

NBER WORKING PAPER SERIES

NATION BUILDING

Alberto Alesina
Bryony Reich

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

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

We thank Tim Besley, Martin Cripps, Jeffrey Frieden, Oded Galor, Terri Kneeland, Mark Koyama, Alessandro Riboni, Enrico Spolaore and participants in seminars at Brown, Cambridge, U of Chicago, UCL, Warwick, a CEPR meeting, an ISNIE meeting, a conference in Berkeley and the NBER Summer Institute for useful comments. Giulia Giupponi and Andrea Passalacqua 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.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Alberto Alesina and Bryony Reich. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Nation Building
Alberto Alesina and Bryony Reich
NBER Working Paper No. 18839
February 2013, Revised April 2015
JEL No. F3

ABSTRACT

Nations stay together when citizens share enough values and preferences and can communicate with each other. Homogeneity amongst people can be built with education, teaching a common language, building infrastructure for easier travel, but also by brute force such as prohibiting local cultures or even genocide. Democracies and dictatorships have different incentives when it comes to choosing how much and by what means to homogenize the population. We study and compare both regimes, and the transition from dictatorship to democracy, in a model where the size of countries and the degree of active homogenization is endogenous. We offer some historical discussions of several episodes which illustrate our theoretical results.

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

Bryony Reich
Department of Economics
University College London
Drayton House
30 Gordon Street
London
WC1H 0AX
b.reich@ucl.ac.uk

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, 1805¹

In 1860 French was still a foreign language to half of all French children.² Outside major cities, France was a country of different languages, dialects and diverse currencies.³ Travel far outside one’s own village was rare, and indifference or hostility to the French state common.⁴ From the French Revolution and throughout the 19th century, French rulers expressed the imperative “to form French citizens”.⁵ Following the unification of Italy (1860), a process led by a Northern elite which then ruled the country, Massimo d’Azeglio (one of the founders of unified Italy) famously remarked: “Italy has been made; now it remains to make Italians.” In 1860 at most 10% of the Italian population spoke what would become the Italian language, there was only one railway line which crossed any of the pre-unification states, and many were openly hostile to the new nation.⁶ 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.” They introduced state controlled education, including compulsory elementary schooling; banned languages other than the “national language” in schools, religious services and administration; introduced compulsory military service often with the explicit aim of integrating and mixing individuals from different parts of the country; and extended road and rail links.

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.” Hobsbawm (1990) notes, “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

¹Quote from Ramirez and Boli (1987).

²Estimate Weber (1979) p67. Hobsbawm (1990) p60 gives a figure of 12 – 13% of the population who spoke French at the French Revolution.

³Weber (1979) in just a few case studies mentions Basque, Béarnais, Catalan, Flemish, Germanic dialects, dialects of Boulongne, Artois, Picardy, and so on. On currency see Weber (1979), p30 – 40.

⁴Weber (1979), p95 – 114; 485 – 496. It is also argued that knowledge of the nation of France itself was not always guaranteed. In 1864 a school inspector in Lozère noted that not a single child could answer questions such as “Are you English or Russian?”, p110. On travel, p195 – 220. Note that 50% of France’s population were estimated to be farmers or peasants in 1870, p8.

⁵Quote from Félix Pécault in 1871 who conducted a general inspection of public education for the French government. See Weber (1979) for many more examples.

⁶Duggan (2007). The railway line was the Piacenza-Bologna line, Schram (1997).

education and other administrative mechanisms.” In contrast, European elites did not enact such policies in their colonies (Michalopoulos and Papaioannou, 2012). Yet once these colonies gained independence in the 1950’s and after, many introduced policies to create a national language and national identity, similar to those of 19th century Europe (Miguel, 2004).⁷ The 20th century also saw dictators and political elites who built homogeneity by prohibiting local cultures and attempting to impose their ideologies, often by odious means, for example the Soviet Union, Nazi Germany, Mao’s China, or Franco’s Spain.⁸ Nation-building continues to remain relevant in the 21st century; in China, a range of nation-building policies are being implemented in peripheral regions which have large minority groups.⁹

Why did 19th century European elites see homogenization as imperative? Why not in their colonies? Why did those colonies undertake nation-building after independence? Why did the Soviet Union and other modern dictatorships undertake harsh methods to impose homogenization? Do these experiences have implications for the long-run heterogeneity and stability of a country?

The goal of this paper is to analyze nation-building in its more or less benevolent forms, across political regimes and in times of transition from various forms of dictatorship to democracy. 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 so that they do not wish to separate from each other.¹⁰ We model a heterogeneous population which may choose to break-up, as in Alesina and Spolaore (1997).¹¹ The equilibrium size of a country emerges from a trade-off between economies of scale in the production of public goods and services or the size of the market and the heterogeneity of the population, which may have different priorities and preferences for shared public goods, languages or institutions. We depart from this, however, in an important way, since we assume that the degree of divergence of preferences amongst the population is endogenous: we explicitly model the choice of the central government regarding how much to homogenize the population.¹²

When and why would a particular regime undertake such homogenization? Let us consider a fully secure ruler (or ruling elite). The ruler simply extracts rents from his territories. He does not care about nation-building since he has the type of government and the location of

⁷Miguel (2004) provides a fascinating comparison between nation-building policies in post-colonial Tanzania and Kenya, with evidence suggestive of a strong effect of Tanzania’s nation-building policies.

⁸For example, Franco declared his aim to create “a single language, Castilian, and a single personality, the Spanish one”, Jones (1976).

⁹In 2014, financial incentives were introduced to encourage inter-ethnic marriage in an area with a large Uighur population (a minority group in China which is largely Muslim and speaks a Turkic language). Similar policies on inter-ethnic marriage exist in Tibet. The same year saw arrests of Uighur intellectuals on charges of “inciting separatism” and restrictions on Uighur dress. In 2014 the Chinese President also proposed tightening state control over religion, improving bilingual education and employment for minorities and encouraging minority group members to move to other parts of China. This is similar to previous policies which encouraged members of the Han majority to migrate to peripheral areas dominated by minority groups. From E. Wong *China Moves to Calm Restive Xinjiang Region*, 30 May 2014, and *To Temper Unrest in Western China Officials Offer Money for Inter-marriage*, 2 September 2014, retrieved from <http://www.nytimes.com/>.

¹⁰Recently, state-building and nation-building have sometimes been used interchangeably. However, state-building generally refers to the construction of state institutions for a functioning state, while nation-building the construction of a national identity, also for a functioning state.

¹¹See Alesina and Spolaore (2003) for a review of the economic literature on country size.

¹²Alesina and Spolaore (2003) in their discussion mention this avenue of possible research but they do not develop it.

the capital that match his preferences.

The incentives of a non-democratic regime when 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, who would remain as part of the population of the new democracy. In addition, a democratic vote may break up the population into more than one country. In general, democratic rule will not produce the most preferred policies of the ruling group.¹³ The threat of democratization motivates the ruler to homogenize for two distinct reasons. First, homogenization and indoctrination, sometimes by brutal means, allow those in charge to better maintain the status quo (their preferred policies and a larger country) even if democracy prevails. Second, 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. In our model, the most extreme episodes of homogenization will be undertaken by non-democratic regimes under threat of democracy.¹⁴

When installed, a democratic regime may break apart the territories of the dictator (i.e. the former Soviet Union) or may stay together (the European Nation states which became democratic in the nineteenth century, e.g. France, England). Even democracies, however, benefit from “benevolent” forms of nation-building. The majority benefits from a certain degree of homogenization, for example, better roads or railways to the capital city improve the individual’s access to resources located there and may avoid distant minorities becoming isolated and disenfranchised; schooling in a common language enables better participation in the democratic process; indoctrination in common values reduces heterogeneity of preferences so that policies and public goods are a better fit.¹⁵ However, since homogenization (schooling, roads, etc.) is costly, the majority chooses to homogenize up to the point at which marginal benefits equal marginal costs. In some cases, the median voter might choose a level of homogenization which avoids an otherwise sure split of the country; a population that would otherwise split, may stay together with a technology of homogenization, for example, road building or learning a common language.¹⁶

We also study another interesting case, in which more homogenization may actually increase

¹³Hobsbawm (1990) writes 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.”

¹⁴We focus on internal factors which motivate governments to implement nation-building policies. Specifically we find that the internal threat of democracy induces non-democratic governments to enact very high levels of nation-building and by harsh means. We do not explore external motives for nation-building, namely the threat of external war (see Aghion, Persson and Rouzet, 2014), but in Section 6 we provide a brief discussion of theories of nation-building, including external wars, and we describe in more detail where our model fits.

¹⁵For instance Michalopoulos and Papaioannou (2012) provide evidence of how national rule, institutions and policies in African countries do not reach isolated ethnicities far from the capital. These ethnicities revert to ethnic based rules, making the country unstable.

¹⁶One could also think of “private” forms of homogenization. For instance a linguistic minority setting up its own private schools to learn the dominant language, or isolated communities building private roads to be more connected to the rest of the country. We concentrate on homogenization by governments and leave this point for future research.

the probability of insurrection. A more homogenous population may communicate better and develop common goals, which may increase the likelihood of coordination in an insurrection attempt. This effect works against the other incentives of the ruler to homogenize. It is a type of “divide and rule” effect. In this case, and only in this case, a ruler may choose to increase heterogeneity in the population. We argue that colonizers, rulers who face a low probability of overthrow, and rulers with limited state capacity, are more likely to implement policies that increase the heterogeneity of the population. As a result, the failure of states which were former colonies may indeed be the result of “divide and rule” policies implemented by former colonizers.

As we discuss in the final section of the paper, our results imply non obvious and “non linear” comparisons between certain public policies in democracies and non-democracies, an insight broadly consistent with Aghion, Persson and Rouzet (2014) and Mulligan, Gil and Sala-i-Martin (2004). Safe dictators homogenize less than democracies, unsafe rulers more than democracies, and by harsher methods. Our results also imply that the type of regime that governed in the past has implications for the long-run heterogeneity of that population, and therefore its likelihood of fragmenting into multiple states. A country ruled by a domestic elite which undergoes a smooth transition to democracy may become more homogenous than a similar population that is first ruled by a colonizer before becoming democratic. Empirical research in economics documents that greater heterogeneity is associated with largely worse outcomes in areas such as growth, public goods and conflict.¹⁷ Potentially then, the history of nation-building within a country can affect the future success of that population. In fact, a wealth of historical literature examines accounts of nation-building as part of the formation of successful states.¹⁸ Our paper also relates to a literature in economics on “state capacity,” as in Besley and Persson (2009, 2010), which examines the development of state institutions in the formation of successful states. The literature on state capacity emphasizes the role of wars as an engine for building the ability of the state to raise taxes and establish law and order. Nation-building, in terms of homogenization of the population, is something more than (and different to) state capacity. In the final section of the paper we analyze the role of wars and democratization as complements in the formation of the modern “state capable nation.”

Our approach relates to work on democratic transition, 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 (2006) argue that democratic transition motivates elites to invest in institutions which allow them to maintain a higher degree of power under democracy and mitigate their economic losses from democratic transition. Besley, Persson and Reynal-Querol (2014) present evidence suggesting that when threatened with loss of power, rulers invest in institutional reform, namely improving executive constraints, to constrain the ability of future regimes to act against their interests. Our model suggests that forward-looking elites also invest heavily in building nations (compulsory schooling, nationalism, national languages and military service) when threatened with democracy.

¹⁷See Alesina and La Ferrara (2005) and Alesina et al. (2013) for further references.

¹⁸See Smith (1998) for a detailed description of and key references in the development of the study of nationalism and Laitin (2007) for a discussion of nationalism, homogenization and state formation.

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 Bourgignon and Verdier (2000). Papers by Gradstein and Justman (2002) and Ortega and Tangelas (2008) examine schooling as a means to improve communication across groups and so increase growth. Our results are particularly related to an argument that proposes that the expected extension of the franchise motivated European elites to introduce mass compulsory schooling, despite its unpopularity with the masses.¹⁹

This paper is organized as follows. Section 2 describes the basic model of nation-building and transition. Section 3 solves the basic model to examine nation-building under a non-democratic regime, determine the effect of a threat of transition, and compare this with a democracy. Section 4 extends the model to allow democratic transition to be endogenous to the nation-building policies of the ruler. In this section we determine what happens when a “divide and rule” effect is present and how this is relevant to rulers who may exit the country should democracy prevail (e.g. colonizers). Section 5 extends our framework to allow for a choice over different homogenization technologies. We determine which technologies will be used by which types of regimes and how this affects the extent of nation-building across different regimes. Section 6 discusses historical examples and the last section concludes.

2 Basic Model Setup

We consider a two period model of nation-building. In the first period, a country is governed by a ruler (dictator or elite, terms used interchangeably). In the second period, that country may become democratic or remain governed by the ruler. In the first part of the paper the probability of democratization is taken as exogenous, below we endogenize it.

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 any time t , an individual, i , resides in a country with a single government, j , that serves the citizens of the country. Individual i 's per period utility function at time t is given by

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

The first term 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 from the government when distance is zero, where d_{ij} is the preference distance of individual i from government j . We think of distance as the geographical distance, language difference, or difference in preferences 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 second term is income y , which is exogenously given, identical for everyone, and identical

¹⁹See Green (1990).

across time periods. The third term is taxes in period t , r_t , which are split equally amongst the population of the country.

We model “homogenization” as a technology which uses state apparatus 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

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

So that, for a country with government j , and for any individual in that country, i , the cost to individual i of facing policies and public goods different to his ideal is reduced by fraction λ_t^j . We refer to this as homogenization of the population. Any homogenization policy, λ_t^j , is applied to the whole population within the country governed by j .²⁰

The simplest way of thinking about homogenization is building roads (or railroads or airports) in order to reduce the costs of distance from the capital. This facilitates access to resources or government services offered in the capital, reducing economic isolation. The second interpretation is one of communication in terms of 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 (literally, reducing the distance between languages) so that individuals can better communicate with the government and access public services.²¹ Neither of these two interpretations of homogenization imply a change in individuals’ preferences; especially if alternative languages and dialects are not prohibited by force. A third interpretation implies changing individual preferences by indoctrination (by more or less “kind” means). That is, convincing individuals far from the type of government chosen that they do not dislike it that much. For instance, one may argue that in schools, say in France or Scandinavia, the benefits of regulation and social welfare are emphasized while in the US and the UK the merits of individualism are stressed more.²² A benevolent interpretation of this “indoctrination” is one that views the latter as a help for individuals to fit in better with accepted social norms. But of course there exist much more malevolent forms of indoctrination. In communist countries indoctrination in schools of Marxist-Leninist ideas was common and other ideologies forbidden.²³ The same applies to fascist regimes or theocracies. Changing preferences can also involve severe repression or elimination of groups with particular preferences (political or otherwise).

One can choose the preferred interpretation of homogenization. In order to maintain all three together, one needs to make the assumption that geographic location, language and preferences are perfectly correlated. We should also assume language or preferences are perfectly correlated with geography to allow for a split of the population (described below). From now

²⁰Homogenization λ_t^j , applied to the whole population, has a greater impact on the utility of those further from the government (i.e. with higher d_{ij}). This captures the intuition that building roads affects those further from the government the most and teaching the language of the government has a bigger impact on those who speak a different language.

²¹In France, in the nineteenth century, knowledge of French was also important for accessing coveted government jobs (Weber, 1979).

²²See Alesina and Glaeser (2004) for a discussion of these cultural differences. See also Aspachs-Bracons et al. (2008) for a study of the effect of compulsory Catalan language education on identity.

²³For instance, Alesina and Fuchs-Schündeln (2007) present evidence of a large amount of indoctrination in East Germany.

on with the term “distance” we summarize either one of the three interpretations above (or a combination of the three) and with the term “homogenization,” a reduction in such distance. Note that, instead of homogenizing to reduce the costs of diversity, diverse countries could transfer resources to citizens further away in geography and preferences from the government. We do not explore this issue here but note that once homogenization occurs it may last forever (say having a common language), while transfers may need to be paid every period and so may not be credible or, in the long run, they may be more expensive for the center (i.e. those closer to the central government).

Homogenization is durable: roads built today remain tomorrow, languages learnt today are not forgotten tomorrow, preferences influenced today by the government influence future preferences.²⁴ To model this, we assume the cost of the homogenization policy λ_t^j , for a country of some 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.²⁵ 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 $\lambda_t^j \geq \lambda_{t-1}^j$.²⁶

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 $\lim_{\lambda_t^j \rightarrow 1} C''(\lambda_t^j) = \infty$.*

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 Section 5.²⁷

In our model income is exogenous. However, at least up to a point, diversity of skills, education, background, and culture may increase productivity.²⁸ 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.

²⁴Alesina and Fuchs-Schündeln (2007) present evidence of differences in preferences of East Germans even after German unification.

²⁵Observe 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 $\lambda_{t-1}^j = 0$. Intuitively, the effort of a government to teach the population one language is redundant if the next government imposes a different national language. In some cases homogenization enacted by one government may actually make homogenization by a different government more costly.

²⁶This assumption implies “reverse homogenization” policies are not available. We relax this later.

²⁷We will allow for a choice over different homogenization technologies. In particular, we introduce a technology in which the costs are unequally split and fall more heavily on those further away from the decision maker. This is one way to capture homogenization technologies that entail harsh personal costs, such as repression of minorities. Allowing for greater choice over homogenization technologies further strengthens our main results.

²⁸On this point see Alesina, Spolaore and Wacziarg (2000) and Alesina, Harnoss and Rappoport (2013).

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 which case the population either forms a 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).²⁹ A single government is 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.³⁰ Note that only a democracy in period 2 would have an incentive to separate, thus the results would not change were we to give the ruler the option to split the country in either period.

The cost of “government” (public goods and policies) in period t in a given country is k .³¹ Since the costs k can be divided amongst all citizens in the country this captures the benefits of forming a single country rather than breaking into two.³² However, 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 geography) to the median individual in that country. This set-up captures the motivation for the break-up of the population: 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 (Alesina and Spolaore, 1997). To summarize, the government budget constraint at time t for a country of mass s is

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

Decision-Making and Timing

We model an initial period in which a single ruler or elite governs, followed by a second period in which democracy may prevail. Period 1 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 (1).

²⁹In a model of endogenous country formation, the interval $[0, 1]$ could be divided into any number of countries of varying sizes (see Alesina and Spolaore, 1997). Our assumption of a maximum of two countries of equal size is made for simplicity. In fact, 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. See Alesina and Spolaore (2003) for a discussion of this case in a model without endogenous homogenization.

³⁰Adding a fixed cost for these changes would add notation with not much additional insight.

³¹Obviously the assumption of a fixed cost is extreme and adopted for simplicity of notation. It could be easily generalized to the case of $k = \alpha + s$ where s is the size of the country and α a fixed cost.

³²Alesina, 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.

Period 1

The population is controlled by a single ruler who is located at the center of the population, with the government at his ideal point.³³ The ruler makes decisions in period 1 to maximize his expected utility. He has to decide how much to invest in homogenization.

Period 2

With probability $1 - p$, the ruler remains in power in period 2. The ruler continues to make decisions as above. With probability p , democracy prevails in period 2. Under a democracy, decisions are made by the population by majority rule, with the timing as follows:

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

Decisions (2) and (3) are made by majority rule by the population of each country. To break any ties, we assume (with no loss of generality) when indifferent between one country or two, a single country is always formed. The order of decision making is realistic since a “government” cannot be chosen before borders are set, and only an established government can choose public policies regarding homogenization. Nevertheless, the qualitative results remain even with a different ordering of decisions.³⁴

3 Homogenization Decisions

We solve the model backwards. We determine homogenization choices by a ruler and a democracy and compare the two.

3.1 A Democracy

If democracy prevails in period 2, the population chooses whether to form single country or split, where to locate the government, and how much to homogenize. This problem is solved in detail Appendix A. We here describe the choices a democracy makes, starting with homogenization.

For individual i the following level of homogenization equates the marginal cost and marginal benefit,

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

³³In the working paper version of this article we analyze explicitly the case of a ruler located anywhere between 0 and 1. In the appendix of the present paper we briefly discuss this more general case. 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.

³⁴For example, the argument driving our results holds if we suppose a population first decides how much to homogenize, then whether to split, and finally the location of the government. Results on this are available from the authors. This example highlights the contrived nature of alternative orderings since with this ordering a decision is made on how much homogenization to undertake without choosing a government location; in other words without choosing *which* language to teach everyone or *to where* to build roads.

The latter depends on the distance of individual i from the government and the former on the cost of the homogenization technology. For now, we assume a technology that benefits those furthest out the most, while sharing the cost equally among the population. Below we generalize this assumption. For example, building roads to the capital, where the cost is shared equally, benefits those who live farther from the capital the most. Thus individuals who are further from the government prefer more homogenization.³⁵ For a given country and government, since preferences over homogenization are single peaked, the level of homogenization chosen by majority rule will be the median preferred homogenization within that country. 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 intuition for homogenization under a democracy is immediate if we interpret homogenization in terms of roads, infrastructure, or public schools teaching a common language. The “preference” interpretation of homogenization, literally speaking, implies that an individual “chooses” a policy that changes his preferences, knowing that after the change he would feel happier in the country in which he 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.³⁶ Strong attachment to cultural values can be captured by very high costs of homogenization.

Choice of government under a democracy is as follows: a democracy will always locate the government at the median ideal point in the population, namely the center of any country. 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.

The choice of a democracy of whether to form a single country or split is analogous to Alesina and Spoloare (1997) and captures the trade-off between the benefits of a larger country and the costs of heterogeneity. Voters located near the center of the population prefer to form a single country since the government well represents their preferences and, *ceteris paribus*, a larger country is better. The preferences of voters located towards $1/4$ and $3/4$ will be better represented after a split and they may be willing to pay the additional cost of forming two countries. *Ceteris Paribus*, the higher is k , the cost of government, the lower is a , the *ex-ante* cost of distance, and the lower is g , the value of government, the more the median voter will prefer a single country. 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 captures

³⁵This is true for a fixed country and a fixed government. The same individuals who prefer high homogenization in a single country may also prefer secession.

³⁶For models related to parents “choosing” values for children see Alesina et al. (2013) and Bisin and Verdier (2000). Algan et al. (2012) discuss the cost of lack of assimilation of Arabs in France and their effort to do so. They document a substantial increase in salaries for children of families which signal assimilation by choosing French rather than Arab first names.

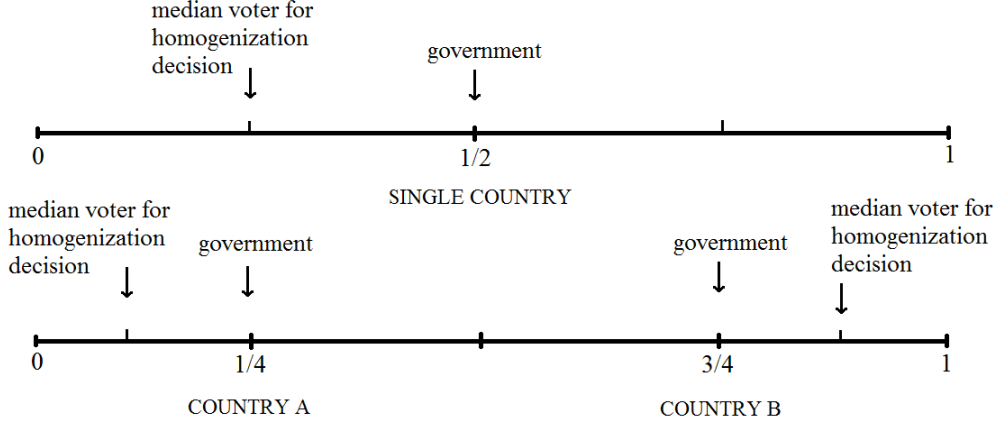


Figure 1: Homogenization and government location under a democracy, for a single country and a split.

the idea of “nation-building”. Nation-building represents a particular kind of transfer from the center of the population towards the periphery to reduce the costs of being located towards the periphery and therefore avoid separation.³⁷

A ruler in period 1 can affect the decisions made by a democracy. We have the following:

Proposition 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. The intuition is as follows. Homogenization by government j reduces the cost of distance from government j . Thus, for all citizens, homogenization by the ruler increases utility from the ideal government of the ruler relative to other government locations. Building communication infrastructures to the capital city reduces the cost of distance and access to that capital. Investing in teaching a national language to the population

³⁷Intuitively, who tends to prefer a single country versus two? Voters near the center of the population, 1/2, prefer to form a single country relative to splitting into two. They would also vote for low homogenization since they do not need it as much as people further away. As we move from 1/2 towards 1/4 and 3/4 on the unit interval, voters begin to place greater value on splitting into two countries since their preferences are better represented in Countries A and B. In a model without any homogenization, voters at the extremes, say with ideal points lower than 1/4 and greater than 3/4 would all prefer two countries to one. However this is not necessarily the case with endogenous homogenization. If, due to the nature of the cost function, homogenization under a single country is substantially greater than homogenization under a split, then it is possible that some voters at the extreme (close to 0 or 1) may prefer a single country with very high homogenization to two countries with a relatively low homogenization. In other words minorities may sometimes prefer to be in a large very homogenized country than in two countries where they would still be far from the center and not very homogenized.

increases the benefit of public goods, services and government provided in that language. Indoctrinating to Marxist-Leninist doctrine increases utility from policies of a communist government. In reducing the cost of distance, homogenization by the ruler also reduces the costs of heterogeneity under his ideal government and so makes secession less attractive.

The threshold, λ^* , takes into account the homogenization choices of a democracy in period 2. If $\lambda^* = 0$, this implies a democracy would form a single country in period 2 and locate the government at the ruler's ideal point, even with no homogenization by the ruler in period 1. Note that, by locating the ruler in the center of the population, the decisions not to split and to locate the government at the ruler's location are perfectly aligned in a democracy. We discuss this further below.

3.2 A dictatorship

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. He does not need to homogenize since the government perfectly matches his preferences and he has no incentive to increase the welfare of the population he rules.

We are interested in the ruler's policy choices in period 1, since period 1 captures the behavior of a forward-looking ruler. A more homogeneous population is of no direct benefit to the ruler in period 1; however, if democracy prevails then homogenization by the ruler in period 1 can ensure the ruler's ideal government persists and avoid the break-up of the country in period 2. We are ready to state the decisions of the ruler in period 1.

Proposition 2 *In period 1 the ruler undertakes a level of homogenization which is (weakly) increasing in the probability of democracy, p .*

If a democracy splits, then the ruler suffers for two reasons: the government under a democracy is no longer the preferred government of the ruler and the taxes he has to pay are higher. Thus the ruler “nation-builds” with a particular agenda: he homogenizes to build a large nation that better reflects his preferences. Homogenization is costly, but it will improve the utility of the ruler should democracy prevail. 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, his ideal government persists and the country avoids secession.

Does a non-democratic ruler homogenize more or less than a democracy? The following corollary compares homogenization under a ruler who faces a threat of democracy $p \in [0, 1]$, and homogenization under a democracy.

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 those at the periphery. Thus a democracy undertakes more nation-building than a safe non-democratic regime. In contrast, when the probability of democracy is high, under certain parameters an “unsafe” ruler will undertake a strictly 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.³⁸

A democracy invests in homogenization up to the optimal point for the median voter. A ruler has stronger preferences for his ideal government compared to the median preferences in a democracy towards any particular government. Thus an unstable non-democratic regime (i.e with a high chance of democratization) may over-invest in homogenization compared to a democracy to ensure the regime’s ideal government is preserved even under democracy. In contrast, under a stable non-democratic regime, a ruler 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 homogenous enough to form a single country, whereas the other is not. The implication is that, as a result of nation-building by elites, today’s democracies which followed a smooth path to democracy, where elites foresaw the advent of democracy, may be more homogenous and bigger than they would be otherwise. In the next section we analyse in more detail how incentives to homogenize vary across different “types” of non-democratic regime and how this can impact long-run heterogeneity of a population.³⁹

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

³⁸Note Proposition 2 (ii) does not say ‘strictly higher.’ This is because total homogenization under a democracy in period 2 always incorporates homogenization undertaken in period 1. In part (ii), total homogenization in period 2 is a result of homogenization undertaken by the ruler in period 1.

³⁹There 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.

large country. In the working paper version of the present paper we analyze in detail the case of a ruler located anywhere on $[0, 1]$; we also provide a brief discussion in the Appendix. 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, then a higher probability of democracy will induce the ruler to homogenize more. A very high probability of democracy may induce him to homogenize by more than would ever be chosen under a democracy. The extent to which any ruler homogenizes more or less than a democracy depends on his location but also the costs of homogenization. In Section 5, we show that relaxing the constraints such that there is a greater choice over homogenization technologies (for example, the government can force more of the costs on minorities), strengthens the ruler's incentives to undertake high levels of nation-building relative to anything that would be seen under a democracy.

4 Endogenous Democratization

Homogenization is relevant not only in affecting the outcome if democracy prevails, but may also be relevant in the probability of democratic transition itself. By building infrastructure a ruler can improve conditions and lessen the isolation of minority groups at the borders, thus reducing opposition to the current regime. Through schooling, non-democratic governments can indoctrinate the next generation to their own ideology. At the extreme end of homogenization, governments can reduce dissent through repression or elimination of particular individuals and groups. Here we model the possibility that homogenization can reduce opposition to the ruler's regime. Note that the promise of future redistribution would not be credible. This point is made in Acemoglu and Robinson (2000) in a model of democratic transition, and by Alesina and Spolaore (1997) in a model of border stability.

Suppose, as above, a revolution opportunity arrives at the beginning of period 2 with probability p . In Section 3, 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 $U_{i2,ruler}$.⁴⁰ Individual i prefers overthrow if

$$U_{i2,dem} - U_{i2,ruler} - L \geq 0. \quad (2)$$

The population chooses to overthrow if a majority prefer overthrow. Note that the median value of (2), a measure of opposition to the ruler's regime, is decreasing in homogenization by the ruler. Proposition 3 describes the choices of a ruler.

⁴⁰Given the choices in period 1, the values of $U_{i2,dem}$ and $U_{i2,ruler}$ are known at the beginning of period 2.

Proposition 3 *In period 1, the ruler undertakes homogenization which is (weakly) increasing in p .*

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 for a proof. Now there is a second incentive to homogenize: to indoctrinate people to be happy with the government and so reduce the threat of democracy. Proposition 3 says that if the ruler homogenizes enough in the initial period he can avoid overthrow. As in Proposition 2, 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 homogenous 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, however, 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, nineteenth century European elites). The motive to reduce the threat of democracy, shown in Proposition 3, applies to all kinds of non-democratic regimes, even harsh dictators who may be kicked out or eliminated should democracy prevail. More on these points below.⁴¹

4.1 Divide and Rule

In some cases revolutions become more likely when a population is homogeneous. A more homogenous population can communicate better and this may make collective action easier. By the same argument, policies that increase diversity and its costs could hinder collective action. After all, the principle of “divide and rule” is meant to capture precisely this effect.⁴² This case can also be analyzed with our model.

To capture the notion of “dividing” the population, 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. We need to update the assumptions on costs to allow for negative homogenization. First, assume

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

⁴¹Note that, for simplicity, we are not allowing the ruler to extract rents from the population when in office. Rents increase his losses if democracy prevails, therefore increasing his incentives to homogenize to reduce the probability of overthrow. However, rents also increase the gain to the population if democracy prevails, and this raises the cost to the ruler of avoiding overthrow.

⁴²Similarly, if homogenization involves education of the population, a more educated population could also increase the probability of successful overthrow.

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 $\lambda_t^j > \lambda_{t-1}^j$.⁴³ 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.⁴⁴

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 $v : [0, 1] \times [-1, 1] \rightarrow (0, 1)$ 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.⁴⁵

Otherwise the framework is exactly as detailed so far. The model 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 4 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, an assumption discussed in more detail in the proof in the Appendix, we obtain Proposition 4.

Proposition 4 *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

⁴³If 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!

⁴⁴The problem is finding a Condorcet winner in 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.

⁴⁵Assuming $v \in (0, 1)$ is not necessary for the results but simplifies the algebra.

may now choose to “divide and rule.” He may undertake strictly negative homogenization. The proof is in the Appendix. In the Online Appendix we also show that the previous results do not change if we allow for negative homogenization. Only when homogenization increases the probability of collective action does the ruler have an incentive to increase the costs of diversity.

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, 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.⁴⁶ 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 .⁴⁷

As above, if the ruler can extract rents, this can exacerbate his incentive to divide and rule. We conclude this section with two comments.

Comment 1: State capacity

An important consideration in choosing between implementing nation-building policies or divide and rule policies is the role of state capacity. Homogenization may require high state capacity, for example, the implementation of compulsory education for all children requires state infrastructure. A ruler with low state capacity may be limited or unable to homogenize. In contrast, divide and rule policies may require different resources, less state capacity, and are likely to be easier and cheaper to implement.⁴⁸ Consider two places with different levels of state capacity in period 1 but otherwise identical populations, and suppose democracy prevails in both places in period 2.⁴⁹ The ruler with low state capacity will choose to divide and rule up to a higher threshold on p than the ruler with high state capacity. That is, the country that started off with lower state capacity may have had negative homogenization implemented by the ruler and may end up more heterogeneous and perhaps even break up, compared to the country which started with higher state capacity which may have been homogenized by the ruler. We refer the reader to Besley and Persson (2010) for an in-depth model of state capacity and note that nation-building could also be related to state capacity such that higher state capacity results in more nation-building which then makes building state capacity easier still.

⁴⁶For some parameters, because we limit assumptions on the function $v(\cdot, \cdot)$, with an analogous intuition to Proposition 2, 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 $\lambda_1^{1/2} \geq \min\{\lambda^*, \bar{\lambda}\}$.

⁴⁷This occurs when $\frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2} \partial p} > 0$ and $\lambda^* = 0$, for example.

⁴⁸In the model we assume symmetric costs to avoid additional notation, but it is straightforward to assume a lower cost for negative versus positive homogenization.

⁴⁹We can model a ruler with low state capacity as facing a positive shock to his costs of (positive) homogenization in period 1. This assumption implies that implementing policies, such as mass education or road building, is more difficult and costly the less state infrastructure there is in place.

Comment 2: Colonizers

Colonizers are different from the domestic dictatorship or elite analyzed thus far. Colonizers leave the country after their regime falls. Consider then the three motives to homogenize captured by the model. Since the colonizer does not care about what happens to the country after he leaves, he would not pay the costs of homogenizing the population to maintain the status quo should democracy prevail. Thus the only motives possibly relevant to the colonizer include homogenization to reduce opposition to the colonizers regime, and negative homogenization to reduce the ability of the population to act collectively to overthrow the colonizer. Homogenization to significantly reduce the population's willingness to overthrow the colonizer might be very expensive (especially given the above discussion on state capacity). In the case of colonized territories with fragmented populations, the policy of divide and rule may be especially attractive, since it may be that divide and rule is relatively inexpensive and makes the transition away from colonization less likely to occur. As a result, ethnic conflict and division within countries may be exacerbated after decolonization.

5 Odious Homogenization

We now allow for two different technologies. One technology spreads the costs evenly across the population. This captures what we term “non-odious” homogenization. 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 “odious” homogenization, spreads the costs differently. Odious homogenization implies a distribution of costs that fall more heavily on those who are further away from the ruling government. The repression of cultures that are different from the leading one would fall into the category of odious homogenization. Allowing for greater flexibility over choice of homogenization policies further strengthens our main results.

We capture the difference between technologies through the cost of homogenization. What we term non-odious homogenization is modeled previously. The cost to individual i of odious homogenization to level λ_t^j 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 $\lim_{\lambda_t^j \rightarrow 1} M_{\lambda_t^j}(\lambda_t^j, d_{ij}) = \infty$. In contrast to non-odious homogenization, $M(\lambda_t^j, d_{ij})$, 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. That is, the cost of any additional amount of homogenization is higher for those who are homogenized by more.⁵⁰

To make comparisons between odious and non-odious homogenization, 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

⁵⁰Thus for each individual i in a population of size s we can write i 's tax burden as $r_t = k/s + [M(\lambda_t^j, d_{ij}) - M(\lambda_{t-1}^j, d_{ij})]$. Of course we can consider these idiosyncratic costs as personal costs rather than increased taxes directly.

technologies. That is,

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

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 3 (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 5 *When both odious and non-odious technologies are available, in period 1 a ruler strictly prefers odious homogenization, while in period 2 a democracy weakly prefers non-odious homogenization.*

See the Online Appendix for a proof. The intuition is simple: odious homogenization costs 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.⁵¹ The result that a democracy is indifferent relies on both the linearity in distance of the odious cost function 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.⁵² The ruler will thus 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, non-odious homogenization is chosen.

Proposition 6 extends Proposition 2 and Corollary 1 to the case where both an odious and non-odious 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 6 *In period 1 the ruler undertakes an amount of odious homogenization which is (weakly) increasing in p .*

(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;*

⁵¹Proposition 5 states that ‘a democracy weakly prefers non-odious homogenization.’ This is because we solve Proposition 5 for any government location. For a government not located at the center, a democracy sometimes strictly prefers non-odious homogenization.

⁵²A limit on what the dictator can do in terms of allocation of cost is related to the possibility of unilateral secession of regions, or insurgencies of specific groups. This extension is left for future research.

(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 non-odious homogenization is available.

See the Online Appendix for a proof. Analogous to Proposition 2 and Corollary 1, homogenization undertaken by the period 1 ruler is increasing in p . For p high enough, under some parameters, homogenization undertaken by the period 1 ruler will be higher than that which would be undertaken by a democracy. Proposition 6 highlights that greater latitude in homogenization technologies (compared to Proposition 2) 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 odious 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. Of course, this is an extreme case of the model, and such technologies likely do not exist, but it illustrates the point.

There is, in fact, a further incentive to homogenize once we allow for different technologies: by homogenizing those at the periphery by odious means, this avoids having to give more expensive transfers in the form of non-odious homogenization should democracy prevail.

6 Historical Examples

In this section we discuss historical examples of nation-building. We focus on several predictions of our model and examine suggestive empirical evidence in light of our results. Our discussion centers around one particular (and important) nation-building policy: primary education. We start with the prediction that *nation-building is increasing in the threat of democracy*.

6.1 Eighteenth and Nineteenth Century Europe.

In the West, nation-building policies were implemented in force during the 19th and early 20th Centuries. This involved large-scale intrusion into the lives of citizens in the form of compulsory primary education for all children, compulsory military service for all young men, and the forced introduction of national languages, amongst other things. The introduction of mass education provides a particularly interesting example. 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 huge 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.⁵³ 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 present three detailed case studies, covering France, Italy and England, and then more systematically examine 19th Century education reforms across 11 European countries.

France

The French Revolution in 1792 is a turning point in European history. 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. 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.

Our model predicts little, if any, nation-building in Europe before the French Revolution, and intensifying nation-building throughout the 19th century as democracy is increasingly viewed as inevitable. While the Ancien Régime is well known for having implemented a highly centralized state, there was little homogenization of the wider population before the French Revolution.⁵⁴ Hobsbawm (1990) estimates that only 12-13% 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.⁵⁵ The French Crown showed “little concern with the linguistic conquest of the regions under its administration.”⁵⁶ In fact, the ruling elites made a point of distinguishing themselves from the masses, using language as a barrier (Gellner, 1983). There was also little interest in increasing geographic communications in France. Roads were just a means of collecting taxes and transporting troops and areas outside of major cities were often isolated. Primary schooling was predominantly provided by the church and was not a public function.⁵⁷

Historian Eugen Weber (1979) writes that “Diversity had not bothered earlier centuries very

⁵³For 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.

⁵⁴Tilly (1980).

⁵⁵The Ordinance of Villers-Cotterêts, made law in 1539, was designed to end the use of Latin in official documents and replace it with French.

⁵⁶Weber (1979).

⁵⁷Katznelson and Weir (1985).

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. Immediately after the French Revolution, the Constitution of 1791 called for the establishment of free public instruction for all.⁵⁸ 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.”⁵⁹ The Jacobins insisted “The unity of the Republic demands the unity of speech.”⁶⁰ 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 a more demanding 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.⁶¹ Weber (1979) describes 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 unassimilated 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.”

Policies of homogenization were, in part, 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.”⁶² The example of southern France is also illuminating. Historian Joseph Strayer describes the (apparently successful) efforts of the state in homogenization writing “Languedoc was very like Catalonia

⁵⁸Ramirez and Boli (1987).

⁵⁹Weber (1979).

⁶⁰Weber (1979).

⁶¹Other 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.

⁶²Report by the rector of the Academy of Rennes, Weber (1979) p100, 313. 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.”

and very unlike north France, yet it finally became thoroughly French.”⁶³

Italy

Italian unification was completed by Northern elites in the 1860s, with virtually no involvement of local populations. Italy, once unified, constituted 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.⁶⁴ The governing elite considered homogenization vital to ensure the internal stability of what was, in reality, an extremely diverse set of states. Duggan (2007) documents, “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.⁶⁵ 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.”⁶⁶ Enough education to homogenize, but not too much to create rebellious masses.

In Italy, active homogenizing policies also included large investment in railroads. Apart from their role as infrastructure, railroads had the political goal of unifying the country, especially connecting the North with the underdeveloped South. The Minister of public works was viewed as the man who was building Italy as a nation state (Schram, 1997). The Italian military was also a force for unifying the population. Conscripts were purposefully sent to regions away from home and regiments formed of soldiers from diverse parts of the country.⁶⁷ As well as unifying the population, military service aimed to mould “Italians” in the shape of those who established the new state. Giuseppe Guerzoni, a friend of Garibaldi, explained at a conference in 1879 that “having made Italy the army is making Italians.”⁶⁸ Nicola Marselli expressed in 1871 “I know, too, that Italy has been reunited for only ten years and is not

⁶³Tilly (1975) p43.

⁶⁴The largest proportion of adult males were enfranchised in Italy in 1912.

⁶⁵Duggan (2007), p289.

⁶⁶Duggan (2007), p280.

⁶⁷Duggan (2007), p288. This policy continued well into the 20th century.

⁶⁸Duggan (2007), p283.

yet established [...] I have always said that even if it had no other purpose, the army would always be a great school of Italian-ness.”⁶⁹

England

Historian Linda Colley 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 the threat of democracy was very high.

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 1860’s the “full democratization of the political realm was seen as inevitable”.⁷⁰ 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 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”⁷¹ 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.”⁷²

Education and Nation Building in Nineteenth Century Europe

We now take a more systematic look at education reforms in eleven European countries between 1800 and 1875.⁷³ Data on reforms are from Flora (1983) and include introductions and extensions of compulsory education as well as major events and laws, for example, bringing education under state control, major introductions and changes in types of school and curric-

⁶⁹Duggan (2007), p274.

⁷⁰Ramirez and Boli (1987).

⁷¹Quote from Marcham (1973). As above, the act enfranchised a part of the male urban working-class population.

⁷²Quote from Young and Handcock (1964).

⁷³The countries are Austria, Belgium, Denmark, France, Germany, Italy, Norway, Netherlands, Sweden, Switzerland, United Kingdom.

ula.⁷⁴ A binary variable, *education reform*, takes the value 1 if any major educational reform takes place in that country and year. Our discrete variable, *number of education reforms*, sums *education reform* across all countries in a given year.

To examine Prediction 1, we need a measure of the perceived threat of democracy. This is undoubtedly difficult, nevertheless there is an obvious possibility. The French Revolution and recurring major uprisings in France that followed during the 19th century had a significant impact both on France and on other European countries in two ways. First, these uprisings scared ruling elites with the prospect that populations could, and were willing, to overthrow the existing order. Second, they arguably sparked uprisings in other countries.⁷⁵ Our variable, $riot(t, t - 1)$, takes a value of 1 if a major uprising occurred in a given year in France, or in the previous year. The variable is constructed from years of major uprisings compiled by Tilly et al. (1980) between 1830 and 1875.⁷⁶ The dates recorded refer to years of major episodes of collective violence involving a large number of people engaging “in seizing and damaging persons or property” across a range of locations.⁷⁷ For example, the year 1831 involves the continuation of smaller disturbances from the July Revolution of 1830, multiple violent demonstrations in a number of large cities including Paris, and a silk workers’ insurrection in Lyon.

Figure 1 shows that education reforms across Europe were largely concentrated in periods of insurrections in France, a proxy for democratic threats in Europe in this period. We argue that this positive relationship between riots and education reform exists because threatened elites imposed mass education on their populations to serve their own interests. 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. Ramirez and Boli (1987) document that in Sweden, around 1810, “bourgeois liberals led a movement to develop mass schooling to provide national unity and purpose,” but it was primarily resistance by the peasantry that slowed the adoption of state-controlled education until 1842. 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 burnings of 1830 and 1831?” In England, 526 parishes responded. The

⁷⁴We do not include those relating only to university education since they are considered irrelevant to mass homogenization.

⁷⁵Examples from a range of countries include the following. A period of reform swept Sweden in the 1830s inspired by the uprisings in France (Ramirez and Boli, 1987). Reactionary politics “swept Austria in the aftermath of the French Revolution” (Ramirez and Boli, 1987). The two major concentrations of violence in Germany in the 19th Century “followed closely upon increased turbulence in neighboring France” Tilly et al. (1980) p209, p247. In England, there was a feeling that events in other European countries could impact unrest in England (Holland, 2005). Hobsbawm and Rudé (1969) *pxxiv*, 62 – 64, regarding the English riots in the early 1830s, “it is doubtful whether it would have occurred on so vast a scale when it did, without the...French and Belgian revolutions abroad.”

⁷⁶Years of revolution: July Revolution 1830; February Revolution 1948; revolution 1870 – 1871. Years of major uprisings which are not considered to be revolutions: 1831, 1832, 1834, 1839, 1840, 1841, 1846, 1847, 1849, 1850, 1851, 1869.

⁷⁷These dates correspond to other data in Tilly et al. (1980) describing different measures of collective violence including number of violent events, participants in collective violence and arrests in collective violence across France.

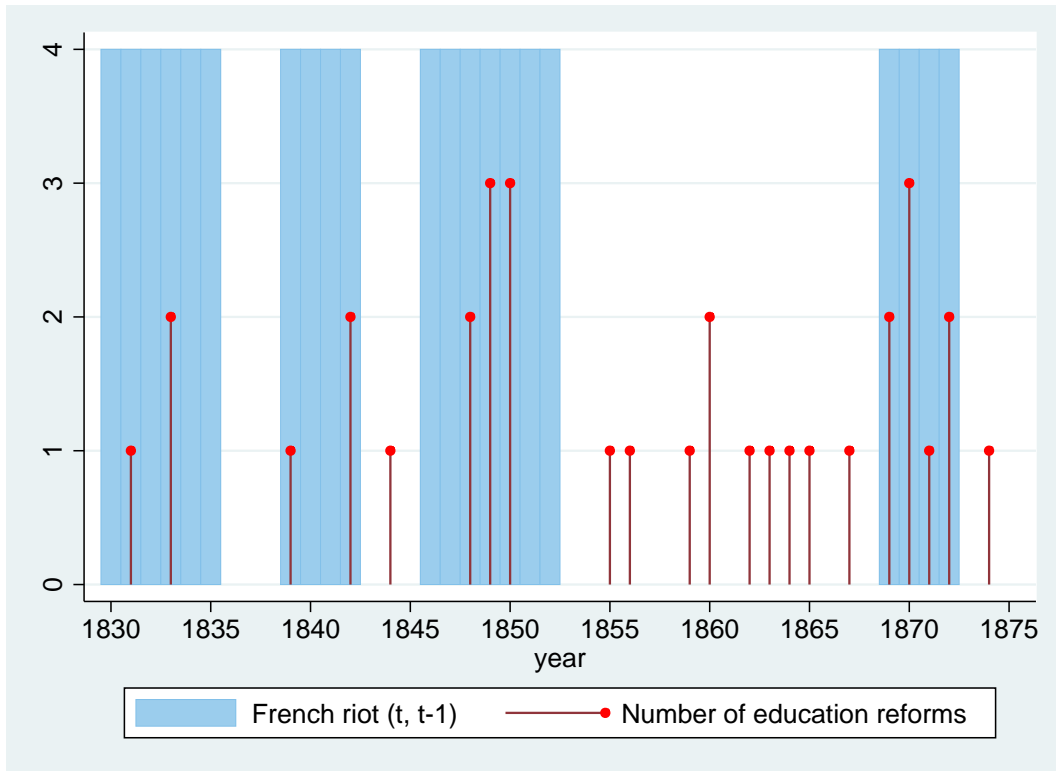


Figure 2: Years of major rioting in France and number of education reforms across Europe

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, Charles Tilly’s detailed study of episodes of collective disturbances in France 1830–1860 provides information on the objective of the group involved in the disturbance.⁷⁸ Education is not mentioned. This is consistent with evidence from modern day Brazil: Bursztyn (2012) shows that the poor prefer cash transfers to subsidies for education and that their assessment of the government is negatively affected when they perceive government funding for public education to have increased but cash transfers to have decreased.

A second way to examine this argument is to observe that if education in the 19th century was provided with a nation-building motive, then we should expect differences in the implementation of education policies compared to clearly redistributive policies, such as social security or health care. Especially since redistributive concerns were closer to population demands. Indeed, there are stark differences in the timing of education reform and redistributive

⁷⁸Tilly (1998).

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).⁷⁹

6.2 “Odious” Nation Building by Twentieth Century Dictators

Intense nation-building was not just a feature of 19th century Europe. Well-known 20th century examples include Germany under the Nazis, the Soviet Union under Stalin, and Spain under Franco. Efforts in the Soviet Union to “Russify” and make loyal a huge population are well documented. One particularly extreme homogenization policy highlights the link between territorial stability and nation-building. Conquest (1970) describes the deportation of eight entire ethnic groups (including the Crimean Tatars, Volga Germans and Chechens) in the Soviet Union in the 1940s. They were exiled to Siberia and Central Asia, their own names removed from the list of Soviet peoples, the names of their original habitats changed and these original habitats repopulated with others more “loyal” to the Soviet regime. In some cases these groups had attempted autonomy and were considered a threat to unity. Other cases were somewhat precautionary, but with a similar motive. The population (which may have been considered to have Turkish sympathies) of a block of land over which Turkey and Russia had fought for more than a century was deported, arguably as a precaution to avoid future trouble in this area with Turkey.

Public education was central to homogenization efforts in the Soviet Union, in particular in forcing a Marxist-Leninist doctrine. A study by Lott (1999) suggests an “over-supply” of education under the Soviet Union: after democratization in former communist countries, public funding for education as a percentage of GDP went down while, in contrast, health spending increased by 70%. Education was also an important tool in Spain as part of Franco’s drive to create a nation “with a single language, Castilian, and a single personality, the Spanish one.” Catalonia, which was seen as a threat to Castilian political rule as well as to the unity of the country, was subject to harsh homogenization measures.⁸⁰ Education measures included “Purifying committees” that sacked teachers involved in Catalanism, the expulsion of teachers thought to be associated with Catalan nationalist parties to posts in distant parts of Spain, and the replacement of teachers in Catalonia with 700 teachers brought in from other parts of Spain and hired mainly for their lack of knowledge of Catalan.⁸¹

⁷⁹Our historical observations appear consistent with the historical discussion in Acemoglu and Robinson (2000) on the extension of the franchise. They suggest that in a number of cases redistributive concessions were not credible and franchise extension was required by the elite to avoid costly overthrow. Welfare reform would then follow franchise extension. They cite Germany as an exception. Indeed Germany is responsible for the earliest welfare reforms.

⁸⁰From Jones (1976). Even before Franco, the economic prosperity and cultural differences of Catalonia had been seen as a threat both to Castilian political rule in Spain as well as to the unity of the country. The demand for greater autonomy of these regions featured heavily in the non-democratic regimes in the first half of the 20th Century. The Spanish dictator Primo de Rivera is reported to have remarked in 1925 : “Regions? Out of the question. A quarter of a century’s silence about regions...and Spain will have been freed from one of her gravest perils.”

⁸¹Apart from education, other policies included huge numbers of political executions, incentivizing Castilian through jobs, and putting a Castilian speaking bishop in charge of the church in Barcelona who was given the task of eliminating Catalan from the

6.3 Cross Country Evidence

Our model implies non-linear comparisons between certain public goods in democracies and non-democracies, in particular we find that *nation-building may be higher in a threatened non-democratic regime relative to a democracy*. A number of studies compare education policies across democratic and non-democratic regimes and we here summarize recent empirical evidence. Aghion, Persson and Rouzet (2014), using annual data on 137 countries from 1830 – 2001, find that autocracies have higher enrolment rates in primary education than democracies. Consistent with this, 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, Burstyn (2012) finds that democracies spend less on public education than non-democracies for below median income countries. This receives further support from a study by Lott (1999) who 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) also examines health care expenditure in two separate data sets and finds either no effect of totalitarianism or a negative effect.

It would of course be interesting to compare the content of education under a democratic regime or a dictatorship. Under the latter we would expect education to have a much higher content of indoctrination, for instance only studying Marxist economics in communist dictatorships, not teaching minority languages or the history of ethnic or religious minorities.

6.4 Divide and Rule

Our model implies that *policies of divide and rule are more likely to be implemented in non-democratic regimes that are ruled by a colonizer*. Colonizers have no interest in homogenizing and building a national identity since they are there simply to extract rents and know, should insurrections prevail, they will leave. Indeed, colonizers of Africa did not make any effort to build cohesive nation states (see Easterly and Levine, 1997; Herbst, 2000; Alesina, Easterly and Matuszeski, 2010; and Michaolopoulos and Papaioannou, 2012, amongst others). Building a national identity could even be counter-productive to a colonizer if it serves to increase support for nationalist movements and independence. In fact, our model predicts that policies designed to fracture a population may be optimal for a colonizer in order to maintain power. Active policies of “divide and rule” were used by colonizers, for example by the British in India (Christopher, 1988). As a result of the lack of nation-building and the implementation of divide and rule, on independence ex-colonial populations may be particularly fractured and lack national unity. Gennaioli and Rainer (2008) show that the lack of nation-building in many countries in Africa had long lasting effects after decolonization, with reversal to tribal based institutions. Our model also predicts that after decolonization, whether as a democracy or “unsafe” non-democratic regime, newly independent countries would enact nation-building

churches in his diocese.

policies. Indeed, after decolonization in Africa and Asia, many leaders of the newly independent countries attempted homogenization policies to unify their populations, more or less successfully.⁸²

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.⁸³ 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 most newly 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.⁸⁴

Another example is India. The British had done little, if anything, to homogenize a diverse population, even using specific policies of divide and rule (Christopher, 1988). On independence, India formed a democracy. A democratic India was unable to force homogenization in the same way as non-democratic regimes. Hobsbawm (1990) claims that the multitude of languages spoken in India made the creation of a single national language impossible since many were unwilling to accept the disadvantage of having a national language that was not their mother tongue. Hindi was the most widely spoken language in India and was Ghandi’s choice for a national language, however, those advocating Hindi as a national language were unable to impose it on the population as a whole.⁸⁵ The Indian National Congress was also “committed to a single united subcontinent” but had to accept its partition into different states.

An important prediction of our model is that the path of rule of a country can have long-run implications for heterogeneity of the population and even its likelihood of fragmenting into multiple states. Countries that moved more gradually through revolutions and other petitions towards modern democracy (arguably Europe) may be larger and more homogenous today than countries which faced a different path. In particular, countries that moved straight from colonization to democracy may be less homogenous. Several fragile states in Africa are an example.

6.5 Alternative Theories of Nation Building

Our theory and other theories of nation-building are not mutually exclusive. One non-mutually

⁸²Smith (2003) and Miguel (2004).

⁸³Marten and Kula (2008).

⁸⁴Marten and Kula (2008) claim that the decision to make English the common language “was seen as the only ‘non-tribal’ alternative available to serve as a vehicle of national unity, an argument often made in post-colonial African language policies.” Of course English was often spoken by the urban, educated and wealthy, so it does not necessarily represent a choice independent of interest groups.

⁸⁵English became the “median of national communication,” while also maintaining state level languages, as it gave no single language group a particular advantage.

exclusive variable is war. Foreign conflicts are considered a major driver of state-building (Tilly, 1975, and Besley and Persson, 2010). However this argument may apply more directly to the building of “state capacity,” in terms of raising taxes to support armies and establish law and order. Wars which motivated ruling elites to invest in state-building did not, at the same time, motivate policies of nation-building. The period of state-building (meaning state capacity to raise taxes) in Europe begins roughly in 1500. Over the following three centuries, European states invested in state-building (see Gennaioli and Voth, forthcoming). In contrast, nation-building policies, including, as we discussed above, education, language, infrastructure for domestic communication, inclusion (by more or less generous means) of minorities, only begin to occur after the French Revolution, once there was a major threat to old aristocracies throughout Europe.

Aghion et al. (2014) provide more insight into the role of war in nation-building. For a panel of countries from 1830-2001, they 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. For the mean value of democracy in their sample, in their baseline regressions the effect of a threat of war on education enrollment is either negative or very small and positive.⁸⁶ That is, for significantly non-democratic countries, the effect of a threat of war is, at best, insignificant. War is relevant to nation-building in democratic countries, but it does not do as well at explaining nation-building in non-democratic regimes. This result makes perfect sense. A brutal 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.

Yet major nation-building occurred, and occurs, in non-democratic regimes, when threatened by fear of uprisings. Right from the beginning of the 19th century, states invested in nation-building; indeed this is the most documented historical period of nation-building. Yet, throughout much of this century, European countries have very low democracy scores in the index used by Aghion et al. (2014). As we discussed above, the threat of war seems to be especially relevant for democracies. This suggests that the threat of war does not fully explain this major period of nation-building in 19th century Europe, and neither does it explain intense episodes of nation-building under other non-democratic regimes. In summary, we see our argument about nation-building for fear of democratization and splitting of countries, and state-building for fear of aggression, as very complementary and not as alternatives.

A second 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.⁸⁷ In contrast, an industrial society based upon markets (as opposed to a stratified agrarian society with local markets) needs better means

⁸⁶From Table 1 columns (2) and (3) in Aghion et al. (2014), using the mean democracy score in the sample and the mean value of the size of the rival army. The measure of democracy they use is a scale from -10 to 10 given by the *polity2* variable from the Polity IV database. The result that in non democratic countries education enrollment is negatively effected by war may be a result of assuming a linear relationship in the regression. Probably for non-democracies the effect of war is zero, and war only has a positive effect on education for significantly democratic countries.

⁸⁷Gellner (1983) writes that political units in agrarian societies “can be divided into two species: local self governing communities and large empires”. Neither of these type of governments represent a modern nation state.

of communication. Universal schooling serves an economic purpose as well, necessary for the development of an industrial society.⁸⁸ In other words, productivity would increase in an industrial society with more homogenization relative to an agrarian one.⁸⁹

The timing of this theory is questionable. Smith (2003) 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 the 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. Industrialization may provide a better explanation for bottom-up homogenization (for example increased homogenization of language simply as a result of working with people from different regions and the need to communicate with them), a topic that we leave for future research.⁹⁰

7 Conclusion

We examined when and to what extent a government chooses policies directed toward homogenizing its population. We offer six 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, a democracy undertakes a positive amount of homogenization in order to improve general access to the public good chosen by majority rule. Three, homogenization by the ruler is increasing in the threat of democracy. Four, 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 the status quo should democracy prevail, as well as to reduce opposition to his regime and so lower the threat of democracy itself. Five, contrary to a democracy, a dictator will always choose odious forms of homogenization that are particularly costly to minorities. Six, in some cases the ruler may choose policies of divide and rule in order to reduce feelings of national identity because the latter may increase the likelihood of national insurgencies. Rulers will tend to implement policies of divide and rule when the threat of democracy is not imminent, and will homogenize when the threat of democracy becomes very high. Divide and rule is also particularly likely to occur when the rulers are foreign colonizers and when rulers face low state capacity. Finally we offer some suggestive historical discussion which is consistent with several of these results.

We do not explore the effectiveness of individual homogenization policies. It may be that

⁸⁸See 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.

⁸⁹See Alesina and La Ferrara (2005) for a survey of models regarding the pros and cons of diversity for productivity and development.

⁹⁰As an example of bottom-up homogenization Weber (1979) documents the department of Vosges in France where the introduction of the cotton industry in the 1870s “all but wiped out the local dialect when country people moved into small industrial centers.”

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.⁹¹ The possible perverse effects of indoctrination is an excellent topic for future research.

8 References

- Acemoglu, D. and J. A. Robinson (2000) “Why did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective.” *Quarterly Journal of Economics*, 115, 1167-99.
- Acemoglu, D. and J. A. Robinson (2008) “Persistence of Power, Elites and Institutions.” *American Economic Review*, 98, 267-93.
- Aghion, P., X. Jaravel, T. Persson and D. Rouzet (2014) “Education and Military Rivalry.” NBER Working Paper No. 18049, National Bureau of Economic Research, Inc.
- Alesina, A., W. Easterly and J. Matuszeski (2011) “Artificial States.” *Journal of the European Economic Association*, 9, 246 – 77, 04.
- Alesina, A. and N. Fuchs-Schündeln (2007) “Goodbye Lenin (or Not?): The Effect of Communism on People’s Preferences.” *American Economic Review*, 97, 1507 – 28.
- Alesina, A. and E. Glaeser (2004) *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press.
- Alesina, A., J. Harnoss and H. Rapoport (2013) “Birthplace Diversity and Economic Prosperity.” NBER Working Paper No. 18699, National Bureau of Economic Research, Inc.
- Alesina A. and E. La Ferrara (2005) “Ethnic Diversity and Economic Performance.” *Journal of Economic Literature*, 43, 721-61.
- Alesina, A. and E. Spolaore (1997) “On the Number and Size of Nations.” *Quarterly Journal of Economics*, 112, 1027 – 56.
- Alesina, A. and E. Spolaore (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, 1276 – 96.
- Algan, Y., T. Mayer and M. Thoenig (2013) “The Economic Incentives of Cultural Transmission: Spatial Evidence from Naming Patterns across France.” CEPR discussion paper No. DP9416.
- Asprachs-Bracons, O., J. Costa-Font, I. Clots-Figueras and P. Masella (2008) “Compulsory Language Educational Policies and Identity Formation.” *Journal of the European Economic Association*, 6, 434 – 44.
- 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, 955 – 88.
- Bourguignon F. and T. Verdier (2000) “Oligarchy, Democracy Inequality and Growth.” *Journal of Development Economics*, 62, 285 – 313.
- Besley, T. and T. Persson (2009) “The Origins of State Capacity: Property Rights, Taxation,

⁹¹Vouka (2014) (see also the historical references provide in her paper) provides a fascinating example of assimilation policies gone wrong in the case of Germans in the US.

- and Politics.” *American Economic Review*, 99, 1218 – 44.
- Besley, T. and T. Persson (2010) “State Capacity, Conflict, and Development.” *Econometrica*, 78, 1-34.
- Besley, T., T. Persson and M. Reynal-Querol (2014) “Resilient Leaders and Institutional Reform: Theory and Evidence.” mimeo.
- Bowen, W.H. (2007) “The Spanish Military during World War Two.” in *A Military History of Modern Spain: from the Napoleonic Era to the International War on Terror*, W.H. Bowen and J.E. Alvarez eds. Praeger.
- Bowles, S. (1998) “Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions.” *Journal of Economic Literature*, 36, 75-111.
- Bursztyn, L. (2013) “Poverty and the Political Economy of Public Education Spending: Evidence from Brazil.” UCLA Working Paper.
- Bolton, P. and G. Roland (1997) “The Breakup of Nations: A Political Economy Analysis.” *Quarterly Journal of Economics*, 112, 1057-90.
- Christopher, A. J. (1988) “‘Divide and Rule’: The Impress of British Separation Policies.” *Area*, 20, 233-40.
- Conquest, R. (1970) *The Nation Killers: the Soviet Deportation of Nationalities*. Macmillan, London.
- Colley, L (1986) “Whose Nation? Class and National Consciousness in Britain 1750 – 1830.” *Past and Present*, 113, 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.” *Quarterly Journal of Economics*, 112, 1203-50.
- Flora, P. et al. (1983) *State, economy, and society in Western Europe 1815-1975: a data handbook in two volumes. Vol.1*. Campus Verlag, Frankfurt.
- Fouka, V. (2014) “Backlash: The Unintended Effects of Language Prohibition in US Schools after World War I.” mimeo.
- Gellner, E. (1983) *Nations and Nationalism*. Cornell University Press.
- Gennaioli, N. and I. Rainer (2007) “The modern Impact of Precolonial Centralization in Africa.” *Journal of Economic Growth*, 12, 185-234.
- Gennaioli, N. and H. Voth (forthcoming) “State Capacity and Military Conflict.” *Review of Economic Studies*.
- Glaeser, E., G Ponzetto and A. Shleifer (2007) “Why does Democracy need Education?” *Journal of Economic Growth*, 12, 77-99.
- Gradstein, M. and M. Justman (2002) “Education, Social Cohesion, and Economic Growth.” *American Economic Review*, 92, 1192-1204.
- Herbst, J. (2000) *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press, Princeton, NJ.
- Hobsbawm, E. J. (1990) *Nations and Nationalism since 1780: Programme, Myth, Reality*. Cambridge University Press.
- Holland, M. (2005) “The Captain Swing Project” in *Swing Unmasked*, M. Holland ed. FACHRS

publications.

- Jones, N. L. (1976) "The Catalan Question since the Civil War" in *Spain in Crisis: the Evolution and Decline of Franco*. Paul Preston ed. Harvester Press.
- Katznelson, I. and M. Weir (1985) *Schooling for All: Class, Race, and the Decline of the Democratic Ideal*. New York: Basic Books.
- Laitin, D (2007) *Nations, States, and Violence*. Oxford University Press.
- Lott, J. R. Jr. (1999) "Public Schooling, Indoctrination, and Totalitarianism." *Journal of Political Economy*, 107, 127-57.
- Lott, J. R. Jr. (1987) "Why is Education Publicly Provided? A Critical Survey." *Cato Journal*, 7, 475-501.
- Marcham, A. J. (1973) "Educating our Masters: Political Parties and Elementary Education 1867 to 1870." *British Journal of Educational Studies*, 21, 180-191.
- Marten, L. and N. Kula (2008) "Zambia: One Zambia, One Nation, Many Languages." in *Language and National Identity in Africa*, Andrew Simpson ed. Oxford University Press.
- Michaopolous, S. and E. Papaioannou (2012) "The Long-Run Effects of the Scramble for Africa." NBER Working Paper No. 17620.
- Miguel, E. (2004) "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania." *World Politics*, 56, 328-62.
- Mulligan C. B., R. Gil and X. Sala-i-Martin (2004) "Do Democracies have Different Public Policies than Nondemocracies" *Journal of Economic Perspectives*, 18, 51-74.
- Ortega, J. and T. P. Tangeras (2008) "Unilingual versus Bilingual Education: A Political Economy Analysis." *Journal of the European Economic Association*, 6, 1078-1108.
- Phiri, B. J. (2006) *A Political History of Zambia: From the Colonial Period to the 3rd Republic*. Africa World Press.
- Ramirez, F.O. and J. Boli (1987) "The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization." *Sociology of Education*, 60, 2-17.
- Schram, A. (1997) *Railways and the Formation of the Italian State in the Nineteenth Century*. Cambridge University Press.
- Smith, A. D. (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*, C. Tilly ed. Princeton University Press.
- Tilly, C. (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.
- Tilly, C., L. Tilly and R. Tilly (1975) *The Rebellious Century 1830 – 1930*. Harvard University Press, Cambridge MA.
- Weber, E. (1979) *Peasants into Frenchmen: Modernization of Rural France, 1870-1914*. Chatto and Windus, London.
- Young, G. M. and W. D. Handcock. (1964) *English Historical Documents: Vol. 9, 1833–1874*. Eyre and Spottiswoode, London.

Appendix

Appendix A - A 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 1 : 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

$$g - ga(1 - \lambda_{1/2}^m)\hat{l}_i + y - k - C(\lambda_{1/2}^m) = g - ga(1 - \lambda_j^m)(1/2 - j - \hat{l}_i) + y - k - C(\lambda_j^m) \quad (3)$$

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

Expressions (3) and (4) 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 (3) and (4) 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. \square

When will a democracy split or form a single country?

For $\lambda_1^{1/2} \in [0, \lambda_{1/2}^m)$, expression (5) 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:

$$\left[g - (1 - \lambda_{1/2}^m)gal_i + y - k - \left(C(\lambda_{1/2}^m) - C(\lambda_1^{1/2}) \right) \right] - \left[g - (1 - \lambda_{1/4}^m)ga(1/4 - l_i) + y - 2k - C(\lambda_{1/4}^m) \right] \quad (5)$$

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

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

Expression (5) is at a maximum when $l_i = 0$ and decreasing until $l_i = 1/4$; while expression (6) 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 (5) and (6) 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

$$l'_i = \frac{(1 - \lambda_{1/2}^m) + (1 - \lambda_{1/4}^m)}{8(1 - \lambda_{1/4}^m)} \quad l''_i = \left(1 + \frac{(1 - \lambda_{1/2}^m) + (1 - \lambda_{1/4}^m)}{8(1 - \lambda_{1/4}^m)} \right).$$

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

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

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

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

Proof of Proposition 1

Expressions (5) and (7) 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 (7) is positive. The threshold, λ^* , is the value of $\lambda_1^{1/2}$ at which (5) (alternatively (7)) is equal to zero for the median voter.

Proof of Proposition 2

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

$$[g + y - k] + p [g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p) [g + y - k]; \quad (9)$$

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

$$[g + y - k - C(\lambda^*)] + p [g + y - k - [C(\lambda_{1/2}^m) - C(\lambda^*)]] + (1 - p) [g + y - k]; \quad (10)$$

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

$$[g + y - k - C(\lambda^*)] + [g + y - k]. \quad (11)$$

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 (9) is decreasing in p at a faster rate than (10) and (11).

Proof of Proposition 3

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

$$[g + y - k] + p[g - (1 - \lambda_{1/4}^m)ga0.25 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k]; \quad (12)$$

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

$$[g + y - k - C(\lambda^*)] + p[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda^*)]] + (1 - p)[g + y - k]; \quad (13)$$

and for $\bar{\lambda}$ is

$$[g + y - k - C(\bar{\lambda})] + [g + y - k]. \quad (14)$$

Observe that (12) is decreasing in p at a faster rate than (13), and (14) 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 $\lambda^* > \bar{\lambda}$.

Proof of Proposition 4

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

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

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 Propositions 1 and 3, these propositions 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

$$[g + y - k - C(\lambda_1^{1/2})] + v(p, \lambda_1^{1/2})[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - v(p, \lambda_1^{1/2}))[g + y - k]. \quad (15)$$

The derivative of (15) with respect to $\lambda_1^{1/2}$ is

$$-C'(\lambda_1^{1/2}) - \frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2}}[(1 - \lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)] \quad (16)$$

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

Expression (15) 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

$$[g + y - k - C(\lambda^*)] + v(p, \lambda^*)[g + y + k - [C(\lambda_{1/2}^m) - C(\lambda^*)]] + (1 - v(p, \lambda^*))[g + y - k] \quad (17)$$

which is decreasing in p at rate

$$-\frac{\partial v(p, \lambda^*)}{\partial p}[C(\lambda_{1/2}^m) - C(\lambda^*)]. \quad (18)$$

If, for some $p = \bar{p}$, λ^* gives higher utility than the value of $\lambda_1^{1/2}$ that maximizes (15), then this is true for all p higher if

$$-\frac{\partial v(p, \lambda_1^{1/2})}{\partial p}[(1 - \lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)] < -\frac{\partial v(p, \lambda^*)}{\partial p}[C(\lambda_{1/2}^m) - C(\lambda^*)], \quad (19)$$

for all $p \geq \bar{p}$. Then (19) 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 (15), 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 3. 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 since his expected utility from $\bar{\lambda}$ is constant in p , this is true for all p higher.

Discussion of a ruler located anywhere

Take p exogeneous. The intuition behind the key results remains wherever the ruler is located. First, note that when p is exogeneous a high enough level of homogenization can improve the democratic outcome for the ruler, whatever his location. Therefore, when the probability of the democratic outcome, p , is higher, the ruler is (weakly) more willing to undertake costly homogenization. Second, a ruler facing a high probability of democracy may undertake homogenization above and beyond a democracy. A simple way to illustrate that the same intuition applies, wherever the ruler is located, is as follows. Suppose 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 odious homogenization technology. Then, when the probability of democracy is sufficiently high, the ruler will always homogenize at least as much as a democracy and will always homogenize enough to avoid secession and make sure his ideal government persists. In contrast, in a democracy the decisive voter over homogenization is the voter at median distance from the government. For the same level of homogenization, he will face more substantial costs (whether using odious or non-odious technology), and therefore he 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.

Online Appendix

Proof of Proposition 5

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 odious homogenization for i is

$$gad_{ij} = \beta'(\lambda_t^j) + \alpha'(\lambda_t^j)d_{ij}. \quad (20)$$

By the above, λ_t^j that satisfies (20) is increasing in d_{ij} . Preferences are single peaked over λ_t^j .

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

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

$$\beta(\lambda_t^j) + \alpha(\lambda_t^j)/4 = C(\lambda_t^j), \quad \beta'(\lambda_t^j) + \alpha'(\lambda_t^j)/4 = C'(\lambda_t^j), \quad \forall \lambda_t^j \in [0, 1]. \quad (21)$$

From (21), for $d_{ij} = 1/4$ the levels of homogenization that satisfy (20) and $gad_{ij} = C'(\lambda_t^j)$ are equal, for $d_{ij} < 1/4$ the level of homogenization that satisfies (20) is strictly higher, and for $d_{ij} > 1/4$ it is strictly lower than the level of homogenization that satisfies $gad_{ij} = C'(\lambda_t^j)$.

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 odious and non odious technologies are equal. For $j \in [0, 1/4)$, the median voter over homogenization is $i = 1/2$ with $d_{ij} = 1/2 - j > 1/4$, thus odious homogenization chosen by majority rule will be lower than non-odious. Each individual evaluates the difference between their utility in the case of non-odious homogenization and their utility in the case of odious homogenization,

$$[g - (1 - \lambda_{nonod})gad_{ij} + y - k - C(\lambda_{nonod})] - [g - (1 - \lambda_{od})gad_{ij} + y - k - M(\lambda_{od}, d_{ij})], \quad (22)$$

where λ_{od} (respectively λ_{nonod}) is the level of odious (respectively non-odious) homogenization chosen by majority rule. For $j \in [1/4, 3/4]$, $\lambda_{od} = \lambda_{nonod}$, expression (22) is increasing in d_{ij} , the median voter when deciding between odious and non-odious homogenization is at $d_{ij} = 1/4$, and he is indifferent between the two technologies. For $j \in [0, 1/4)$, $\lambda_{od} < \lambda_{nonod}$, expression (22) is increasing in d_{ij} and the median valuation of (22) is also $d_{ij} = (1/2 - j)$. It follows that the median voter $d_{ij} = (1/2 - j)$ must prefer non-odious homogenization since for any level of odious homogenization he can homogenize to the same level instead using non-odious methods and do strictly better.

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

Proof of Proposition 6

The proof is a repeat of previous arguments.

Allowing for $\lambda_t^j \in [-1, 1]$ does not change the results.

From the proof of Proposition 4, for a democracy preferences over homogenization $\lambda_t^j \in [-1, 1]$ remain single peaked and Proposition 1 extends to negative homogenization. To show that Proposition 2 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 2 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

$$[g + y - k] + p[g + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k]. \quad (23)$$

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

$$[g + y - k - C(\lambda_{j,1})] + p[g + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k]. \quad (24)$$

Expression (23) is strictly lower than (24) (and so is any $\lambda_{j,1} < \lambda^* \leq 0$). If $\lambda^* > 0$, then from Proposition 4, 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 $\lambda_1^{1/2} < 0$.