time periods. This required considerable efforts because only independent nation-states systematically collect information on their economies and societies. Available datasets thus do not allow us to understand why such states emerged in the first place, which is perhaps the main reason why quantitatively minded scholars have so far shied away from a more systematic evaluation of existing theories of nation-state formation (but see the work of Strang and Roeder, to be discussed in later sections). The new dataset introduced here includes independent states, colonies, and imperial dependencies over two centuries, and contains almost the entire universe of nation-state creations. It provides information on 145 of today's states from 1816 until the years they achieved nation-statehood (or 2001 if they did not). Many of the variables in this dataset—for example, the length of railways, government expenditures, and literacy rates—had to be assembled by extracting

information from secondary sources, such as country histories. Despite its limitations, this new, global dataset allows us to assess the plausibility of major theories of the nation-state from a global, comparative perspective.

Many comparative historical sociologists are deeply skeptical of the use of quantitative techniques and large datasets because these seem to imply a disregard for contingency, context, and complexity. We show that at least some of these concerns can be addressed within a quantitative framework. A focus on recurrent patterns does not rule out the importance of contingency, and enough cases will be “off the regression line” to motivate in-depth case studies based on historical methods. Context can be transformed into cause by using dummy variables that indicate whether a group of cases or a particular time period differs from the rest (Collier and Mazzuca 2006) or by using subsample analysis (Young 2009). Causal complexity can be addressed, for example, by exploring whether the effects of one variable depend on the values of another. In other words, some of the configurational logic of inquiry that underlies traditional comparative historical sociology can be translated into a quantitative research design, while keeping the advantage of being able to generalize across a large number of cases. This is particularly important if the outcome of interest recurs across the globe—as is the case with nation-state formation.

HYPOTHESES AND EXISTING QUANTITATIVE STUDIES

Alas, a quantitative approach to historical processes comes at a price. Not only must we content ourselves with proxy variables that measure the hypothesized processes imperfectly, we also cannot address the rich arguments offered by past scholarship in an as nuanced way as one would wish. Effectively, we test crucial elements of theories—the correlation between core conditions and outcomes—but

not whether the postulated mechanisms linking conditions to outcomes are actually at work. Our empirical analysis therefore does not pretend to submit whole theories to a sort of Popperian falsification test. Rather, we focus on key arguments that can be evaluated for a wide range of territories over long periods.

What are these theories theories *of*? They do not mainly concern the emergence of the nation-state model in the United States, France, or perhaps earlier in Britain, but rather its subsequent proliferation across the world. While sharing this common focus, many classical authors are somewhat ambiguous as to whether their primary aim is to explain nationalism as a political movement, the spread of national consciousness among a population (i.e., nation-building), or the shift in the institutional set-up of the state (i.e., the creation of a nation-state). They all concur, however, that these three processes are closely related to each other, if through different pathways. In Anderson's account, nationalism leads to nation-building and eventually a nation-state, while according to Gellner, nationalists form nation-states that then build their nations. World polity theorists such as Meyer, by contrast, consider neither nationalism nor nations to be a necessary condition for nation-states to emerge. Our own historical institutionalist approach assumes that nationalists create nation-states, whether or not nations have already been built. All of these arguments contain the nation-state as a central element in the analytic tableau; the emergence of nation-state institutions therefore represents an appropriate dependent variable for this study.

Economic Modernization

According to Gellner (1983), the epochal shift from an agricultural to an industrial society brings about nationalism and eventually the nation-state. In the agricultural

empires of the past, the economic system contained many highly specialized niches reproduced through on-the-job training in the specific skills demanded. The industrial mode of production, by contrast, needs a mobile and flexible labor force. A rationalized, standardized education in a common language provides workers with the generic skills to shift from job to job and communicate effectively with strangers. The educational apparatus of a nation-state eventually provides the new, standardized, and homogenized culture that industrial societies need.

Gellner's (1983) functionalist analysis is complemented by a subtle study of four historical pathways through which industrial society's needs were met (we will discuss only the two most important). First, uneven industrialization drew rural peasants into industrialized centers, where their ascent and prospects remained limited if their language and culture did not correspond to the center's high culture. Resentment fed into nationalism and eventually led to the creation of nation-states, as in the Balkans and the peripheries of the Habsburg Empire. Second, a similar process unfolded in the colonial world, where skin color was associated with unequal power, unleashing anti-colonial nationalisms as soon as industrialization set in and delegitimized the colonial hierarchy.

These different trajectories describe how industrial society, arriving at different times in different parts of the world, led to the reorganization of political boundaries along cultural lines and the formation of nation-states. Focusing on this general association, rather than the mechanisms that bring it about, we can state the simple hypothesis that the likelihood of nation-state creation should increase with industrialization (Hypothesis 1).

Political Modernization

Tilly's (1994), Mann's (1995), and Hechter's (2000) political modernization theories shift our attention to the system of governance.

Starting in the sixteenth century, permanent war between competing European states made techniques of governmental control and resource extraction ever more effective and efficient. Indirect rule via regional elites and notables was replaced by direct rule through a unified and hierarchically integrated bureaucracy. From there, two major pathways led to the nation-state. In autonomous states (e.g., France), state elites gradually homogenized the population over the course of the nineteenth century and developed an assimilatory nationalism to legitimize their rule (Hechter 2000; Tilly 1994). In Mann's (1995) related, yet differently accented, account of this process, nationalism emerged from below to justify the public's demands for democratic representation vis-à-vis the increasingly interventionist military state.

Far more frequent than the transition to the nation-state within existing boundaries, however, is the second, secession trajectory. In the multiethnic empires of the Habsburgs and the Ottomans (and according to Hechter also in Yugoslavia and beyond), the shift to direct rule led to nationalist mobilization by peripheral elites who resented being governed by ethnic others and sought to re-establish self-rule. Whether such state-seeking nationalists are successful depends on additional (including international) factors and forces. Simplifying these accounts by subsuming these additional factors and forces under a *ceteris paribus* clause, we can derive Hypothesis 2: the more directly a territory is ruled, the more likely nation-state formation should be.

Cultural Modernization

Anderson's (1991) theory of nationalism distinguishes between three mechanisms that combine in different ways across four different waves of nation-state creation. First, reformation, state bureaucratization, and, most importantly, the rise of print capitalism

enabled and propelled literacy in vernacular languages, replacing complex elite languages such as Latin. The emerging reading public thus shared a narrative cosmos and soon imagined itself as a national community of common origin and future political destiny.

Mass literacy was less important for the first wave of nation-state creation than for subsequent waves. Overall literacy levels were still low when the first wave rolled over Latin America, but Anderson nevertheless sees the emergence of newspapers and reading publics as crucial. Mass literacy then became central to the empowerment of second wave linguistic nationalisms in nineteenth-century Europe, as well as for the third wave, when dynastic rulers sought to contain nationalism by adopting it as a state doctrine themselves. Mass literacy remained a central causal force during the fourth wave, leading to decolonization after World War II. Anderson's first mechanism of nation-state formation thus leads to Hypothesis 3: an increase in the literacy rate in vernacular language should make nationalism, nation-building, and ultimately a transition to the nation-state more likely.

Anderson's second mechanism comes into play during the first and fourth waves of nation-state creation. Why did Bolivia, the Ivory Coast, and Vietnam become independent states, rather than the whole territory of Spanish Latin America, French West Africa, or French Indochina as one would expect in view of the popular literacy argument? Low-level colonial administrators recruited from the local population could not aspire to positions above the provincial levels, Anderson argues, which led to resentment and growing nationalist dissent. Being confined to the provincial bureaucratic space laid the groundwork for imagining the nation along provincial, rather than linguistic, lines. During the fourth wave, the European colonies' vastly expanded educational system not only helped fuel nationalism by spreading literacy, but it also reinforced the provincial

segmentation of these nationalisms, especially where educational and administrative boundaries overlapped, as in Indonesia (Anderson 1991). This suggests Hypothesis 4: a territory that corresponds to a province or a state should be more likely to see nationalism arise and more likely to eventually become a nation-state.

The third, and perhaps least crucial, mechanism in Anderson's account relates to global diffusion processes, which are especially important for the last wave of nation-state formation in the former colonies, as well as in Japan, Thailand, and Switzerland. Such global processes are at the heart of Meyer's world polity approach.

World Polity Theory

Meyer's diffusion theory emphasizes external influences rather than domestic modernization. Meyer and his co-authors show that the nation-state template is part of a world culture that emerged over the past 200 years and eventually became institutionalized in the United Nations. This world culture gradually forced state elites and political challengers alike to adopt nationalism as a template of political legitimacy and the nation-state as the most legitimate form of statehood (Meyer et al. 1997).

World polity theory offers a cross-sectional and a longitudinal argument. First, the more linkages a territory maintains to the centers of global culture and power, the more its elites are exposed to world culture and the more likely they will adopt world-cultural templates and create a nation-state (Hypothesis 5). Second, the likelihood of a transition to the nation-state should increase the more territories of the world have already adopted the nation-state (Hypothesis 6).

Historical Institutionalism

Historical institutionalism (Lachmann 2009; Pierson and Skocpol 2002) emphasizes the

power configurational and political factors overlooked by modernization and global diffusion theories. Transition from one form of state organization to another is seen as the outcome of a struggle between various politically organized segments of society. The balance of power between these actors determines which vision of a legitimate political order and which institutional principles will prevail. In this view, proximate causes, most importantly the power configurations between actors, trump slowly moving structural forces, and external diffusion is more important than endogenous modernization.

Historical institutionalism usually takes the emergence of new templates of political legitimacy as exogenously given. It is a theory of selection, as it were, not of mutation. The global diffusion of nationalism is thus outside the purview of the historical institutionalist argument we develop here. It is enough to assume an imitation process through which a variety of political movements across the world adopt nationalist ideologies (Greenfeld 1992). Because the first nation-states (i.e., Great Britain, the United States, and France) happened to be the most powerful states in the world from the eighteenth century to today, these political movements "pirated" nationalism, to use Anderson's felicitous term, hoping they would one day preside over states that matched the military glory, political might, and cultural prestige of these powerful nation-states. This imitation process proceeds along established networks of political and cultural relations: African nationalists were inspired by the might of France or Great Britain, Turkish and Japanese nationalists resented these two imperial powers and looked at their German nemesis for inspiration, Kurdish and Arab nationalists oriented themselves on Turkish models, and so on. This diffusion process is neither driven by the hegemonic power of a uniform world culture, nor by domestic modernization process, but rather follows the logic of a decentralized contagion process. If this accounts for

nationalism's global appeal as a political ideology and its subsequent diffusion across the world, under what conditions are nationalists able to establish nation-states?

We propose the following set of hypotheses that refer to various domestic and international aspects of the power configuration. A power shift in favor of nationalism is more likely when nationalists are able to convert existing elites to their cause or reach out to larger segments of the population, beyond the intellectual circles, army factions, clergymen, and colonial bureaucrats who are often the first supporters of nationalist movements (Hroch [1969] 2000). This process of empowerment has political and symbolic aspects. Nationalists need to build networks of political organizations and alliances and effectively scandalize the existing regime as an instance of "alien rule" or as a sclerotic and fragmented *ancien régime* unable to withstand domination by powerful nation-states. This effectively undermines the legitimacy of the ethnopolitical hierarchy that characterizes many empires and dynastic states. Disregarding short-term cycles of popular mobilization and demobilization (Beissinger 2002), we assume that nationalists' political and symbolic power increase monotonically. The more time nationalists have to propagate their worldview and establish networks of followers, the more powerful they will be vis-à-vis non-nationalist forces and the more likely they will succeed in eventually establishing a nation-state (Hypothesis 7).

The power balance also depends on the strength of the established regime—its capacity to resist nationalist forces and to avoid conversion to the nationalist cause, institutional reform in the direction nationalists propose, abdication, or the loss of territory to nationalist secession. Following Skocpol's (1979) lead, we assume that wars fought either on the territory in question or elsewhere in an empire reduce the established elites' staying power and make

a nationalist revolution possible. The likelihood of nation-state creation thus increases with an increase in the number of wars fought on a territory or within an empire (Hypotheses 8 and 9). Similarly, a pre-nationalist state's global military and economic standing should influence its capability to co-opt or suppress nationalist movements and thus maintain the status quo, making nation-state creation on its territories unlikely (Hypothesis 10, as proposed by Strang [1990]).

Finally, diffusion of the nation-state within empires and between neighboring territories can also shift the power balance in favor of nationalists.¹ Empires establish a communicative field within which peripheral political elites observe and imitate each other's nation-building projects or mirror that of the imperial center (Fieldhouse 1966). Recently established nation-states within the same imperial domain demonstrate that independence is feasible and that the center is no longer willing or able to uphold the status quo. The new nation-states not only provide a model to follow, but also new alliance partners in the imperial political arena, and thus empower nationalist movements and further delegitimize the pre-national regime. This leads to Hypothesis 11: the more territories within an empire that have become nation-states, the more likely that other territories will follow. Note that this imperial diffusion effect is theoretically distinct from global diffusion because the source of the external influence is different (i.e., the empire versus the world), as is the mechanism through which it operates (i.e., contagion versus imposition).²

Within neighborhoods, a new nation-state may propel its own replication in adjacent territories through similar demonstration and alliance effects. It might also take the more bellicose form of a domino effect through competition over ethnically mixed or ill-defined territories (Weiner 1971; Wimmer 2002): if one territory is organized

as a nation-state and demands inclusion of its ethnic kin population across the border, this increases pressure in the adjacent territory to adopt the nation-state model as well. Hypothesis 12 states that the likelihood of nation-state formation increases with the number of nation-state creations in a neighborhood.

Existing Datasets and Quantitative Findings

To date, only two quantitative studies have explored the dynamics of nation-state formation, both with a different focus and a much smaller universe of cases. Strang's (1990) work attempts to understand the conditions under which colonial dependencies became independent states from 1870 to 1987. He finds support for Wallerstein's hegemonic cycles theories (decolonization is more likely when a global hegemon rules), world polity theory (decolonization accelerated after the UN general assembly adopted an anti-colonial statement in 1960), balance of power arguments (colonies governed by a metropole with strong naval capability are less likely to become independent), and imperial diffusion effects. Strang (1991) arrived at similar findings in a related study with fewer variables but a longer time span. This second study also reports a global diffusion effect, measured by the number of colonies that have already achieved independence. We build on this endeavor by enlarging the empirical horizon to include autonomous states, dependencies of land-based empires (e.g., of the Ottomans, Romanov, and Habsburgs), as well as the former Soviet Union and Yugoslavia. Our dataset also improves on data quality and adds variables that are relevant to the classic literature on nationalism and the nation-state.

In the second existing study, Roeder uses a global dataset to test his institutional capacity argument. A large degree of institutional

autonomy, he maintains, allows provincial elites to establish cultural and ethnic hegemony within their territories and provides them with the political resources necessary to successfully challenge the metropolis and establish an independent state. This argument parallels Anderson's provincial confinement hypothesis and can thus be tested, in an approximate way to be sure, with Hypothesis 3 (i.e., territories that correspond to a province or an autonomous state should be more likely to become nation-states).

Roeder's (2007) dataset uses autonomous provinces, including colonial dependencies, as units of observation and provides information on 336 units from 1901 to 2000. He shows that the likelihood of nation-state formation increases if a substate unit is self-governing, if central elites are weakened by internal strife and political turmoil, if the provincial population is excluded from political participation or is linguistically and religiously different from the core population, and if the province experienced independent statehood prior to incorporation into the current state.

Roeder's institutional capacity model advances the theoretical understanding of nation-state formation by revealing the importance of the balance of power between nationalist contenders and representatives of the *ancien régime*. Many of his arguments parallel our own understanding of the rise of the nation-state. However, his quantitative analysis displays weaknesses that raise doubts about the validity of this specific version of the balance-of-power argument. Most importantly, there are serious sample selection problems. His dataset misses many provinces that never became nation-states, such as German *Bundesländer*, U.S. states, and Swiss cantons, although these enjoyed as much autonomy as the states of India that do appear in his list. The Ottoman Empire has only Bulgaria, Crete, and Samos listed as substate units, and none of the Ottoman villayets, which never developed into nation-states.

DATASET AND MODELING APPROACH

The two existing quantitative studies thus leave room for improvement. Each excludes parts of the world that were candidates for making the transition to the nation-state: the land-based Eurasian empires stretching from Vienna to Vladivostok, or the nineteenth-century waves of nation-state creations in the Americas, the Balkans, and Western Europe. Furthermore, both studies include a list of variables that is mostly unrelated to classical theories of nation-state formation, perhaps because corresponding data is not easily available.

Units of Observation

Our dataset contains information on 145 territories from 1816 until a nation-state was created. By 2001, 139 of these territories had made the transition to the nation-state, while the others were still governed as absolutist monarchies. All territories refer to the geographic boundaries of countries that existed in 2001. Data coverage is almost complete; we miss only mini-states with less than 20,000 km² surface plus eight larger states for which no literacy data are available.³ We also exclude early nation-state creations in Great Britain, France, Paraguay, and Haiti because they occurred before our data series starts in 1816. In line with most classical theories of nation-state formation, our analysis is not about the origins of the nation-state but the general mechanisms that might help explain its subsequent proliferation across the world.

Some notes on the units of observation may be in order. To clarify, our approach leads us to combine data from various polities in the case of territories that did not correspond to a political unit at a given point in time. Data for Poland in the 1870s, to give an extreme example, are proportionally combined from the Russian, German, and

Austrian empires that controlled pieces of what is today Poland. The disadvantage of creating units that do not conform to political entities is outweighed by the advantage that these units actually experienced the event: it is Poland that became a nation-state, not any of these Russian, German, or Austrian provinces on their own. Medical studies of mortality take a similar approach: they use individuals, rather than families or couples, as units of observation because it is individuals, not families, who die. Constant territorial units also have the advantage of defining a stable risk set. They allow us to pursue "Poland" and other territories throughout history, rather than having to deal with a different set of units every time a state ceases to exist or comes into being, or every time provincial boundaries are redrawn, as would be the case if the dataset was composed of provinces or states.

Does choosing today's countries as units mean that we select on the dependent variable because we include only successful instances of nation-state creation and not failed ones? Note that the grid of states defines our units in 2001, independent of whether these are nation-states. Most are nation-states simply because the nation-state has proliferated so widely, not because our research design excludes non-national states.

To be sure, we define the risk set retrospectively; there is no Kurdistan, Tibet, West-Sahara, or Southern Sudan in our dataset, although they could eventually become nation-states. This potential selection problem is less severe than it might first appear because we know from past research (Wimmer and Min 2006) that few secessionist states break away from nation-states and most emerge from imperial polities. No such empires are left today. We thus do not expect a large number of nation-states to be created in the foreseeable future. We believe our risk set, while certainly not complete, nevertheless captures the overwhelming majority of possible events.⁴ Furthermore, there is no reason to believe that the dynamic of future

nation-state creations will differ from that of the past (i.e., there is no sorting mechanism biasing our results systematically).

Dependent Variable

To use event history methods, we need to identify the particular point in time at which a territory became a nation-state.⁵ We code year of nation-state creation when sovereignty shifted from kings, emperors, or theocrats to the nation. Sovereignty has a domestic and an external component. Domestically, a written constitution claims a nationally defined community of equal citizens as the political (and moral) foundation of the state and foresees some institutional representation of this community (not necessarily a freely elected parliament). Internal sovereignty thus stands in opposition to dynasticism, theocracy, feudal privilege, and mass slavery. Externally, national sovereignty means control over foreign policy decisions that affect the nation, and it stands in opposition to foreign rule of all sorts. These two conditions must be cumulatively fulfilled. The definition and most data were adopted from Wimmer and Min (2006).

Note that we focus on principles of political legitimacy, regardless of how far these have been realized in practice. For example, we code autocracies as nation-states as long as the dictator claims to rule “in the name of the people.” The transition to a nation-state is therefore not always associated with democratization—sometimes a nation-state is as autocratic as the regime it replaced. Less than 50 percent of independent states were democracies by 2000 (Gleditsch and Ward 2006), while over 95 percent were nation-states. The extent to which citizenship rights are granted to all citizens is also not of concern here. We disregard the fact that property restrictions on voting rights and denial of citizenship rights to women usually last many decades into nation-statehood.⁶

Following these coding principles, 24 territories experienced more than one episode of nation-state formation, which we treat as independent from each other. The Baltic territories, for example, were independent national states from 1918 to World War II, when they were swallowed again by the Soviet Union, and they regained nation-state status in 1991. In other cases, a territory achieved nation-state status as part of a larger state, which subsequently broke apart into smaller nation-states. Such was the case for Gran Colombia (which later gave birth to Colombia, Venezuela, Ecuador, and eventually Panama), the Central American Republic (which fragmented into Guatemala, Honduras, Costa Rica, Nicaragua, and El Salvador), Yugoslavia, and Czechoslovakia. To check for selection bias in this definition of the risk set, we ran all our models with a restricted definition of the outcome that excludes these repeated events (see Table 2). Our results remain almost identical. A list of nation-state creations per territory can be found in the online supplement (<http://asr.sagepub.com/supplemental>).

Independent Variables

To test the economic modernization argument, we code the length of railway tracks (in km) per 1,000 square kilometers. We adopt data from the monumental compendium of historical statistics assembled by Mitchell (various years) and code other data from primary sources, which are remarkably rich thanks to the enthusiasm that the history of railways has sparked among lay and professional scholars (see the online supplement for sources). Is this an adequate proxy to measure Gellner’s notion of industrialization? A flexible labor market is the crucial element in his analysis of industrial society, while the manufacturing of industrial goods is not given much consideration. It therefore makes sense to include fully commercialized agricultural and extractive economies into his definition, both of which are historically associated

with railway construction. This is also justified by Gellner's treatment of African colonies as representing early stages in industrial society's development—despite the almost complete absence of manufacturing.

Certainly, a more direct measurement of the degree of industrialization or the flexibility of the labor market would be preferable. The railroad variable, however, offers the possibility of full data coverage for the entire dataset, while it would be unthinkable to collect global data on the sectoral distribution of the labor force, for example. The railroad variable is also very precise; it is possible, for instance, to find out how many kilometers long the railways of colonial Burma were in 1880, but it is quite impossible to learn how many people were employed in which professions. We did use existing historical data on energy consumption, urbanization, iron and steel production, and the percentage of the labor force employed outside agriculture as alternative measurements, available for only 1,000 to 2,000 pre-nation-state observations, and we arrived at similar results (not shown here) as in the models using the railway variable.⁷

Does the rise of the automobile, which increasingly replaced railcars in the twentieth century, make the length of railway tracks a poor proxy variable for industrialization? Most industrialized countries that ceased to build railways after World War I or II achieved nation-statehood before they started to rely on automobiles. While some Communist dependencies and colonies let their railway systems decline, many more continued to build railways after World War II and well into the 1970s and beyond. We ran models for the pre-World War I period alone and found that results do not change significantly.

To test Anderson's main argument about the role of literacy in generating nationalist imaginaries, we assembled data on adult literacy rates for all territories in the database, relying on published country studies, government censuses, historical research on

particular regions, and existing quantitative datasets (see the online supplement for sources). In all cases, we found estimations within a 10-year range from the year of nation-state creation; to interpolate, we combined these data with the best estimates available for the beginning of the time series in 1816, as well as several later data points. The reliability and comparability problems characteristic of literacy estimates (see Reis 2005), especially before census taking became widespread in the 1870s (Kaestle 1985), also haunt our efforts. In view of the substantial variation of estimated literacy rates over time and across territories, however, we believe that the quality of the data is sufficient to justify inclusion of this variable in the study.

To test Anderson's "provincial horizon of identity" argument and Roeder's theory of institutional capacity, we code for each year whether a territory *corresponds to an autonomous state or a substate unit* (1) or not (0). Given that our territorial grid is fixed, while political boundaries might change over time, we effectively test whether nation-state creation is more likely if a territory previously enjoyed some degree of institutional autonomy as a province or a state. We count provinces, colonies, mandate territories, *vilayats* and *sanjaks*, and Russian governorates, for example, as relevant substate units, and we took into account dozens of reorganizations of colonial and imperial provinces. We code pre-colonial territories that were divided between various indigenous states or were stateless as 0.

The degree of directness of rule plays a crucial role in Tilly's and Hechter's theories of nationalism and the nation-state. It can be approximated by calculating *government expenditure* for a particular territory, assuming that the more a government spends on a territory, the deeper administrative penetration will be, increasing the government's interference in local affairs and its ability to circumvent local power brokers. We again use data from Mitchell (various years) and

complement this with additional sources for the Ottoman and Spanish empires and the Soviet Union (see the online supplement). We give imperial dependencies (but not colonies) the same values throughout an empire, assuming that land-based empires were more uniform than sea-born colonial empires in modes of territorial control. Pre-colonial territories that had not developed indigenous states are assigned a value of 0 (because, by definition, they cannot be ruled directly). All other polities (including pre-colonial states such as Dahomey, Burma, and Morocco) are given their proper values or coded as missing values.⁸ These coding rules generate many missing values; we therefore test this variable in separate models.⁹

Is this variable a good proxy for directness of rule? Lange (2005) developed a more specific measurement for British colonies, using the percentage of court cases handled by traditional courts as a proxy. This variable shows a high correlation with our measurement of direct rule via government expenditure ($r = .82$ for 19 data points, using non-standardized values). In addition, our measurement seems to capture historical shifts in directness of rule quite adequately. For example, the decentralization of the Hapsburg monarchy after the *Ausgleich* in 1867—which effectively created two federal states, one under Hungarian and one under Austrian control—is faithfully reflected in a dramatic decrease in Vienna’s government expenditure for the Hungarian lands.

We generated our diffusion variables by counting the *number of territories governed as nation-states in the neighborhood, within the same imperial domain, or in the world*. We then created a variable that reflects the number of nation-states established during the previous five years because imitation and domino effects are best captured by a dynamic coding. All results reported below hold when using a total count or the percentage of nation-states in the world, empire, or neighborhood. To test the cross-sectional version of the global diffusion argument,

we code *number of memberships in international governmental organizations* by the polity to which a territory belonged. We assign dependencies the same value as their centers, assuming that world cultural values spread from a colonial or imperial metropolis to its peripheries. All data are adopted from the Correlates of War (COW) project.

To evaluate whether *wars in the territory or the empire* affect the creation of nation-states, we use Wimmer and Min’s (2009) dataset of wars in all territories of the world from 1816 to 2001. This dataset allows us to distinguish between inter-state wars, civil wars, and nationalist wars of independence. We code a large number of different war variables to test whether specific types of wars are more effective in weakening the political center and thus shifting the balance of power in favor of nationalists.

Strength of nationalist challengers is proxied by years elapsed since the foundation of the first national organization. To count as *national organization*, its membership must be defined formally (thus excluding clientelist networks and informal factions) and leadership roles must be institutionalized independent of individuals (thus excluding a political leader’s personal followings). In addition, an organization had to claim to represent the national community in the name of which the territory eventually became governed—without being necessarily nationalist in the strict sense of the term.¹⁰ To statistically distinguish the monotonically increasing strength of nationalist challengers from the presence or absence of nationalism as such, we include a dummy variable that codes 1 for *all years since the first national organization* was founded. The existence of nationalism should make nation-state creation much more likely, given that our whole historical institutionalist argument is premised on this antecedent.

To take the political center’s *capacity to resist* nationalist movements into account, we rely on the COW dataset’s “composite index of national capabilities,” which

combines energy consumption, military expenditure, number of soldiers, steel production, urbanization, and population size (Singer 1987). The index reflects a country's share of total economic and military power available in the world (ranging from 0 to 100 percent). All dependent territories are assigned the value of the imperial or colonial center; pre-colonial, stateless territories are coded as 0. Unfortunately, we have to code autonomous states not listed in the COW dataset as missing values. Table 1 presents correlation matrix and summary statistics.

Modeling Approach and Time Specification

We designed the dataset as an unbalanced panel with territories as units. All independent variables are lagged one year to avoid endogeneity problems. We use discrete-time event history models, estimated via a logistic regression analysis of territory years, to analyze the effect of covariates on the likelihood of nation-state creation. Following standard practice, we cluster standard errors on territories to account for the non-independence of observations within territories. Given the chronic instability of cross-national regression results (see Young 2009), we ran all models with multiple specifications of dependent and independent variables and different combinations of variables. We report only results that are robust to all different model specifications.

One of the major challenges in event history analysis is how to conceptualize effects of time. Much has changed between 1816 and 2001 that our independent variables cannot capture. The simplest possibility is to include a linear time trend in the statistical model (as is done in most diffusion models, see Strang [1991]). This tests whether the passing of time itself affects the outcome, each year being more (or less) likely to see nation-states created than the one preceding it. One can also use discrete time periods,

such as decades, to see whether specific periods are particularly likely to produce nation-states. We also use natural cubic splines (the standard in comparative political research [see Beck, Katz, and Tucker 1998]), which allow us to control for nonlinear trends. We ran all models with all three time specifications and rely only on results that remain robust across these different specifications (Tables 2 and 3 show the results with cubic splines).

Substantively, the splines describe a constant hazard rate until World War II and a sharply increasing baseline risk thereafter, similar to the smoothed hazard rates in Figure 1. This trend could be interpreted in line with Strang's world polity argument by attributing it to an increase in the nation-state's legitimacy after the United Nations' founding in 1945 (although Strang [1990] himself points to an anti-colonial UN declaration of 1960 as a critical turning point).

We are uncomfortable with the post-hoc-ergo-propter-hoc nature of these interpretations. The post-1945 surge in the baseline hazard rate might be due to the increased global power of the United States, champion of decolonization and self-determination. Or it may relate to the unprecedented growth of the global economy that made many more nation-state projects economically feasible. Or it may capture the decreasing popularity of the colonial project in France and Britain—which may or may not be related to the decreasing legitimacy of colonialism in the world as a whole.¹¹ In conclusion, period dummies and general time trends rarely provide conclusive evidence in support of a particular substantive argument.

RESULTS

The correlation matrix in Table 1 shows that many of the developmental variables associated with modernist theories of nation-state formation are strongly correlated. We

Table 1. Means, Standard Deviations, and Correlations among Variables

	R	L	Y	E	S ₁	S ₂	I	NS _W	NS _E	NS _N	P	W _E	W _T
R: Length of railway tracks per km ²													
L: Percent literate among adults	.555		.542	.025	.320	.298	.140	.299	-.009	.091	.010	.025	-.006
Y: Years since first national organization founded		.667		.292	.449	.526	.253	.522	.041	.101	.132	-.007	.0025
E: Central government expenditure for territory			.218		.435	.539	.266	.507	.054	.120	.051	-.019	.053
S ₁ : 1st cubic spline on year				.494	.419	.494	.299	.5021	.0842	.0527	-.005	.011	.041
S ₂ : 2nd cubic spline on year				.897		.897	.753	.9103	.2221	.1281	-.045	.096	.008
I: Center's membership in IGOs						.726		.9526	.2383	.1341	-.070	.048	.016
NS _W : Total number of nation-states in world							.6589		.3978	.0796	-.006	.220	.008
NS _E : Number of nation-states created in the empire in past five years								.217		.128	-.054	-.003	.006
NS _N : Number of nation-states created in neighborhood in past five years										.205	-.036	.191	.031
P: Center's share of global power											-.074	.065	.108
W _E : Number of wars fought in an empire												.254	.041
W _T : Number of wars fought in a territory													.102
Mean	5.671	20.563	11.655	.447	1888.442	21.581	15.203	40.360	.252	.087	7.372	.539	.056
Standard deviation	16.044	27.161	24.185	1.249	46.625	30.221	20.803	26.898	1.076	.361	8.033	1.134	.252
N	17,500	17,667	17,667	9,821	17,667	17,667	17,522	17,667	17,667	17,667	16,488	17,522	17,522

therefore introduce them in separate models to test their explanatory power independent from each other. We maintain variables in subsequent models if results are statistically significant and substantially plausible until we arrive at what we consider the best fitting explanation of nation-state creation. We include continental dummies in all our main models to account for unobserved heterogeneity across world regions.

Main Findings

Model 1 in Table 2 includes the length of railway tracks as a proxy to test the core hypothesis derived from Gellner's work. The variable fails to achieve standard levels of significance, even if we restrict the sample to pre-1914 years, during which the length of railways is a better proxy for industrialization than afterward. The main reason why we find no general association between industrialization and nation-state formation (according to additional analyses not reported here) is that early nation-states in Latin America were created in a preindustrial environment and many weakly industrialized African territories achieved nation-statehood in the 1960s, while the highly industrialized Soviet and Yugoslav provinces had to wait another generation to accomplish nation-statehood. This combination of historical developments might explain—if we leave the possibility of systematic measurement error aside for a moment—the lack of association between industrialization and nation-state creation.

This hints at the importance of power configurational factors that are largely missing from Gellner's account and which we will explore further below: The Soviet Union had the power to keep nationalist movements in its highly industrialized provinces in check for generations, not least by co-opting and controlling minority elites into the governments of the republics. Spain, preoccupied with the Napoleonic invasion and a subsequent

civil war, failed to achieve the same in its far-away colonial empire, long before it had been touched by industrialization.

Model 2 contains variables associated with Anderson's (and Roeder's) approach. Territories that correspond to the boundaries of provinces or states are not more likely to become nation-states. The literacy variable is negative and significant (indicating that more literate societies are *less* likely to become nation-states). This result is not robust to other time specifications, however, which all produce insignificant results. We therefore do not consider it reliable. What happens if we look more precisely at only the types of territories for which Anderson thought mass literacy or being a province would be most effective in bringing about a nation-state? The political entity variable is also insignificant for the colonial dependencies of Europe, where it should matter most (results not shown). The literacy variable also fails to reach standard levels of significance if we analyze only European territories, that is, the domains of second- and third-wave nationalisms for which high literacy levels should be a catalyst (results not shown).

Why does the literacy variable not behave as Anderson expected? If we again exclude the possibility of systematic measurement error for the moment,¹² additional analysis (not shown here) suggests the following: After World War I, literacy was promoted heavily by Communist regimes all over Eastern Europe and the colonial empires, especially in Africa, long before the creation of nation-states in these areas of the world. Perhaps these regimes were able to keep nationalists at arm's length by inducing the population to imagine—at least temporarily—other, non-national communities, such as a world-spanning family of subjects loyal to her Majesty, or the revolutionary working classes of the world. The relationship between mass literacy and nationalism might therefore be less straightforward than Anderson's account suggests.

Table 2. Parameter Estimates for Logistic Regression Models of Nation-State Creation

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6 (fixed effects)	Model 7 (first NSCs only)
Length of railway tracks per km ²	.005 (.004)						
Territory corresponds to state or province		-.300 (.222)					
Percent literate among adults		-.016** (.005)					
Central government expenditure for territory			-.213* (.084)				
Center's membership in IGOs				.008 (.005)			
Total number of nation-states in world				-.025** (.008)			
Number of nation-states created in the empire in past five years				.165** (.038)	.124** (.034)	.304** (.063)	.103** (.035)
Number of nation-states created in neighborhood in past five years				.780** (.121)	.512** (.124)	.419** (.141)	.659** (.149)
Existence of national organization					1.087** (.292)	3.60** (1.135)	1.294* (.346)
Years since first national organization					-.007* (.003)	.075* (.036)	.008* (.004)
Number of wars fought in the empire					.297** (.051)	.585** (.073)	.315** (.054)
Number of wars fought in the territory					.481** (.182)	.604* (.29)	.490* (.213)
Center's share of global power					-.061* (.029)	-.063 (.162)	.053 (.031)
Center's share of global power × dependency					-.108** (.03)	.047 (.168)	-.117** (.038)
Dependent territory					.406 (.287)	-3.082** (1.196)	.502 (.376)
Middle East		-2.839** (.468)	-2.459** (.524)	-3.165** (.442)	-2.298** (.524)		-2.301** (.537)
Eastern Europe		-2.362** (.312)	-2.640** (.452)	-2.187** (.292)	-2.048** (.410)		-1.461** (.442)

(continued)

Table 2. Continued

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6 (fixed effects)	Model 7 (first NSCs only)
Africa	-1.864** (.271)	-3.147** (.475)	-1.898** (.359)	-2.437** (.280)	-1.682** (.421)		-1.670** (.436)
Asia	-1.728** (.294)	-2.905** (.477)	-1.901** (.371)	-1.850** (.257)	-1.037** (.389)		-983* (.399)
Oceania	-957** (.347)	-1.071** (.329)	-1.035* (.521)	-854* (.345)	-037 (.483)		.082 (.544)
Latin America	.265 (.474)	-378 (.525)	-1.309 (.748)	-339 (.439)	.135 (.377)		-029 (.414)
1st cubic spline on year	-007 (.005)	-002 (.005)	.020 (.011)	-010 (.005)	-015** (.005)	.095** (.028)	-014* (.006)
2nd cubic spline on year	.039** (.006)	.045** (.006)	.018 (.012)	.062** (.010)	.042** (.006)	-0233 (.028)	.039** (.007)
Constant	8.644 (9.857)	1.457 (9.628)	-42.428* (21.375)	14.131 (9.784)	21.959* (9.773)		19.33 (10.976)
Observations	17,500	17,522	9,821	17,522	16,488	15,533	15,351
Likelihood Ratio	266(9)**	280(10)**	185(9)**	375(12)**	448(17)**	636(11)**	431(17)**
BIC	-178	-182	-102	-258	-283	-582	-267

Note: Robust standard errors in parentheses.
* $p < .05$; ** $p < .01$ (two-tailed tests).

Model 3 integrates our measurement of directness of rule, which reduces the number of observations by almost half because we could not find data for most pre-colonial territories. The variable is robustly significant but with a negative coefficient, indicating that the more directly a territory is ruled by the political center, the less likely it is to eventually become a nation-state. This is contrary to expectations derived from Hechter's and Tilly's arguments. It seems that more directly ruled states are more capable of resisting the pressure to shift to the nation-state model and can keep nationalists in check by "buying" the population's consent. However, this result is mostly due to the Middle Eastern kingdoms that have not yet experienced a transition to the nation-state despite very high government expenditures per capita. This is the only such dependence on a few observations in all the results that we report in this article.

To analyze the full dataset again, we exclude the government expenditure variable from subsequent models and note that its effects are largely the same when additional variables are added to the equation (results not shown). Model 4 introduces diffusion variables. The term for global diffusion is negative and significant because of collinearity with the cubic splines. This result is therefore substantially meaningless. In contrast to Strang's (1991) analysis of decolonization,¹³ we thus do not find any support for the longitudinal version of the world polity argument. A political center's number of IGO memberships—that is, the variable to test its cross-sectional aspects—also fails to achieve significance. A territory that is more integrated into the world polity is not more likely than a more isolated territory to become a nation-state.

We do, however, get strong and meaningful evidence for diffusion at the imperial and neighborhood levels. The number of nation-states founded during the past five years within an empire and within a territory's neighborhood increases the likelihood of

nation-state creation substantially, pointing toward imitation and domino mechanisms.¹⁴ Our findings thus parallel research on the transition to democracy, which has been shown to diffuse within networks of related states (Torfason and Ingram 2010) and among geographical neighbors (Gleditsch and Ward 2006), rather than within a uniform global space.

Examples of cases that underlie these results include the increasing pressure on Bolivia's remaining royalists after the Bolivarian spirit of nationalist revolution gained a foothold in most neighboring territories; the impetus to create a modern nation-state and to abandon the Hamidian search for imperial restoration felt by the Committee for Union and Progress when one Ottoman province in Rumelia after the other became an independent state; and the powerful demonstration effect that Indian independence had on many nationalist movements in the British Empire.

The competing nation-state-building projects on the Spanish peninsula, born out of nationalist resistance against Napoleon's occupation, provide examples of neighborhood diffusion effects. The Portuguese liberal revolutions, and thus the creation of a modern Portuguese nation-state, were directed as much against the absent emperor (who fled to Brazil) as against the state's traditional rival, Spain, whose newly formed mass army—modeled after the French *peuple en arme* it had just defeated in the world's first guerilla war—threatened Portugal's independence.

Model 5 contains variables that should produce a shift of power from pre-national elites to nationalists. Our proxy for the strength of nationalist movements—years that have passed since the foundation of the first national organization—has a significant effect on the likelihood of nation-state creation. Because we also include a dummy for the period after the foundation of the first national organization, this effect is net of the existence of nationalism per se. However,

Table 3. Additional Tables: Time Periods and Imperial Contexts

	Model 1 Before 1915	Model 2 After 1914	Model 3 Decades ^a	Model 4 Empires ^b	
Number of nation-states created in the empire in past five years	.297** (.104)	.134** (.037)	.110** (.040)	.135** (.049)	
Number of nation-states created in neighborhood in past five years	.496* (.232)	.634** (.158)	.486** (.123)	.630** (.124)	
Years since first national organization	.025** (.005)	.010** (.003)	.018** (.003)	.019** (.003)	
Share of global power × dependency	-.150* (.059)	-.028 (.017)	-.056** (.014)	-.047 (.026)	
Number of wars fought in the empire	.521** (.097)	.239** (.051)	.289** (.047)	.318** (.043)	
Number of wars fought in the territory	.818** (.230)	.394 (.224)	.661** (.173)	.555** (.183)	
1821 to 1840			1.580 (1.013)	.650 (.530)	Spanish
1841 to 1855			.636 (1.068)	.080 (.403)	Hapsburg
1856 to 1870			.414 (1.073)	.461 (.400)	Romanov
1871 to 1885			.303 (1.102)	-.391 (.513)	Ottoman
1886 to 1900			-1.084 (1.408)	-.588 (.370)	Yugoslav
1901 to 1915			.868 (1.062)	-.777 (.505)	Soviet
1916 to 1930			.820 (1.041)	.444 (.347)	French
1931 to 1945			-.153 (1.119)	.455 (.458)	British
1946 to 1955			1.379 (1.025)	.168 (.373)	Dutch
1956 to 1965			2.456* (1.008)	-.732 (.569)	Portuguese
1966 to 1975			1.906 (1.032)	.748 (.607)	Other empires
1976 to 1985			-.411 (1.409)	.168 (.299)	Independent states
1986 to 1995			3.129** (1.026)		
1996 to 2001			2.761* (1.28)		
Observations	11,116	5,372	16,488	16,421	
Likelihood Ratio	112(8)**	164(8)**	406(20)**	374(20)**	

Note: Robust standard errors in parentheses. Cubic splines are included in all models except Model 3. The coefficients of the splines and the constant are omitted from the table.

^aYears 1816 to 1820 are the reference category.

^bPre-colonial territories are the reference category.

* $p < .05$; ** $p < .01$ (two-tailed tests).

the coefficients of the two variables suggest that the existence of nationalism is itself a strong factor. This is in line with the basic power configurational argument, which assumes that nationalist forces first need to emerge before they can eventually take over an existing state or found a new one.¹⁵

Another core hypothesis associated with the power configurational argument is that nationalists will be more successful if the center is weakened by wars—similar to Skocpol's analysis of revolutions. Indeed, both the number of wars fought within a territory during the past year and the number of wars fought within an empire (excluding those fought on a territory) significantly affect the likelihood of nation-state creation. Examples of the latter include the civil war in Bolivia between royalists and Bolivarists, the nationalist wars of liberation that helped bring about the Baltic republics' independence, and the Russian-Turkish war that allowed Bulgaria to become an independent nation-state in 1879 after Ottoman forces had crushed a Bulgarian rebellion four years earlier.

World War I, which debilitated the Habsburg and Soviet empires and enabled a wave of nation-state creations in their domains, provides a prime example of the role of wars in other parts of an empire. Other cases include the Mau-Mau rebellion in Kenya and the Malaysian anti-colonial Communist insurgency, which decreased the British Empire's willingness to hold on to its imperial possessions and helped accelerate Ghana's independence—the first on the continent. Similarly, the bloody struggles in Algeria and French Cameroon weakened France's capacity and willingness to erect further obstacles against decolonization in its West-African domains in 1960.¹⁶

According to our theory, a center's power to resist nationalist challengers also depends on its international standing. The center's global power variable is negatively related to the likelihood of a transition to nation-statehood. The effect is different, however, for self-ruled territories and for dependencies.

An interaction term is significant and negative, while the sign of the coefficient of the non-interacted term is positive and significant. This latter result disappears if we use decades or a linear time trend instead of cubic splines to specify the effects of time (see Model 3 in Table 3); thus, we do not rely on this result. We conclude that imperial states that are powerful players in the international arena can more easily co-opt, control, or suppress nationalist movements and prevent the establishment of nation-states in their dependent territories. As an example for the opposite case, we can point to Spain, which could not contain or co-opt Creole nationalists in its New World possessions or fight independence movements effectively after its fleet was dramatically decimated in the famous battle of Trafalgar and its attention was further diverted by Napoleon's occupation of her lands.

Model 5 therefore lends strong and consistent support to the various hypotheses associated with our historical institutionalist model of nation-state creation. Model 6 reports results of a territory fixed-effect model, which takes into account that many time-invariant characteristics of individual territories, which are not adequately captured by our set of variables, might influence the likelihood of nation-state creation. To put it simply, fixed-effect models give much more weight to within territory changes over time than do standard models. With the exception of the center's share of global power variable, to which we turn in a moment, all variables remain statistically significant in the expected direction. Model 7 excludes instances of repeated nation-state creation on the same territory (as in the Central American republics), which might bias our results; they stay almost identical, however. The same holds true if we consider only the last events in cases of repeated nation-state creations (results not shown). Note also that Model 5 stays largely unchanged if we drop all observations of one continent from the sample (also not shown).

Context and Contingency

Do our findings hold across time and across the different waves of nation-state creation that rolled over the modern world? Is not the story of the dissolution of the Habsburg Empire quite different from that of the break-up of the Soviet Union? Should we expect variables to affect outcomes in different ways in the early-nineteenth century than in the late-twentieth century? Table 3 shows the results of some additional tests meant to answer these questions. Model 4 demonstrates that none of the various imperial domains—for example, Ottoman, Romanov, or British—is significantly different from the others or from territories that remained autonomous throughout history, such as Japan or Switzerland. No decade stands out as particularly prone to nation-state creation (see Model 3), with the exception of the decades starting in 1956, when Africa was decolonized, and the 1990s, when the Soviet Union dissolved and Yugoslavia disintegrated (more on this below). If we look only at observations either before 1914 (Model 1) or after 1914 (Model 2), roughly the midpoint in our data series, we discover that the results are fairly similar. After World War I, wars fought in a territory no longer significantly affect the likelihood of nation-state creation, and the center's share of global power variable is only borderline significant.¹⁷

We took a closer look at how the effects of the share of global power variable change over time. It is negative (making nation-state creation less likely) until the 1970s, but it has a positive effect thereafter (making nation-state creation more likely).¹⁸ This is because many nation-states were created within the domains of the Soviet Union in the nineties, while Moscow still commanded an extraordinary share of global military and economic power, especially compared with some small Gulf states that continued to resist the siren songs of nationalism and thus remain in the risk set throughout the 1990s.

According to our model and data, Moscow would have had the capacity to fight or co-opt independence movements—yet nation-states popped up all over its domains. Similarly, the dissolution of the French and British empires in Sub-Saharan Africa were engineered in advance (for British Africa, see Flint [1983]) and in the end supported, rather than fought, by the imperial center (on French West Africa, see Chafer [2002]). Does the lack of willingness to uphold and defend an imperial domain explain why so many territories achieved nation-statehood in 1960 and 1991, such that the two corresponding decades stand out compared to all others in Model 3? Does contingency play a role in these two waves of nation-state creations—contingent, that is, from the point of view of a power configurational argument that does not foresee a lack of willingness to use one's power?

Alternatively, we might have a problem of data quality, rather than theoretical misspecification and contingency. The following analysis focuses on the Soviet case, but similar arguments could be made with regard to African decolonization. First, it might be that the Kremlin was already too weak domestically to use its global military and economic power against nationalist independence movements, even if it had wanted to. Our data are not fine-grained enough to capture power relations between the Russian president Boris Yeltsin—who famously stood on a tank in the center of Moscow amid thousands of supporters—and the putsch generals who wanted to roll back the nationalist movements and strengthen the USSR's control over the empire, including over Russia.

Second, the yearly resolution of our data does not capture fast moving possible diffusion effects. For example, the inspiration that Baltic declarations of sovereignty (in 1988) and independence (in September 1991) provided to leaders of the Caucasian republics (who declared independence in early December of 1991); or the fact that

Yeltsin's Russian nationalism and declaration of independence in mid-December provoked the collapse of the remaining Soviet Union and left Central Asian republics no choice but to embrace independence themselves, which they did later that month.¹⁹

With better data and a more fine-grained analysis, we could gain a deeper understanding of the complex dynamics of political mobilization, contestation, repression, diffusion, and imitation that change the balance of power between nationalists and existing elites within days or weeks (see Beissinger's [2002] superb analysis). Our global dataset, stretching from the fall of Napoleon's empire to the beginning of the twenty-first century, is not well equipped to handle this task. It can, however, highlight the waves of nation-state formation in which such fast-moving dynamics are particularly likely to have been consequential.

CONCLUSIONS

Past comparative historical scholarship has explored various routes of nation-state formation: reform from above as in Japan; gradual transition into nation-statehood as in Sweden and Thailand; the overthrow of an *ancien régime* through revolution as in Russia or through civil war as in the United States; nationalist secession as in Yugoslavia and Mexico; and unification movements such as in Germany and Yemen. Independent of which of these routes a territory travels down, our analysis suggests that the success of nationalist projects is determined by the constellation of power relating nationalist movements and factions to imperial centers, *ancien régimes*, or other factions of the ruling elites.

These nationalist movements emerge through an imitation process driven by the extraordinary success and global dominance of the first nation-states. Nation-states are subsequently created around the world wherever a power shift allows nationalists to

overthrow or absorb the established regime, quite independent of whether domestic modernization processes have readied a society for nation-building. Our analysis shows that such a power shift is more likely when nationalists have had ample time to mobilize the population and delegitimize the old regime or when the established regime is weakened by wars. Diffusion of nation-states among neighbors or within the same empire also empowers nationalists by providing a model to follow and new alliance partners on which to rely. On the other hand, nationalists who struggle against an imperial center are at a disadvantage when the empire has considerable global military and economic power that it can use to fight independence movements.

We thus integrate balance-of-power and diffusion mechanisms into a simple power configurational model that includes domestic and international dimensions as well as military, political, and symbolic aspects of the balance of power. In contrast to political modernization arguments, this model emphasizes more proximate political factors, such as war and the political standing of imperial elites. With regard to Gellner's and Anderson's classical theories, our analysis suggests that the slow moving forces of economic and cultural modernization did not play a crucial role in the rise of the nation-state across the world.²⁰ In contrast to world polity theory, we find diffusion effects operating within neighborhoods and imperial domains, while the growing global hegemony of the nation-state template—certainly a historical fact worth underlining—is not a good predictor of individual instances of nation-state creation. This hints at the possibility that this global hegemony might result from the worldwide rise of the nation-state rather than causing it.

This global outcome—the almost universal adoption of the nation-state form—therefore emerges from local and regional processes that are not coordinated or causally produced by global social forces. As in

epidemiology, processes of contagion follow established networks of political relationships and cultural similarity that span the entire world. The logic of contagion is purely regional and produces a decentralized pattern of diffusion, all the while generating the illusion of a systemic process when seen from a global point of view. Future research should attempt to theoretically decipher and empirically capture such processes with more precision than we have been able to do here, perhaps by revitalizing some of Gabriel Tarde's anti-Durkheimian models of imitation and diffusion.

We conclude with a note on what our analysis implies for the contemporary salience of nationalism and the future of the nation-state. It is unlikely that many more countries will experience the kind of fragmentation and division into a series of nation-states that characterized the dissolution of the Soviet, British, and Ottoman empires. To be sure, secessions from established nation-states will continue to occur, as the recent creations of Kosovo, East Timor, and Montenegro illustrate. And the few existing non-national states in the Middle East and elsewhere might experience a constitutional revolution in the future. Overall, however, we do not expect new waves of nation-state creations to sweep over the world. The nationalist dream of organizing the world into a series of states that provide a roof for each culturally defined people, to use a Gellnerian metaphor, has come close to being realized. History, however, refuses to ever come to an end. It represents a trail traversed in the past, not a compass to determine its future direction. Generations to come will certainly imagine other communities than the nation and reshape the world's political landscape according to tectonic principles that we cannot possibly imagine today.

Acknowledgments

The authors wish to thank Nurullah Ardic, Philippe Duhart, Wesley Hiers, Hazem Kandil, Yana Kucheva, Joy Morgen, Anoop Sarbahi, Frank Yen, Jeremy Yobe,

and Sarah Zingg Wimmer for superb research assistance. Mauro Guillén, Robert Mare, Brian Min, Gabriel Rossman, Art Stinchcombe, and Don Treiman provided help and advice regarding research methods and data management. We are especially grateful for several rounds of discussions with Robert Mare. Angus Maddison helped with the standardization of economic data. We received advice regarding the early literacy rates of individual territories from Harvey Graff, Aaron Althouse, Nola Jean Cook, Keith Weller Taylor, Ólöf Gardarsdóttir, Loftur Guttormsson, Luca Godenzi, Alois Gehart, Jörg Baten, Brian Min, and Gustav Brown. Previous versions of the article were presented at departmental lecture series of the sociology departments of Northwestern University, UCLA, the University of Washington, and Cambridge University; the political science departments of the University of Chicago and George Washington University; at the conference "Comparing Past and Present" organized by the Comparative Historical Section of ASA in 2009; the Social Science History Association Annual Meeting of 2009; and the Association for the Study of Nationalism meeting in 2010. We thank conveners and audiences for stimulating comments and challenging criticisms. Rogers Brubaker offered written comments and suggestions on a first draft. The article greatly profited from the helpful advice and critical encouragement by outstanding *ASR* reviewers.

Notes

1. Gleditsch and Ward (2006) describe a similar relation between regional diffusion and local empowerment in their analysis of democratization.
2. Our argument incorporates several of the diffusion mechanisms identified by Dobbin, Simmons, and Garrett (2005). The diffusion of nationalism corresponds to the "follow the leader" effects emphasized by constructivist sociologists. Imperial diffusion is similar to the "copying between similar countries" mechanism underlined by sociologists or the "channeled learning" between networked actors studied by political scientists. Finally, neighborhood diffusion contains aspects of the copying mechanism, but also of the competition effects studied by economists.
3. Due to missing literacy data, we exclude Albania, Belize, Djibouti, Equatorial Guinea, Greenland, Iceland, Lesotho, and Namibia. We do include Gambia, Kuwait, Cyprus, Bahrain, Qatar, and Mauritius, all of which control less than 20,000 km² of territory.
4. A brief discussion of the advantages and disadvantages of other possible definitions of the risk set might be in order. First, one could choose the states of 1816 as fixed units of observation, avoiding the

- problem of “coding history backward.” However, with the exception of Germany, Italy, and Yemen, these units are much larger than those that actually made the transition to the nation-state, making it difficult to identify a single event for these units. Second, one could code the political units that exist in any given year, rather than fixed territorial units. While all independent variables would be coded in relation to meaningful political arenas, the outcome (i.e., nation-state creation) often would not be attributed to units that actually experienced it (e.g., in the case of empires that dissolved into a series of nation-states). Furthermore, these political units (states or provinces) would produce different risk sets over time, depending on how states appear and disappear from the political map and how provincial boundaries are redrawn, making causal inference more difficult. A third alternative research design would be to choose constant spatial units of observation, for example, using a $100 \times 100 \text{ km}^2$ grid. This would have the advantage of generating a risk set completely independent of the outcome. The problem is that these units correspond neither to arenas within which political processes unfold nor to areas that eventually experienced the nation-state event. The question would then arise of whether to code a separate nation-state event for all units that eventually formed a nation-state together, and how to deal with the fact that these events are obviously dependent on each other. Finally, one could argue that today’s federal states (e.g., Nigeria and India) contain units that are still at risk of becoming nation-states of their own, especially if the federal provinces are populated by ethnic minorities. We redefined our risk set accordingly (using data from Christin and Hug [2009]) and found no major differences in the results. This research design obviously increases measurement error because we assign the same values for all independent variables to all substate units, and it is based on an ex-post definition of which states contain federal subunits.
5. While this is problematic for gradual processes of nation-state formation that had few historical turning points or dramatic ruptures (e.g., Sweden), it is quite unproblematic for most cases in our dataset. In any case, the alternative of coding “degrees of nation-stateness” for each year and each territory would be utterly unfeasible.
 6. We do not define states as modern nation-states that, in their constitutions, exclude segments of the population from the citizenry (e.g., the slaveholding United States before the Civil War). We do not consider subsequent repeals of citizenship status for segments of the population (e.g., Jews in Nazi Germany) as cases of reversal into a pre-nation-state situation. Changing this handful of coding decisions does not affect the results.
 7. Still, railway track length is correlated .48 with the percentage employed outside agriculture based on 196 data points from Vanhanen (2000), and .65 for the years prior to 1970. The correlations of percentage employed outside agriculture with other measurements of industrialization contained in the Correlates of War (COW) dataset are similar: .68 with energy consumption per capita, .49 with total energy consumption, and .47 with iron and steel production per capita (all correlations based on only 633 observations).
 8. We code these variables in various ways (e.g., by assigning 0 to all pre-colonial and pre-modern autonomous states, assuming that direct rule necessarily means modern, bureaucratic rule). None of the results depend on these coding decisions.
 9. We standardized all units as per capita figures, converted into U.S. constant dollars using time-varying conversion rates, and then additionally adjusted for purchasing power differences using Maddison’s (2003) GDP estimations. We also estimated our models with unadjusted data and found no significant differences in the results.
 10. Most of the information used for this variable is based on Woronoff (Various years).
 11. It is interesting to note, in this context, that Britain had already started to prepare for decolonization in the late 1930s (Flint 1983), while Nazi Germany’s imperial project was in full swing.
 12. One could argue that the foundation of the first national organization would be a better dependent variable to test Anderson’s mass literacy theory. We find no significant association between levels of literacy and the foundation of a national organization as soon as continental dummies are included in the model (results not shown); the same holds true for the railroad length and government expenditure variables, while diffusion mechanisms within empires and among neighbors seem to be at work again.
 13. Note, however, that Strang’s (1991) finding does not hold up if observations after 1955 are dropped, that is, without African decolonization. Furthermore, we find no robust effects for Strang’s (1990, 1991) other global-level variable—Wallerstein’s period of hegemony (results not shown). Also contrary to Strang’s findings, nation-states are not more likely to emerge among dependencies of democratic centers or in settler societies. Our results do confirm, however, his findings regarding imperial diffusion effects and the international power of an imperial or colonial center.
 14. Are these really imitation and domino effects operating at the imperial and neighborhood levels, or are these regional and local manifestations of the pressure to adopt the nation-state emanating from the *global* system? First, these variables are significant for the pre-1915 subsample as well, when global pressures

- were arguably much weaker (see Table 3). Second, additional analysis shows that the absolute number of nation-states in an empire or a neighborhood increases the likelihood of nation-state creation in the remaining territories. One would not expect such a cumulative effect if we were dealing simply with local manifestations of a global adoption pressure.
15. The years since the foundation of a national organization variable is sensitive to inclusion of a small group of territories. If we exclude the dummy variable from the model, however, this sensitivity disappears entirely.
 16. All results relating to the war variables are similar if we code the average number of wars fought over the past five years. Are these wars simply one way that nationalists achieve an independent nation-state, as shown by Wimmer and Min (2006), rather than an independent factor that explains when nationalists are strong enough to gain power? First, our coding of the imperial war variable excludes wars fought on the territory in question. The imperial war variable is thus not directly connected to independence struggles in the territory itself. Second, while nationalist wars of independence within an empire are significant on their own, so are non-nationalist wars (results not shown). The same holds true for wars in the territory. The war variables therefore capture two different mechanisms: first, a specific avenue through which nationalists achieve nation-statehood (thanks to successful nationalist wars of independence in a territory, in line with Wimmer and Min's argument); second, a tilting of the balance of power in favor of nationalists through wars that are unrelated to the nationalist struggle in a particular territory but that weaken established state elites and thus facilitate a nationalist revolution (similar to Skocpol's account of peasant revolutions).
 17. Share of global power and wars in the territory variables lose significance from 1880 onward. The model is much more robust to right-hand truncation (which, in contrast to left-hand truncation, does not create an incomplete, and thus problematic, risk set). All covariates remain significant in subsamples that exclude years after 1880 or beyond; only the neighborhood diffusion variable is borderline or insignificant when years after 1850 are dropped.
 18. We interacted all independent variables with linear time to determine whether their effect changes linearly across history. This is the case only for the center's share of global power.
 19. For empirical support of these diffusion effects based on weekly data, see Hale (2000).
 20. We should mention two interesting non-results here. First, the size of the largest ethnic group (data are from Fearon [2003] and for the 1990s) has no effect on the likelihood of nation-state creation. This sheds doubts on Anthony Smith's (1990) argument

that nationalist mobilization is easier to achieve within territories with a demographically dominant ethnic group that can form the ethnic core of a future nation. This result is all the more remarkable because the obvious reverse causation problem in this analysis works in favor of the hypothesis: early nation-state creation could cause ethnic homogeneity over the long run through assimilation or ethnic cleansings, rather than being caused by it. Second, some scholars argue that Protestantism provided an ideal breeding ground for the nation-state, mostly due to its ideological affinities to nationalism (Gellner 1983; Smith 2003). We find, however, that Catholicism is robustly associated with the likelihood of nation-state creation (using data from Barrett [1982]).

References

- Anderson, Benedict. 1991. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London, UK: Verso.
- Barkey, Karen and Mark von Hagen. 1997. *After Empire: Multiethnic Societies and Nation-Building: The Soviet Union and the Russian, Ottoman, and Habsburg Empires*. Boulder, CO: Westview.
- Barrett, David. 1982. *World Christian Encyclopedia: A Comparative Study of Churches and Religions in the Modern World, 1900–2000*. Oxford, UK: Oxford University Press.
- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *Journal of Political Science* 42:1260–88.
- Beissinger, Mark. 2002. *Nationalist Mobilization and the Collapse of the Soviet State*. Cambridge, UK: Cambridge University Press.
- Bendix, Reinhard. 1964. *Nation-Building and Citizenship: Studies in Our Changing Social Order*. New York: John Wiley.
- Breuilly, John. 2005. "Dating the Nation: How Old is an Old Nation?" Pp. 15–39 in *When is a Nation? Towards an Understanding of Theories of Nationalism*, edited by A. Ichijo and G. Uzélac. London, UK: Routledge.
- Chafer, Tony. 2002. *The End of Empire in French West Africa: France's Successful Decolonization?* Oxford, UK: Berg.
- Christin, Thomas and Simon Hug. 2009. "Federalism, the Geographic Location of Groups, and Conflict." University of Geneva, Geneva, Switzerland. Unpublished manuscript.
- Collier, Ruth Berins and Sebastián Mazzuca. 2006. "Does History Repeat?" Pp. 472–89 in *Contextual Political Analysis*, edited by R. E. Goodin and C. Tilly. Oxford, UK: Oxford University Press.
- Dobbin, Frank, Beth Simmons, and Geoffrey Garrett. 2007. "The Global Diffusion of Public Policies:

- Social Construction, Coercion, Competition, or Learning?" *Annual Review of Sociology* 33:449–71.
- Fearon, James D. 2003. "Ethnic and Cultural Diversity by Country." *Journal of Economic Growth* 8: 195–222.
- Fieldhouse, David Kenneth. 1966. *The Colonial Empires*. New York: Dell.
- Flint, John. 1983. "Planned Decolonization and its Failure in British Africa." *African Affairs* 82:389–411.
- Gellner, Ernest. 1983. *Nations and Nationalism*. Ithaca, NY: Cornell University Press.
- Gleditsch, Kristian Skrede and Michael D. Ward. 2006. "Diffusion and the International Context of Democratization." *International Organization* 60:911–33.
- Greenfeld, Liah. 1992. *Nationalism: Five Roads to Modernity*. Cambridge, MA: Harvard University Press.
- Hale, Henry. 2000. "The Parade of Sovereignties: Testing Theories of Secession in the Soviet Setting." *British Journal of Political Science* 30:31–56.
- Hechter, Michael. 2000. *Containing Nationalism*. Oxford, UK: Oxford University Press.
- Hroch, Miroslav. [1969] 2000. *Social Preconditions of Patriotic Groups among the Smaller European Nations*. New York: Columbia University Press.
- Kaestle, Carl F. 1985. "The History of Literacy and the History of Readers." *Review of Research in Education* 12:11–53.
- Lachmann, Richard. 2009. "Greed and Contingency: State Fiscal Crisis and Imperial Failure in Early Modern Europe." *American Journal of Sociology* 115:39–73.
- Lange, Matthew. 2005. "British Colonial State Legacies and Development Trajectories: A Statistical Analysis of Direct and Indirect Rule." Pp. 117–40 in *States and Development: Historical Antecedents of Stagnation and Advance*, edited by M. Lange and D. Rueschemeyer. New York: Palgrave Macmillan.
- Maddison, Angus. 2003. *The World Economy: Historical Statistics*. Paris, France: OECD.
- Mann, Michael. 1995. "A Political Theory of Nationalism and its Excesses." Pp. 44–64 in *Notions of Nationalism*, edited by S. Periwal. Budapest: Central European University.
- Meyer, John, John Boli, George M. Thomas, and Francisco O. Ramirez. 1997. "World Society and the Nation-State." *American Journal of Sociology* 103:144–81.
- Mitchell, Brian R. Various years. *International Historical Statistics*. New York: Palgrave.
- Pierson, Paul and Theda Skocpol. 2002. "Historical Institutionalism in Contemporary Political Science." Pp. 693–721 in *Political Science: State of the Discipline*, edited by I. Katznelson and H. V. Milner. New York: Norton.
- Reis, Jaime. 2005. "Economic Growth, Human Capital Formation and Consumption in Western Europe before 1800." Pp. 195–226 in *Living Standards in the Past*, edited by R. C. Allen, T. Bengtsson, and M. Dribe. Oxford, UK: Oxford University Press.
- Roeder, Philip G. 2007. *Where Nation-States Come From: Institutional Change in the Age of Nationalism*. Princeton, NJ: Princeton University Press.
- Roshwald, Aviel. 2001. *Ethnic Nationalism and the Fall of Empires: Central Europe, Russia and the Middle East, 1914–1923*. London, UK: Routledge.
- Saideman, Stephen and William Ayres. 2008. *For Kin and Country: Xenophobia, Nationalism, and War*. New York: Columbia University Press.
- Singer, J. David. 1987. "Reconstructing the Correlates of War Dataset on Material Capabilities of States, 1816–1985." *International Interactions* 14:115–32.
- Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia and China*. Cambridge, UK: Cambridge University Press.
- Smith, Anthony D. 1990. "The Supersession of Nationalism?" *International Journal of Comparative Sociology* 31:1–31.
- . 2003. *Chosen Peoples: Sacred Sources of National Identity*. Oxford, UK: Oxford University Press.
- Spruyt, Hendrik. 2005. *Ending Empire: Contested Sovereignty and Territorial Partition*. Ithaca, NY: Cornell University Press.
- Strang, David. 1990. "From Dependency to Sovereignty: An Event History Analysis of Decolonization 1879–1987." *American Sociological Review* 55:846–60.
- . 1991. "Global Patterns of Decolonization, 1500–1987." *International Studies Quarterly* 35:429–54.
- Tilly, Charles. 1975. "Western State-Making and Theories of Political Transformation." Pp. 601–686 in *The Formation of National States in Western Europe*. Princeton, NJ: Princeton University Press.
- . 1994. "States and Nationalism in Europe, 1492–1992." *Theory and Society* 23:131–46.
- Torfason, Magnus Thor and Paul Ingram. 2010. "The Global Rise of Democracy: A Network Account." *American Sociological Review* 75:355–77.
- Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy, 1810–1998." *Journal of Peace Research* 37:251–65.
- Weiner, Myron. 1971. "The Macedonian Syndrome: An Historical Model of International Relations and Political Development." *World Politics* 23:665–83.
- Wimmer, Andreas. 2002. *Nationalist Exclusion and Ethnic Conflicts: Shadows of Modernity*. Cambridge, UK: Cambridge University Press.
- . 2008. "How to Modernize Ethnosymbolism." *Nations and Nationalism* 14:9–14.
- Wimmer, Andreas and Brian Min. 2006. "From Empire to Nation-State: Explaining Wars in the Modern

- World, 1816–2001.” *American Sociological Review* 71:867–97.
- . 2009. “The Location and Purpose of Wars around the World: A New Global Dataset, 1816–2001.” *International Interactions* 35:390–417.
- Woronoff, Jon. Various years. *Historical Dictionaries*. Lanham, MD: Scarecrow Press.
- Young, Cristobal. 2009. “Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth.” *American Sociological Review* 74:380–97.
- Andreas Wimmer** is Professor of Sociology at the University of California, Los Angeles. His research aims to understand the dynamics of nation-state formation, ethnic boundary making, and political conflict from a comparative perspective.
- Yuval Feinstein** is a PhD candidate in the Department of Sociology at the University of California, Los Angeles. He is studying why international wars fought by democracies tend to be supported by the public during “rally-round-the-flag” periods.