Comparative Political Studies
1–39

DOI: 10.1177/0010414015592645
cps.sagepub.com
SAGE

Is Diversity Detrimental? Ethnic Fractionalization, Public Goods Provision, and the Historical Legacies of Stateness

Andreas Wimmer[1]

Abstract

Existing research has shown that highly diverse countries tend to provide less public goods. This article argues, by contrast, that the relationship is spurious: both contemporary ethnic heterogeneity and low public goods provision represent legacies of a weakly developed state capacity inherited from the past. Classical theories of state formation are then tested to show that favorable topography and climate, high population densities, as well as a history of warfare are conducive to state formation. Using an instrumental variable approach, I show that previous ethnic diversity is not consistently an impediment to the formation of indigenous states and thus to contemporary public goods provision. Empirically, this article uses three different measurements of public goods provision and data on pre-colonial levels of state formation in Asia and Africa to test these various hypotheses.

Keywords

state building, race, ethnicity and politics, public administration

[1]Columbia University, NJ, USA

Corresponding Author:
Andreas Wimmer, Department of Sociology, Columbia University, Knox Hall, 606 West 122nd Street, MC 9649, New York, NY 10027, USA.
Email: aw2951@columbia.edu

In politics, academia, and business, ethnic and racial diversity is now embraced as a positive aspect of human life, with the exception of some radical anti-immigrant movements. Fostering diversity is seen as tool for enhancing a vibrant, post-nationalist society by governments favoring a liberal immigration regime; universities seek to create a faculty and student body that mirrors the diverse ethnic and racial composition of the population at large; business organizations believe that enhancing diversity will lead to creativity and innovation—and help with minority costumers.

On the other hand, however, social scientists find that ethnic and racial diversity is detrimental for social trust, economic development, public goods provision, political peace, and more. Several different micro-mechanisms have been proposed of why diversity should have such detrimental consequences, from coordination problems that arise if diverse publics hold different preferences, to an intrinsic difficulty to trust ethno-racial others or sanctioning them for violating norms of co-operation.

This article seeks to re-evaluate these empirical claims. It zooms in on one particular aspect of the overall debate: whether or not countries with a more diverse ethnic make-up of the population are indeed less able or willing to provide its citizens with public goods such as education, health, and physical infrastructure. It does not address whether diversity has detrimental consequences for public goods provision at the level of cities, villages, organizations, and other sub-national communities.

The core contribution of this article is to look at ethnic diversity and the provision of public goods from a longer term, political development perspective that spans many generations. Rather than treating ethnic diversity as an exogenous variable that is unaffected by political developments, similar to climate or topography, it should be viewed as the consequence of the history of state formation and nation building that at the same time shapes contemporary capacity to provide public goods. In other words, diversity and the state's capacity to public goods may *both* result from a previous history of state formation.

Strong states offered incentives for minorities to adopt the language and culture of the dominant *Staatsvolk*, thus decreasing diversity over the generations. Well-developed states also left a legacy of indigenous bureaucratic capacity, on which colonial rule often rested, which shapes contemporary ability to provide public goods in path-dependent ways. The association between today's diversity and public goods provision is therefore brought about by this shared historical legacy and not by a direct causal connection. To evaluate this argument empirically, I proceed in three steps.

The first step uses data on adult literacy rates, infant mortality, and railroad density as proxies for public goods provision by the state, in line with the recent quantitative literature on public good provision. To be sure, these are not "pure" public goods in the strict sense of the term as originally defined by economists (Olson, 1965) because they are not completely non-rival (if everyone used health care facilities at the same time, everyone gets less attention) and they are not entirely non-excludable either (not everybody gets the same chance at an education, see Kramon & Posner, 2012). But they are the kinds of outcomes that the political economy literature on government provided public goods has focused upon in the past.

Data on these three outcomes are available for almost all states of the world from 1945 or the year of independence onward (reliable data on infant mortality rates start in 1960, however). Regression analysis reveals that the association between ethno-linguistic diversity and low public goods provision disappears once we integrate a measurement of late 19th century levels of state building, which is available for the 76 countries of Africa and Asia.

In the second step, I show that ethno-linguistic fractionalization, measured in the early 1960s by Soviet ethnographers, is also closely associated with 19th century levels of stateness, controlling for other factors, including subsequent colonial experiences. If 19th century state formation is such a powerful predictor of both contemporary ethnic diversity and public goods provision, how can we in turn understand why it varies across countries?

In a third step, I pursue this historical question by testing a series of classical arguments about the rise of centralized states. This part of the analysis is based on cross-sectional data of varying quality. The analysis suggests that states were more likely to emerge where wars were more frequent (as maintained by Charles Tilly), geography conscribed the population (the classical theory of Carneiros), the climate was favorable (à la Sachs), and population density high (as argued by Jeffrey Herbst). Instrumental variable analysis shows that preceding ethnic diversity does not consistently shape late 19th century state formation, thus reducing the possibility that the analysis suffers from endogeneity problems.

There is no ambiguity, however, about the crucial finding that contemporary ethnic diversity is not systematically related to current public good provision once historical levels of state building are taken into account. The article thus calls for revisiting the linkage between diversity and public goods provision and to embed its study into a longer term, historically informed perspective that allows disentangling and specifying their endogenous relationships. It thus parallels other research that has shown that diversity in and of itself is not detrimental to peace (Fearon & Laitin, 2003; Wimmer,

Cederman, & Min, 2009), public goods provision (Baldwin & Huber, 2011; Glennerster, Miguel, & Rothenberg, 2013), or democracy (Gerring, Zarecki, & Hoffman, 2013). The article is organized in a straightforward way. The first section outlines the detrimental diversity argument in more detail, followed by the "endogenous diversity" perspective advocated here. The next three sections follow the three steps outlined above, with data sources, measurements, and modeling strategies discussed separately for each step. A short final section concludes.

Detrimental Diversity

Easterly and Levine's (1997) influential article found that ethnically diverse countries experience economic growth rates of up to 2% less per year than homogeneous states. Compounded over decades, this relationship is supposed to explain a large part of the "growth tragedy" afflicting many developing countries, especially in Africa. Many subsequent studies have found a similar correlation between ethno-linguistic fractionalization and growth rates (Alesina, Devleeschauwer, Easterley, Kurlat, & Wacziarg, 2003; Alesina & La Ferrara, 2005; Montalvo & Reynal-Querol, 2005; Rodrik, 1999; Sala-i-Martin, Dopperlhoffer, & Miller, 2004). Before too long, economists and political scientists probed into other detrimental consequences of diversity, including generalized trust (Bjornskov, 2004; Glennerster et al., 2013; Knack & Keefer, 1997; Soroka, Banting, & Johnston, 2002) and social capital (Alesina & La Ferrara, 2000; Putnam, 2007), or welfare state development (Alesina & Glaeser, 2004; also Desmet, Weber, & Rotuno-Ortin, 2010; for the most recent overview, see Gerring, Thacker, Lu, & Huang, 2015).

According to the seminal study by Alesina and co-authors, ethnic diversity is also associated with lower provision of public goods (Alesina, Baqir, & Easterley, 1999)—the focus of this article. They foresee two mechanisms. First, individuals might not want to share public goods with ethnic others, which results in lower overall levels of public goods provision (for U.S. cities, see Poterba, 1997). I call this the "ethnic egotism" mechanism (for U.S.-based evidence, see Trounstine, n.d.). By logical implication, this should be especially true if not all ethnic groups hold a share of government power in accordance with their population size, in other words, in situations of ethno-political inequality in which only the coalition of ethnic groups represented in government will be provided with an adequate level of public goods (for Kenyan evidence, see Kramon & Posner, 2012).

Second, different ethnic groups could hold divergent preferences, independent of whether or not ethnicity is politically relevant, which increases collective action and coordination problems. Low level of public goods

provision overall are the result. Alesina and co-authors offered a variety of reasons and mechanisms through which the association between ethnicity and preferences emerges: high levels of residential segregation along ethno-racial lines may lead to different needs for public infrastructure; speakers of different languages may advocate their own as official language of instruction (Alesina et al., 1999, p. 1251; also Easterly & Levine, 1997, pp. 1214-1216).

Whatever the precise mechanisms, a negative association between ethnodemographic diversity and public goods provision has been found at the level of U.S. cities (Alesina et al., 1999; Goldin & Katz, 1999; Vigdor, 2004; but see the conclusion from a more dynamic analysis by Hopkins, 2011), Kenyan villages (Miguel & Gugerty, 2005), Indian villages (Banerjee, Lakshimi, & Somanathan, 2005),[1] as well as at the country level, with which this article is concerned. Using a global data set, La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1999) show that linguistic fractionalization is associated with higher infant mortality and lower alphabetization rates, their measurements for public good output (similarly Gerring et al., 2015, Table A7; Ahlerup, 2009, who control for possible endogeneity; Mahzab, Atiq, & Devrariani, 2013).[2]

So much has the relationship between diversity and under-provision of public goods be taken for granted that subsequent research mainly focused in identifying the various possible mechanisms foreseen in Alesina's framework and beyond—mostly with a focus on the diverse preferences mechanism, rather than ethnic egotism. Baldwin and Huber (2011) showed empirically on the basis of the analysis of 42 countries that economic inequality along ethnic lines may lead to different preferences for public goods and thus coordination problems and under-provision overall. Lieberman and McClendon (2011) show that such preference divergence exists in 18 African countries especially if ethnicity is politicized and if wealth disparities between groups are high (in line with Baldwin & Huber, 2011). Habyarimana, Humphreys, Posner, and Weinstein (2007), however, find no correspondence between policy preference and ethnic group membership in their experiments in a slum of Kampala. Their results suggest that co-ethnics chose—most likely following normative expectations that have become routinized—to cooperate with each other and punish defectors more than heterogeneous teams of individuals. Similarly, Algan, Hémet, and Laitin (2011) find that ethnic heterogeneity in French social housing complexes, to which residents are allocated randomly, leads to fewer public goods because sanctioning across ethnic divides is more difficult, as is social mobilization for petitioning such goods from state authorities. In how far these local experiments tell us anything about public goods provision by the local governments yet alone by the state, remains obviously an open question.

Endogenizing Diversity

This article argues that before we further explore possible mechanisms, we need to revisit the relationship between diversity and public goods provision itself. It does so by focusing exclusively on the country level because different mechanisms might operate at the local level (for evidence, see Gerring et al., 2015). Existing research often takes ethnic diversity as an exogenously given, demographic variable not affected by public goods provision in the past. It is treated as a naturally given feature of the social world, rather than a product of history, as Alesina and La Ferrara (2005, p. 788f.) acknowledge. Although one recent working paper moves beyond these assumptions to test whether diversity results from low public goods provision or economic growth, using an instrumental variable approach (Ahlerup, 2009), the literature has not considered that ethnic diversity and public goods provision might be spuriously related to each other because *both* depend on slowly evolving state capacity.

More precisely, states that are capable to provide public goods across the entire territory and that are politically strong enough to disempower and integrate local elites provide incentives for minorities to adopt the language of the dominant groups and to eventually adopt the core identity. Ordinary citizens shift their primary language to communicate more easily with state officials, to demand services, participation, and recognition more effectively, or to become a civil servant themselves. Linguistic heterogeneity should thus decrease the stronger the capacity of the state to interfere in the daily lives of its citizens and provide them with public goods. From a long-term historical perspective, therefore, ethno-linguistic diversity is not exogenously given, but results from slow-moving, generation-spanning processes of linguistic assimilation (and dissimilation). Following recent arguments in sociology, we can assume that religions and that racial phenotypes are more resistant to such assimilation, the first because it is more closely associated with cultural norms and practices (and thus more costly to change; Brubaker, 2013) and the second because it cannot be changed at the individual level (though it can, of course, be re-interpreted, see Loveman & Muniz, 2006).

In France, to give an example for the process I have in mind, the king had started to extend a uniform and integrated bureaucracy, personified by the royal intendant, since the 17th century. From the 18th century onward, it also provided public goods by financing local police, postal services, construction and repair projects, and the "hospitals" that cared for and at the same time confined and controlled orphans, the poor, and the sick. In the Third Republic, finally, a massive boost in state capacity occurred, and the center now mandated and financed public schooling for the entire

population, new hospitals for the poor and sick, policing in every commune of the country, and so on (for details, the online appendix to Kroneberg & Wimmer, 2012). The result of this increase in public goods provision was, as Eugene Weber (1979) has shown in a seminal historical study, that fewer and fewer speakers of minority languages identified as Provençale, Aquitaine, Occidental, and so forth, rather than as French. And fewer and fewer were able to speak these languages, given the impact of the school system and of universal conscription into an army commanded in standard French.[3]

Contrast this with mainland Tanzania, which did not know, throughout its modern history, any political entity centralized above the level of clusters of village and tribal segments. The Zanzibari Sultans militarily dominated the mainland from the early 18th century onward. Its slave raiders were operating throughout the area, but the Sultanate never administered the inland population directly—let alone provided any public goods. By the end of the 19th century, during the period of the Third Republic in France, Tanzania had come under the rule of Germans, who controlled the territory by military force, rather than by building a strong state infrastructure. Conformingly, Tanzania today represents one of the most ethno-linguistically heterogeneous countries of the world (for other explanations of linguistic homogeneity, see Ahlerup & Olsson, 2012; Michalopoulos, 2012; Nunn, 2008).

But did not the colonial period profoundly alternate or even reverse processes of ethnic homogenization? First, we should note that processes of linguistic homogenization were rarely reversed during the colonial period, but maintained even where pre-colonial political units, such as a traditional kingdom, were destroyed by the colonial rulers (on the Matabele in colonial Rhodesia, see Ranger, 1966; on the Bakongo in Congo, Lemarchand, 1964, p. 193f.). Second, some cultural assimilation processes proceeded within political units that preceded colonization and were maintained by the colonial rulers (on Yoruba ethnogenesis, see Peel, 1989). Third, other colonial historians have observed the continued amalgamation and integration of smaller ethnic units into larger ones. Various reasons can be distinguished: larger units were often more appropriate for competing for power and status in the new political arena established by colonial governments (Peel, 1989; on the Fang of Gabon and Cameroon Fernandez, 1966). Smaller political units were brought into larger colonial administrative, often ethno-regional units within which assimilation processes unfolded (see the "montagnards" in the highlands of Vietnam described by Tefft, 1999; on the Katanga, Young, 1965). Elsewhere, missionaries undertook linguistic standardization and homogenization (on the Tsonga in South Africa, see Harries, 1989). Migration within colonial labor markets often produced assimilation into local majorities (see

the classic study on Rhodesia's mining towns by Mitchell, 1974; but sometimes, anti-assimilation movements emerged, see Chai, 1996).

In short, I maintain that processes of homogenization during the colonial period remained either strongly influenced by pre-colonial levels of state formation or accelerated and were enhanced, but rarely reversed. One could argue that in how far additional homogeneity was produced by colonial rule depended on the degree to which it took on more direct, interventionist forms (on colonial styles, see Young, 1994). As Gerring, Ziblatt, van Gorp, and Arévalo (2011) have shown, however, colonial rulers often chose indirect rule where indigenous state capacity was high—a clear indication of the historically obvious fact that colonial government was shaped by pre-colonial political realities (but see for Latin America Mahoney, 2010). Still, to capture intervening, colonial modifications of both state capacity and ethnic diversity, we will need to control for style of colonial rule, or in absence of corresponding data, for the colonial master in the regression analyses on pre-colonial state centralization and ethno-linguistic diversity below.[4]

Second, these slow-moving processes of domestic state formation not only shaped ethnic heterogeneity but also influenced contemporary capacity to provide public goods through a path-dependency effect of sort (in line with Bockstette, Chanda, & Putterman, 2002; Englebert, 2000; Gennaioli & Rainer, 2007). Contemporary China, to shift examples, looks back on a history of at least a thousand years of bureaucratic state making, administrative centralization, and effective intervention in the daily life of ordinary Chinese. In Zaire, by contrast, the departing Belgian colonial administration, after having destroyed the small indigenous kingdoms of the interior and the coast, took everything with them—from knowledge of how to organize a state all the way to the typewriters. There was neither the physical infrastructure nor the human capital or organizational routines for the independent state to provide its citizenry even with a minimal level of public goods. Centralized pre-colonial states also provided institutional capacity on which colonial administrations often built (as in Korea, Botswana, or Uganda) and which allowed for the infrastructural integration of the territory and an effective provision of public goods by the post-colonial state (for African evidence of such path-dependency,[5] see Gennaioli & Rainer, 2007).

If past bureaucratic capacity and levels of state building influence both contemporary ethnic diversity and public goods provision, the relationship between the latter might be spurious. Not taking this into account therefore creates the appearance that ethnic diversity—if conceived as an ahistorical, exogenous force—is making public goods provision more difficult today.

This argument is evaluated empirically in the following three sections, the first showing that the association between public goods provision and ethnic diversity is spurious once we control for pre-colonial levels of state centralization. The second section demonstrates that pre-colonial state centralization also determines levels of ethnic diversity in the 1960s. A third step will then seek to understand variation in historical levels of state building, using two different measures of inherited levels of stateness.

The First Step: Explaining Public Goods Provision

Measurements and Data

How could we measure contemporary state capacity to provide public goods? I will rely on two commonly used indicators of public goods provision and introduce a new third one. Among the more commonly used are adult literacy rates, which are supposed to be strongly influenced by public school systems as well as state-led alphabetization campaigns. The data were assembled from various sources (see Wimmer & Feinstein, 2010). They refer to the proportion of alphabetized adults in the overall population and are available for most countries of the world since the early 19th century. For this article, I will use the post-1945 data only.[6]

The second commonly used indicator is infant mortality rates per 1,000 live births. I rely on the World Development Indicators assembled by the World Bank, which are available from 1960 onward only. Although it is obvious that mortality also depends on climate and disease prevalence as well as the general standard of living of a population, many researchers (Gennaioli & Rainer, 2007; La Porta et al., 1999) believe that it is so strongly influenced by government-run immunization programs and basic health care infrastructure that it represents a good enough proxy for state capacity to provide public goods.

A third, less commonly used indicator of public goods provision is the length of railroad tracks per square kilometer (for data sources see online appendix for Wimmer & Feinstein, 2010).[7] Railroads often represent a public good in themselves, provided and maintained by the state—though some railroads also served military purposes or to transport agricultural goods to the coast on tracks maintained by private companies. Still, railway length comes as close to a good measurement of public goods provision as do the other two indicators. The advantage of railway length as a measurement of public good provision is that of data quality. It is obviously much easier to measure railways lengths for each year than to measure the provision of public goods directly (such as state-run hospitals or schools).

How have other scholars measured public goods provision? La Porta et al.'s (1999) well-known study introduced the detrimental diversity argument to the economics of public goods provision. They used three proxy variables in their cross-country, cross-sectional analysis: child mortality, adult literacy, and average years of schooling (which is not associated with ethnic diversity, however). To check for robustness of results obtained in the following analysis, the online appendix will offer a replication study using la Porta et al.'s data set with all their dependent and control variables.

Next, I introduce the independent variables used in the pooled time series data set assembled for this project. The detrimental diversity argument, as discussed above, foresees two basic possibly mechanisms: ethnic egoism and preference heterogeneity. According to the ethnic egoism mechanism, the unwillingness to share public goods with ethnic others leads to lower overall public goods provision. As mentioned above, this should be especially the case if the ethnic governing coalition is small and if large segments of the population remain excluded from representation at the highest level of government. I rely on the Ethnic Power Relations data set (Wimmer et al., 2009) to measure the share of the population that is not represented in government on the basis of their ethnic background. To avoid endogeneity problems (after all, public goods provision might affect the level of ethnic inclusiveness, as argued by Kroneberg & Wimmer, 2012; Wimmer, 2014), I measure ethnic exclusiveness in the first year of data available (1945 or the year after independence).

The second mechanism, according to the detrimental diversity argument, is that ethnic groups diverge in their policy preferences, which makes the provision of public goods more difficult. To measure ethno-linguistic diversity, on which the country-level literature has focused so far, the several possible indices are available. I use the earliest available data, which was assembled by Soviet ethnographers in the 1950s and 1960s (data adopted from Fearon and Laitin [2003]), which reduces possible reverse causation problems because most of the data on the dependent variables concern years after 1960. Two other data sets on ethno-linguistic fractionalization produce substantially identical results, as robustness analyses in the online appendix will demonstrate.

Finally, how could we measure inherited levels of state formation? For the first step of the analysis, I rely on the Human Relations Area Files that were assembled by anthropologists on the basis of thousands of ethnographies referring to pre-colonial economic, social, political, and cultural features. This rich data have been aggregated to the country-level by Müller (1999; for other recent use of the Human Relations Area File [HRAF] data, see Gennaioli & Rainer, 2007; Gerring et al., 2011). They mapped the pre-colonial ethnic

groups onto the grid of today's states and estimated the population shares of these groups in 1960. Based on these data, I calculate the percentage of today's population that was ruled in pre-colonial times by some form of state with a minimal degree of bureaucratic development and hierarchical political integration.

For example, the HRAF contains information on number of ethnic groups in Ghana, and for each of them coded the number of levels of political authority above the local community. Three or more levels are considered states, following traditional political anthropology definitions of statehood. Müller and co-authors then calculated the population share in 1960 of each of these ethnic groups (earlier data are not available or unreliable). This allows to calculate the share of the population in 1960 that was governed, in the pre-colonial era, by states.

Unfortunately, these data are available for only half of the countries of the world, excluding the Americas and Europe. We therefore will have to carefully assess whether the association between public goods and ethnic fractionalization disappears once we include this measurement of pre-colonial stateness because the sample scope changes or because of the effects of the pre-colonial stateness variable itself. The data limitations represent, however, an advantage from a conceptual point of view: Almost all of Africa and Asia was colonized by European powers in the late 19th century, and the data on pre-colonial institutions therefore refer to a similar period, which would not be the case if we included the Americas (and obviously, Europe). This does not mean, of course, that these institutions developed entirely independent of European influence or independent of the world system centered on Europe since the early 16th century, as historically minded anthropologists have shown (Wolf, 1982). But it clearly refers to the situation before direct European conquest and the rise of the colonial state.

A set of common controls that affect all three measurements of public goods provision is considered. First, recent work in political science has shown that democracies are more likely to provide public goods because rulers have incentives to curry the favor of their voters (see Golden & Min, 2013: 75). Following standard approaches, I use the combined autocracy and democracy score from the Polity 4 project to control for regime type. Second, the resource curse literature (Ross, 2012) suggests that oil-rich countries provide less public goods because their rulers are prone to rent seeking and/or gain legitimacy by not raising any taxes, rather than by providing public goods. More evidently, we need to control for GDP per capita, which should be associated with all three measurements of public goods provision. Fourth, the "artificiality" of contemporary states, measured with the number of years with constant borders since 1816, is believed by some authors to affect

contemporary state capacity (Bockstette et al., 2002; Englebert, 2000; data are from Wimmer & Min, 2006). Fifth, a simple chronological year variable captures possible time trends in the provision of public goods. Sixth, the demand for public goods is also influenced by the sectorial composition of the economy—independent of the state capacity to provide them. Industrialized economies create a much higher demand for public goods, whereas the opposite should be true for agricultural economies, where families or village communities still assume many state functions. I introduce a variable that measures agriculture's share of GDP (from the World Bank Development Indicator data suit), which is unfortunately only available from 1960 onward. I will thus run all models both with and without this control variable to make sure that results are not affected by list-wise deletion.

For each of the three dependent variables, one additional control is added. Literacy rates are possibly affected by the size of the population, either through economies of scales or, quite the opposite, because very large populations might discourage a complete alphabetization. For railroad density, we need to add a measure of topography, and the one most effective turned out to be steepness of terrain, measured as the difference between highest and lowest elevations in a country (these data are from Fearon & Laitin, 2003). Infant mortality, finally, is affected not only by the state's efforts at lowering it, but by climate and disease prevalence. I add a measurement of the risk of being infected by fatal malaria in 1990 as a control (Sachs, 2003).

Modeling Strategy

All three models with public good provision as outcome variables will run with pooled time-series data sets. This is an appropriate modeling strategy because the outcomes vary over time within countries. I cluster robust standard errors at the country level to take the non-independence of subsequent observations in the same country into account. As railways and infant mortality are continuous outcomes, a standard ordinary least squares (OLS) regression model will be adequate. Literacy rates represent proportions, bound by 0 and 1. The dependent variable is not over-dispersed, and the appropriate statistical model is therefore a general linear regression with a logistic link function and the specification of a binomial distribution of the dependent variable.

I pursue a four-step (nested) modeling strategy to enhance transparency and account for the possibility that the estimates are sensitive to sample definitions. We are forced to list-wise delete data, as mentioned previously, as information on pre-colonial stateness is available only for Asia and Africa and those on agricultural share of GDP only from the 1960 onward. I therefore

proceed in four steps. In the first step, I run models with all covariates that do not have missing values as well as another model with the agricultural share of GDP added to the equation. In a second step, I reduce the sample to those countries that do not have missing values on the pre-colonial stateness variable without, however, adding it at this point. This is to make sure that our estimates for ethnic fractionalization do not change because of a changed sample definition. These models are again both run with and without the agricultural share of GDP variable. In the third step, the equations include the pre-colonial stateness variable but not ethnic fractionalization. This model will tell us whether it is significantly associated with public goods provision when not controlling for fractionalization. The last and analytically crucial step includes both fractionalization and pre-colonial stateness—hoping that we can thus demonstrate the spuriousness of the association between fractionalization and public goods provision.

Results

Before focusing on the core variables of interest, a brief discussion of the control variables is in order. As Tables 1 through 3 show, the most consistently significant control variables are the combined democracy/autocracy score, the oil production measurement, levels of economic development, the share of agriculture in the GDP, as well as the time trend. As expected, more democratic states have more literate populations (though not consistent across model specifications), railway systems with longer tracks, and lower infant mortality rates, whereas oil-rich countries show the opposite characteristics. Agriculturally based economies are governed by states that provide much fewer public goods, though the variable is insignificant when railroads are the outcome variable. Over time, more public goods are being provided, again except in the models on railway length, where time is generally insignificant. GDP is systematically related to all outcome measures again except to railway track length, perhaps because some very rich countries have ceased to build railways. Continuity of state borders are only associated with literacy (in the full sample). Among the additional control variables specific to each outcome, malaria risk is a (very powerful) predictor of child mortality, whereas neither topography nor population size seem to matter consistently for the building of railways or the alphabetization of the population.

Moving to the variables of core theoretical interest for this article, the ethnic egoism mechanism seems not to affect public goods provision at all. Countries that excluded large proportions of their populations from central government power on the basis of their ethnic background do not subsequently provide fewer public goods overall. The coefficient of the initial

Table 1. Generalized Linear Models of Proportion of Adult Literates.

	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
Dependent variable: Proportion adult literates, Wimmer and Feinstein	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Combined autocracy and democracy score, Polity 4	0.0063**	0.0056**	0.0074***	0.0037	0.0065**	0.0027	0.0068**	0.0027
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Oil production per capita (averaged Humphreys, BP, IHS)	−0.0840***	−0.0657***	−0.0257*	−0.0260**	−0.0266*	−0.0241**	−0.0255*	−0.0236**
	(0.015)	(0.013)	(0.014)	(0.011)	(0.014)	(0.010)	(0.014)	(0.010)
GDP per capita, lagged, Penn World Tables	0.2301***	0.1719***	0.0840**	0.0744**	0.0907***	0.0700**	0.0845**	0.0671**
	(0.036)	(0.033)	(0.036)	(0.031)	(0.033)	(0.029)	(0.033)	(0.029)
No. of years since 1816 with constant borders, Wimmer and Min	0.0024	0.0035**	0.0035*	0.0036	0.0016	0.0017	0.0019	0.0017
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Year	0.0172***	0.0326***	0.0328***	0.0330***	0.0362***	0.0361***	0.0363***	0.0360***
	(0.003)	(0.004)	(0.005)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Population size in thousands, Gleditsch	0.0000	0.0000**	0.0000***	0.0000***	0.0000	0.0000	0.0000	0.0000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Agriculture's share of economy (% of GDP), Word Bank		−0.0116**		−0.0125**		−0.0145**		−0.0143**
		(0.006)		(0.006)		(0.006)		(0.006)

(continued)

Table 1. (continued)

	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
Dependent variable: Proportion adult literates, Wimmer and Feinstein	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Prop. of excluded population at first year of available data, EPR	−0.4155	−0.3051	−0.4086	−0.2306	−0.3854	−0.1615	−0.3049	−0.1377
	(0.308)	(0.290)	(0.338)	(0.311)	(0.383)	(0.345)	(0.378)	(0.347)
Linguistic fractionalization, Soviet Atlas, Fearon and Laitin	−1.2186***	−0.8616***	−0.8448**	−0.5493			−0.4765	−0.2162
	(0.320)	(0.306)	(0.365)	(0.346)			(0.425)	(0.382)
% population ruled by states before colonization, Müller					0.7817***	0.7011***	0.6124**	0.6348**
					(0.261)	(0.267)	(0.300)	(0.294)
Constant	−34.0030***	−64.5222***	−65.2555***	−65.4129***	−72.6296***	−71.9041***	−72.4180***	−71.5798***
	(5.996)	(8.449)	(8.942)	(9.997)	(8.261)	(8.898)	(8.228)	(8.851)
Observations	6,538	4,103	3,557	2,513	3,557	2,513	3,557	2,513
AIC	0.781	0.749	0.903	0.905	0.897	0.897	0.895	0.897

Robust standard errors in parentheses. BP = British Petroleum; IHS = Information Handling Services; EPR = Ethnic Power Relations data set; AIC = Akaike Information Criterion.
*$p < 0.1$. **$p < .05$. ***$p < .01$.

Table 2. OLS Regression on Length of Railway Tracks.

	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
Dependent variable: Length of railway tracks (in km), Wimmer and Feinstein	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Combined autocracy and democracy score, Polity 4	0.0555	0.0780*	0.0864***	0.0636**	0.0769**	0.0525*	0.0793**	0.0527*
	(0.049)	(0.041)	(0.031)	(0.028)	(0.029)	(0.026)	(0.030)	(0.027)
Oil production per capita (averaged Humphreys, BP, IHS)	−1.0553***	−1.0010***	−0.3859*	−0.5624*	−0.3955*	−0.5631*	−0.3873*	−0.5585*
	(0.210)	(0.193)	(0.216)	(0.321)	(0.228)	(0.327)	(0.218)	(0.323)
GDP per capita, lagged, Penn World Tables	2.3090***	1.9895***	0.8920	1.2455	0.9464	1.2452	0.8902	1.2233
	(0.532)	(0.426)	(0.588)	(0.851)	(0.626)	(0.879)	(0.587)	(0.859)
No. of years since 1816 with constant borders, Wimmer/Min	−0.1344***	−0.0862***	0.0002	−0.0016	−0.0207	−0.0222	−0.0175	−0.0219
	(0.043)	(0.032)	(0.022)	(0.023)	(0.023)	(0.024)	(0.024)	(0.024)
Year	−0.2634***	−0.0458	−0.0518	−0.0778	−0.0237	−0.0470	−0.0229	−0.0479
	(0.090)	(0.065)	(0.058)	(0.052)	(0.064)	(0.052)	(0.062)	(0.053)
Difference between highest and lowest elevation (in m), Fearon and Laitin	−0.0013*	−0.0010	0.0005	0.0006*	−0.0001	0.0000	0.0001	0.0001
	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Agriculture's share of economy (% of GDP), Word Bank		−0.1029		−0.0270		−0.0345		−0.0342
		(0.088)		(0.062)		(0.062)		(0.062)

(continued)

Table 2. (continued)

Dependent variable: Length of railway tracks (in km), Wimmer and Feinstein	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Prop. of excluded population at first year of available data, EPR	−5.1221	0.7239	−3.2361	−1.4481	−2.1922	0.0222	−1.6177	0.1301
	(5.465)	(3.620)	(2.378)	(1.977)	(2.721)	(2.141)	(2.504)	(2.114)
Linguistic fractionalization, Soviet Atlas, Fearon and Laitin	−15.4302*	−9.7390	−9.1709**	−5.5500			−4.9782	−1.3799
	(8.253)	(6.068)	(4.075)	(3.566)			(4.224)	(3.062)
% population ruled by states before colonization, Müller					8.8506***	7.7875***	6.8673**	7.2809***
					(2.404)	(2.653)	(2.619)	(2.685)
Constant	554.7683***	114.4439	111.1834	160.0065	50.2364	95.9503	51.3702	98.6135
	(177.414)	(129.892)	(113.901)	(100.438)	(124.911)	(100.292)	(122.765)	(101.350)
Observations	6,458	4,088	3,492	2,513	3,492	2,513	3,492	2,513
R^2	.382	.463	.271	.345	.294	.390	.306	.391

Robust standard errors in parentheses. OLS = ordinary least squares; BP = British Petroleum; IHS = Information Handling Services; EPR = Ethnic Power Relations data set.
*$p < 0.1$. **$p < .05$. ***$p < .01$.

Table 3. OLS Regression on Infant Mortality Rate.

	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
Dependent variable: Infant mortality, WDI	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Combined autocracy and democracy score, Polity 4	−0.2638***	−0.1896***	−0.2789***	−0.1889***	−0.2544***	−0.1628***	−0.2746***	−0.1668***
	(0.066)	(0.055)	(0.076)	(0.065)	(0.077)	(0.059)	(0.074)	(0.058)
Oil production per capita (averaged Humphreys, BP, IHS)	0.5400***	0.4795***	0.4938*	0.6050***	0.5018*	0.6078***	0.4971*	0.5931***
	(0.134)	(0.095)	(0.265)	(0.212)	(0.280)	(0.206)	(0.271)	(0.210)
GDP per capita, lagged, Penn World Tables	−2.0097***	−1.3817***	−2.1859***	−1.9144***	−2.3023***	−2.0152***	−2.2193***	−1.9070***
	(0.312)	(0.292)	(0.695)	(0.675)	(0.711)	(0.635)	(0.710)	(0.666)
No. of years since 1816 with constant borders, Wimmer/Min	−0.0100	−0.0127	0.0167	−0.0062	0.0375	0.0290	0.0287	0.0289
	(0.039)	(0.037)	(0.069)	(0.065)	(0.067)	(0.063)	(0.066)	(0.063)
Year	−1.3233***	−1.2228***	−1.7600***	−1.3827***	−1.7981***	−1.4415***	−1.7762***	−1.4292***
	(0.118)	(0.133)	(0.182)	(0.199)	(0.180)	(0.193)	(0.176)	(0.190)
Fatal malaria risk probability in 1994, Sachs	40.6332***	26.6427***	34.5395***	24.5698**	39.6502***	23.5426**	33.2519***	19.4740*
	(6.519)	(6.351)	(9.279)	(9.576)	(10.517)	(11.220)	(10.088)	(10.684)
Agriculture's share of economy (% of GDP), Word Bank		0.7604***		0.7028***		0.7673***		0.7836***
		(0.140)		(0.167)		(0.194)		(0.192)
Prop. of excluded population at first year of available data, EPR	15.6947	9.5660	11.0582	5.9740	12.7886	4.3746	10.0638	2.7293
	(10.668)	(9.812)	(11.791)	(10.844)	(12.128)	(11.495)	(11.942)	(11.430)

(continued)

Table 3. (continued)

	1	2	3	4	5	6	7	8
			Without missing values on pre-colonial state centralization					
	Full model		With ethnic fractionalization		With pre-colonial state centralization		With both	
Dependent variable: Infant mortality, WDI	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture	Without % agriculture	With % agriculture
Linguistic fractionalization, Soviet Atlas, Fearon and Laitin	28.2957***	22.7812***	22.7072*	17.8383*			20.6322	13.1578
	(8.783)	(7.471)	(12.011)	(9.741)			(12.522)	(10.522)
% population ruled by states before colonization, Müller					−8.9551	−15.4265	−4.7174	−13.1039
					(10.247)	(9.456)	(10.330)	(9.713)
Constant	2,673.7212***	2,463.4118***	3,549.3492***	2,789.8680***	3,636.0671***	2,919.5356***	3,584.9099***	2,889.0938***
	(234.107)	(265.252)	(358.463)	(392.517)	(354.548)	(381.845)	(346.537)	(375.875)
Observations	5,102	4,000	2,910	2,419	2,910	2,419	2,910	2,419
R^2	.689	.745	.564	.617	.556	.621	.565	.625

Robust standard errors in parentheses. OLS = ordinary least squares; WDI = World Development Index; BP = British Petroleum; IHS = Information Handling Services; EPR = Ethnic Power Relations data set.
*$p < 0.1$. **$p < .05$. ***$p < .01$.

share of the excluded population variable is never significant in Tables 1 through 3.

However, much seems to speak in favor of the preference heterogeneity mechanisms according to which ethnic heterogeneity negatively affects a state's capacity to provide public goods. Both in the full sample (Models 1 and 2) and in the reduced sample of Asian and African countries (Models 3 and 4), ethno-linguistic fractionalization is significantly associated with low public goods provision and with a quite large coefficient, thus reproducing the results of previous research (the one exception is Model 2 of Table 2; in Models 4 of Tables 1 and 2, the variable misses standard levels of significance with a *p* value of .113 or .123, respectively). For example, decreasing the chances that two randomly chosen individuals speak the same language by 28% (a standard deviation) decreases the length of railways tracks by one fifth of a standard deviation, infant mortality by 7 children per 1,000 live births, and adult literacy rates by 10% (these are standardized coefficients based on Model 1 in Tables 1 to 3).

The picture changes quite dramatically, however, as soon as we control for levels of pre-colonial stateness. This variable is significantly associated with all three measurements of public good provision when we disregard the ethno-linguistic composition of the population (Models 5 and 6), except in the models on infant mortality (Model 5 and 6 in Table 3). With *both* pre-colonial stateness and ethno-linguistic fractionalization in the equation (Models 7 and 8), ethnic fractionalization is no longer significantly associated with *any* of the three outcomes, lending strong support to the idea that diversity is endogenous to, rather than representing an impediment to the formation of capable states. Meanwhile, levels of pre-colonial stateness are significantly associated with two of the three outcomes even when controlling for ethnic diversity, the exception again being infant mortality. Increasing the share of the population that was governed by states in the pre-colonial period by 40% (one standard deviation) is associated with almost a third of a standard deviation more railways and a reduction of illiteracy by 6% (standardized coefficients based on Models 7 in Tables 1 and 2).

Some Robustness Checks

Online appendix Table 1 shows that these results are upheld if we control for the intervening style of imperial rule—proxied, in the absence of better data, for example, on the indirectness of rule, by whether a territory was governed by the Ottomans, Portuguese, French, or British (with all others serving as the omitted category). The results are substantially identical to the ones presented

in Tables 1 through 3, including that the pre-colonial centralization variable fails to reach standard levels of significance in models with infant mortality as the dependent variable. None of the colonial master dummies reach standard levels of significance consistently across all measurements of public goods provision.

As an additional robustness check, online appendix Table 2 offers a series of replication models based on de la Porta's et al. data universe and modeling approach. This is to make sure that the above findings can be reproduced within their framework. I proceed in the same four steps as above, this time regressing on average school achievement as well to replicate their definition of public goods provision one by one. The results are generally consistent with the ones discussed above and thus support the idea that we need to take the long-term history of state formation into account when trying to evaluate how diversity affects public goods provision.

As a final robustness check, I used two different coding of linguistic diversity besides the one based on the Soviet Atlas in Tables 1 to 3 above: a fractionalization index based on Roeder's (2007) list of ethno-linguistic groups as well as the often used linguistic fractionalization index introduced by Alesina et al. (2003). As online appendix Table 3 shows, the results are substantially very similar to those presented above. In one of the six models (Model 6), however, Alesina et al.'s linguistic fractionalization index remains marginally significant when regressed on infant mortality rate despite including the pre-colonial stateness variable.

Online appendix Table 3 also contains two additional sets of models that use religious instead of linguistic fractionalization indices (offered by Alesina et al., 2003; Fearon, 2003). In line with theoretical expectations, they are insignificant in all six models without the pre-colonial stateness variable—indicating that religious diversity has none of the supposed negative consequences that linguistic diversity has, even though religion should be more closely associated with public policy preferences and/or produce stronger in-group bias. In two of the six models (Models 20 and 22), the coefficient for the religious diversity index even becomes positively significant when the pre-colonial stateness variable is included—for reasons that one would need to explore more fully by inspecting the cases that drive this result.

The Second Step: Understanding Ethnic Fractionalization

In a second set of analyses, I evaluate whether linguistic heterogeneity is itself associated with historically achieved levels of state formation, as argued

in the theory section. The following models take ethnic diversity as a dependent variable and the level of pre-colonial stateness as the key independent variable to test the endogenous diversity model.

Measurements, Data, and Modeling Strategy

A different set of controls is added as ethnic fractionalization is now the dependent variable. According to Michalopoulos (2012), a range of ecological factors should be associated with diversity because they encouraged, in the very remote past, a differentiation of economic survival strategies and thus contemporary ethnic diversity. I include the four variables, all from his data set, that are consistently significant in his models on linguistic fractionalization: the variability in suitability of a territory for agriculture, variability in average precipitation, average precipitation, and distance from the Ocean.

Second, past ethnic and nationalist wars could affect ethnic diversity through ethnic cleansing, which could produce more homogeneous countries in the present. I use a cumulative count of the number of ethnic or nationalist wars from 1816 to 1900 (data are from Wimmer & Min, 2006). GDP per capita proxies for levels of development, which could be associated with ethnic diversity either because diversity is bad for growth (as argued by the detrimental diversity school), or, conversely, because low development implies less migratory mobility across a territory and thus less linguistic assimilation (à la Deutsch, 1953).

The models are run as a general linear regression models since 0 and 1 bound the outcome. Because the data are not over-dispersed, the appropriate model specification is again a logistic link function and the specification of a binomial distribution of the dependent variable. All models are cross-sectional because the dependent variable does not vary over time. I again pursue a nested modeling strategy, first analyzing how controls are associated with diversity in the full sample of countries, then proceeding to a model with the same controls but for a subsample of Asian and African countries only, and finally integrating the pre-colonial stateness variable.

As a robustness test, I use a second indicator of inherited stateness. The "state antiquity" index developed by Bockstette et al.'s (2002) combines levels of centralization, control by the local population (as opposed to outside conquerors), and territorial continuity. In short, it measures inherited levels of centralization and continuity of indigenous states. This index is available with different rates at which previous history is discounted (the information starts in the year 1 and goes to 1950). It covers 133 countries of the world.

Results

Table 4 reports the results. None of the control variables show the expected associations with linguistic diversity across model specifications. The pre-colonial stateness variable, however, strongly influences linguistic diversity in the early 1960s (Model 3). This association is also substantially important: A standard deviation increase in percentage of the population ruled by states in pre-colonial period decreases the chances that two randomly chosen individuals speak different languages by half of a standard deviation or roughly 15%.

In Model 4, I check for the robustness of these results by using the index of state antiquity. It is again strongly associated with linguistic fractionalization in 1960, and again with a substantial, though considerably weaker effect: Decreasing levels of inherited state centralization by one standard deviation will produce a 6% lower chance that two randomly chosen individuals will speak a different language in the 1960s.

Note that none of the other control variables, including the exogenous geographical, topographical, and climatic variables, are consistently associated with linguistic heterogeneity; only GDP per capita (which might be, according to the detrimental diversity argument, endogenous to diversity) and the variability of suitability for agriculture are at least significantly associated with diversity in three of the four models.

Given the long lag between these two measurements of inherited stateness and the data on linguistic diversity, would we not expect that the intervening colonial experience profoundly modified the linguistic landscape shaped by previous processes of state centralization? Online appendix Table 4 shows that some colonial/imperial domain variables are indeed associated with more or less diversity in 1960. More specifically, Ottoman domains both in models with the pre-colonial centralization variable and those with the index of state centralization show less linguistic diversity—for reasons beyond the scope of this article. A comparison of models with and without these colonial/imperial dummies demonstrates, however, that the size and significance of the coefficients of both variables measuring historically achieved state capacity remains largely identical. This lends some support to the conjecture made in the theory section above: that colonial rule modified, but rarely radically changed the diversity that had emerged previously under various levels of state centralization.

The Third Step: Determinants of Inherited Stateness

So far, we have seen that levels of pre-colonial stateness are associated with both public goods provision in the present—through a path-dependency

Table 4. General Linear Models on Ethnic Fractionalization.

Labels	(1) Baseline	(2) Without countries lacking data on pre-colonial stateness	(3) With pre-colonial stateness	(4) With inherited state centralization
GDP per capita, Penn World Table	−0.0541***	−0.0572**	−0.0275	−0.0534***
	(0.013)	(0.028)	(0.024)	(0.013)
Dispersion in elevation across regions of country, Michalopoulos	−0.0110	−0.1346	0.0564	0.0347
	(0.085)	(0.120)	(0.135)	(0.094)
Dispersion in agricultural suitability across regions of country, Michalopoulos	0.7847*	1.0167*	0.8962	1.2842***
	(0.415)	(0.615)	(0.616)	(0.489)
Average monthly precipitation 1961-1990 in 1,000s of mm, Michalopoulos	0.0015	0.0043*	0.0049*	0.0003
	(0.002)	(0.002)	(0.003)	(0.002)
Distance from coast in km, Michalopoulos	0.5542*	0.8727	0.8256*	0.3072
	(0.298)	(0.588)	(0.479)	(0.298)
Cumulative no. of ethno-nationalist wars fought between 1816 and 1900, Wimmer and Min	−0.1604	0.4466	0.5152	−0.1542
	(0.120)	(0.294)	(0.347)	(0.124)
In pre-colony, % pop. ruled by states, Müller			−1.5326***	
			(0.315)	
Cumulative index of state centralization since 100 BC with 5% discount rate, Putterman				−1.2162***
				(0.423)
Constant	−0.8090**	−1.0372**	−0.6339	−0.5184
	(0.349)	(0.527)	(0.671)	(0.439)
	144	75	75	133

Robust standard errors in parentheses.
*$p < 0.1$. **$p < .05$. ***$p < .01$.

effect—and with ethnic homogeneity through intervening processes of assimilation. These findings raise the crucial question of how we can account for such different levels of pre-colonial stateness. Why does it vary so dramatically between Tanzania and China, to come back to the previous examples? Some of the most important arguments put forward in the literature can be evaluated with existing data—albeit with reduced number of observations because only cross-sectional data are available and with considerable measurement uncertainty for the outcome variables. The following analysis therefore remains rather tentative.

To evaluate Tilly's (1975) seminal study on the mutually reinforcing relationship between war-making and state-building, we can test whether pre-colonial state centralization or state antiquity are related to the number of civil and interstate wars fought between 1816 and 1900 (excluding wars of conquest; data are again from Wimmer & Min, 2006). This relationship is probably endogenous, a fact well captured by Tilly's famous dictum that "wars made states and states made war."

Second, I hypothesize that mountainous terrain (data from Fearon & Laitin, 2003) is associated with higher levels of historical stateness. This would be in line with Carneiros' (1970) classical theory of "environmental conscription" as a condition for the rise of the pre-modern agrarian state. Such "conscription" by mountain ranges (as in highland Mexico and Oaxaca, where the Aztec and Mitla states emerged) or deserts (as in Iraq, the center of Babylonia, or ancient Egypt) prevented the population from escaping the control of state builders. The mountainous terrain variable thus captures only one of several possible modes of confinement.

Third, state formation might also be affected by geography and climate, in line with environmentalist arguments that have recently resurged in the economic development literature (Sachs, 2003). Hot temperatures and disease prevalence near the equator would make state building more difficult to organize and sustain. Indeed, recent research has shown that geography influences economic growth through an indirect effect—because countries closer to the equator are characterized by less well-developed state institutions (Rodrik, Subramanian, & Trebbi, 2004). To evaluate this geographical argument, we use latitude as a variable, in line with this literature on economic growth.[8]

Fourth, I evaluate Herbst's (2000) argument according to which low population densities, climatic diversity associated with elevation difference, and difficult transport conditions explain Africa's low levels of state development. Data for the population density in the year 1500—long before levels of stateness are measured, to avoid endogeneity problems[9]—are from Putterman and Weil (2010). Data on elevation differences would be available from

Michalopoulos (2012), but they all produce collinearity problems and will thus not be used as a control variable (but as an instrument for linguistic diversity, see below). The difficulty of transport condition can be approximated with the minimum kilometer distance to the Ocean as well as with the mountainous terrain variable—this time, however, we would expect a negative association with levels of state centralization.

Finally, we need to discuss whether levels of pre-colonial stateness might not, in turn, be influenced by pre-existing ethnic heterogeneity. We thus include the linguistic heterogeneity variable in all equations. In Models 1 and 5, I use the standard linguistic diversity variable referring to the 1960s, well knowing that this will not allow us to determine the direction of a causal association: Higher level of stateness in the late 19th century might produce lower levels of diversity in 1960s through processes of assimilation, as argued above, or because high levels of diversity in the early 19th century impeded subsequent state building and thus remained at a similar level up to the 1960s—both processes producing a negative relationship between the two variables. Models 2 to 4 and 6 to 8 therefore instrument linguistic fractionalization with a variety of climatic, ecological, and topographical variables provided by Michalopoulos (2012; a similar approach is pursued by Ahlerup & Olsson, 2012).

To construct an instrumental variable model, I first regress Michalopoulos' variables on linguistic diversity, avoiding collinearity problems by running different models with different sets of these variables. I retained those variables as instruments that are significantly associated with linguistic diversity. To ensure that they affect pre-colonial stateness only indirectly, through diversity, but not directly, I ran a second set of models with pre-colonial centralization and state antiquity as dependent variables and the ecological-topographical variables as independent variables, excluding the significant ones as instruments for linguistic diversity. The instruments thus satisfy the "exclusion" criterion. They slightly differ in the state antiquity models and the pre-colonial centralization models. Models 2 and 6 include only the instrumental variables thus identified as well as the covariates that test the various classical theories of state formation discussed above. Models 3 and 7 add those ecological-topographical-climatic variables that were identified as being significantly associated with the measure of stateness, whereas Models 4 and 8 retain only significant variables.

All instrumental variable regression models use a two-stage least square estimator. Other model specifications (such as with generalized method of moments [GMM] estimators) produce substantially identical results. The first stage of the instrumental variable models is not shown. The *T* value for the first stage is around 10 for the state antiquity models, and considerably

higher (above 30) for the pre-colonial centralization models. In both sets of models, the instrumental variables are thus relevant at or well above standard thresholds.

Table 5 displays results for pre-colonial state centralization in Models 1 to 4, and for state antiquity in Models 5 to 8. I start the discussion with the variables that are supposed to test classical theories of state building. The results are generally consistent, even if not all variables achieve standard levels of significance in both sets of models. Wars are positively associated with state antiquity but not with pre-colonial levels of centralization (though the *p* value in Model 1 is .13). If we dichotomized the variable and created a dummy for territories with one or more wars fought between 1816 and 1900, this variable would be consistently significant in the expected direction in seven of the eight models (results not shown).[10] Mountainous terrain is consistently positively and significantly related to both measures of stateness—thus lending support for the environmental conscription argument, rather than Herbst's conjecture that transportation difficulties impede state formation. The geographic distance from equator variable is consistently significant (except in Model 2), thus offering evidence for the idea that hot climate and disease prevalence make state building around the equator more difficult.[11] This is further supported in Models 3 and 4 as well as 7 and 8 with additional climate variables, of which, however, only average temperature is robustly significant (with a positive sign of the coefficient—controlling for latitude) in Models 7 and 8. Population density in 1500 is consistently and positively influencing levels of stateness, this time in line with Herbst's reasoning, whereas distance from the sea again does not support his transportation difficulty argument.

Next, I explore how ethno-linguistic diversity is associated with stateness. In the GLM models in which ethnicity is not instrumented, we arrive at a highly significant and large negative coefficient for linguistic diversity in the model with pre-colonial centralization as the dependent variable, whereas the variable is entirely insignificant (with a *p* value greater than .6) in the model with state antiquity as the outcome. The possibly endogenous relationship between state building and linguistic diversity is explored with a series of six instrumental variable regressions, to which I now turn.

Models 2 to 4 again show a negative association between the instrumented ethnic fractionalization variable and pre-colonial stateness. We would be tempted to conclude that diversity impedes state centralization. The models with state antiquity as a dependent variable lead to a different conclusion, however. In Models 6 and 7, the coefficient of the instrumented linguistic diversity variable turns positive and comes close to standard levels of significance in Model 6. In Model 8, from which non-significant terms were

Table 5. GLM and Two-Stage Least Squares Instrumental Variable Models on Inherited Levels of Stateness.

	DV: Proportion of population governed by states before colonization				DV: State antiquity			
	1	2	3	4	5	6	7	8
	GLM model	IV model (first stage not shown)	IV model (first stage not shown)	IV model (first stage not shown)	GLM model	IV model (first stage not shown)	IV model (first stage not shown)	IV model (first stage not shown)
Linguistic fractionalization	−1.8512**				−0.1468			
	(0.783)				(0.319)			
Linguistic fractionalization, instrumented with av. temperature, dispersion of elev. across regions, dispersion of suitability for agric. across regions		−0.6135**	−0.6614***	−0.4962**				
		(0.237)	(0.235)	(0.194)				
Linguistic fractionalization, instrumented with av. monthly precipitation, dispersion of suitability for agric. across regions						0.4960	0.3917	0.4080**
						(0.321)	(0.340)	(0.174)
Cumulative No of wars (without conquests) 1816-1900	0.2109	0.0341	0.0361		0.0630**	0.0190**	0.0142	0.0183**
	(0.140)	(0.034)	(0.034)		−0.03	−0.009	−0.013	(0.009)
ln (Mountainous terrain)	0.3375**	0.0627**	0.0504*	0.0736***	0.1264**	0.0430**	0.0537*	0.0611***
	(0.146)	(0.027)	(0.027)	(0.022)	−0.05	−0.017	−0.029	(0.014)
Latitude	0.0657***	0.0075	0.0118**	0.0098***	0.0158***	0.0067***	0.0156***	0.0160***
	(0.025)	(0.005)	(0.005)	(0.004)	−0.004	−0.002	−0.003	(0.003)

(continued)

Table 5. (continued)

	DV: Proportion of population governed by states before colonization				DV: State antiquity			
	1	2	3	4	5	6	7	8
	GLM model	IV model (first stage not shown)	IV model (first stage not shown)	IV model (first stage not shown)	GLM model	IV model (first stage not shown)	IV model (first stage not shown)	IV model (first stage not shown)
ln (Population density in 1500 AD)	0.5556***	0.0780***	0.0534*	0.0845***	0.3736***	0.0869***	0.0873***	0.0892***
	(0.185)	(0.029)	(0.032)	(0.026)	−0.052	−0.012	−0.011	(0.011)
Distance from coast	−0.0495	0.0466	0.1571		0.0968	−0.1002	−0.0394	
	(0.555)	(0.101)	(0.105)		−0.225	−0.088	−0.069	
Average monthly precipitation 1961-1990			0.0016*					
			(0.001)					
Average temperature 1961-1990							0.0217***	0.0224***
							−0.005	(0.005)
Dispersion in elevation across regions							0.0184	
							−0.039	
Constant	−1.5906**	0.4132*	0.2234	0.3274*	−1.3021***	−0.0955	−0.7417**	−0.7848***
	(0.789)	(0.218)	(0.239)	(0.172)	(0.251)	(0.194)	(0.297)	(0.250)
Observations	74	74	74	74	132	132	132	132
R^2		.475	.496	.484		.177	.350	.331

Robust standard errors in parentheses. GLM = Generalized Linear Model.
*$p < 0.1$. **$p < .05$. ***$p < .01$.

excluded, the coefficient of the diversity variable remains positive and becomes significant with a coefficient about as large as that of Model 4 in which the sign was negative. The same would result if we would exclude the insignificant distance from sea variable in models 6 or 7 (results not shown) and in many other model specifications. We thus do not arrive at a robust result that would hold across model specifications and with both measurements of historically achieved stateness.

I conclude that there is no consistent support for the idea that linguistic diversity impedes the emergence of strongly centralized states. Depending on how we measure stateness[12] and which other covariates we include, we find that diversity impedes, is indifferent to, or even enhances the formation of states. To illustrate this inconclusiveness, we might briefly compare the histories of Somalia and Botswana. Both have the reputation of being among the most linguistically homogeneous countries in Africa.

In Botswana this is the result of the strongly centralizing mini-states, dominated by Tswana-speaking kings and their entourage, who assimilated, over the 19th and 20th centuries, many of their subject population into the Tswana majority language and identity. To demonstrate the extent of this process, we can calculate the population share of groups of Tswana tribal origin in the seven traditional kingdoms of Botswana in 1946, thanks to the monumental work of anthropologist Schapera (1952) who listed the historical origins (and thus original language) of each ward of each village and town of the entire country. Comparing this to the population who listed Tswana as their "mother tongue" in the 2001 census (the only one to ever ask such a question), I find that the Tswana origin population (which supposedly always spoke Tswana) was 55% in 1946 (which would produce a high linguistic fractionalization index), whereas 78% of the population mentioned Tswana as mother tongue in 2001. This is also, one should add, the result of highly capable Botswana state that emerged after independence (Samatar, 1999) and its policy of linguistic dominance and assimilation.

Somalia, on the other hand, was arguably a much more homogeneous society in the early modern period, with the large majority of the population speaking the same, Somali language and being of the same historical origin—with the exception of the Southern, riverine areas where some Bantu-speaking peasants had already settled before Somali nomads moved into and subsequently dominated that area (Lewis, 1994, Chapter 6). This high level of homogeneity did not, however, enhance the building of a strongly unified and centralized state system. To be sure, states such as the Ajuran sultanate emerged in the riverine South before the 17th century and trading cities developed into small states along the coast. Most of the latter were nominally part of the Ottoman Empire but de-facto independent well into the 20th

century (Lewis, 1988). The strongly centralized riverine sultanate decayed later on, whereas the coastal states' authority never extended to the hinterland (Lewis, 1988). Post-independence Somalia was notorious—even before its descent into a decade long civil war and its subsequent fragmentation into independent regional mini-states and competing warlord domains—for its failure to develop an integrated bureaucracy that would rule uniformly across the entire territory (Laitin & Samatar, 1987). Linguistic (and one might add: religious) homogeneity did not help, it seems, to prevent this development.

Beyond this suggestive comparison, we cannot rule out nor confirm that linguistic heterogeneity impedes state building, as shown in the conflicting findings in Table 5 (see also Ahlerup, 2009). However, we do know for certain, thanks to a temporal lag of roughly 60 years between the two measurements, that centralized indigenous states or a long history of indigenous stateness subsequently produced linguistic homogeneity à la Botswana (Table 4, Models 3 and 4). The possibly endogenous relationship between state formation and linguistic heterogeneity, to be sure, does not undermine the major finding of this article: that contemporary diversity is not directly related to public goods provision once we take into account that both are influenced by historically inherited levels of stateness.

Conclusion

The overall picture of the relevant causal relationships can be summarized in Figure 1. In the first step of the analysis, I showed that state capacity to deliver public goods in post-war countries of the world is not systematically associated with their linguistic heterogeneity, once we take into account that contemporary capacity also depends on historically achieved levels of state centralization. Tanzania is different from Korea, among other things, because Korea looks back on a thousand years of state building while mainland Tanzania knew little indigenous stateness. In a second step, I showed that such past levels of stateness also influence contemporary heterogeneity because strong centralized states such as Korea (or Botswana) were able to homogenize the population linguistically, whereas heterogeneity remained high in stateless societies such as mainland Tanzania. Therefore, the negative association between heterogeneity and public goods provision that past research has brought to light is spurious. Diversity might not be detrimental to public goods provision because members of different ethnic groups cannot agree on what goods the state should provide or because they do not want to share such goods with ethnic others. Rather, states that were weakly developed in the past left a legacy of both high diversity and limited capacity of public goods provision in the present.

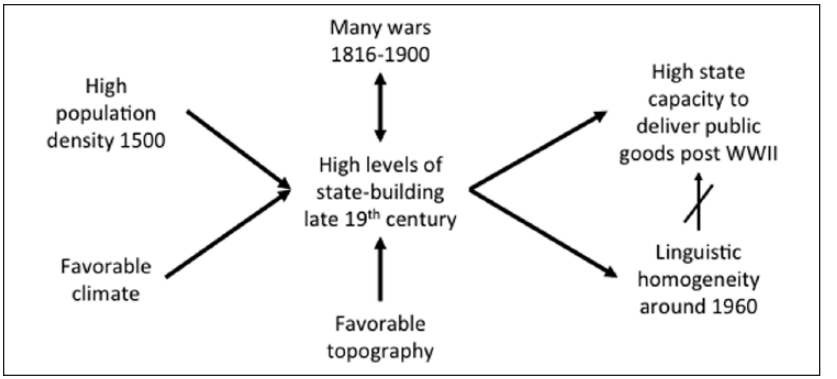

Figure 1. Summary.

In line with classic theories of state formation, this article also indicated, in a somewhat more inconclusive way, what might have led to such weak states: low population density and thus a reduced capacity to carry a non-productive, administrative-political elite; a topography that makes it easy to evade the efforts of state builders by retreating into the bush or forest; a climate unfavorable to urban civilization because of heat, humidity, and disease; and a lack of state-building warfare (though evidence is less conclusive for this mechanism).

Instrumental variable regressions showed no consistent results regarding the question whether high levels of linguistic heterogeneity in the past should be added to this list of factors that make state-building less likely. Using one measurement of state centralization, diversity clearly impedes state formation, whereas with the other measurement, it has no or even state enhancing effects. This question cannot be answered in more conclusive terms without data on early 19th century linguistic diversity, without in other words, a massive effort at adding temporal depth to the analysis of how state formation relates to linguistic diversity and vice versa.

Other data limitations are quite obvious: It would be helpful to have a measurement of late 19th century state centralization that would include the entire world—rather than the 74 or 133 states to which I had to confine the above analysis. Public goods provision by the state is proxied with outcome variables that are affected by other processes as well, rather than with a more direct measurement of government activities such as the number of government-run schools, hospitals, and railways.

Despite its limitations, this article suggests that the relationship between public goods provision and diversity can and should not only be addressed in

field experiments or through surveys—the two dominant methodological trends in contemporary political science—but through a more historical approach that explores their possibly endogenous relationship over the long run with appropriate data. Much can also be learned from already available, qualitative case studies of nation building through assimilation or violence to understand how diversity is shaped by the emergence of bureaucratic states and has in turn influenced further trajectories of state building. We have just only begun to ask these important questions within the framework of a quantitative and global perspective.

I conclude with a rather speculative remark on how this research speaks to questions relevant to policy makers. Quite obviously, past state capacity cannot be engineered to create a historical legacy favoring contemporary public good provision. However, we should not forget that the relationship is probabilistic and that there is plenty of room for political craftsmanship and engineering—and of course for fortuitous political contexts, historical contingencies, and other important factors that cannot possibly be caught with the widely spun net of a quantitative research. An analysis of deviance between observed and predicated values, that is, of those countries that offer more public goods than expected, given levels of stateness achieved previously, contemporary economic development, and so on, does not offer itself to any obvious conclusions.

With regard to at least two of the three public goods outcomes, Japan, North Korea, Taiwan, and South Korea provide more public goods than predicted, but also Tunisia, Libya, Kuwait, or Bangladesh. Most of these countries were led over decades by strongmen (Kim Jong of North Korea, Gadhafi of Libya, the military dictators of South Korea, the emir of Kuwait, Tunisia's Ben Ali) or one party regimes (the LDP in Japan) with a strong commitment to nation building and perhaps less corruption than in neo-patrimonial regimes of Africa and the Middle East, many of which appear on the list of countries with less than predicted public goods provision. Also noteworthy is the political stability that many of these countries displayed—at least until the recent wave of democratization and the associated turmoil. Maybe such political commitment can overcome—or at least offset partially—the disadvantages that a short history of stateness brings about for the provision of public goods, in line with Miguel's (2004) comparison of Tanzania and Kenya?

Note that this would not imply a plead for autocratic rule: The above analysis of deviance from predicted values is of course net of the Polity4 score of all countries, meaning that the democratic deficits of these countries and the handicap this implies for public goods provision are already taken into account. From a policy-making point of view, then, the strong and consistent

association between democratic rule and public goods provision brought to light in the above analysis remains the most important and encouraging message this article has to offer.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

Notes

1. The one dissenting finding regards participation in community efforts to provide public goods in the villages of Sierra Leone (Glennerster, Miguel, & Rothenberg, 2013). In a related study, Miguel (2004) finds that diversity affects local public goods provision only in Kenya that lacks the history of nation-building efforts of neighboring Tanzania.
2. They also find, however, that diversity has sometimes positive impact on public goods provision at the sub-national level in a sample of 34 countries (Gerring, Thacker, Lu, & Huang, 2015).
3. The argument outlined above assumes that capable states are aiming at homogenizing their populations in ethnic and linguistic terms. This is, obviously, a strong assumption since some states (such as Switzerland) pursued a policy of maintaining, rather than eliminating linguistic diversity. On the international configurations that influence a state's choice of minority policies, see Mylonas (2012).
4. Data on missionary activity have been collected by Woodberry (2012); it is not associated with diversity once we control for which was the former colonial power.
5. For a more precise and restricted understanding of path-dependency, see discussion in Mahoney (2000).
6. I should note here that literacy is obviously influenced by other factors as well, most importantly by alphabetization through non-state actors such as churches (as historically in Sweden) or missionary schools (as in most sub-Saharan African countries). For the post-independence and post-War period, however, literacy should be a reasonably good indicator for public goods provision by the state.
7. Railroad density is highly correlated with road density ($r = .75$), data that are available for recent cross-sections only. Railroads thus might be a good measurement of public goods provision even after the introduction of cars.
8. We could also use a measurement of the contemporary prevalence of malaria (data from Sachs, 2003), which is at least partly endogenous to state capacity, however. Results are broadly consistent.

9. The same results are obtained if we use data for 1820 or 1870 from Maddison (2003), which are only available for 38 of the countries in Africa and Asia. See already the findings by Bates (1983:35) on Africa.
10. I should note here that this variable is afflicted by considerable measurement problems for the African and Asian states covered in Models 1 to 4 because historical records on pre-colonial warfare are sparse and rarely contain the precise information on battle-deaths needed. See the discussion in Wimmer and Min (2009).
11. Malaria risk in the 1990s and settler mortality in the 19th century produce similar results. The former is at least in part influenced by contemporary state capacity, and thus less suited as a measurement than the exogenous latitude variable. The latter is available for only slightly more than half of the countries covered in Models 1 to 4.
12. Note that when restricting the sample of Model 8 to observations that correspond to those of Models 1 to 4, linguistic fractionalization remains positive and insignificant (results not shown). The difference between the two sets models is therefore not due to different sample universes, but rather because the measurements of the outcome variables differ.

Supplemental Material

The online appendices are available at http://cps.sagepub.com/supplemental

References

Ahlerup, P. (2009, September). *The causal effects of ethnic diversity: An instrumental variable approach* (Working Papers in Economics, 386). Gothenburg, Sweden: University of Gothenburg.

Ahlerup, P., & Olsson, O. (2012). The roots of ethnic diversity. *Journal of Economic Growth*, *17*(2), 17-102.

Alesina, A., Baqir, R., & Easterley, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics*, *114*, 1243-1284.

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

Alesina, A., & Glaeser, E. L. (2004). *Fighting poverty in the US and Europe: A world of difference*. Oxford, UK: Oxford University Press.

Alesina, A., & La Ferrara, E. (2000). Participation in heterogeneous communities. *The Quarterly Journal of Economics*, *115*, 847-904.

Alesina, A., & La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, *63*, 762-800.

Algan, Y., Hémet, C., & Laitin, D. (2011, October). *Diversity and public goods: A natural experiment with exogenous residential allocation* (Institute for the Study of Labor Discussion Paper, 6053). Retrieved from http://ftp.iza.org/dp6053.pdf

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

Banerjee, A., Lakshimi, I., & Somanathan, R. (2005). History, social divisions, and public goods in rural India. *Journal of the European Economic Association*, *3*, 639-647.

Bates, R. H. (1983). *Essays on the political economy of rural Africa*. Cambridge: Cambridge University Press.

Bjornskov, C. (2004). Determinants of generalized trust: A cross-country comparison. *Public Choice*, *130*, 1-21.

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

Brubaker, R. (2013). Language, religion, and the politics of difference. *Nations and Nationalism*, *19*, 1-20.

Carneiros, R. L. (1970). A theory of the state. *Science*, *21*, 733-738.

Chai, S.-K. (1996). A theory of ethnic group boundaries. *Nations and Nationalism*, *2*, 281-307.

Desmet, K., Weber, S., & Rotuno-Ortin, I. (2010). Linguistic diversity and redistribution. *Journal of the European Economic Association*, *7*, 1291-1318.

Deutsch, K. W. (1953). *Nationalism and social communication: An inquiry into the foundations of nationality*. Cambridge, MA: MIT-Press.

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *The Quarterly Journal of Economics*, *112*, 1203-1250.

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

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth*, *8*, 195-222.

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

Fernandez, J. W. (1966). Folklore as an agent of nationalism. In I. Wallerstein (Ed.), *Social change and the colonial situation* (pp. 585-591). New York, NY: John Wiley.

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

Gerring, J., Thacker, S., Lu, Y., & Huang, W. (2015). Does diversity impair human development? A multi-level test of the diversity debit hypothesis. *World Development*, *66*, 166-188.

Gerring, J., Zarecki, D., & Hoffman, M. (2013). *Ethnic diversity and democracy*. Unpublished Manuscript, Department of Political Science, Boston University.

Gerring, J., Ziblatt, D., van Gorp, J., & Arévalo, J. (2011). An institutional theory of direct and indirect rule. *World Politics*, *63*, 377-433.

Glennerster, R., Miguel, E., & Rothenberg, A. D. (2013). Collective action in diverse Sierra Leone communities. *The Economic Journal*, *123*, 285-316.

Golden, M., & Min, B. (2013). Distributive politics around the world. *Annual Review of Political Science*, *16*, 73-99.

Goldin, C., & Katz, L. F. (1999). The shaping of higher education: The formative years in the United States, 1890-1940. *Journal of Economic Perspectives*, *13*, 37-62.

Habyarimana, J., Humphreys, M., Posner, D. N., & Weinstein, J. M. (2007). Why does ethnic diversity undermine public goods provision? *American Political Science Review*, *101*, 709-725.

Harries, P. (1989). Exclusion, classification and internal colonialism: The emergence of ethnicity among the Tsonga-speakers of South Africa. In L. Vail (Ed.), *The creation of tribalism in Southern Africa* (pp. 82-110). London, England: James Currey.

Herbst, J. (2000). *States and power in Africa. Comparative lessons in authority and control*. Princeton, NJ: Princeton University Press.

Hopkins, D. J. (2011). The limited impacts of ethnic and racial diversity. *American Politics Research*, *39*, 344-379.

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

Kramon, E., & Posner, D. N. (2012). *Ethnic favoritism in primary education in Kenya*. Los Angeles, CA: Department of Political Science, University of California, Los Angeles.

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

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1999). The quality of government. *The Journal of Law, Economics, & Organization*, *15*, 222-279.

Laitin, D., & Samatar, S. S. (1987). *Somalia: Nation in search of a state*. Boulder, CO: Westview Press.

Lemarchand, R. (1964). *Political awakening in the Congo*. Berkeley: University of California Press.

Lewis, I. M. (1988). *A modern history of Somalia: Nation and state in the horn of Africa*. Boulder, CO: Westview Press.

Lewis, I. M. (1994). *Blood and bone: The call of kinship in Somali society*. Lawrenceville, NJ: The Red Sea Press.

Lieberman, E. S., & McClendon, G. C. H. (2011). *Endogenous ethnic preferences*. Princeton, NJ: Department of Political Science.

Loveman, M., & Muniz, J. (2006). How Puerto Rico became white: Boundary dynamics and inter-census racial classification. *American Sociological Review*, *72*, 915-939.

Maddison, A. (2003). *The world economy: Historical statistics*. Paris, France: OECD.

Mahoney, J. (2000). Path dependency in historical sociology. *Theory and Society*, *29*, 507-548.

Mahoney, J. (2010). *Colonialism and postcolonial development: Spanish America in comparative perspective*. Cambridge, UK: Cambridge University Press.

Mahzab, M., Atiq, M., & Devrariani, N. (2013). *Comparative analysis of ethnic diversity measures on provisioning of basic public goods: Cross country assessment*. Unpublished Manuscript, Center for Development Economics, Williams College

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

Miguel, E. (2004). Tribe or nation? Nation building and public goods in Kenya versus Tanzania. *World Politics*, *56*, 327-362.

Miguel, E., & Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, *89*, 2325-2368.

Mitchell, J. C. (1974). Perceptions of ethnicity and ethnic behaviour: An empirical exploration. In A. Cohen (Ed.), *Urban ethnicity* (pp. 1-35). London, England: Tavistock.

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic polarization, potential conflict and civil wars. *American Economic Review*, *95*, 796-816.

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

Mylonas, H. (2012). *The politics of nation-building: Making co-nationals, refugees, and minorities*. Cambridge, UK: Cambridge University Press.

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

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Cambridge, MA: Harvard University Press.

Peel, J. D. Y. (1989). The cultural work of Yoruba ethnogenesis. In E. Tonkin, M. McDonald, & M. Chapman (Eds.), *History and ethnicity* (pp. 964-982). London, England: Routledge & Kegan Paul.

Poterba, J. M. (1997). Demographic structure and the political economy of public education. *Journal of Policy Analysis and Management*, *16*(1), 48-66.

Putnam, R. (2007). E Pluribus Unum: Diversity and community in the twenty-first century. *Scandinavian Political Studies*, *30*, 137-174.

Putterman, L., & Weil, D. N. (2010). Post-1500 population flows and the long-run determinants of economic growth and inequality. *Quarterly Journal of Economics*, *125*, 1627-1682.

Ranger, T. (1966). Traditional authorities and the rise of modern politics in Southern Rhodesia, 1898-1930. In E. Stokes & R. Brown (Eds.), *The Zambezian past: Studies in central African history* (pp. 94-136). Manchester, UK: University of Manchester Press.

Rodrik, D. (1999). Where did all the growth go? External shocks, social conflict, and growth collapses. *Journal of Economic Growth*, *4*, 385-412.

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, *9*, 131-165.

Roeder, P. G. (2007). *Where nation-states come from: Institutional change in the age of nationalism*. Princeton, NJ: Princeton University Press.

Ross, M. (2012). *The oil curse. How petroleum wealth shapes the development of nations*. Princeton, NJ: Princeton University Press.

Sala-i-Martin, X., Doppelhoffer, G., & Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *American Economic Review*, *94*, 813-835.

Sachs, J. D. (2003). *Institutions don't rule: Direct effects of geography on per capita income* (NBER Working Paper, 9490). Retrieved from http://www.nber.org/papers/w9490

Samatar, A. I. (1999). *An African miracle. State and class leadership and colonial legacy in Botswana Development*. Portsmouth, NH: Heinemann.

Schapera, I. (1952). *The ethnic composition of Tswana tribes*. London, England: London School of Economics and Political Science.

Soroka, S., Banting, K., & Johnston, R. (2002). Ethnicity, trust, and the welfare state. In F. Kay & R. Johnston (Eds.), *Diversity, social capital and the welfare state* (pp. 95-132). Vancouver, Canada: University of British Columbia Press.

Tefft, S. K. (1999). Perspectives on panethnogenesis: The case of the Montagnards. *Sociological Spectrum*, *19*, 387-400.

Tilly, C. (1975). Western state-making and theories of political transformation. In C. Tilly (Ed.), *The formation of national states in Western Europe* (pp. 601-638). Princeton, NJ: Princeton University Press.

Trounstine, J. (n.d.). *One for you, two for me: Support for public goods investment in diverse communities*. Unpublished manuscript, University of California, Merced.

Vigdor, J. L. (2004). Community composition and collective action: Analyzing initial mail response to the 2000 census. *The Review of Economics and Statistics*, *86*(1), 303-312.

Weber, E. (1979). *Peasants into Frenchmen: The modernization of rural France, 1870-1914*. London, England: Chatto & Windus.

Wimmer, A. (2014). Nation building. A long-term perspective and global analysis: *European Sociological Review*, *30*(6),

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

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

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

Wimmer, A., & Min, B. (2009). The location and purpose of wars around the world. A new global dataset, 1816-2001. *International Interactions*, *35*, 390-417.

Wolf, E. (1982). *Europe and the people without history*. Berkeley: University of California Press.

Woodberry, R. D. (2012). The missionary roots of liberal democracy. *American Political Science Review*, *106*, 244-274.

Young, C. (1965). *Politics in the Congo: Decolonization and independence*. Princeton, NJ: Princeton University Press.

Young, C. (1994). *The African colonial state in comparative perspective*. New Haven, CT: Yale University Press.

Author Biography

Andreas Wimmer studies state formation, nationalism, and ethnic conflict from a comparative historical point of view. His two most recent book publications are *Waves of War* (CUP 2013) and *Ethnic Boundary Making* (OUP 2013).