NBER WORKING PAPER SERIES

NATION-BUILDING AND EDUCATION

Alberto Alesina
Paola Giuliano
Bryony Reich

Working Paper 18839
http://www.nber.org/papers/w18839

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013, Revised January 2019

Previously circulated as “Nation Building.” For helpful feedback and comments, the authors thank Tim Besley, Martin Cripps, Jeffrey Frieden, Oded Galor, Terri Kneeland, Mark Koyama, Alessandro Riboni, Enrico Spolaore as well as participants at various seminars and conferences. Nicolas Dubost, Giulia Giupponi, Andrea Passalacqua and Mikhail Poyker provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Nation-Building and Education
Alberto Alesina, Paola Giuliano, and Bryony Reich
NBER Working Paper No. 18839
February 2013, Revised January 2019
JEL No. F3

ABSTRACT

Nations stay together when citizens share enough values and preferences and can communicate with each other. Democracies and dictatorships have different incentives when it comes to choosing how much and by what means to homogenize the population, i.e. "to build a nation". We study and compare nation-building policies under the transition from dictatorship to democracy in a model where the location and type of government and the borders of the country are endogenous. We find that the threat of democratization provides the strongest incentive to homogenize. We focus upon a specific nation-building policy: the provision of mass primary education. As a motivation, we offer historical discussions of several episodes in the nineteenth century and suggestive correlations for a large sample of countries over the 1925-2014 period.

Alberto Alesina
Department of Economics
Harvard University
Littauer Center 210
Cambridge, MA 02138
and IGIER
and also NBER
aalesina@harvard.edu

Paola Giuliano
Anderson School of Management
UCLA
110 Westwood Plaza
C517 Entrepreneurs Hall
Los Angeles, CA 90095-1481
and IZA
and also NBER
paola.giuliano@anderson.ucla.edu

Bryony Reich
Northwestern University
Kellogg School of Management
GLOBAL Hub 4263
2211 Campus Drive
Evanston, IL 60208
bryony.reich@kellogg.northwestern.edu

1 Introduction

> *"There cannot be a firmly established political state unless there is a teaching body with definitely recognized principles. If the child is not taught from infancy that he ought to be a republican or a monarchist, a Catholic or a free-thinker, the state will not constitute a nation; it will rest on uncertain and shifting foundations; and it will be constantly exposed to disorder and change."* Napoleon I, 1805[1]

From the French Revolution and throughout the 19th century, French rulers expressed the imperative "to form French citizens."[2] Following the unification of Italy (1860), a process led by a Northern elite, Massimo d'Azeglio (one of the founders of unified Italy) remarked: "Italy has been made; now it remains to make Italians." During the 19th and early 20th centuries, those who governed France and Italy implemented a range of policies with the aim of building commonality among the population and "forming" what they determined to be "Frenchmen" and "Italians." A major policy to this end was the introduction of state-controlled education, including compulsory elementary schooling. Other nation-building policies included the introduction of a "national language" in schools, religious services and administration; and the introduction of compulsory military service, which often had the explicit aim of integrating and mixing individuals from different parts of the country.

France and Italy are just two examples. History has witnessed a multitude of efforts to "nation-build."Tilly (1975) observes that "almost all European governments eventually took steps which homogenized their populations: the adoption of state religions, expulsion of minorities,[...] institution of a national language, eventually the organization of mass public instruction." According to Hobsbawm (1990) "states would use the increasingly powerful machinery for communicating with their inhabitants, above all the primary schools, to spread the image and heritage of the 'nation' and to inculcate attachment to it," and that "the official or culture-language of rulers and elites usually came to be the actual language of modern states via public education and other administrative mechanisms." A vast body of work has documented the nation-building motives for the development of compulsory state education systems across European states (Weber, 1979; Ramirez and Boli, 1987).

Why did 19th century European elites see nation-building and the introduction of mass education as imperative? The goal of this paper is to analyze nation-building through education across political regimes and in times of political transitions. We define "nation-building" as a process which leads to the formation of countries in which the citizens feel a sufficient amount of commonality of interests, goals and preferences that they do not wish to separate from each other. The terms state-building and nation-building are sometimes used interchangeably. However, state-building generally refers to the construction of state institutions for a functioning state, one able to collect revenues etc., while nation-building implies the construction of a national identity, which also helps state institutions. We model a heterogenous population and assume that the degree of

[1] Quote from Ramirez and Boli (1987).

[2] Quote from Félix Pécault in 1871 who conducted a general inspection of public education for the French government. See Weber (1979) for many more examples.

divergence of preferences amongst the population is endogenous, in the sense that it can be affected by nation-building policies, which we explicitly model.

Let us consider first a fully secure non-democratic ruling elite (the "ruler" for short). The ruler only extracts rents from his territories. He builds the type of government and adopts policies which match his preferences. He has no interest in nation-building. The incentives of a non-democratic ruler facing a substantial probability of overthrow and the establishment of a democracy are different. A democratic government may choose public goods and policies that differ from the preferences of the ruler or elite. In addition, when installed, a democratic regime may break apart the territories of the dictator (e.g. the former Soviet Union). Thus, homogenization and indoctrination allow those in charge to better maintain their preferred policies and a larger country if democracy prevails. In addition, more homogenization, if it reduces distaste towards the existing government, may reduce the incentive of the population to overthrow the ruler. Both of these incentives to homogenize work in the same direction: a higher threat of democracy induces more homogenization. In more colorful terms: rulers threatened by overthrow will indoctrinate people in order to teach them to "enjoy" the current regime and the current borders of the country and not break away from them. In this paper, we focus on internal factors which motivate governments to implement nation-building policies. Aghion, Jaravel, Persson, and Rouzet (2015) and Alesina, Reich, and Riboni (2017) study the importance of external motives for nation-building, namely the threat of external wars. Internal and external motives to nation-build may coexist as we will show below.

Education facilitates nation-building in several ways. It can change individual preferences by indoctrination. That is convincing individuals far from the ruling government that they do not dislike it that much. For instance, one may argue that schools, say in France or Scandinavia, emphasize the benefits of regulation and social welfare while in the UK and the US the merits of individualism are stressed more (Alesina and Glaeser, 2005). Cantoni et. al. (2017) show that a Chinese education reform, introduced with the explicit aim of shaping ideology, shifted the attitudes of students towards the ideological position of the government in aspects such as their view of free market economics and the political system. Mass education can also facilitate nation-building by teaching a common language. Imagine that the further an individual is from the government the more his or her language will differ. Reducing distance in this case can be interpreted as teaching a common language so that individuals can better communicate with the government and access public services. You (2018) studies the effect of the Chinese reform in 1960 which enforced the use of Mandarin in all schools in China with the explicit goal of reducing diversity. Interestingly this reform was implemented before the country moved toward more economic and political freedom. Clots-Figueras and Masella (2013) show the effect of compulsory Catalan language education on encouraging Catalan identity.

Well-functioning democracies also have reasons to promote homogenization of their citizens, up to a point. Our model implies non-linear comparisons between mass education in democracies and non-democracies: nation-building is lowest in a "safe" dictatorship but may be *higher in a threatened non-democratic regime relative to a democracy.* We show that this novel implication of the model is consistent with the data using historical examples from the 19th century and econometric evidence on a large sample of countries for the period for which we have the necessary data, from 1925 to 2014. We also discuss under which conditions it may be optimal for the ruler to "divide and rule" rather than

homogenize.

Our paper is related to several stands of the literature. One is about education policies across democratic and non-democratic regimes. Aghion, Jaravel, Persson, and Rouzet (2015), using annual data on 137 countries from $1830-2001$, find that autocracies have higher enrollment rates in primary education than democracies. Consistent with this finding, Mulligan, Gil, and Sala-i-Martin (2004) examine cross-country data from 1960-1990 and find that there is no evidence that democracies spend more on public education than non-democratic regimes. Looking at the same data set, Bursztyn (2016) finds that democracies spend less on public education than non-democracies for below median income countries. Lott (1999) also examines education expenditure data from 99 countries in the period 1985-92 and finds that an increase in totalitarianism increases education spending, again with the strongest effects for lower income countries. As a comparison with other public policies, Lott (1999) examines health care expenditure, finding either no effect of totalitarianism or a negative effect.

The second strand is the work on border and country size and separations by Alesina and Spolaore (1997, 2003) and Bolton and Roland (1997). These authors take diversity of preferences amongst individuals as given, whereas in our model the degree of divergence of preferences amongst the population is endogenous.

The third strand is the work on democratic transitions, showing that forward-looking rulers and elites may act to mitigate, not only the threat of democracy, but also the democratic outcome itself. Acemoglu and Robinson (2008) argue that democratic transitions motivate elites to invest in institutions which allow them to maintain a higher degree of power under democracy and mitigate their economic losses from democratic transitions. Besley, Persson, and Reynal-Querol (2016) present evidence that rulers facing a greater threat of loss of power invest in institutional reforms, namely improving executive constraints, to limit the ability of future regimes to act against their interests. Our model suggests that forward-looking elites also invest heavily in building nations through compulsory schooling when threatened with democracy.

The fourth strand is the literature on "state capacity,"as in Besley and Persson (2009, 2010), which examines the development of state institutions in the formation of successful states. This work emphasizes the role of war as an engine for building the ability of the state to raise taxes and establish law and order. Alesina et al. (2017) discuss how indoctrination may motivate soldiers during wars and become part of state-building. The role of wars and democratization as complements in the formation of the modern "state capable nation" will be discussed throughout the paper.

Finally, our paper is connected to the literature on the need for education for the better functioning of institutions, as in Glaeser, Ponzetto, and Shleifer (2007) or Bourguignon and Verdier (2000). Papers by Gradstein and Justman (2002) and Ortega and Tangers (2008) examine schooling as a means to improve communication across groups and so increase growth. Our results are particularly related to the argument that the expected extension of the franchise motivated European elites to introduce mass compulsory schooling, despite its unpopularity with the masses (see Green 1990).

This paper is organized as follows. Section 2 describes three historical examples that speak to the relationship between mass education, nation-building, and the threat of democratization under non-democratic regimes. Section 3 discusses systematic correlations between mass education and the probability of a regime being overthrown for a

large sample of 172 countries over the 1925-2004 period. Section 4 presents the basic model and Section 5 solves it to examine nation-building via education under different regimes. Section 6 extends the model to allow for democratic transition to be endogenous to the nation-building policies of the ruler, and it also provides extensions to examine the importance of divide and rule policies and of state capacity. The last section concludes.

2 Historical Examples

In the West, public policies to educate the population were implemented in force during the 19th and early 20th centuries. During the 19th century, European countries moved from little to no government intervention in schooling (and generally low participation rates) to centralized full-time primary schooling which was compulsory for all children within the nation. This was a significant shift in government policy over a short period of time, made all the more interesting because in many cases it occurred decades before similar welfare interventions and was generally unpopular with the masses.[3] We document that such education reforms followed periods of unrest and were implemented by governments with the stated aim to mitigate the effects of democratization.

We illustrate, in this section, some historical examples (France, Italy and England) which suggest a relationship between the provision of mass education, nation-building, and the threat of democratization. In the next section, we present suggestive and more systematic evidence on a large sample of 172 countries.

France

Although something approaching democracy was almost a century (or more) away in most Western European countries, the 19th century marks the period during which democracy became a major threat for the elites. The French Revolution in 1792 was a turning point in this respect. Hobsbawm (1990) writes of this period, "it became increasingly manifest that the democratization, or at least the increasingly unlimited electoralization of politics, were unavoidable." Hobsbawm sums up the resulting conundrum of elites, observing that it became "obvious, at least from the 1880s, that wherever the common man was given even the most nominal participation in politics as a citizen...he could no longer be relied on to give automatic loyalty and support to his betters or to the state." The resulting effect was to place "the question of the 'nation', and the citizen's feelings towards whatever he regarded as his 'nation', 'nationality' or other center of loyalty, at the top of the political agenda." This is where nation-building comes in.

While the Ancien Régime was a very centralized state, there was little homogenization of the population before the French Revolution (Tilly, 1975). Hobsbawm (1990) estimates that only 12-13% of the population spoke French at the time of the French Revolution. Although the Ancien Régime aimed to centralize administration and imposed French at the highest administrative level, there was little, if any, effort to foster more widely a nation of French-speakers. Weber (1979) writes that the French Crown showed "little concern with the linguistic conquest of the regions under its administration."In fact, the

[3]For example, the first compulsory social insurance system implemented in Europe was a Health Insurance bill in 1883 in Germany. In contrast, public education was already well developed. Even in the first half of the 19th Century, large numbers of German children attended compulsory state-provided primary schools. By 1870, 70% of German $5-14$ year old's attended public primary schooling.

ruling elites made a point of distinguishing themselves from the masses, using language as a barrier (Gellner, 1983). Primary schooling was predominantly provided by the church and was not a public function (Katznelson and Weir, 1985).

Weber (1979) writes that "Diversity had not bothered earlier centuries very much[...] But the Revolution had brought with it the concept of national unity as an integral and integrating ideal at all levels." Schooling was one way to homogenize and, after the Revolution, schools became a key concern of elites. The Constitution of 1791 called for the establishment of free public instruction for all. A major role for schooling was to enforce a national language. The Convention (the legislative assembly from September 1792 to October 1795) decreed that in the Republic children should learn to "speak, read and write in the French language" and that "instruction should take place only in French."The Jacobins insisted "The unity of the Republic demands the unity of speech."[4] Weber (1979) notes that "Linguistic diversity had been irrelevant to administrative unity. But it became significant when it was perceived as a threat to political - that is, ideological - unity."

The first serious attempt to implement mass schooling was made in 1833 following a period of major rebellion (the "July Revolution", $1830-32$). In France, as elsewhere in Europe, the emergence of state intervention in schooling was in no way a concession to the demands of the population; state-provided schooling was, at least into the last quarter of the 19th century, largely unpopular (Katznelson and Weir, 1985; Weber, 1979). What was perhaps the most intense period of schooling reform followed the establishment of the Third Republic in 1870. Hobsbawm (1990) describes this period as one in which the inevitability of a shift of power to the wider population became clear. Schooling was regarded as a key tool in moving the values and way of life of the population towards those of the elite. Weber (1979) highlights the chasm between the way of life and culture of the urban elite and that of the rural masses throughout much of the 19th century. He writes of the perceived need after the Revolution to integrate this part of the population and to make it "French": "the not assimilated rural masses had to be integrated into the dominant culture as they had been integrated into an administrative entity." Weber notes "the village school, compulsory and free, has been credited with the ultimate acculturation process that made the French people French - finally civilized them, as many nineteenth-century educators liked to say." Other nation-building measures by the French government included the suppression of other languages: as late as 1890 a ministerial decree banned religious instruction in Flemish and in 1902 the government banned Breton language sermons.

Policies of homogenization were also motivated by concerns of secession, as highlighted by the case of Brittany. A report on the Breton departments in the 1880s noted that "Brittany, which was not willingly joined to France, which never wholeheartedly accepted its annexation, which still protests" had still to be merged into the nation. The report urged the use of education to "Frenchify Brittany as promptly as possible[...] integrate western Brittany with the rest of France," and that only schooling could "truly unify the peninsula with the rest of France and complete the historical annexation always ready to dissolve."[5] The example of southern France is also illuminating. Historian Joseph Strayer describes the (apparently successful) efforts of the state in homogenization, writ-

[4]Both quotations Weber (1979).

[5]Report by the rector of the Academy of Rennes, Weber (1979).

ing "Languedoc was very like Catalonia and very unlike Northern France, yet it finally became thoroughly French" (Tilly, 1975). Ensuring French was spoken was considered a vital component in integrating the French population and avoiding secessionist threats. Indeed, use of languages other than French were viewed as a particular threat to the stability of the French state: in 1891, the Minister of the Interior argued that preaching in local dialects "may endanger French unity."

Italy

Italian unification was completed by Northern elites in the 1860s, with virtually no involvement of local populations. Italy, once unified, included a diverse population speaking a range of very different languages and dialects. At best, 10% of the population spoke what would become Italian. This was a time of increasing pressure for more democracy (the largest proportion of adult males were enfranchised in Italy in 1912). The governing elite considered homogenization vital to ensure the internal stability. Duggan (2007) documents that "during the 1860s the government had embarked on extensive discussions about what form of Italian should be adopted as the national language. There was a strong feeling in official circles that linguistic centralization was needed to complement political unity." Tuscan was chosen. Linguistic homogenization was to be achieved mainly through schooling and, despite the frequent lack of popularity within the population, "the official line remained that Italian should as far as possible be enforced, with 'Italian' texts being used in schools and dialect literature (of which there was a distinguished tradition in many regions) being discouraged."

In Italy, the link between the introduction of compulsory schooling and the threat of democratization can be read directly from statements of politicians of the time. Francesco Crispi, the Italian Prime Minister from $1887-1891$ and $1893-1896$ wrote "I do not know if we should feel regret at having broadened the popular suffrage before having educated the masses." Politician Nicola Marselli claimed that Italy had introduced freedom before educating the masses, omitting to learn lessons from countries like Britain which had educated first. Michele Coppino, the author of the 1877 Italian compulsory education reform, declared that primary schooling should ensure the masses were "content to remain in the condition that nature had assigned to them" and that the aim of elementary education should be to "create a population[...] devoted to the fatherland and the king." Enough education to homogenize, but not too much to create rebellious masses.

Holding the country together and avoiding a break up was also a major goal of the rulers. Southern regions saw reunification more as a conquest from the North. Cultural differences and animosity across regions persisted in Italy for decades. Even today a political party calls for the separation of some Northern regions.

England

Colley (1986) argues that in England "dividing and ruling seemed a more attractive strategy than state-sponsored nationalism" and that "only after the 1870s did Britain's governing elite commit itself to a patriotic, blatantly nationalist appeal. Not accidentally, this coincided with a massive extension of the suffrage and the introduction of compulsory public education." The fear that nationalism might increase demands by the population meant that nation-building policies were enacted in Britain only once it became clear

that the population as a whole would have a greater say in things.

Public education first appeared in minimal form in 1833, following three years of widespread rioting in rural England and the Great Reform Act of 1832. With further political reform in the 1860s the "full democratization of the political realm was seen as inevitable" (Ramirez and Boli, 1987). Green (1990) writes that the "Education Act of 1870, which established a quasi-national system, was a result, as much as anything, of the desire to control the political effects of the extension of the franchise in 1867 to the skilled working class."

Again, the driving force of democratization behind the introduction of mass education can be read directly from English political debate of the time. The desire to protect the status quo is explicitly stated. Robert Lowe, a British politician and later Home Secretary and Chancellor of the Exchequer, in an address in 1867, highlighted the urgency for education reform following the 1867 Reform Act: "we cannot suffer any large number of our citizens, now that they have obtained the right of influencing the destinies of the country, to remain uneducated [...] it is a question of self preservation - it is a question of existence, even of the existence of our Constitution"[6] In 1870 when W.E. Forster put forward the bill for his education act in Parliament, his speech included the following: "Upon this speedy provision [of elementary education] depends also, I fully believe, the good, the safe working of our constitutional system. To its honour, Parliament has lately decided that England shall in future be governed by popular government [...] now that we have given [the people] political power we must not wait any longer to give them education."[7] Bandiera, Mohnen, Rasul, and Viarengo (forthcoming) highlight this side of nation-building in the context of the United States. Americans introduced compulsory education, in large part, to 'civilize' and instill common civic and other values in migrants, in order to influence their participation in American life.

These three examples suggest that elites imposed mass education on their populations to serve their own interests when threatened with democracy. Of course, an alternative explanation is that rioters demanded public education and the latter was a concession under duress on the part of the rulers. Rioters, however, did not demand education. As noted previously, state-run mandatory schooling was unpopular and opposed by peasantry for much of the 19th century in France. In England violent and non-violent protest spread across the country in the first years of the 1830s. The Royal Commission into the Poor Laws in 1834, that was set up in part in response to this unrest, asked the following question: "Can you give the commissioners any information respecting the causes and consequences of the agricultural riots and burning of 1830 and 1831?" In England, 526 parishes responded. The only causes cited by more than 30 parishes were labor concerns (unemployment, wages, and mechanization of jobs that previously provided employment), subsidies for the poor (poor law) and beer shops (where it is believed many of the protests were organized). Not a single response considered demand for education or anything related to education as a cause of the unrest (Holland, 2005). Similarly, Tilly (1998) provides a detailed study of episodes of collective disturbances in France $1830 - 1860$ with information on the objective of the group involved in the disturbance. Education is not mentioned.

If education in the 19th century was provided with a nation-building motive not as

[6] Quote from Marcham (1973). The 1867 act enfranchised a part of the male urban working-class population.

[7] Quote from Young and Handcock (1964).

a redistributive device, then we should expect differences in the implementation of education policies compared to other welfare policies, such as social security or health care, especially since direct redistributive concerns were closer to population demands than education. Indeed, there are stark differences in the timing of education reform and redistributive policies. The earliest European non-voluntary government insurance system was introduced in 1883 and the first voluntary system in 1871; in contrast, most countries had compulsory universal education by the time welfare reforms were introduced and in some countries it was highly developed (e.g. France and Germany). This is consistent with the historical discussion in Acemoglu and Robinson (2000) on the extension of the franchise. They suggest that in many cases redistributive concessions were not credible before franchise extension (Germany being an exception). Welfare reform then tended to follow franchise extension. In contrast, education reform preceded it.

3 Cross-country evidence

In this section, we show that in a large set of countries mass education reforms are preceded by threats of democratization.

3.1 Data and Specification

3.1.1 Sources and Variable Definitions

Education. We use an unbalanced panel with ten year averages data for 172 countries between 1925 and 2004 with data on primary educational enrollment per capita. Our measure of imputed reform is a binary variable set equal to one if enrollment grew by more than 20 per cent over the previous 10 year period.[8] In performing the analysis, we collapse the data into 10-year averages so as to minimize measurement error. The variable on primary enrollment is defined according to the UNESCO criteria and expressed per 10,000 inhabitants. The underlying data are drawn from the CNTS Data Archive of Banks (2011). For a reduced sample of 14 European countries, we also use a dummy indicating whether new education reforms were adopted. The adoption of education reforms is based on any new law which extended compulsory education, lowered the cost of education (by abolishing school fees or providing for free primary education), or increased the number of schools (by making it compulsory for each municipality to set up at least one primary school). The source for this variable is Flora (1983).

Political Regimes. The autocracy variable is constructed from the polity2 variable taken from the Polity IV database. This variable ranges from -10 to 10, where a higher score means that the country is more democratic. The variable is based on information on constraints on the executive, the openness and competitiveness of the executive recruitment, and the competitiveness of political participation. We define autocracy when the polity2 variable is lower than zero.

Threat to the current regime. Data on threats to the current regime are taken from the CNTS database. We use three different variables, all of which should proxy for the likely probability of threatening the current government:

[8] In Table 1 we also report the robustness of our results using a binary variable set equal to one if enrollment grew by more than 10 per cent over the previous 10 year period

- *major government crises*: document any rapidly developing situation that threatens to bring the downfall of the present regime;
- *revolutions*: documents any illegal or forced change in the top government elite, any attempt at such a change, or any successful or unsuccessful armed rebellion whose aim is independence from the central government;
- *weighted conflict average (WCI)*: the dataset contains also a weighted conflict average (WCI) which is a weighted average of all the conflicts indicators contained in the dataset;[9]

Control variables. Since our measure of educational reform is based upon enrollment per capita, rather than enrollment per school-age child, we control for population growth to mitigate the concern that our measure is affected by shifts in the demographic structure of the population. Aghion, Jaravel, Persson, and Rouzet (2015) show that mass education is associated with the country being involved in an external war in the previous 10 years. The external war variable is taken from the Correlates of War database. We also control for fiscal capacity, measured as revenue and expenditure over GDP (taken from the CNTS dataset), and for GDP per capita, taken from Madison.

Descriptive statistics for all our variables are provided in the Appendix, Part E (Table A1). As the table shows, 26 percent of the countries in our sample experienced an increase in the per capita enrollment in primary school larger than 20 percent, compared to the previous decade; 35 percent experienced at least an increase equal to 10 percent. When we looked at legal education reforms, for the sample limited to European countries, almost 50 percent of them implemented a legal reform in education. In the overall sample, around half of the time period was under autocratic regimes.[10]

3.1.2 Empirical specification

Our baseline regression equation is expressed as:

$$\begin{aligned} educational\ reform_{it} = \alpha_0 + \alpha_1 autocracy_{i,t-1} + \alpha_2 threat\ to\ regime_{i,t-1} + \\ + \alpha_3 autocracy_{i,t-1} \cdot\ threat\ to\ regime_{i,t-1} + \alpha_4 X_{i,t-1} + \delta_i + \gamma_t + \epsilon_{it}. \quad (1) \end{aligned}$$

where $educational\ reform_{i,t}$ is a dummy indicating whether educational enrollment has increased by more than 10 or 20 per-cent in the last 10 years. Our coefficient of interest

[9] The CNTS dataset contains various measures of domestic conflict. In addition to the ones mentioned above it also contains the following variables. *Assassinations*, records the occurrence of any politically motivated murder or attempted murder of a high government official or politician. *General strikes* lists strikes of 1,000 or more industrial or service workers that involve more than one employer and that are aimed at national government policies or authority. *Guerrilla warfare* gives information about armed activities, sabotage, or bombings carried on by independent bands of citizens or irregular force and aimed at the overthrow of the present regime. *Purges* identifies any systematic elimination by jailing or execution of political opposition within the ranks of the regime or the opposition. *Riots* records the occurrence of any violent demonstration or clash of more than 100 citizens involving the use of physical force. *Anti-government demonstrations* record any peaceful public gathering of at least 100 people for the primary purpose of displaying or voicing their opposition to government policies or authority, excluding demonstration of a distinctly anti-foreign nature. We do not consider any of these as part of our analysis as some of them are not related to the probability of the regime being overthrown (riots, anti-government demonstrations, riots and general strikes). Guerrilla warfare could also be relevant but does not refer to a desire of regime overthrow from the general population, whereas assassination refers to the assassination of any high government official and not only to the assassination of the ruler.

[10] This fraction is much smaller in the European sample used for the legal reform measure, where only 4 percent was under autocratic regimes.

is α_3, which indicates that more unstable autocracies are likely to implement education reforms. All our specifications include country (δ_i) and year (γ_t) fixed effects, and population growth, to account for varying shares of school-age children in total population. We also test the robustness of our results to a larger set of controls, $X_{i,t-1}$, including the level of development, fiscal capacity and whether the country was involved in a war in the previous 10 years. The standard errors are clustered at the country level.

Table 1
Educational reform and threats to democracy

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	Educational reform: 10% threshold			Educational reform: 20% threshold			Legal reform, based on Flora (1983)		
Autocracy*gov. crises	**0.174***			**0.161**			**0.460****		
	(0.090)			**(0.100)**			**(0.208)**		
Autocracy*revolutions		**0.169****			**0.225*****			**1.796****	
		(0.079)			**(0.075)**			**(0.697)**	
Autocracy*all internal conflicts			**0.055****			**0.052****			**0.163***
			(0.022)			**(0.023)**			**(0.079)**
Autocracy	-0.083	-0.071	-0.103*	-0.006	-0.011	-0.026	-0.493	-0.296	-0.472
	(0.057)	(0.055)	(0.059)	(0.066)	(0.058)	(0.064)	(0.376)	(0.305)	(0.416)
Pop. growth	0.314	0.295	0.301	0.026	-0.007	0.012	0.082	0.352	0.103
	(0.210)	(0.214)	(0.211)	(0.165)	(0.163)	(0.162)	(0.989)	(0.922)	(0.928)
Gov. crises	-0.156**			-0.117**			0.019		
	(0.060)			(0.054)			(0.081)		
Revolutions		-0.015			-0.005			0.262	
		(0.065)			(0.057)			(0.175)	
All internal conflicts			-0.026			-0.014			0.026
			(0.017)			(0.016)			(0.041)
Number of countries	172	172	172	172	172	172	14	14	14
Country fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Year fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	977	977	977	977	977	977	110	110	110
R-squared	0.453	0.452	0.452	0.406	0.414	0.408	0.641	0.643	0.645

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 years country averages, for the 1925-2004 period. Educational reforms is a dummy if primary per capita school enrollment increased more than 10% (20%) from the previous 10 years (columns 1-3, and columns 4-6). Legal reform is a dummy if the country experienced at least a legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All the dependent variables are lagged. Standard errors are clustered at the country level

Table 1 shows the results for our baseline estimation. The first three columns show the results when the education reform is defined as an increase in primary enrollment higher than 10 percent from the previous 10 years, whereas Columns 4-6 report the results with the 20 percent threshold. Columns 7-9 use the definition of reform constructed by Flora (1983) and it is limited to a sample of 14 European countries. Our coefficient of interest, α_3, is always positive and significant, indicating that the threat to the regime is associated with nation-building when autocracy is the prevalent form of government.

In Table 2, we control for potential confounders that could be driving the results. In this table we use the 20 percent threshold as a measure of education reform. In the Appendix, Table A2, we show the robustness to the 10 percent threshold. Aghion, Jaravel, Persson, and Rouzet (2015) show that the threat of war is associated with increased primary education enrollment (considered as a measure of nation-building), but that the threat of war may only be relevant when countries are sufficiently democratic. This result would be consistent with our model as well: a dictator can "force" armies to fight

by fear, in a more democratic regime it may be more difficult to do so and teaching nationalism may be more compelling and necessary. To take this into account, we add to our specification a variable indicating whether the country was involved in an external war in the previous 10 years, and an interaction term with the fraction of years spent under autocratic regimes (columns 1-3). Consistent with Aghion, Jaravel, Persson, and Rouzet (2015) we find that education reforms responds more positively to military threats in democracies, however the interaction term between threat to democracy and the presence of autocratic regimes remains significant and of similar magnitude. We see our argument about nation-building for fear of democratization and splitting of countries, and state-building for fear of aggression, as complementary and not as alternatives.

Table 2

Educational reform and threats to democracy, robustness to additional controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Educational reform: 20% threshold					
Autocracy*gov. crises	**0.171***			**0.176**		
	(0.097)			**(0.112)**		
Autocracy*revolutions		**0.235*****			**0.215*****	
		(0.075)			**(0.078)**	
Autocracy*all internal conflicts			**0.057****			**0.063*****
			(0.023)			**(0.022)**
Autocracy*international war	-0.135*	-0.139*	-0.142*	-0.150	-0.144	-0.157*
	(0.081)	(0.079)	(0.080)	(0.098)	(0.092)	(0.093)
Autocracy	0.010	0.005	-0.012	-0.029	-0.025	-0.055
	(0.066)	(0.058)	(0.065)	(0.071)	(0.062)	(0.070)
Pop. growth	0.027	-0.007	0.012	-0.080	-0.105	-0.095
	(0.169)	(0.168)	(0.167)	(0.154)	(0.154)	(0.152)
Gov. crises	-0.120**			-0.102*		
	(0.054)			(0.061)		
Revolutions		-0.006			0.002	
		(0.057)			(0.064)	
All internal conflicts			-0.014			-0.011
			(0.016)			(0.017)
International war	0.020	0.012	0.013	0.002	-0.002	-0.003
	(0.053)	(0.054)	(0.054)	(0.058)	(0.058)	(0.058)
Log(revenue)				0.087**	0.088**	0.093**
				(0.042)	(0.040)	(0.040)
log(GDP per capita)				-0.045	-0.033	-0.039
				(0.071)	(0.072)	(0.072)
Number of countries	172	172	172	144	144	144
Observations	977	977	977	781	781	781
R-squared	0.410	0.418	0.412	0.470	0.477	0.476

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 years country averages, for the 1925-2004 period. Educational reforms is a dummy if primary per capita school enrollment increased more than 20% from the previous 10 years. All the dependent variables are lagged. Standard errors are clustered at the country level.

The second confounding effect is related to the building of "state capacity," in terms of raising taxes and establishing law and order. It could be that states view nation-building as a necessity or complement in being able to build state capacity. However, the timing of state-building versus nation-building does not suggest the motives for the two are completely interlinked. In Europe, the period of state-building begins roughly in 1500. Over the following three centuries European states invested in state-building. In contrast, nation-building policies based on education only begin to occur after the French Revolution, once there was a major threat to old aristocracies throughout Europe. We nevertheless control for this theory by including revenue as a proxy for state capacity (columns 4-6).

A third prominent theory is that industrialization prompted governments to undertake significant nation-building. Gellner (1983) argues that agrarian societies have no need for a "nation" in the modern sense of the word. In contrast, an industrial society based upon markets (as opposed to a stratified agrarian society with local markets) needs better means of communication. Universal schooling serves an economic purpose as well, necessary for the development of an industrial society.[11] In other words, productivity would increase in an industrial society with more homogenization relative to an agrarian one. The timing of this theory is questionable. Smith (1998) and Green (1990) argue that education reforms were not implemented country by country in a way that is consistent with industrialization acting as a major driver of reforms. In many continental European countries there was no industrial development when nationalism and the beginnings of mass education first emerged, while in England, education reforms arrived long after the industrial revolution. Also inconsistent with the argument that education was provided as a result of industrialization, Green (1990) suggests that state education, when implemented, did not furnish children with the appropriate technical skills. Consistent with this argument, the inclusion of per-capita GDP as a proxy of industrialization, does not alter our main findings (columns 4-6).

4 A Model of Nation-Building

We propose a model of nation-building to examine homogenization policies across regimes and in times of regime change. The model provides an explanation for the correlations and the historical discussion. We consider a two period model in which governments can choose to nation-build. In the first period, a country is governed by a ruler (dictator or elite, terms used interchangeably). In the second period, the country either becomes democratic or remains governed by the ruler. The probability of democratization is taken as exogenous for now, but we relax this assumption in Section 6.

Homogenization and distance

The population is composed of a continuum of individuals of mass 1 with heterogenous ideal points distributed uniformly on the segment $[0, 1]$. At time t an individual i resides in a country with a single government j that serves the citizens of the country. Individual

[11] See also Bowles (1998) on this point and for a survey of other models in which preferences are endogenous and can be influenced by various institutions.

i's per period utility function at time t is given by

$$u_{it} = g(1 - a_t^j d_{ij}) + y - r_t. \tag{2}$$

The first term $g(1 - a_t^j d_{ij})$, measures the value of the government to individual i. By "government" we refer to a set of public goods and policies provided by an authority. The parameter g is the maximum utility an individual receives from the government when distance is zero, and d_{ij} is the preference distance of individual i from government j. The value of the government to individual i falls with his distance from the government. We think of distance as the language, cultural, ideological, or preference differences between individual i and the public goods and policies provided by government j. The value a_t^j measures the cost of this distance. The remaining terms are income y, which is exogenously given, identical for everyone, and identical across time periods, and taxes in period t, r_t, which are split equally amongst the population of the country.

The government can choose to implement a mass education policy to homogenize the population. We model "homogenization" as a technology which uses state education to reduce the cost of distance from the government. Specifically, government j at time t implements a homogenization policy $\lambda_t^j \in [0, 1]$ such that

$$a_t^j = (1 - \lambda_t^j)a.$$

This reduces the cost to individual i of facing policies and public goods j that are different to his ideal. Since we consider mass education, any homogenization policy, λ_t^j, is applied across the whole population within the country governed by j. Education homogenizes preferences. From now on with the term "distance" we summarize any difference in preferences and with the term "homogenization,"a reduction in such a distance through education policies. To allow for a split of the population (described below) we assume that preferences are perfectly correlated with geography.[12]

Homogenization is durable: languages learned today are not forgotten tomorrow, preferences influenced today by schooling influence future preferences. To model this, we assume the cost of the homogenization policy λ_t^j, for a country of mass s, is

$$s[C(\lambda_t^j) - C(\lambda_{t-1}^j)],$$

where λ_{t-1}^j is homogenization of this population by government j in the previous period.[13] That is, homogenization by government j in the previous period persists so that the cost of homogenization this period covers any additional homogenization. For now, we also assume $\lambda_t^j \geq \lambda_{t-1}^j$.

Assumption 1 *The function $C(\cdot)$ is strictly increasing, strictly convex and twice continuously differentiable as λ_t^j increases from 0 to 1. With $C(0) = 0$, $C'(0) = 0$ and $\lim_{\lambda_t^j \to 1} C'(\lambda_t^j) = \infty$.*

[12] See Alesina and Spolaore (2003) for a discussion and justification of this assumption.

[13] Observe that homogenization by previous governments is redundant if the "location" of the government changes. If, in the previous period, the population of mass s has government $j' \neq j$, then $\lambda_{t-1}^j = 0$. In the working paper version of this article we analyze the complementary case, where homogenization in a different location is not redundant. The results are analogous.

The cost of the homogenization policy is paid with period t taxes. Since we assume taxes are split evenly, this implies the cost of homogenization is split equally among the population of the country. We relax this assumption on equal costs in the Appendix, Part D.

In our model, income is exogenous. However, at least up to a point, diversity of skills, education, background, and culture may increase productivity. In this case a reduction in diversity would have costs and benefits. The latter are already modeled. The former would include not only the costs modeled above but also a reduction in productivity, therefore of income. Given that income enters linearly in the utility function and taxes are lump sum, this reinterpretation of the costs and benefits of diversity would be immediate.

Country Formation

In period 1, the population is ruled by a dictator located at 1/2. In period 2 either the dictator continues to rule the population, or democracy prevails. In the latter case, the population maintains the borders of the single country or splits into two equal-sized countries, A and B, comprising the intervals of ideal points $[0, 1/2]$ and $(1/2, 1]$ respectively. We adopt the restriction of having at most two equal-sized countries to keep the analysis simple while still allowing for endogenous country size (secession).[14] A single government is also located at some j inside each country. Borders and the location of the government can be altered by a democracy at the beginning of period 2 at no cost.

The cost of government (public goods and policies) in period t in a given country is k. Since the cost k can be divided amongst all citizens in the country this captures the benefits of forming a single country rather than breaking into two.[15] In a democracy the voters face a trade off between homogeneity and costs of government. In fact, when a population splits into two countries, the separate countries are more homogeneous and so the government provided in those countries is closer (in language, ideology or preferences) to the median individual in that country. Some individuals in the population may prefer to break up into two countries and face higher costs, rather than be part of a single country with a government that poorly represents their preferences, others may have the opposite preferences (Alesina and Spolaore, 1997). Thus only a democracy in period 2 would have an incentive to separate. A dictator would never split the country since he would lose rents having to provide two governments.

The government budget constraint at time t for a country of mass s is thus

$$sr_t = k + s[C(\lambda_t^j) - C(\lambda_{t-1}^j)].$$

The model allows for diversity within a country to be influenced through two different channels: a choice over government homogenization and a choice to split into more homogeneous entities.

[14] Alesina and Spolaore (1997), in a model of country formation without homogenization, show that a "stability" condition of indifference at the border delivers countries of equal size. We do not allow for unilateral secessions, namely a situation in which without any majority vote a group of citizens form a third country.

[15] Alesina, Spolaore, and Wacziarg (2000) and Alesina and Spolaore (2003) investigate sources of benefits of size, like the dimension of the market and diversity of inputs in productivity. See Bolton and Roland (1997) for a discussion about separatist movements due to income differences.

Decision-Making and Timing

We model an initial period in which the ruler governs, followed by a second period in which democracy may prevail. In period 1, the expected utility for individual i is $U_{i1} = u_{i1} + E[u_{i2}]$ and period 2 utility is $U_{i2} = u_{i2}$, with u_{it} given by (2).

Period 1

The ruler is located at 1/2 and has the government at his ideal point.[16] He decides how much to invest in homogenization in period 1 to maximize his expected utility.

Period 2

With probability $1 - p$, the ruler remains in power in period 2. With probability p, democracy prevails in period 2. Under a democracy, decisions are made by majority rule with the order of voting as follows:

(1) the population decides whether to form a single country or split into two;
(2) the population of each country decides where to locate the government in that country;
(3) the population of each country decides the homogenization policy in that country.

For tie-breaking we assume that when indifferent between one country or two, a single country is formed. We solve the model backwards.

5 Homogenization Decisions

5.1 A Democracy

If democracy prevails in period 2, the population chooses whether to form a single country or split, where to locate the government, and how much to homogenize. This problem is solved in detail in the Appendix, Part A. Here we summarize the basics. For individual i the level of homogenization which equalizes the marginal cost and marginal benefit is given by:

$$gad_{ij} = C'(\lambda_t^j).$$

The optimal level of homogenization for individual i depends upon the distance of individual i from the government and the cost of the homogenization technology. Since for now we assume that the cost of homogenization falls equally on those close to and far from the government, homogenization is a kind of transfer from the center to the periphery, since the latter benefits more. A technology that reduces distance to public goods may be especially beneficial to people with distant preferences. For example, a common language taught in school can help minorities to get access to and increase the

[16] In the Appendix Part B of the present paper we briefly discuss the more general case of a ruler located anywhere along $[0, 1]$. Modeling a dictator as a single agent (technically speaking of measure zero) can be easily generalized by allowing for an elite group to rule the population. The elite group is represented by a group of mass δ with ideal point 1/2. Results on this point are available from the authors. Such an extension complicates notation and algebra with little advantage in terms of insight.

benefits from government public goods provided in that language. Of course, distant minorities may also resist homogenization attempts; distant minorities may be made to pay more for homogenization via a higher tax bill, or may be the target of more extensive homogenization or more painful repressive policies.[17] In the Appendix, Part D, we model unequal costs of homogenization.

A democracy makes a choice not just about homogenization but also government location and borders. Individuals further away from the center of a nation may strictly prefer secession and a new a government that better represents their preferences, rather than be part of a single country subject to high homogenization. The model captures both aspects.

Preferences over homogenization are single peaked, thus a democracy homogenizes up to the point at which the marginal cost of homogenization equals the marginal benefit for the individual at median distance from the government. If homogenization by the ruler in period 1 exceeds this amount, then no additional homogenization will be undertaken by a democracy in period 2. The "preference" interpretation of homogenization, literally speaking, implies that an individual "chooses" a policy that changes her preferences, knowing that after the change she would feel happier in the country in which she lives. This argument becomes more plausible if we think of a dynamic extension in which parents transmit values and educate their children in such a way which makes them fit better in the country in which they live by adopting certain social norms and types of behavior.[18] This is not contradictory to strong attachment to cultural values which can be captured by very high costs of homogenization.

A democracy locates the government at the median ideal point in the population, namely the center of the one or two countries. Thus, in a single country the government is located at $j = 1/2$. In Countries A and B the government is located at $j = 1/4$ and $3/4$ respectively. These results are illustrated in Figure 1 below.

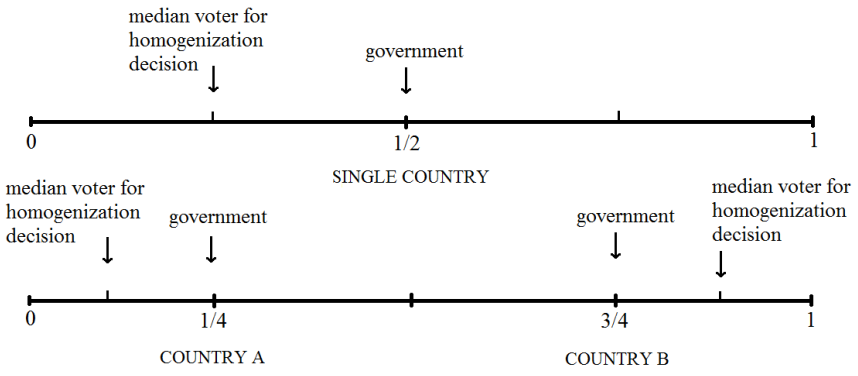

Figure 1: Homogenization and government location under a democracy, for a single country and a split.

[17] Fouka (2016) examines how language restrictions in elementary school in the US, instead of facilitating the assimilation of German immigrant children, instigated a backlash, increasing the sense of cultural identity among them. Affected individuals were indeed less likely to volunteer in WWII and more likely to marry within their ethnic group and to choose decidedly German names for their offspring.

[18] For models related to parents "choosing" values for children see Bisin and Verdier (2000). Algan, Mayer, and Thoenig (2013) discuss the costs of lack of assimilation in France. They document a substantial increase in salaries for children of Arabic families who signal assimilation by choosing French rather than Arabic first names.

The choice of a democracy of whether to form a single country or split captures the trade-off between the benefits of a larger country and the costs of heterogeneity. In our model, however, a democracy also has the option to homogenize. It is perfectly possible that without the option of homogenization ($\lambda_t^j = 0$) a democracy would decide to split into two countries, but the option of choosing $\lambda_t^j \in [0,1]$ would lead a democracy to homogenize somewhat and form a single country. This is "nation-building" even within a democracy. It represents a particular kind of transfer from the center of the population towards the periphery to reduce the costs of being located far from the government, and may therefore avoid separation.

The period 1 ruler can influence the decisions made by a democracy. The first lemma shows how the choice of homogenization by the ruler in the first period will influence the decision of a democracy.

Lemma 1 *There exists a level of homogenization $\lambda^* \in [0,1)$ such that:*

(i). if the ruler homogenizes by at least λ^* in period 1, a democracy would form a single country in period 2 and locate the government at the ruler's location;

(ii). if the ruler homogenizes less than λ^* in period 1, a democratic population in period 2 will choose to split and locate the new governments at $1/4$ and $3/4$ respectively;

where λ^* depends on g, a, k, and $C(\cdot)$.

The proof is in the Appendix, Part A. The intuition is as follows. Homogenization implemented by the ruler in period 1 is costly in period 1, but this homogenization persists in period 2. Since it changes the citizens' relative payoffs from different types of government, it therefore potentially changes the citizens' period 2 choices under democracy in the direction of the ruler. For all citizens in period 2, homogenization by the ruler (in period 1) increases utility from the ideal government of the ruler at $1/2$ relative to other government locations. Homogenization by the ruler reduces the costs of heterogeneity under his ideal government and so makes separation less attractive.

The threshold, λ^*, takes into account the homogenization choices of a democracy in period 2. We discuss this further below.

5.2 A dictatorship

We next consider the choices a non-democratic ruler will make in period 1 and, if still in power, in period 2. Period 2 is the final period and so, if a ruler is in power in period 2, he faces no threat of democracy. He will undertake no additional homogenization in period 2. In period 1 the ruler has to be forward-looking. A more homogeneous population is of no direct benefit to the ruler in period 1, however, homogenization by the ruler in period 1 affects the outcome of period 2 if a democracy prevails (including whether there is a break up of the country). If in a democracy the country does not split the government is located at $1/2$, still at the ideal of the ruler, if the country splits the ruler is located at the border between the two countries, thus at the farthest point form the government.

We are ready to state the decisions of the ruler in period 1.

Proposition 1 *In period 1 the ruler undertakes a level of homogenization which is (weakly) increasing in the probability of democracy, p.*

If democracy prevails, the country may split, so the ruler is located at the periphery (far from the government) and the taxes he has to pay are higher relative to the case of a single country. The worst possible outcome for the ruler. Thus the ruler nation-builds with a particular agenda: he homogenizes to maintain a single country if a democracy prevails. The higher the probability of democracy, the more willing the ruler is to invest in costly homogenization. When the probability of democracy is sufficiently high and homogenization is not too costly, then the ruler will nation-build to ensure that, under a democracy, there will be no break up of countries.

5.3 Comparing regimes

The following corollary compares homogenization under a ruler who faces a threat of democracy $p \in [0, 1]$, and homogenization under a democracy. The incentives to homogenize not only differ between a non-democratic ruler and a democracy, but also differ depending on the type of non-democratic regime. Namely, the incentives for a non-democratic ruler to homogenize depend crucially on whether the non-democratic regime is secure or faces a threat of democratic overthrow.

Corollary 1 *There exists a threshold $\bar{p} \in (0, 1]$ such that:*

(i). for $p \leq \bar{p}$, the period 1 ruler chooses a strictly lower level of homogenization than will be chosen by the period 2 democracy (should democracy prevail);

(ii). for $p > \bar{p}$, the period 1 ruler chooses a higher level of homogenization than will be chosen by the period 2 democracy (should democracy prevail);

where $\bar{p}$ depends on g, a, k, and $C(\cdot)$.

When the probability of democracy is low, a "safe" ruler has little incentive to homogenize. A safe ruler has his ideal government, faces little threat of overthrow and break-up, and has no concern for general welfare, so he is largely unconcerned with the heterogeneity of the population. In contrast, a democracy homogenizes to improve the welfare of its citizens, particularly those at the periphery. Thus a democracy undertakes more nation-building than a relatively safe non-democratic regime.

When the probability of democracy is high, an "unsafe" ruler will undertake a higher level of homogenization than would be chosen by a democracy. Under some parameters, a ruler will even homogenize in period 1 to such an extent as to avoid secession and ensure his ideal government persists in period 2; whereas, without any homogenization by the ruler, a democracy in period 2 would choose less homogenization, split, and opt for governments representing preferences very different from the ruler's.[19][20] Thus an unstable non-democratic regime (i.e. with a high chance of democratization) may over-invest in homogenization compared to a democracy in order to ensure the regime's ideal government is preserved. In contrast, under a stable non-democratic regime, a ruler

[19] Note Corollary 1 part (ii) does not say 'strictly higher.' This is because total homogenization under a democracy in period 2 necessarily incorporates homogenization undertaken in period 1 (that is, $\lambda_t^j \geq \lambda_{t-1}^{j-1}$).

[20] There are two situations under which $\bar{p} = 1$ in Corollary 1, implying a ruler undertakes less homogenization than a democracy whatever the probability of overthrow. These situations are straightforward to interpret. These occur: 1. when homogenization is extremely costly and the ruler cannot preserve his ideal government without a very large cost, and 2. when the ruler's ideal policies are preserved anyway with very little or no homogenization.

under-invests in homogenization compared to a democracy, since he expects to be in power next period and so have his ideal government preserved anyway.

Corollary 1 has the implication that two initially identical populations may both end up as democracies, but a population that has been controlled by an unsafe non-democratic regime may be homogenized by more than one controlled by a safe non-democratic regime. Possibly so much so, that the democracy previously controlled by an unsafe regime is homogeneous enough to form a single country, whereas the other is not. The implication is that today's democracies which followed a smooth path to democracy (where elites foresaw the advent of democracy) may be more homogeneous and bigger than they would be otherwise as a result of nation-building by those elites.

Note that the above results hold for the case of a dictator located at 1/2. Locating the ruler at the center aligns the two incentives of maintaining his preferred government and a large country. In the Appendix, Part B, we discuss the case of a ruler located anywhere in $[0, 1]$. Changing the ruler's location will vary his incentives to homogenize because it varies how different the democratic government is from his ideal (both when a democracy splits or forms a single country). But the intuition remains the same. If homogenization can improve the outcome for a ruler when democracy prevails (by preserving his ideal government and perhaps a larger country), then a higher probability of democracy will induce the ruler to homogenize more.

6 Extensions

6.1 Endogenous Democratization

Homogenization in period 1 may also affect the probability of democratic transition itself by reducing opposition to the ruler's regime. Through schooling, non-democratic governments can indoctrinate in order to lower the value of overthrowing that regime.

Suppose, as above, a revolution opportunity arrives at the beginning of period 2 with probability p. In Section 5, a revolution opportunity always results in democratic transition. If a revolution opportunity arises, we now allow the population governed by the ruler to decide whether or not to overthrow him and install a democratic regime. If the population chooses overthrow, then democracy prevails in period 2; if not, the ruler continues to hold power. The known cost of overthrowing the ruler is L. If a democratic regime is installed, then the utility attained by individual i in period 2 is denoted $U_{i2,dem}$, and if a non-democratic regime is in power in period 2, then the utility attained by individual i is denoted $U_{i2,ruler}$. Given the choices in period 1, the values of $U_{i2,dem}$ and $U_{i2,ruler}$ are known at the beginning of period 2. Individual i prefers overthrow if

$$U_{i2,dem} - U_{i2,ruler} - L \geq 0. \tag{3}$$

The population chooses overthrow if a majority prefer overthrow. The median value of (3), a measure of opposition to the ruler's regime, is decreasing in homogenization by the ruler. Our aim in modeling revolution is to capture, albeit in an abstract way, the idea that higher disutility from dictatorship relative to democracy may be more likely to lead to overthrow. Proposition 2 describes the choices of a ruler.

Proposition 2 *In period 1:*

(i). the ruler undertakes homogenization which is (weakly) increasing in p, and

(ii). there exists a threshold, $\bar{\lambda} \in (0,1)$, which depends upon g, a, k, $C(\cdot)$ and L, such that if the ruler homogenizes to $\bar{\lambda}$ or above the population will choose not to overthrow the ruler.

See the Appendix, Part A, for a proof. Proposition 2 says that if the ruler homogenizes enough in the initial period he can avoid overthrow. As in Proposition 1, homogenization is increasing the higher the probability of a revolution opportunity. The positive relationship between homogenization and threat of democracy occurs for two reasons. The first is the motivation to preserve the status quo should democracy occur, discussed in the previous section. The second reason is to reduce the probability of democracy occurring, and the associated losses.

Let us briefly compare the two motives of rulers to homogenize. In both cases rulers indoctrinate people in order to teach them to "enjoy" the current regime defined by the type of government. The motive to do so in each case is slightly different. One motive is to reduce the threat of democracy. The other is to build a more homogeneous nation that reflects the rulers preferences so that, if democracy prevails, the population will anyway choose to maintain the status quo. Both motives work in the same direction, but the relevance of each motive may vary depending on the "type" of non-democratic regime. The motive to maintain the status quo under democracy applies to domestic elites that expect to stay in the country after democratization (for example, 19th century European elites). The motive to reduce the threat of democracy, shown in Proposition 2, applies to all kinds of non-democratic regimes, even harsh dictators who may be kicked out or eliminated should democracy prevail.

6.2 Divide and Rule

We now study the case in which homogenization continues to reduce the cost of distance to the government, but at the same time it also increases the ability of the population to act collectively and so increases the probability of overthrow. If revolutions become more likely when a population is homogenous, by the same argument, policies that increase diversity and its costs could hinder collective action: the principle of "divide and rule".

To capture this notion, we relax our previous assumption and permit both positive and negative homogenization, $\lambda_t^j \in [-1,1]$, in any period and by any regime. As a result, we also need to update the assumption on costs to allow for negative homogenization (see Appendix, Part C).

A more homogenous population makes collective action easier and this increases the probability of a successful revolution opportunity arising. To model this, we assume the probability of a revolution opportunity now depends both on $p \in [0,1]$, which measures exogenous factors affecting the likelihood of a revolution opportunity (as before, but now p does not directly determine probability), and on homogenization undertaken by the ruler in period 1, denoted $\lambda_1^{1/2}$. The probability of a revolution opportunity is given by $v(p, \lambda_1^{1/2})$, where the function $v : [0,1] \times [-1,1] \to (0,1)$ is twice differentiable, strictly

increasing in p, and strictly increasing and convex in $\lambda_1^{1/2}$. That is, a higher exogenous threat of a revolution opportunity and higher homogeneity both increase the probability of a revolution opportunity occurring. Convexity in $\lambda_1^{1/2}$ ensures a unique optimal homogenization policy.[21] Otherwise the framework is exactly as detailed so far. The model now captures three possible effects of homogenization together: the direct effect of reducing the cost of distance to the ruler's ideal government, the effect this can have on reducing willingness to overthrow, and the effect of increasing the ability of the population to act collectively. Proposition 3 describes the homogenization choices of the ruler under this richer model. Under a sufficient condition which implies that the marginal effect of p on the revolution opportunity is not too sensitive to homogenization, we obtain Proposition 3.

Proposition 3 *There exists a threshold $\hat{p}$, such that in period 1:*

(i). if $p \leq \hat{p}$ the ruler undertakes weakly negative homogenization (divide and rule policies);

(ii). if $p > \hat{p}$ the ruler undertakes strictly positive homogenization which is (weakly) increasing in p;

where $\hat{p}$ depends on g, a, k, $C(\cdot)$, L, and $v(\cdot,\cdot)$.

A democracy never chooses strictly negative homogenization. A ruler in period 2 still always chooses zero additional homogenization since it is the final period. However, a period 1 ruler may now choose to divide and rule. He may undertake strictly negative homogenization. The ruler has an incentive to increase the costs of diversity only when homogenization increases the probability of collective action. The proof is in the Appendix, Part C.

Instead of both forces acting in the same direction, the ruler faces two conflicting forces. On the one hand, if he implements the divide and rule policy he makes collective action more difficult and reduces the probability of a revolution opportunity. However, with low (or even negative) homogenization, if democracy prevails, the country may be unstable and split and the new government may not reflect the ideal of the ruler, the outcome that the rulers like the least. When conditions make democracy unlikely, p low, the incentive to divide and rule dominates; when conditions favor democracy, p high, the incentive to homogenize dominates.[22] In fact, under some conditions on $v(\cdot,\cdot)$, when $p \leq \hat{p}$, not only is homogenization negative but it is also decreasing in p, while for $p > \hat{p}$ homogenization is positive and increasing in p.[23]

6.2.1 Colonizers who divide and rule

Our analysis can shed light on the policies of some colonizers. Colonizers are different from domestic rulers because if overthrown they can leave the country and go home where they maintain a "high status" as part of the elite.[24] This possibility implies that

[21] Assuming $v \in (0,1)$ is not necessary for the results but simplifies the algebra.

[22] For some parameters, because we limit assumptions on the function $v(\cdot,\cdot)$, with an analogous intuition to Proposition 1, we can have the degenerate cases: for all $p \in [0,1]$ the ruler implements strictly negative homogenization, or for all $p \in [0,1]$ the ruler chooses $\lambda_1^{1/2} \geq \min\{\lambda^*, \bar{\lambda}\}$.

[23] This occurs when $\frac{\partial v(p,\lambda_1^{1/2})}{\partial \lambda_1^{1/2} \partial p} > 0$ and $\lambda^* = 0$, for example.

[24] Even a domestic rulers can go in exile if overthrown but this is of course very different. They would be in exile and not member of a ruling elite.

a colonizer would not pay the costs of homogenizing the population to maintain a large country should democracy prevail. For a colonizer, divide and rule has the benefit of reducing the possibility of a revolution, but no costs in terms of increasing the likelihood of separation if a democracy prevails (since this cost is not internalized by the colonizers). Zambia, a British colony from the 19th century to its independence in 1964, adheres to this pattern. Colonization was a "take-the-money-and-run affair" with education mainly provided by missionaries. Colonization exacerbated differences among the Zambian population (Marten and Kula (2008) on language; Phiri (2006) on regional divisions). On independence, a multitude of languages were spoken, with English existing as the main language of commerce and administration. Kenneth Kuanda, the first president of Zambia, claimed that although nationalism had led to independence, national identity in Zambia was completely lacking. Phiri (2006) writes that "Zambia's experience in the first eight years of independence is a typical example of how mostly new independent African countries grappled with the need to create a sense of national identity." In this period the national motto "One Zambia, One Nation" was adopted and English became the official language. More generally, colonizers of Africa did not make an effort to build cohesive nation states (Easterly and Levine, 1997; Herbst, 2000; Alesina, Easterly, and Matuszeski, 2011; Michalopoulos and Papaioannou, 2016). Active policies of divide and rule were also used by British colonizers in India. The British had done little, if anything, to homogenize a diverse population, even using specific policies of divide and rule (Christopher, 1988).

7 Conclusion

We examined when and to what extent a government chooses policies directed toward homogenizing its population. We offer four key findings. One, when the probability of democracy is low a dictator undertakes no homogenization. He allows the population to remain heterogeneous since he faces little threat of overthrow and does not care about population welfare. Two, homogenization by the ruler is increasing in the threat of democracy. Three, a ruler who faces a high probability of overthrow may undertake the highest levels of homogenization, beyond anything that would be undertaken by a democracy. He does this in order to better preserve his ideal government and a large country should democracy prevail, as well as to reduce opposition to his regime and so lower the threat of democracy itself. Four, if dividing the population makes the organization of a revolt harder, under certain conditions the ruler may choose a policy of divide and rule. Finally we offer some suggestive historical discussion and evidence which is consistent with some of these results.

We do not explore the effectiveness of individual nation-building policies. It may be that certain policies are effective while others not, depending on the situation. In some cases an attempt by a democracy to nation-build may even be counter productive. These are excellent topics for future research.

References

Acemoglu, D., and J. A. Robinson (2000): "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *The Quarterly Journal*

of Economics, 115(4), 1167–1199.

——— (2008): "Persistence of Power, Elites, and Institutions," *American Economic Review*, 98(1), 267–293.

Aghion, P., X. Jaravel, T. Persson, and D. Rouzet (2015): "Education and Military Rivalry," NBER Working Papers 18049, National Bureau of Economic Research, Inc.

Alesina, A., W. Easterly, and J. Matuszeski (2011): "Artificial States," *Journal of the European Economic Association*, 9(2), 246–277.

Alesina, A., and E. Glaeser (2005): *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press, Oxford UK.

Alesina, A., B. Reich, and A. Riboni (2017): "Nation-Building, Nationalism and Wars," *Mimeo*.

Alesina, A., and E. Spolaore (1997): "On the Number and Size of Nations," *The Quarterly Journal of Economics*, 112(4), 1027–1056.

——— (2003): *The Size of Nations*. MIT Press, Cambridge MA.

Alesina, A., E. Spolaore, and R. Wacziarg (2000): "Economic Integration and Political Disintegration," *American Economic Review*, 90(5), 1276–1296.

Algan, Y., T. Mayer, and M. Thoenig (2013): "The Economic Incentives of Cultural Transmission: Spatial : Spatial Evidence from Naming Patterns across France," Working Papers 2013-25, CEPII research center.

Bandiera, O., M. Mohnen, I. Rasul, and M. Viarengo (forthcoming): "Nation-Building Through Compulsory Schooling During the Age of Mass Migration," *Economic Journal*.

Besley, T., and T. Persson (2009): "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 99(4), 1218–1244.

——— (2010): "State Capacity, Conflict, and Development," *Econometrica*, 78(1), 1–34.

Besley, T., T. Persson, and M. Reynal-Querol (2016): "Resilient Leaders and Institutional Reform: Theory and Evidence," *Economica*, 83(332), 584–623.

Bisin, A., and T. Verdier (2000): "'Beyond the Melting Pot': Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits," *The Quarterly Journal of Economics*, 115(3), 955–988.

Bolton, P., and G. Roland (1997): "The Breakup of Nations: A Political Economy Analysis," *The Quarterly Journal of Economics*, 112(4), 1057–1090.

Bourguignon, F., and T. Verdier (2000): "Oligarchy, Democracy, Inequality and Growth," *Journal of Development Economics*, 62(2), 285–313.

Bowles, S. (1998): "Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions," *Journal of Economic Literature*, 36(1), 75–111.

Bursztyn, L. (2016): "Poverty And The Political Economy Of Public Education Spending: Evidence From Brazil," *Journal of the European Economic Association*, 14(5), 1101–1128.

Cantoni, D., Y. Chen, D. Y. Yang, N. Yuchtman, and Y. J. Zhang (2017): "Curriculum and Ideology," *Journal of Political Economy*, 125(2), 338–392.

Christopher, A. J. (1988): "'Divide and Rule': The Impress of British Separation Policies," *Area*, 20, 233–240.

Clots-Figueras, I., and P. Masella (2013): "Education, Language and Identity," *Economic Journal*, 123(570), 332–357.

Colley, L. (1986): "Whose Nation? Class and National Consciousness in Britain 1750-1830," *Past and Present*, 113(1), 97–117.

Duggan, C. (2007): *The Force of Destiny: A History of Italy since 1796*. Penguin Books, London.

Easterly, W., and R. Levine (1997): "Africa's Growth Tragedy: Policies and Ethnic Divisions," *The Quarterly Journal of Economics*, 112(4), 1203–1250.

Flora, P. (1983): *State, Economy, and Society in Western Europe 1815 - 1975. A Data Handbook. Vol. 1. The Growth of Mass Democracies and Welfare States*. Campus Verlag, Frankfurt am Main.

Fouka, V. (2016): "Backlash: The Unintended Eects of Language Prohibition in US Schools after World War I," Discussion paper.

Gellner, E. (1983): *Nations and Nationalism*. Cornell University Press, Ithaca.

Glaeser, E., G. Ponzetto, and A. Shleifer (2007): "Why does Democracy Need Education?," *Journal of Economic Growth*, 12(2), 77–99.

Gradstein, M., and M. Justman (2002): "Education, Social Cohesion, and Economic Growth," *American Economic Review*, 92(4), 1192–1204.

Green, A. (1990): *Education and State Formation in England, France and the USA*. Palgrave Macmillan, UK.

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

Hobsbawm, E. (1990): *Nations and Nationalism since 1780: Programme, Myth, Reality*. Cambridge Univeristy Press, Cambridge UK.

Holland, M. (2005): "The Captain Swing Project," in *Swing Unmasked: The Agricultural Riots of 1830 to 1832 and Their Wider Implications*, ed. by M. Holland. FACHRS publications, UK.

Katznelson, I., and M. Weir (1985): *Schooling for All: Class, Race, and the Decline of the Democratic Ideal*. Basic Books, New York.

Laitin, D. (2007): *Nation, States, and Violence*. Oxford University Press, Oxford UK.

Lott, Jr., J. R. (1987): "Why Is Education Publicly Provided? A Critical Survey," *Cato Journal*, 7(2), 475–501.

——— (1999): "Public Schooling, Indoctrination, and Totalitarianism," *Journal of Political Economy*, 107(6), 127–157.

Marcham, A. (1973): "Educating our Masters: Political Parties and Elementary Education 1867 to 1870," *British Journal of Educational Studies*, 21(2), 180–191.

Marten, L., and N. Kula (2008): "Zambia: One Zambia, One Nation, Many Languages," in *Language and National Identity in Africa*, ed. by A. Simpson. Oxford University Press, Oxford.

Michalopoulos, S., and E. Papaioannou (2016): "The Long-Run Effects of the Scramble for Africa," *American Economic Review*, 106(7), 1802–1848.

Mulligan, C. B., R. Gil, and X. Sala-i-Martin (2004): "Do Democracies Have Different Public Policies than Nondemocracies?," *Journal of Economic Perspectives*, 18(1), 51–74.

Ortega, J., and T. Tangers (2008): "Unilingual Versus Bilingual Education: A Political Economy Analysis," *Journal of the European Economic Association*, 6(5), 1078–1108.

Phiri, B. J. (2006): *A Political History of Zambia: From the Colonial Period to the 3rd Republic*. Africa World Press, Inc.

Ramirez, F., and J. Boli (1987): "The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization," *Sociology of Education*, 60(1), 2–17.

Smith, A. (1998): *Nationalism and Modernism: A Critical Survey of Recent Theories of Nations and Nationalism*. Routledge, London.

TILLY, C. (1975): "Reflections on the History of European State-Making," in *The Formation of National States in Western Europe*, ed. by C. Tilly. Princeton University Press, Princeton NJ.

——— (1998): *Disturbances in France, 1830-1860 and 1930-1960: Intensive Sample*. Inter-University Consortium for Political and Social Research, Ann Arbor, MI. doi:10.3886/ICPSR00051.v1.

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

YOU, Y. (2018): "Unified Language, Labor and Ideology," *Mimeo*.

YOUNG, G., AND W. HANDCOCK (1964): *English Historical Documents: Vol. 9, 1833-1874*. Eyre and Spottiswoode, London.

Appendix

Part A

The Case of Democracy

Suppose at $t = 1$ the ruler homogenizes by some $\lambda_1^{1/2} \in [0, 1]$. This section details the choices made by a democracy in period 2.

Lemma $1A$: **A democracy will locate the government at the center of the country.**

First examine the case where $\lambda_1^{1/2} = 0$ and a democracy forms a single country. For any government location $j \in [1/4, 3/4]$, the level of homogenization chosen by majority rule, denoted λ_j^m, satisfies $ga/4 = C'(\lambda_j^m)$, since the median voter is at distance $d_{ij} = 1/4$. Therefore, locating the government at $j = 1/2$ beats all other $j \in [1/4, 3/4]$ in a pairwise vote. For $j \in [0, 1/4)$ (the argument for $j \in (3/4, 1]$ is symmetric), the level of homogenization chosen by majority rule, λ_j^m, satisfies $ga(1/2 - j) = C'(\lambda_j^m)$. To show that $j = 1/2$ beats all $j \in [0, 1/4)$ in a pairwise vote, let l_i denote the distance of individual i from the center of the population. Let $\hat{l}_i$ denote this distance for individual $i \in [0, 1/2]$ who is indifferent between a government at some fixed $j \in [0, 1/4)$ and a government at the center. Similarly denote by $\hat{\hat{l}}_i$ the distance of the individual that satisfies the same condition on the interval $i \in [1/2, 1]$. It is straightforward to see that relative utility from $j = 1/2$ versus some fixed $j \in [0, 1/4)$ is strictly decreasing in l_i. Thus $\hat{l}_i$ and $\hat{\hat{l}}_i$ are unique and $\hat{l}_i + \hat{\hat{l}}_i$ is the proportion of the population who vote for $j = 1/2$ in a pairwise vote. Observe that necessarily $\hat{\hat{l}}_i > 1/4$, so if $\hat{l}_i \geq 1/4$ then at least half the population prefer $j = 1/2$. It remains to examine the possibility that $\hat{l}_i < 1/4$. In this case $\hat{l}_i$ and $\hat{\hat{l}}_i$ satisfy

respectively

$$g - ga(1-\lambda_{1/2}^m)\hat{l}_i + y - k - C(\lambda_{1/2}^m) = g - ga(1-\lambda_j^m)(1/2 - j - \hat{l}_i) + y - k - C(\lambda_j^m) \quad (4)$$

$$g - ga(1-\lambda_{1/2}^m)\hat{\hat{l}}_i + y - k - C(\lambda_{1/2}^m) = g - ga(1-\lambda_j^m)(1/2 - j + \hat{\hat{l}}_i) + y - k - C(\lambda_j^m). \quad (5)$$

Expressions (4) and (5) can be rearranged to find

$$\hat{l}_i + \hat{\hat{l}}_i = \frac{1}{ga}\left(\frac{2(1-\lambda_{1/2}^m)}{(1-\lambda_{1/2}^m)^2 - (1-\lambda_j^m)^2}\right)[C(\lambda_j^m) - C(\lambda_{1/2}^m) + ga(1-\lambda_j^m)(1/2-j)].$$

Since $C(\cdot)$ is a convex continuously differentiable function on $(0,1)$ then $C(\lambda_j^m) - C(\lambda_{1/2}^m) \geq C'(\lambda_{1/2}^m)[\lambda_j^m - \lambda_{1/2}^m] = (\lambda_j^m - \lambda_{1/2}^m)ga/4$ and since we examine $j < 1/4$ we have $(1 - \lambda_j^m)ga(1/2-j) > (1-\lambda_j^m)ga/4$. Using these inequalities it can be seen that $\hat{l}_i + \hat{\hat{l}}_i > 1/2$.

Next examine the case where $\lambda_1^{1/2} \in (0,1]$ and a democracy forms a single country. A democracy will choose $j = 1/2$. This follows by the argument above, noting that, for all i, the utility from locating $j = 1/2$ weakly increases compared to (4) and (5) while the utility from any $j \neq 1/2$ is the same.

Next examine the case where $\lambda_1^{1/2} \in [0,1]$ and a democracy chooses to split. If the governments are located at $1/4$ and $3/4$ respectively, then homogenization satisfies $ga/8 = C'(\lambda_j^m)$, denoted $\lambda_{1/4}^m$ and $\lambda_{3/4}^m$ respectively. By the same argument as a single country, $j = 1/4$ beats all other locations $j \neq 1/2$ in a pairwise vote. Additionally, we need to show that $j = 1/4$ necessarily beats $j = 1/2$, even if $\lambda_1^{1/2} > 0$. By contradiction, suppose the population splits and a majority in country A prefer a government at $j = 1/2$. Then that same majority must strictly prefer a single country with the government located at $j = 1/2$ to a split country with any government $j \in [0, 1/2]$. By symmetry, a majority in Country B must also prefer a single country with $j = 1/2$ to a split. A contradiction. □

When will a democracy split or form a single country?

For $\lambda_1^{1/2} \in [0, \lambda_{1/2}^m)$, expression (6) gives the period 2 utility of individual i at distance $l_i \in [0, 1/4]$ from the center if a single country is formed minus his utility from a split:

$$\Big[g - (1-\lambda_{1/2}^m)gal_i + y - k - \Big(C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})\Big)\Big] - \big[g - (1-\lambda_{1/4}^m)ga(1/4 - l_i) + y - 2k - C(\lambda_{1/4}^m)\big] \quad (6)$$

For individual i at distance $l_i \in [1/4, 1/2]$ this is

$$\Big[g - (1-\lambda_{1/2}^m)gal_i + y - k - \Big(C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})\Big)\Big] - \big[g - (1-\lambda_{1/4}^m)ga(l_i - 1/4) + y - 2k - C(\lambda_{1/4}^m)\big] \quad (7)$$

Expression (6) is at a maximum when $l_i = 0$ and decreasing until $l_i = 1/4$; while expression (7) is increasing from the same value at $l_i = 1/4$ to a maximum at $l_i = 1/2$. Thus there exist uniquely two individuals, $l_i' \in [0, 1/4]$ and $l_i'' \in [1/4, 1/2]$, with the same value

of (6) and (7) respectively and such that $l_i' + (0.5 - l_i'') = 1/4$. Then l_i' solves

$$
\begin{aligned}
&- \left[(1-\lambda_{1/4}^m) + (1-\lambda_{1/2}^m)\right] gal_i' + (1-\lambda_{1/4}^m)ga/4 + k - C(\lambda_{1/2}^m) + C(\lambda_1^{1/2}) + C(\lambda_{1/4}^m) \\
&= \left[(1-\lambda_{1/4}^m) - (1-\lambda_{1/2}^m)\right] ga(1/4 + l_i') - (1-\lambda_{1/4}^m)ga/4 + k - C(\lambda_{1/2}^m) + C(\lambda_1^{1/2}) + C(\lambda_{1/4}^m),
\end{aligned}
$$

and individuals at distance l_i', l_i'' have the median valuation of a single country or split

$$
l_i' = \frac{(1-\lambda_{1/2}^m) + (1-\lambda_{1/4}^m)}{8(1-\lambda_{1/4}^m)} \qquad\qquad l_i'' = 1/4 + \frac{(1-\lambda_{1/2}^m) + (1-\lambda_{1/4}^m)}{8(1-\lambda_{1/4}^m)}.
$$

Similarly, for $\lambda_1^{1/2} \geq \lambda_{1/2}^m$, i's utility from a single country versus a split is

$$
[g - (1-\lambda_1^{1/2})gal_i + y - k] - [g - (1-\lambda_{1/4}^m)ga(1/4 - l_i) + y - 2k - C(\lambda_{1/4}^m)] \tag{8}
$$

where $l_i \in [0, 1/4]$. The median voter is at distance

$$
l_i = \frac{(1-\lambda_{1/4}^m) + (1-\lambda_1^{1/2})}{8(1-\lambda_{1/4}^m)}. \tag{9}
$$

Proof of Lemma 1

Expressions (6) and (8) evaluated for the respective median voters are equal at $\lambda_1^{1/2} = \lambda_{1/2}^m$ and both strictly increasing in $\lambda_1^{1/2}$. At $\lambda_1^{1/2} = 1$, expression (8) is positive. The threshold, λ^*, is the value of $\lambda_1^{1/2}$ at which (6) (alternatively (8)) is equal to zero for the median voter. □

Proof of Proposition 1

The ruler chooses either $\lambda_1^{1/2} = 0$ or $\lambda_1^{1/2} = \lambda^*$, since his expected utility from any other level of homogenization is strictly lower than one of 0 or λ^*. Expected utility from $\lambda_1^{1/2} = 0$ is

$$
[g + y - k] + p\left[g - (1-\lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)\right] + (1-p)\left[g + y - k\right]; \tag{10}
$$

from $\lambda_1^{1/2} = \lambda^* \leq \lambda_{1/2}^m$ is

$$
[g + y - k - C(\lambda^*)] + p\left[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda^*)]\right] + (1-p)\left[g + y - k\right]; \tag{11}
$$

and from $\lambda_1^{1/2} = \lambda^* \geq \lambda_{1/2}^m$ is

$$
[g + y - k - C(\lambda^*)] + \left[g + y - k\right]. \tag{12}
$$

At $p = 0$, $\lambda_1^{1/2} = 0$ is optimal. By the optimality of $\lambda_{1/2}^m$ for $d_{ij} = 1/4$, $(1-\lambda_{1/4}^m)ga/4 + C(\lambda_{1/4}^m) > (1-\lambda_{1/2}^m)ga/4 + C(\lambda_{1/2}^m)$. Thus (10) is decreasing in p at a faster rate than (11) and (12). □

Proof of Proposition 2

Using results from the previous proofs and going through the different cases, it is straightforward to show that $U_{i2,dem} - U_{i2,ruler}$ is increasing in $\lambda_1^{1/2}$. That there exists a $\bar{\lambda}$ follows by noting that there exists some $\lambda_1^{1/2} \in [0,1)$ at which $U_{i2,dem} - U_{i2,ruler} = 0$. To show that $\lambda_1^{1/2}$ is weakly increasing in p observe that all other choices of $\lambda_1^{1/2}$ are strictly dominated for the ruler by 0, λ^* or $\bar{\lambda}$. The ruler's utility from $\lambda_1^{1/2} = 0$, when $\lambda^* > 0$, is

$$[g + y - k] + p[g - (1 - \lambda_{1/4}^m)ga0.25 + y - 2k - C(\lambda_{1/4}^m)] + (1-p)[g + y - k]; \qquad (13)$$

from $\lambda_1^j = \lambda^* \leq \lambda_{1/2}^m$ is

$$[g + y - k - C(\lambda^*)] + p[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda^*)]] + (1-p)[g + y - k]; \qquad (14)$$

and for $\bar{\lambda}$ is

$$[g + y - k - C(\bar{\lambda})] + [g + y - k]. \qquad (15)$$

Observe that (13) is decreasing in p at a faster rate than (14), and (15) does not change with p. The result follows by noting that if $\lambda^* > \bar{\lambda}$ then expected utility from $\bar{\lambda}$ is strictly higher than λ^* for all p, and if $\lambda^* > \lambda_{1/2}^m$ then $\lambda^* > \bar{\lambda}$. □

Part B

Discussion of a ruler located anywhere

The intuition behind the main results continues to hold for a ruler at a location other than 1/2.

Take p exogenous. If a ruler at a location other than 1/2 homogenizes by a high enough amount, then in the second period, a majority in a democracy will choose to form a single country and will choose to locate the government at the ruler's location. The reason for this is exactly the same as a ruler located in the center: if the ruler homogenizes enough, he makes his own location more attractive compared to other locations for enough of the population. It follows that when the probability of the democratic outcome, p, is higher, the ruler is (weakly) more willing to undertake costly homogenization.

The intuition that rulers under threat of democracy will tend to undertake more extreme homogenization than democracies also continues to hold for a ruler at a location other than 1/2. A simple way to see this is to suppose that the costs of homogenization can be placed on minorities such that the government (either a democratic government or a dictator) faces negligible costs. This can be captured within our definition of an unequal homogenization technology. When the cost of homogenization is low enough and the probably of democracy is sufficiently high, the ruler will for certain homogenize enough to avoid secession and make sure his ideal government persists. In contrast we have seen that in a democracy, the decisive voter over homogenization is the voter at median distance from the government. He faces more substantial costs (whether using unequal or equal-cost technology), and will be less willing to undertake high homogenization. A democracy has to take into account the views of everyone, so a situation where very high homogenization is undertaken by a particular government is less likely to get agreement from the median voter.

Part C

Divide and rule

In this part of the Appendix we extend the model and results to capture a situation where homogenization might increase the probability of overthrow. That is, revolutions become more likely when the population is more homogenous. We start by permitting both positive and negative homogenization, $\lambda_t^j \in [-1, 1]$, in any period and by any regime.

As a result we need to update the assumptions on costs to allow for negative homogenization. First, assume

$$C(\lambda_t^j) = C(|\lambda_t^j|), \text{ for all } \lambda_t^j \in [-1, 1],$$

where $C(\cdot)$ is defined in Assumption 1 for $\lambda_t^j \in [0, 1]$. Second, the cost of homogenization policy λ_t^j for a population of mass s is as before, $s[C(\lambda_t^j) - C(\lambda_{t-1}^j)]$, unless policy λ_t^j "reverses" policy λ_{t-1}^j to some degree, in which case the cost is $sC(|\lambda_t^j - \lambda_{t-1}^j|)$. By reversal, we mean when $\lambda_{t-1}^j > 0$ then $\lambda_t^j < \lambda_{t-1}^j$ and when $\lambda_{t-1}^j < 0$ then $\lambda_t^j > \lambda_{t-1}^j$.[25] To avoid a technical complication that arises when allowing for negative homogenization policies, we also make the assumption that, following negative homogenization, if democracy prevails, a democratic government is always located at the center of any democratic country.[26]

A more homogenous population makes collective action easier and this increases the probability of a successful revolution opportunity arising. To model this, we assume the probability of a revolution opportunity now depends both on $p \in [0, 1]$, which measures exogenous factors affecting the likelihood of a revolution opportunity (as before, but now p does not directly determine probability), and on homogenization undertaken by the ruler in period 1, denoted $\lambda_1^{1/2}$. The probability of a revolution opportunity is given by $v(p, \lambda_1^{1/2})$, where the function $v : [0, 1] \times [-1, 1] \to (0, 1)$ is twice differentiable, strictly increasing in p, and strictly increasing and convex in $\lambda_1^{1/2}$. That is, a higher exogenous threat of a revolution opportunity and higher homogeneity both increase the probability of a revolution opportunity occurring. Convexity in $\lambda_1^{1/2}$ ensures a unique optimal homogenization policy.[27] Otherwise the framework is exactly as detailed in the main text. Proposition 3 in the text describes the homogenization choices of the ruler under this richer model.

Proof of Proposition 3. Suppose the ruler homogenizes in period 1 by $\lambda_1^{1/2} \in [-1, 1]$. Examine the choices of a democracy in period 2. If $\lambda_1^{1/2} \geq 0$, there is no change to previous results. If $\lambda_1^{1/2} < 0$, and a democracy forms a single country then the democracy will choose $\lambda_2^{1/2} \geq \lambda_1^{1/2}$ with utility under a democracy equal to

$$u_{i2} = g - ga(1 - \lambda_2^j)d_{ij} + y - k - C(|\lambda_2^j - \lambda_1^j|).$$

[25] If λ_t^j reverses policy λ_{t-1}^j, the period t cost $s[C(\lambda_t^j) - C(\lambda_{t-1}^j)]$ does not make sense. To see this, note that if $\lambda_{t-1}^j < 0$, $\lambda_t^j > 0$, and $|\lambda_{t-1}^j| > |\lambda_t^j|$ then the cost of period t homogenization is negative!

[26] The problem is finding a Condorcet winner for the choice of government location after negative homogenization has been implemented in period 1. To see this, observe that $j = 1/2 - \epsilon$, where $\epsilon > 0$ is small enough, beats $j = 1/2$ in a pairwise vote and $j = 1/2 - \epsilon'$ beats $j = 1/2 - \epsilon$, where $\epsilon > \epsilon' > 0$. This occurs because homogenization does not persist when the location of the government changes.

[27] Assuming $v \in (0, 1)$ is not necessary for the results but simplifies the algebra.

Thus a democracy will choose homogenization $\lambda_2^{1/2} = \lambda_{1/2}^m + \lambda_1^j$. If a democracy splits there is no change to previous results. Similarly, using analogous arguments to those in the proof of Lemma 1 and Proposition 2, the analogous results continue to hold but with $\lambda^* \in [-1, 1)$ and $\bar{\lambda} \in [-1, 1)$. Next, determine the ruler's choice of $\lambda_1^{1/2} \in [-1, 1]$. When $\bar{\lambda} \leq 0$ the ruler chooses $\lambda_1^{1/2} = 0$ for all p. When $\bar{\lambda} > 0$ and $\lambda^* \geq 0$, then optimal $\lambda_1^{1/2} < \min\{\lambda^*, \bar{\lambda}\}$ maximizes

$$[g+y-k-C(\lambda_1^{1/2})]+v(p,\lambda_1^{1/2})[g-(1-\lambda_{1/4}^m)ga/4+y-2k-C(\lambda_{1/4}^m)]+(1-v(p,\lambda_1^{1/2}))[g+y-k]. \tag{16}$$

The derivative of (16) with respect to $\lambda_1^{1/2}$ is

$$-C'(\lambda_1^{1/2}) - \frac{\partial v(p,\lambda_1^{1/2})}{\partial \lambda_1^{1/2}}[(1-\lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)] \tag{17}$$

Expression (17) is negative for all $\lambda_1^{1/2} \geq 0$, positive as $\lambda_1^{1/2} \to -1$, and decreasing in $\lambda_1^{1/2}$ since $\frac{\partial^2 v(p,\lambda_1^{1/2})}{\partial \lambda_1^{1/2}}^2 \geq 0$. Thus there is a unique optimal value of $\lambda_1^{1/2} < \min\{\lambda^*, \bar{\lambda}\}$ which is negative. Expression (16) is decreasing in p at rate

$$-\frac{\partial v(p,\lambda_1^{1/2})}{\partial p}[(1-\lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)].$$

For $\lambda_1^{1/2} \geq \min\{\lambda^*, \bar{\lambda}\}$, λ^* and $\bar{\lambda}$ dominate all other levels of homogenization in this range. Expected utility from $\lambda_1^{1/2} = \lambda^*$ is

$$[g+y-k-C(\lambda^*)]+v(p,\lambda^*)[g+y+k-[C(\lambda_{1/2}^m)-C(\lambda^*)]]+(1-v(p,\lambda^*))[g+y-k] \tag{18}$$

which is decreasing in p at rate

$$-\frac{\partial v(p,\lambda^*)}{\partial p}[C(\lambda_{1/2}^m) - C(\lambda^*)]. \tag{19}$$

If, for some $p = \bar{p}$, λ^* gives higher utility than the value of $\lambda_1^{1/2}$ that maximizes (16), then this is true for all p higher if

$$-\frac{\partial v(p,\lambda_1^{1/2})}{\partial p}[(1-\lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)] < -\frac{\partial v(p,\lambda^*)}{\partial p}[C(\lambda_{1/2}^m) - C(\lambda^*)], \tag{20}$$

for all $p \geq \bar{p}$. Then (20) is satisfied for all $p \geq \bar{p}$ if $\frac{\partial v(p,\lambda^*)}{\partial p}$ is not too large relative to $\frac{\partial v(p,\lambda_1^{1/2})}{\partial p}$, evaluated at the value of $\lambda_1^{1/2}$ that maximizes (16), for all $p \geq \bar{p}$. A sufficient condition for this is that $\frac{\partial v(p,\lambda_1^{1/2})}{\partial p}$ is not increasing too fast in $\lambda_1^{1/2}$ for all $p \geq \bar{p}$. For $\lambda_1^{1/2} = \bar{\lambda}$, expected utility is

$$[g + y - k - C(\bar{\lambda})] + [g + y - k]$$

The result then follows as in the proof of Proposition 2. When $\bar{\lambda} > 0$ and $\lambda^* < 0$, then, by above, the ruler chooses either $\lambda_1^{1/2} \leq 0$ or $\bar{\lambda}$. If for any p, $\bar{\lambda}$ gives strictly higher utility than any $\lambda_1^{1/2}$, then this is true for all p higher. $\square$

Allowing for negative homogenization does not change the prior results.

From the proof of Proposition 3, for a democracy, preferences over homogenization $\lambda_t^j \in [-1, 1]$ remain single peaked and Lemma 1 extends to negative homogenization. To show that Proposition 1 continues to hold and a ruler will never choose negative homogenization, we show that a ruler always does strictly better by choosing zero homogenization than negative homogenization. Then since his total expected utility from any $\lambda_1^{1/2} \geq 0$ does not change, Proposition 1 does not change. Suppose in period 1 the ruler forms a single country with $j = 1/2$ and undertakes homogenization $\lambda_1^{1/2}$. A period 2 ruler will continue to form a single country with zero homogenization. If $\lambda^* \leq 0$, the ruler's expected utility from $\lambda_1^{1/2} = 0$ is

$$[g + y - k] + p[g + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k]. \tag{21}$$

His expected utility from $0 > \lambda_{j,1} \geq \lambda^*$ is

$$[g + y - k - C(\lambda_{j,1})] + p[g + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k]. \tag{22}$$

Expression (21) is strictly lower than (22) (and so is any $\lambda_{j,1} < \lambda^* \leq 0$). If $\lambda^* > 0$, then from Proposition 3, any $\lambda_1^{1/2} < 0$ results in a split so expected utility for $\lambda_1^{1/2} = 0$ is

$$[g + y - k] + p[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k].$$

and expected utility for any $\lambda_1^{1/2} < 0$ is

$$[g + y - k - C(\lambda_{j,1})] + p[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k].$$

Thus $\lambda_1^{1/2} = 0$ always gives strictly higher utility than $\lambda_1^{1/2} < 0$. □

Part D

Different Homogenization Technologies

We now allow for two different homogenization technologies. The technology studied thus far spreads costs evenly across the population. Such a technology can be considered as a permanent transfer from the center (which benefits from its closeness to the government) to the periphery (which suffers from its distance). A second technology, which we term an unequal-cost technology, spreads the costs differently. It implies a distribution of costs that fall more heavily on those who are further away from the ruling government. Unequal-cost technologies can capture an unequal tax burden, but can also be interpreted more broadly, as methods of homogenization which result in higher personal costs for more distant minorities. For example, repression of cultures that are different from the leading one would fall into the category of an unequal technology. Allowing for greater flexibility over choice of homogenization policies strengthens our main results.

We capture the difference between technologies through the cost of homogenization. What we now refer to as an equal-cost technology is modeled previously. The cost to individual i of homogenization to level λ_t^j using an unequal-cost technology is $M(\lambda_t^j, d_{ij}) -$

$M(\lambda_{t-1}^j, d_{ij})$, where $M(\lambda_t^j, d_{ij})$ is strictly increasing, strictly convex and twice continuously differentiable as λ_t^j increases from 0 to 1, with $M(0, d_{ij}) = 0$, $M_{\lambda_t^j}(0, d_{ij}) = 0$ and $\lim_{\lambda_t^j \to 1} M_{\lambda_t^j}(\lambda_t^j, d_{ij}) = \infty$. In contrast to equal-cost technologies, $M(\lambda_t^j, d_{ij})$, is linearly increasing in d_{ij}, the distance of the individual from the government; that is, the cost of homogenization is higher for those who are homogenized by more. We also assume the marginal cost of homogenization, $M_{\lambda_t^j}(\lambda_t^j, d_{ij})$, is increasing in distance from the government.[28]

To make comparisons between these two technologies, we assume that the total cost of homogenizing a country to λ_t^j (assuming $\lambda_{t-1}^j = 0$) is the same under both technologies. That is,

$$\int_{i \in country} C(\lambda_t^j) di = \int_{i \in country} M(\lambda_t^j, d_{ij}) di,$$

when the government is located in the center of the country. Clearly this may not hold, but it is useful for comparisons. The framework is exactly the same as Section 5 (with an exogenous probability of democracy, p), but allows whoever is in power the choice between the two technologies. The homogenization technology is chosen, followed by the amount of homogenization, after borders and governments have been determined.

Proposition 4 *When both an unequal and equal-cost technology are available, in period 1 a ruler strictly prefers to homogenize using the unequal-cost technology, while in period 2 a democracy weakly prefers to homogenize using the equal-cost technology.*

Proof of Proposition 4

Without loss of generality write $M(\lambda_t^j, d_{ij}) = \beta(\lambda_t^j) + \alpha(\lambda_t^j) d_{ij}$. Since $M_{\lambda_t^j}(\lambda_t^j, d_{ij})$ is increasing in d_{ij}, then $\alpha'(\lambda_t^j) > 0 \ \forall \lambda_t^j > 0$. Since $M(\lambda_t^j, 0) = \beta(\lambda_t^j)$ is increasing in λ_t^j then $\beta'(\lambda_t^j) > 0 \ \forall \lambda_t^j > 0$. The expression that equates the marginal cost and marginal benefit of unequal homogenization for i is

$$gad_{ij} = \beta'(\lambda_t^j) + \alpha'(\lambda_t^j) d_{ij}. \tag{23}$$

By the above, λ_t^j that satisfies (23) is increasing in d_{ij}. Preferences are single peaked over λ_t^j.

Since total costs of homogenization under different technologies are equalized for $j = 1/2$

$$2 \int_0^{0.5} \left[\beta(\lambda_t^j) + \alpha(\lambda_t^j) x \right] dx = C(\lambda_t^j), \qquad \forall \lambda_t^j \in [0,1]; \qquad \text{and hence}$$

$$\beta(\lambda_t^j) + \alpha(\lambda_t^j)/4 = C(\lambda_t^j), \qquad \beta'(\lambda_t^j) + \alpha'(\lambda_t^j)/4 = C'(\lambda_t^j), \quad \forall \lambda_t^j \in [0,1]. \tag{24}$$

From (24), for $d_{ij} = 1/4$ the levels of homogenization that satisfy (23) and $gad_{ij} = C'(\lambda_t^j)$ are equal, for $d_{ij} < 1/4$ the level of homogenization that satisfies (23) is strictly higher, and for $d_{ij} > 1/4$ it is strictly lower than the level of homogenization that satisfies $gad_{ij} = C'(\lambda_t^j)$.

[28] Thus for each individual i in a population of size s we can write i's tax burden as $r_t = k/s + [M(\lambda_t^j, d_{ij}) - M(\lambda_{t-1}^j, d_{ij})]$. As mentioned, we can also consider these idiosyncratic costs as personal costs rather than increased taxes directly.

Suppose a democracy forms a single country with the government at some $j \in [0,1]$ and $\lambda_1^j = 0$. For $j \in [1/4, 3/4]$, the median voter over homogenization is $d_{ij} = 1/4$, thus the levels of homogenization chosen by majority rule under unequal and equal-cost technologies are the same. For $j \in [0, 1/4)$, the median voter over homogenization is $i = 1/2$ with $d_{ij} = 1/2 - j > 1/4$, thus unequal homogenization chosen by majority rule will be lower than equal-cost homogenization. Each individual evaluates the difference between their utility in the case of equal-cost homogenization and their utility in the case of unequal homogenization,

$$[g - (1 - \lambda_{eq})gad_{ij} + y - k - C(\lambda_{eq})] - [g - (1 - \lambda_{uneq})gad_{ij} + y - k - M(\lambda_{uneq}, d_{ij})], \quad (25)$$

where λ_{uneq} (respectively λ_{eq}) is the level of unequal (respectively equal-cost) homogenization chosen by majority rule. For $j \in [1/4, 3/4]$, $\lambda_{uneq} = \lambda_{eq}$, expression (25) is increasing in d_{ij}, the median voter when deciding between unequal and equal-cost homogenization is at $d_{ij} = 1/4$, and he is indifferent between the two technologies. For $j \in [0, 1/4)$, $\lambda_{uneq} < \lambda_{eq}$, expression (25) is increasing in d_{ij} and the median valuation of (25) is also $d_{ij} = (1/2 - j)$. It follows that the median voter $d_{ij} = (1/2 - j)$ must prefer equal-cost homogenization since for any level of unequal homogenization he can homogenize to the same level instead using equal-cost technologies and do strictly better.

The same argument applies for any $\lambda_1^j \in (0, 1]$ and similarly for Country A and B.□

The intuition is simple: unequal-cost technologies cost less to the dictator. The burden of homogenization shifts towards the rest of the population, at an increasing rate the more distant individuals are from the dictator himself. For a government located at the center, a democracy is indifferent between the two homogenization technologies.[29] The result that a democracy is indifferent relies on both the linearity in distance of the cost function for the unequal-cost technology and on the fact that the population is distributed uniformly. Allowing for any type of distribution of costs, as well as any distribution of the population, would make the problem intractable. Our modeling device is meant to capture the fact that, in general, a dictator has more latitude in the allocation of costs, while a democracy must consider (to a greater degree, at least) the views of the whole population and this may place limits on what technologies are chosen.[30] The ruler will choose homogenization technologies that place the costs on others, while a democracy will tend to choose technologies where the costs are more equally spread. Thus we assume that, when indifferent, the equal-cost technology is chosen.

Proposition 5 extends Proposition 1 and Corollary 1 to the case where both an unequal and equal-cost technology are available. The qualitative results remain, but the means of homogenization undertaken by the ruler is harsher and the level of homogenization is higher.

Proposition 5 *In period* 1 *the ruler uses the unequal-cost technology to undertake an amount of homogenization which is (weakly) increasing in p.*

[29] Proposition 4 states that a democracy weakly prefers homogenization via the equal-cost technology. This is because we solve Proposition 4 for any government location. For a government not located at the center, a democracy sometimes strictly prefers the equal-cost technology.

[30] A limit on what the dictator can do in terms of allocation of cost is also related to the possibility of unilateral secession of regions, or insurgencies of specific groups.

(i). There exists a threshold $\tilde{p} \in (0, 1]$ such that

(a) when $p \leq \tilde{p}$, the period 1 ruler chooses a strictly lower level of homogenization than will be chosen by the period 2 democracy (should democracy prevail);

(b) when $p > \tilde{p}$, the period 1 ruler chooses a higher level of homogenization than will be chosen by the period 2 democracy (should democracy prevail);

where $\tilde{p}$ depends on g, a, k, $C(\cdot)$ and $M(\cdot, \cdot)$.

(ii). The amount of homogenization undertaken by the ruler is weakly higher, and for some parameters strictly higher, than when only the equal-cost technology is available.

Proof of Proposition 5

The proof is a repeat of previous arguments. □

Analogous to Proposition 1 and Corollary 1, homogenization undertaken by the period 1 ruler is increasing in p. For p high enough, homogenization undertaken by the period 1 ruler will be higher than that which would be undertaken by a democracy. Proposition 5 highlights that greater latitude in homogenization technologies (compared to Proposition 1) will induce the ruler to homogenize more because he can use technologies which place the costs on minorities. It increases the set of parameters under which a ruler homogenizes more than a democracy. This becomes obvious when we think about technologies where minorities face almost all the costs and the dictator almost none (this can be captured within our definition of an unequal technology). In this case, when the probability of overthrow is high, the dictator will always homogenize to ensure a large state that represents his preferences. The same is not true of a democracy. This would be an extreme case of the model, but it illustrates the point.

Unequal-cost technologies place high costs on those further from the center, lowering the level of homogenization that is desired by those at the periphery. This further highlights the slightly subtle observation that, while distant minorities may have the most to gain from reducing the cost of distance to the government, they may not necessarily welcome homogenization from the center and instead may prefer a split.

We should point out one further incentive to homogenize once we allow for different technologies: by homogenizing those at the periphery by unequal means, this avoids having to give more expensive transfers in the form of equal-cost homogenization should democracy prevail.

Part E

Additional Tables

Table A1
Descriptive statistics

Variable	Obs.	Mean	St. dev.	Variable	Obs.	Mean	St. dev.
Educational reform: 10% threshol	977	0.350	0.477	Legal reform (Flora, 1983)	110	0.464	0.501
Educational reform: 20% threshol	977	0.255	0.436	Autocracy	110	0.038	0.164
Autocracy	977	0.514	0.461	Government crises	110	0.441	0.622
Government crises	977	0.224	0.358	Revolutions	110	0.049	0.214
Revolutions	977	0.214	0.353	All internal conflicts	110	0.839	1.631
All internal conflicts	977	1.080	1.451	Autocracy*government crises	110	0.025	0.195
Autocracy*government crises	977	0.077	0.202	Autocracy*revolutions	110	0.004	0.033
Autocracy*revolutions	977	0.127	0.277	Autocracy*all internal conflict	110	0.079	0.540
Autocracy*all internal conflicts	977	0.534	1.050	Population growth	110	0.054	0.039
Autocracy*international war	977	0.087	0.271				
Population growth	977	0.187	0.152				
International war	977	0.179	0.384				
Log(revenue)	781	10.026	1.952				
Log(GDP per capita)	781	7.922	1.008				

Table A2
Nation building and threats to democracy, robustness to additional controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Educational reform: 10% threshold					
Autocracy*gov. crises	**0.187****			**0.205****		
	(0.089)			**(0.099)**		
Autocracy*revolutions		**0.180****			**0.189****	
		(0.078)			**(0.080)**	
Autocracy*all internal conflicts			**0.060*****			**0.069*****
			(0.022)			**(0.022)**
Autocracy*international war	-0.166*	-0.162*	-0.168*	-0.165	-0.150	-0.166
	(0.088)	(0.088)	(0.089)	(0.105)	(0.103)	(0.102)
Autocracy	-0.065	-0.053	-0.087	-0.066	-0.049	-0.093
	(0.059)	(0.058)	(0.061)	(0.065)	(0.066)	(0.071)
Pop. growth	0.314	0.295	0.301	0.259	0.246	0.243
	(0.215)	(0.220)	(0.216)	(0.225)	(0.235)	(0.229)
International war	0.031	0.019	0.022	0.024	0.015	0.015
	(0.063)	(0.065)	(0.065)	(0.074)	(0.075)	(0.074)
Gov. crises	-0.159**			-0.143**		
	(0.061)			(0.063)		
Revolutions		-0.016			-0.047	
		(0.065)			(0.064)	
All internal conflicts			-0.026			-0.025
			(0.017)			(0.017)
Log(revenue)				0.078*	0.079*	0.082*
				(0.045)	(0.044)	(0.043)
log(GDP per capita)				0.035	0.047	0.041
				(0.074)	(0.074)	(0.073)
Number of countries	172	172	172	144	144	144
Observations	977	977	977	781	781	781
R-squared	0.457	0.457	0.456	0.500	0.499	0.502

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 years country averages, for the 1925-2004 period. Educational reforms is a dummy if primary per capita school enrollment increased more than 10% from the previous 10 years. All the dependent variables are lagged. Standard errors are clustered at the country level.

Table A3
Nation building and threats to democracy, robustness to additional controls

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Legal reform, based on Flora (1983)					
Autocracy*gov. crises	0.660***			0.804***		
	(0.216)			(0.237)		
Autocracy*revolutions		3.285**			4.602***	
		(1.316)			(1.277)	
Autocracy*all internal conflicts			0.230**			0.293**
			(0.085)			(0.099)
Autocracy*international war	-0.958**	-1.055**	-0.907**	-1.111***	-1.395***	-1.105***
	(0.355)	(0.486)	(0.381)	(0.334)	(0.393)	(0.329)
Autocracy	-0.041	0.209	-0.049	-0.150	0.181	-0.142
	(0.361)	(0.316)	(0.377)	(0.356)	(0.301)	(0.367)
Pop. growth	-0.277	-0.037	-0.243	-0.427	-0.161	-0.379
	(1.055)	(0.974)	(1.012)	(1.357)	(1.380)	(1.384)
International war	0.045	0.032	0.043	0.003	-0.009	-0.000
	(0.149)	(0.147)	(0.146)	(0.158)	(0.157)	(0.156)
Gov. crises	0.021			0.016		
	(0.073)			(0.081)		
Revolutions		0.247			0.327	
		(0.170)			(0.184)	
All internal conflicts			0.023			0.024
			(0.040)			(0.043)
Log(revenue)				-0.003	0.043	0.009
				(0.140)	(0.139)	(0.134)
log(GDP per capita)				-0.079	-0.109	-0.067
				(0.246)	(0.197)	(0.237)
Number of countries	14	14	14	13	13	13
Observations	110	110	110	104	104	104
R-squared	0.658	0.661	0.661	0.653	0.662	0.656

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are country averages for 10 years, for the period from 1925-2004. Legal reform is a dummy if the country experienced at least a legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All the dependent variables are lagged. Standard errors are clustered at the country level.