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.

Nation Building
Alberto Alesina and Bryony Reich
NBER Working Paper No. 18839
February 2013, Revised October 2014
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

Nation-building*

Alberto Alesina
Harvard and IGIER Bocconi

Bryony Reich
University College London

First Draft: October 2012
Latest Revision: October 2014

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.

1 Introduction

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

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

*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.

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

[2] Estimate Weber (1979) p67. Hobsbawm (1990) p60 gives a figure of $12-13\%$ of the population who spoke French "correctly" at the French Revolution.

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

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), 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 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]

[4] Weber (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, Weber (1979) p8.

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

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

[7] Miguel (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.

[8] Franco declared his aim to create "a single language, Castilian, and a single personality, the Spanish one", Jones (1976).

[9] In 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

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 one regime to another. 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? First we consider a democracy. Within a country the population only has access to one "government,"a catch-all term for what a public sector does. However, people disagree on which "government" they prefer, i.e. the "location" of the government which can be interpreted either geographically or in terms of preferences. The majority benefits from a certain degree of homogenization, for example, better roads or railways to the capital city improve the individuals' 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.[13] 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.[14]

---

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 Intermarriage*, 2 September 2014, retrieved from http://www.nytimes.com/.

[10] Recently, 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.

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

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

[13] For 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.

[14] One could also think of "private" forms of homogenization. For instance a linguistic minority setting up its own private

By comparison a non-democratic regime, which is in full control of the population and faces little probability of being overthrown, has different incentives. A dictator or ruling elite will have a "government" which matches their own preferences and will not choose to break-up the population since they want to tax the maximum possible number of people. Since a secure regime faces little prospect of secession, has no concern for the welfare of the population, and has in place policies and public goods that already match the preferences of those in charge, then homogenization of the wider population is unnecessary.

The incentives of a non-democratic regime which faces a substantial probability of overthrow (and the establishment of a democracy) are different again. As above, the ruler has a government which perfectly matches his preferences and he will not split the population. However, this regime faces a significant possibility of being overthrown by a democratic movement and therefore the prospect of living under a democracy in the future. The democratic government may choose public goods that differ from the preferences of the ruler or elite. 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 policy of the ruling group.[15] The threat of democracy motivates rulers 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.

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

Historical examples highlight the different forms that homogenization can take: road building is very different from repression of minority cultures. To capture this we model two forms of homogenization. A "benevolent" form, which distributes the costs equally amongst the population, and an "odious" form, which concentrates the costs on "distant" minorities, i.e. on individuals very far from the government's preferences and location. We show how dictators will always chose odious means of homogenization.

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.

[15] Hobsbawm (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."

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 (2012) 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.[16] Potentially then, the history of nation-building within a country can affect the future success of that population.

Related to this, a wealth of historical literature examines accounts of nation-building as part of the formation of successful states.[17] 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. In contrast, nation-building refers to the construction of a national identity. Historically, nation-building occurred relatively late in the history of European state formation (compared with the commencement of state-building), largely in the late 19th and early 20th Centuries. It is associated with the advent of mass education, mass military conscription and nationalism. Our model proposes a theory for the relatively late emergence of nation-building: the coming of democracy.

We are not aware of any formal model directly related to endogenous homogenization. In this paper we 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, 2012), 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.

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 Tangeras (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.[18]

This paper is organized as follows. Section 2 describes the static model and, to ground ideas, examines nation-building under a democratic and a non-democratic regime in the static setting. Section 3 extends our framework to two periods to model the possibility of transition

[16] See Alesina and La Ferrara (2005) and Alesina et al. (2013) for further references.

[17] See 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.

[18] See Green (1990).

to democracy. Section 3 contains our main results since it allows us to compare homogenization across regimes and to determine the effect of a threat of transition. Section 4 examines 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 of 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 The Static Model

We first present the static model of nation-building. Consider a heterogenous population composed of a continuum of individuals of mass 1 distributed uniformly on the segment $[0, 1]$. This population either forms a single country or splits into two equal-sized countries, A and B, comprising the intervals $[0, 1/2]$ 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).[19] Each country has a government which is located at some point $j \in [0, 1]$ inside that country. By "government" we mean a set of public goods and policies provided by an authority.

Individual i, in a country with the government located at j, has utility

$$U_i = g(1 - ad_{ij}) + y - t. \tag{1}$$

The first term measures the value of the government to individual i. Denote by $d_{ij} = |i - j|$ the "distance" of individual i, who is located at $i \in [0, 1]$, from government j in his country. The parameter a measures the cost of this distance. We think of distance as the geographical distance or difference in preferences between individual i and the public goods and policies provided by government j. We let g denote the maximum value of the government when distance is zero. The benefit that individual i receives from his government is decreasing in distance ad_{ij}; it is decreasing in the difference between his ideal and what is provided by the government. The second term is income y, exogenously given and identical for everyone, and the third term is taxes t which are split equally amongst the population of the country.[20]

The cost of public goods in a given country, funded by taxes, is k.[21] Since the costs k can be divided amongst all citizens in the country this captures the benefits of forming a single

[19] In 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.

[20] See Bolton and Roland (1997) for a discussion about separatist movements due to income differences.

[21] Obviously 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.

country rather than breaking into two.[22] 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 average and 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.

The above set-up is used to model the trade-offs in country formation and secession (see Alesina and Spolaore, 1997). We now depart from this since we assume that the costs of heterogeneity, measured by the parameter a, are endogenous. We model "homogenization" as a technology which uses state apparatus to reduce the cost of distance from the government. Specifically, the cost of diversity within a given country can be reduced by the government j by fraction $\lambda_j \in [0,1]$ to $a' = (1-\lambda_j)a$. So that, for a country with government j, and for any individual in that country, i, the cost of the difference between the policies of government j and i's ideal policies is reduced by fraction λ_j, from ad_{ij} to $(1-\lambda_j)ad_{ij}$. We refer to this as homogenization of the population. We restrict the options such that any degree of homogenization, λ_j, must be applied to the whole population within the country governed by j.[23] Homogenizing a population of size s by λ_j costs $sC(\lambda_j)$, where

Assumption 1 *The cost of homogenization, $C(\cdot)$, is strictly increasing, strictly convex and twice continuously differentiable as λ_j increases from 0 to 1. With $C(0) = 0$, $C'(0) = 0$ and $\lim_{\lambda_j \to 1} C'(\lambda_j) = \infty$.*

For the moment we assume the costs of homogenization are split equally across the population. The government budget constraint is then $st = k + sC(\lambda_j)$.

What follows is a discussion of the interpretation of the model and the assumptions.

Homogenization and distance

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. 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,

[22] Alesina, Spolaore and Wacziarg (2000) and Alesina and Spolaore (2003) investigate sources of benefits of size, like the dimension of the market and diversity of inputs in productivity.

[23] Observe that homogenization λ_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 effects 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.

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.[24] 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.[25] 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 the split of the country. From now 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 be kept together by transferring resources to the 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 refers to a reduction in the cost of heterogeneity. The option of negative λ_j implies a policy that increases the costs of heterogeneity. This is irrelevant for most of the paper, as negative λ_j would never be chosen under any type of regime until we study the case of "divide and rule" below in Section 4. To minimize notation we allow the possibility of a negative λ_j only in Section 4 and in the online appendix we show that allowing for negative homogenization does not change the other results.

Costs of homogenization

For the main section we assume the costs of homogenization are split equally across the population. We relax this assumption on equal costs in Section 5. In particular we will allow for forms of homogenization 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.

In our model income is exogenous. However, at least up to a point, diversity of skills, education, background, and culture may increase productivity.[26] In this case a reduction in

[24] See 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.

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

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

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/consumption enters linearly in the utility function and taxes are lump sum, this reinterpretation of the costs and benefits of diversity would be immediate.

Decision making

Decision-making proceeds in this order: 1) whether to form a single country or split into two; 2) where to locate the government; 3) to what extent to homogenize. Utility is then realized. We examine these choices across different regimes.[27] We study the case of a democracy first, where the decisions are made by majority rule, and then the case of a non-democratic regime, where the decisions are made by a single ruler or elite group. 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.[28] The ordering of decisions is not a crucial assumption.

2.1 A Democracy

We have three decisions in this model. The decision of whether to form a single country or split, the decision of where to locate the government, and the decision of how much to homogenize. In a democracy each decision is made by the whole population by majority rule, with the timing of votes described above. We solve this backward.

First, examine the choice of homogenization in a democracy. The optimal level of homogenization for person i in a fixed country of size s with a fixed government at j satisfies:

$$\arg\max_{\lambda_j \in [0,1]} (g - (1-\lambda_j) gad_{ij} + y - k/s - C(\lambda_j)).$$

The first order condition,

$$gad_{ij} = C'(\lambda_j),$$

implies that the marginal cost of homogenization has to equal the marginal benefit. 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. 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

[27] A further decision that could be made is on the size of the government, g. This would involve other considerations in the provision of public goods across regimes that are not the focus of this paper.

[28] For 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. To implement such an ordering we have to make this further assumption.

(higher λ_j). 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.

The intuition 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.[29] Strong attachment to cultural values can be modeled as very high costs of homogenization; homogenization would therefore be very low in equilibrium.

Second, let us examine the choice of government. Clearly, the government will be located at the center of any country, namely the median preferences in that country. See Lemma 1 in the appendix for a proof.[30] Thus in a single country the government is located at $j = 1/2$ and homogenization satisfies $ga/4 = C'(\lambda_{1/2})$ which we denote by $\lambda^m_{1/2}$. In Countries A and B the government is located at $j = 1/4$ and $3/4$ respectively, and homogenization satisfies $ga/8 = C'(\lambda_j)$, denoted $\lambda^m_{1/4}$ and $\lambda^m_{3/4}$ respectively.[31]

Third, each individual in a democracy evaluates whether he would be better off under a single country or if the population splits into A and B. Utility from forming a single country minus utility from splitting into two countries for individual $i \in [0, 1/2]$ (this is symmetric for $i \in [1/2, 1]$) is then:

$$\left[g - (1 - \lambda^m_{1/2})gad_{i1/2} + y - k - C(\lambda^m_{1/2})\right] - \left[g - (1 - \lambda^m_{1/4})gad_{i1/4} + y - 2k - C(\lambda^m_{1/4})\right]. \quad (2)$$

Expression (2) encompasses the considerations for a voter when deciding whether he prefers a single country or secession. This depends on how well his preferences are represented when the population splits compared to when a single country is formed, and the extra costs, k, of forming two smaller countries rather than one large country. Homogenization reduces his distance from the government in both cases, the more so when a single country is formed (since it is larger and more heterogeneous). He also takes into account the costs of homogenization.

The following proposition describes precisely the choices that will be made by a democracy. From here onwards, to break any ties, we assume that when indifferent between one country

[29] For 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.

[30] Since the location of the government determines the amount of homogenization chosen by majority rule, the result that majority rule will locate the government at the center is not immediate. Preferences are not necessarily single peaked nor necessarily satisfy single-crossing when policies and voters are ordered by their location on the unit interval.

[31] Homogenization chosen by majority rule is the optimal homogenization for the individual at median distance from the median: at $i = 1/4, 3/4$ in a single country, at $i = 1/8$, $i = 3/8$ in A, and at $i = 5/8$, $i = 7/8$ in B.

or two, a single country is always formed. Let

$$i^m = \frac{\lambda^m_{1/2} - \lambda^m_{1/4}}{8(1 - \lambda^m_{1/4})}. \tag{3}$$

Proposition 1 *A democracy chooses to organize itself as one country, locate the government at 1/2, and homogenize to degree $\lambda^m_{1/2}$ when expression (2) is positive as evaluated for i^m. A democracy chooses to split into two countries, locate the governments at the centers, 1/4 and 3/4, and homogenize to degree $\lambda^m_{1/4}$ and $\lambda^m_{3/4}$ when expression (2) is negative as evaluated for i^m. Moreover, $\lambda^m_{1/2} > \lambda^m_{1/4} = \lambda^m_{3/4}$.*

The proof of this proposition is in the appendix. The case of a static democracy highlights the relationship between nation-building and the break-up of a population. It is perfectly possible that without the option of homogenization ($\lambda_j = 0$) this population would decide to split into two countries, but the option of choosing $\lambda_j \in [0, 1]$ would lead the population to homogenize somewhat and form a single country.[32] This case captures the idea of "nation-building". This population would otherwise split, but will stay together with a technology of homogenization. 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.[33]

2.2 A Non-Democratic Regime

Suppose now that this population is controlled by a single ruler (or dictator or elite, terms used synonymously here) who is located at the center of the population. In 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.[34]

The government is located at the ruler's location. The ruler chooses whether to form a single country or split and the degree of homogenization of the population he rules. He maximizes

[32] See the online appendix $A1$ for a formal proof.

[33] Intuitively, who tends to prefer a singe country versus two? Voters near the center of the population, 1/2, prefer to form a single country relative to splitting into two (see equation (2)). 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 since $\lambda^m_{1/2} > \lambda^m_{1/4}$. If, due to the nature of the cost function, $\lambda^m_{1/2}$ is substantially greater than $\lambda^m_{1/4}$, then it is possible that some voters at the extreme (close to 0 or 1) may prefer a single country with a very high $\lambda^m_{1/2}$ to two countries with a relatively low $\lambda^m_{1/4}$. 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.

[34] The elite group is represented by a group of mass δ with ideal point 0.5. Results on this point are available from the authors. Such an extension complicates notation and algebra with little advantage in terms of insight.

his utility given by (1). For completeness, we also specify what happens to Country B should the ruler choose to split the population.[35]

Proposition 2 *The ruler forms a single country with the government at his location, 1/2, and he undertakes no homogenization.*

The ruler forms a single country to minimize tax (and maximize rent extraction were we to include rent in the model), with a government that is ideal for himself. He faces no incentives to break up the country. He undertakes no homogenization since the government perfectly matches his preferences and location and, in contrast to a democracy, he has no incentive to increase the welfare of the population by improving its access to the public good. He is unconcerned with the heterogeneity of the population he rules.

3 Transition to Democracy

Historically, many countries have moved from dictatorships and rule by elite to democratic regimes. We now study how a potential transition from a non-democratic regime to a democracy affects the pattern of homogenization. To model this transition, we introduce two periods to the framework. In the first period a country is controlled by a ruler or elite and in the second period that country may become democratic.

We show that the threat of transition provides two motives for a ruler to homogenize. The first is that by homogenizing the ruler can better preserve the status quo should transition occur. The second, shown in Section 3.1, is that by homogenizing, the ruler can reduce opposition and lower the probability of transition itself. These two motives work in the same direction and may be present individually or both at the same time. The two period framework also allows us to compare across different regimes and examine how the path of rule of a population affects the long-run heterogeneity of that population. Thus:

Period 1 The ruler is in power. He knows that with probability p democracy will prevail in period 2.

Period 2 With probability p democracy prevails and the population as a whole now makes all decisions; with probability $1-p$ democracy is not realized and the ruler stays in power.

In period 1, the ruler, as defined in Section 2.2, is in power and decision-making proceeds as detailed in Section 2.2. In period 2 all decisions are made by majority rule if democracy prevails, as detailed in Section 2.1; if democracy does not prevail, decisions are made by the ruler. Note that, should democracy prevail, the ruling elite remains in the country as "normal" citizens. For the moment p is exogenous; we will endogenize it below.

[35] If the ruler splits the population, since he is located at $j = 1/2$ he will be in Country A and so Country A is subject to the decisions made by the ruler as described above, with the utility of any i in Country A given by (1). We have to define what happens to Country B. We assume no decisions are made in Country B and utility is zero. The results are not sensitive to this assumption. If we assume that Country B becomes a democracy and makes decisions accordingly, the results do not change.

Borders and the location of the government can be altered in each period at no cost. Adding a fixed cost for those changes would add notation with not much additional insight. Homogenization is persistent: roads built today remain tomorrow, languages learnt today are not forgotten tomorrow, preferences influenced today by the government influence future preferences.[36] Let $\lambda_{j,1}$ (respectively $\lambda_{j,2}$) denote the level of homogenization undertaken by government j in period 1 (respectively period 2). If no government is located at j in period 1 then clearly $\lambda_{j,1} = 0$, respectively for $\lambda_{j,2}$. Individual i is subject to the homogenization undertaken by the government of his own country in any period, therefore let $\lambda^i_{j,1}$ denote homogenization by government j that individual i is subject to in period 1. Let $\lambda^i_{j,2}$ denote homogenization by government j that individual i is subject to in period 2. Utility in each period is as given by (1) but takes into account any previous homogenization. Thus, utility in period 1 for individual i in country of size s with government at j and homogenization $\lambda_{j,1}$ is equal to

$$U_i = g(1 - (1 - \lambda^i_{j,1})ad_{ij}) + y - t, \tag{4}$$

where taxes t continue to be split equally among the population of the country and cover the cost of government in period 1, k, and the cost of period 1 homogenization. Utility in period 2 for individual i in country of size s with government at j and homogenization $\lambda_{j,2}$ is

$$U_i = g(1 - (1 - \lambda^i_{j,1} - \lambda^i_{j,2})ad_{ij}) + y - t, \tag{5}$$

where taxes t continue to be split evenly among the population of the country and cover the cost of government in period 2, k, and the cost of period 2 homogenization.[37] The cost of period 1 homogenization follows Assumption 1. The cost of period 2 homogenization must now take into account any previous homogenization:

Assumption 2 *Following homogenization $\lambda_{j,1}$ of a population of size s, the cost of homogenizing that population by a further $\lambda_{j,2}$ in period 2 is $sC(\lambda_{j,1} + \lambda_{j,2}) - sC(\lambda_{j,1})$.*

We assume no discounting of future periods. As above, we also specify what happens to Country B if a ruler splits the population.[38]

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

[37] Observe that homogenization by the previous government is redundant if the "location" of the government changes. Intuitively, the effort of a government to teach the population one national language is redundant if the next government imposes a different national language. In some cases homogenization enacted by a certain government may actually make homogenization by a different government more costly.

[38] If the ruler splits the population in any period, the outcome for that period is as specified in Section 2.2. We make the additional assumption that, once split, the population remains split forever. This avoids making assumptions on homogenization and its costs in the strange situation where a ruler splits the population in period 1 and then he or a democracy re-forms a single country in period 2.

In any period, the ruler maximizes his total expected utility. Working backwards, a period 2 democracy votes as described in Section 2 but now taking into account homogenization undertaken in period 1. Homogenization by the ruler in period 1 results in a more homogenous population at the start of period 2. In the following proposition we detail the effect of such homogenization on the democratic outcome.

Proposition 3 *Suppose in period 1 the dictator forms a single country with the government at* $1/2$. *Given* g, a, k, *and* $C(\cdot)$, *there exists a level of homogenization* $\lambda^* \in [0,1]$ *such that, if the ruler homogenizes by at least* λ^* *in period 1, a democracy would form a single country in period 2, and if he homogenizes less than* λ^*, *a democratic population would split.*

The proof (as well as details of the democratic outcome that follows any level of homogenization in period 1) is found in the appendix. Proposition 3 determines the amount of homogenization necessary to avoid the break-up of the population in period 2 should democracy prevail. This threshold, λ^*, takes into account the homogenization choices of a democracy in period 2. If the ruler does not undertake homogenization in period 1 then in period 2 a democracy behaves as detailed in Proposition 1; therefore, when a democracy would choose to avoid secession anyway we have $\lambda^* = 0$, meaning that a ruler does not need to homogenize to avoid secession.

Proposition 4 describes the optimal decisions of the ruler in period 1. Note that period 2 is the final period and so the ruler, if in power, maximizes by behaving the same as in the static case.[39] Period 1 is the case we are interested in since it models a situation where the ruler faces some threat of democracy occurring next period. From Proposition 2, we know that 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 avoid the break-up of the country in period 2 and ensure the ruler's ideal government persists. As with a democracy, to break any ties, we assume that when indifferent between homogenizing to ensure a single country or the population splitting into two, the ruler chooses the former.

Proposition 4 *In period* 1 *the ruler forms a single country with the government at 1/2. Given* g, a, k, *and* $C(\cdot)$, *there exists a threshold* $0 < \bar{p} < 1$ *such that*

(i). When $p < \bar{p}$, *the period 1 ruler undertakes no homogenization;*

(ii). When $p \geq \bar{p}$, *if* $C(\lambda^*) \leq \bar{C}$ *the period 1 ruler undertakes homogenization to* λ^* *to avoid future secession;*

(iii). When $p \geq \bar{p}$, *if* $C(\lambda^*) > \bar{C}$ *avoiding future secession is too costly and the period 1 ruler undertakes no homogenization;*

where $\bar{C} = (1 - \lambda^m_{1/4})ga/4 + k + C(\lambda^m_{1/4})$.

The intuition is as follows. If a democracy splits then the ruler suffers for two reasons: the new governments of the two countries are no longer the preferred government of the ruler, in

[39] He forms a single country, the government is located at his ideal and he undertakes no further homogenization in period 2.

fact the ruler is the furthest away from the two new governments at 1/4 and 3/4, and the taxes he has to pay are higher. By homogenizing in period 1 the ruler can ensure a large state and that his preferred government persists in period 2, even if democracy prevails. The ruler "nation-builds" with a particular agenda: he homogenizes to build a large, stable nation that better reflects his preferences.

Nevertheless, homogenization is costly. Proposition 4 (i) says that a ruler who faces a low probability of democracy undertakes no homogenization. He faces a low probability that democracy will prevail and so has no need to homogenize. Those dictators however, who are not as safe in office, homogenize up to the point at which even if democracy prevails the country does not break up. 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. Thus when the probability of democracy is high and avoiding secession is not "too costly" then the ruler will nation-build to avoid secession (Proposition 4 (ii)). If avoiding secession is too costly, the ruler will do nothing (Proposition 4 (iii)).

Does a non-democratic ruler homogenize more or less than a democracy? Clearly a "safe" ruler homogenizes less than a democracy, since for low p a dictator does not homogenize at all while a democracy would homogenize in period 2 should democracy prevail. A democracy homogenizes to improve the welfare of those at the periphery; a safe ruler has his ideal government, faces little threat of overthrow and break-up, has no concern for general welfare, and so he is largely unconcerned with the heterogeneity of the population.

In contrast, an unsafe ruler will homogenize to avoid break-up under a larger set of parameters than a democracy. Under certain parameters a ruler will homogenize in period 1 to avoid secession in period 2 should democracy prevail, whereas, without any homogenization by the ruler, a democracy in period 2 would choose to split. By homogenizing, the ruler can ensure his ideal government will persist in a democracy and he can ensure a large country. Thus Proposition 4 shows that a ruler under threat may undertake more extreme levels of homogenization than will be undertaken by a democracy and he does so to form a large stable nation that reflects his preferences.

Note that all of the above results hold for the case of a dictator located at 1/2. 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). 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 discussion in the appendix. The main point to note is that, wherever the ruler is located, when the ruler does badly from the democratic outcome then he will undertake more extreme nation-building than would be seen under a democracy. How "badly" does he have to do to want to undertake a level of nation-building above that of a democracy? This depends on the costs of homogenization and is discussed further in the appendix. In Section 5, when we relax the constraints and the ruler has more choice over homogenization technologies (i.e. he can force more of the costs on minorities) then the ruler's incentives to undertake high levels of nation-building become even stronger.

3.1 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.[40]

Proposition 5 establishes the effect nation-building has on general discontent under a non-democratic regime. Suppose in period 1 the ruler forms a single country with the government at 1/2 and undertakes some level of homogenization. We then know the outcome in period 2 if democracy prevails or if dictatorship persists. We denote by $U_{i2,dem}$ the utility of individual i in period 2 under the democratic outcome, and by $U_{i2,ruler}$ the utility of individual i if the dictator maintains power. Then $U_{i2,dem} - U_{i2,ruler}$ measures the relative "discontent under dictatorship" of individual i in period 2 following the above actions in period 1. Proposition 5 examines the median value of $U_{i2,dem} - U_{i2,ruler}$ within the population, $i \in [0,1]$, and shows that nation-building can reduce discontent under the non-democratic regime.

Proposition 5 *Given g, a, k, and $C(\cdot)$, the median "discontent under dictatorship", $U_{i2,dem} - U_{i2,ruler}$, is decreasing in the level of homogenization undertaken by the ruler in period 1.*

To model the effect of nation-building in reducing the probability of overthrow, we let the probability of democracy be a function $p(\lambda_{j,1}, v)$, where $\lambda_{j,1}$ is the level of homogenization undertaken by the ruler in period 1 and the term $v \in [0, V]$ measures factors which affect the probability of democracy but are exogenous to the model. A higher v represents conditions associated with a higher threat of democracy.

Assumption 3 *The probability of democracy in period 2, $p(\lambda_{j,1}, v)$, is*

(i). Twice differentiable, strictly convex and strictly decreasing in $\lambda_{j,1}$ and twice differentiable, strictly convex and strictly increasing in v. With $\frac{\partial^2 p(\lambda_{j,1},v)}{\partial \lambda_{j,1} \partial v} = \frac{\partial^2 p(\lambda_{j,1},v)}{\partial v \partial \lambda_{j,1}} = 0$.

(ii). $0 < p(\lambda_{j,1}, v) < 1$, for all $\lambda_{j,1} \in [0,1]$ and for all $v \in [0, V]$.

The probability of democracy is decreasing in the amount of nation-building undertaken by the ruler in period 1 and is increasing in conditions associated with a higher threat of democracy. The final part of assumption 3(ii) ensures the probability of overthrow in period 2 is between zero and one.[41]

[40] Observe that, if rents are included in the model, promising to redistribute future rents would not have the same effect since promising future rents would not be credible. This point is made in Acemoglu and Robinson (2000) in a model of democratic transition.

[41] Assuming strict inequalities is not necessary for the results, but it simplifies the algebra.

Proposition 6 *Given g, a, k, $C(\cdot)$ and $p(\cdot,\cdot)$, in period 1 the ruler forms a single country with the government at $1/2$ and always undertakes a strictly positive amount of homogenization, which is weakly increasing in the exogenous threat of democracy, v.*

See the online appendix for a proof. There is a second important incentive to homogenize: to indoctrinate people to be happy with the government and so reduce the threat of democracy. As with Proposition 4, homogenization is increasing the higher the general threat of democracy, now measured by the parameter v. 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. We can also say which motive is relevant to different types of non-democratic regime: Proposition 4 applies to domestic elites that expect to stay in the country after democratization; the motive to reduce the threat of democracy, shown in Proposition 6, applies to all kinds of non-democratic regimes, even harsh dictators who may be kicked out or eliminated should democracy prevail. Note 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 this would increase the incentive to homogenize in order to reduce the probability of overthrow.

4 Divide and Rule

In some cases revolutions become more likely and more successful when a population is homogeneous. A more homogenous population can communicate better and this may make collective action easier. After all, the principle of "divide and rule" is meant to capture precisely this effect. Similarly, if homogenization involves education of the population, a more educated population could also increase the probability of successful overthrow. By the same argument, policies that increase diversity and its costs ($\lambda_j < 0$) could reduce the threat of overthrow. This case can be analyzed with our model.

Suppose homogenization increases the probability of overthrow. To examine the choices a ruler will make in this situation, we introduce Assumption 4 which replaces Assumption 3.

Assumption 4 *The probability of democracy in period 2, $p(\lambda_{j,1}, v)$, is*

(i). Twice differentiable, strictly convex and strictly increasing in $\lambda_{j,1} \in [-1,1]$ and in $v \in [0,V]$. With $\frac{\partial^2 p(\lambda_{j,1},v)}{\partial \lambda_{j,1}\partial v} = \frac{\partial^2 p(\lambda_{j,1},v)}{\partial v \partial \lambda_{j,1}} = 0$.

(ii). $0 < p(\lambda_{j,1}, v) < 1$, for all $\lambda_{j,1} \in [-1,1]$ and for all $v \in [0,V]$.

Positive or negative homogenization, $\lambda_j \in [-1, 1]$, is now permitted in any period and by any regime. We assume the cost of negative homogenization is symmetric $C(-\lambda_j) = C(\lambda_j)$. We also update assumption 2 to allow for negative homogenization: following homogenization $\lambda_{j,1}$ of a population of size s at cost $sC(\lambda_{j,1})$, the cost of homogenizing that population by $\lambda_{j,2}$ in period 2 is $sC(\lambda_{j,1} + \lambda_{j,2}) - sC(\lambda_{j,1})$ if $\lambda_{j,1}$ and $\lambda_{j,2}$ are the same sign; and is $sC(\lambda_{j,2})$ if $\lambda_{j,1}$ and $\lambda_{j,2}$ are different signs.[42] Finally, to avoid a technical complication (discussed in the following footnote), we make the assumption that, following negative homogenization, if democracy prevails a democratic government is always located at the center of any democratic country.[43]

A democracy never chooses negative homogenization. A ruler in period 2 still always chooses zero homogenization since it is the final period. However, a period 1 ruler may now choose to "divide and rule". That is, to undertake negative homogenization. This is detailed in Proposition 7. Observe that Proposition 3 is extended once we allow for negative homogenization: the threshold level of homogenization that a ruler must undertake to avoid the future break-up of the population, still denoted λ^*, is in the range $[-1, 1]$.

Proposition 7 *In period* 1 *the ruler forms a single country with* $j = 1/2$. *Given* g, a, k, $C(\cdot)$ *and* $p(\cdot, \cdot)$, *there exists a threshold* $\bar{v}$ *such that*

(i). If $v < \bar{v}$ *the ruler implements strictly negative homogenization;*

(ii). If $v \geq \bar{v}$ *the ruler chooses* $\lambda_{j,1} \geq \lambda^*$ *to avoid future secession.*

The proof is in the online appendix. For some parameters, because we limit assumptions on the function $p(\cdot, \cdot)$ and by an analogous intuition to Proposition 4, we can have the degenerate cases: for all $v \in [0, V]$ the ruler implements strictly negative homogenization, or for all $v \in [0, V]$ the ruler chooses $\lambda_{j,1} \geq \lambda^*$ to avoid future secession. In the online appendix A3 we also show that the previous results do not change if we allow for negative homogenization. Only when homogenization increases the probability of overthrow 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 overthrow. 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, v low, the incentive to divide and rule dominates, when conditions favor democracy, v high, the incentive to homogenize dominates. Analogously to Section 3.1, if the ruler could extract rents this would exacerbate his incentive to divide and rule.

[42] Observe that if if $\lambda_{j,1}$ and $\lambda_{j,2}$ are different signs we cannot assume period 2 cost is given by $sC(\lambda_{j,1} + \lambda_{j,2}) - sC(\lambda_{j,1})$. To observe this note that if $\lambda_{j,1} < 0$ and $\lambda_{j,2} > 0$ and $|\lambda_{j,1}| > |\lambda_{j,2}|$ then the cost of period 2 homogenization is negative.

[43] The 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.

4.0.1 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.[44] Consider two places with different levels of state capacity but otherwise identical populations and suppose democracy prevails in both places in period 2.[45] 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.

4.0.2 Colonizers

Colonizers are different from the domestic dictatorship or elite analyzed thus far. Colonizers leave the country after their regime falls.[46] Consider then the case in which the probability of democratization of a colonized country is exogenous. Contrary to the domestic dictators analyzed above, the colonizer would not pay the costs of homogenizing the population. In fact, the colonizer does not care about what happens to the country after he leaves and he cannot affect the probability of democratization by homogenizing. If, by homogenizing, the colonizer can reduce the probability of being overthrown, then he would. However, in the case of colonized territories with fragmented populations, the policy of divide and rule may be especially attractive, since it may be the case that divide and rule 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

Thus far we restricted the homogenization technology such that the cost is spread equally across the population. This assumption captures "non-odious" forms of nation-building. In fact, this is a technology that essentially involves a permanent transfer from the center (which benefits from its closeness to the government) to the periphery (which suffers from its distance). Some types of nation-building policies clearly do not fit this technology. We now relax this

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

[45] We could model a country with low state capacity as one which faces high costs of homogenization. This assumption implies that implementing policies, such as mass education or road building, is more difficult the less state infrastructure there is in place.

[46] In the existing model we could include some probability that the ruler exits the country at democratization and gets a fixed payoff.

assumption and provide governments with a choice over different nation-building technologies. This further strengthens our main results. To do this, we introduce an "odious" nation-building technology, which 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.

We capture the difference between technologies through the cost of homogenization. The cost to individual i of odious homogenization by λ_j is denoted $M(\lambda_j, d_{ij})$ and as before is strictly increasing, strictly convex and twice continuously differentiable as λ_j increases from 0 to 1, with $M(0, d_{ij}) = 0$, $M_{\lambda_j}(0, d_{ij}) = 0$ and $\lim_{\lambda_j \to 1} M_{\lambda_j}(\lambda_j, d_{ij}) = \infty$. The cost of odious homogenization, $M(\lambda_j, d_{ij})$, 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, $M_{\lambda_j}(\lambda_j, d_{ij})$, 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.

To make comparisons between odious and non-odious homogenization we assume that the total cost of homogenizing a country to a given degree is the same under both technologies. That is,

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

when the government is located in the center of the country.[47] 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 for the choice between the two technologies.[48] The homogenization technology is chosen, followed by the amount of homogenization, after borders and governments have been determined. We assume that when indifferent, non-odious homogenization is chosen.

Proposition 8 *When both odious and non-odious technologies are available, a ruler strictly prefers odious homogenization, while a democracy is indifferent.*

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. The result that a democracy is indifferent between the two homogenization technologies 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

[47] The online appendix *A*2 gives an example of such a cost function and highlights details of this type of cost function.

[48] Following homogenization $\lambda_{j,1}$ of a population of size s (undertaken by odious or non-odious technology), the cost of homogenizing that population by $\lambda_{j,2}$ by odious means in period 2 is analogous to the cost using non-odious technology. It is $sM(\lambda_{j,1} + \lambda_{j,2}, d_{ij}) - sM(\lambda_{j,1}, d_{ij})$ for individual i.

a democracy must consider (to a greater degree, at least) the views of the whole population.[49] 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.

Proposition 9 extends Proposition 4 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. Note that, following Proposition 8 and the discussion, a democracy always chooses non-odious homogenization in period 2.[50]

Proposition 9 *In period* 1 *the ruler forms a single country with the government at 1/2. Given* g, a, k, $C(\cdot)$ *and* $M(\cdot,\cdot)$,

(i). There exists a threshold $0 \leq \bar{\bar{p}} < 1$ *such that*

- *(a) When* $p < \bar{\bar{p}}$, *the period 1 ruler undertakes no homogenization;*
- *(b) When* $p \geq \bar{\bar{p}}$, *if* $M(\lambda^*, 0) \leq \bar{C}$, *the period 1 ruler undertakes odious homogenization of at least* λ^* *to avoid future secession;*
- *(c) When* $p \geq \bar{\bar{p}}$, *if* $M(\lambda^*, 0) > \bar{C}$, *avoiding future secession is too costly and the period 1 ruler undertakes no homogenization;*

where $\bar{C} = (1 - \lambda^m_{1/4})ga/4 + k + C(\lambda^m_{1/4})$.

(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 4, under some parameters a ruler will homogenize to avoid secession whereas, without any homogenization by the ruler, a democracy would otherwise split. Proposition 9 highlights that greater latitude in homogenization technologies (compared to Proposition 4) will induce the ruler to avoid secession under a larger range of parameters because he can use technologies which place the costs on minorities. This becomes obvious when we think about technologies where minorities face almost all the costs and the dictator almost none. In this case, when the probability of overthrow is high, the dictator will always homogenize to ensure a large state that represents his preferences. 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.

[49] A 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.

[50] Similarly, since period 2 is the final period, as before, a ruler undertakes no homogenization whatever technologies are available.

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. Because of the wealth of information on education, this focus allows for fruitful comparisons across countries and regimes.

Prediction 1 *Nation-building is increasing in the threat of democracy.*

The model predicts little, if any, nation-building by dictators or ruling elites who face a low threat of democracy. As the threat of democracy increases, so too will the amount of nation-building implemented by such regimes. The threat of democracy induces nation-building by dictators and ruling elites for three reasons. 1. To mitigate secessionist aspirations by parts of the population and avoid future break-up of the population as it becomes more democratic. 2. To ensure the preferences of the elite or dictator persist as the population gets more of a say in things. 3. To reduce opposition to the current regime by making it more palatable relative to a potential democratization.

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.[51] 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.

Eighteenth and Nineteenth Century Europe.

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."

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

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.

France

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.[52] 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.[53] The French Crown showed "little concern with the linguistic conquest of the regions under its administration."[54] 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.[55]

After the French Revolution, and increasingly throughout the 19th Century (following further revolutions and uprisings), it became clear that more and more power would be transferred from French elites to the wider population. Prediction 1 suggests we should observe increasing homogenization in this period. Schooling was one way to homogenize and, indeed, major schooling reforms by the elite followed periods of major unrest. Immediately after the French Revolution, the Constitution of 1791 called for the establishment of free public instruction for all.[56] 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."[57] The first serious attempt to actually implement mass schooling was made in 1833 following a period of major rebelion (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,

[52] Tilly (1980).

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

[54] Weber (1979).

[55] Katznelson and Weir (1985).

[56] Ramirez and Boli (1987).

[57] Weber (1979).

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.[58]

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".[59] 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 and very unlike north France, yet it finally became thoroughly French."[60]

Italy

Italian unification was completed in the 1860*s* accompanied by an increase in pressure for more democracy.[61] 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. 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.[62] Michele Coppino, the author of the 1877 Italian compulsory education reform, declared that

[58] Other nation-building measures by the French government included the suppression of other languages: as late as 1890 a ministerial decree banned religious instruction in Flemish and in 1902 the government banned Breton language sermons.

[59] Report by the rector of the Academy of Rennes, Boudoin, Weber (1979), p100. 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."

[60] Tilly (1975) p43.

[61] The largest proportion of adult males were enfranchised in Italy in 1912.

[62] Duggan (2007), p289.

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."[63] 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.[64] 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."[65] Nicola Marselli expressed in 1871 "I know, too, that Italy has been reunited for only ten years and is not 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."[66]

England

The establishment of English public education also coincided with a greater threat of democracy. Public education first appeared 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".[67] 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 also 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"[68] 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

[63] Duggan (2007), p280.
[64] Duggan (2007), p288. This policy continued well into the 20th century.
[65] Duggan (2007), p283.
[66] Duggan (2007), p274.
[67] Ramirez and Boli (1987).
[68] Quote from Marcham (1973). As above, the act enfranchised a part of the male urban working-class population.

that we have given [the people] political power we must not wait any longer to give them education."[69]

A Closer Look at Education Reform in Europe

We now take a more systematic look at education reforms in eleven European countries between 1800 and 1875.[70] 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 curricula.[71] 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.[72] 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.[73] 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.[74] 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

[69] Quote from Young and Handcock (1964). In England it is argued that "divide and rule" motives played a role in nation-building until the threat of democracy became especially high. Historian Linda Colley argues that to Britain's governors "nationalism was like Pandora's box: something which was best left alone." Colley (1986) argues that "dividing and ruling seemed a more attractive strategy than state-sponsored nationalism". The fear that nationalism might increase demands by the population meant that nation-building policies were enacted late in Britain: "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." This is consistent with Proposition 7 whereby an elite will opt for divide and rule policies up to the point at which the threat of overthrow is high.

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

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

[72] Examples 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 (see also p247). In England, there was a feeling that events in other European countries could impact unrest in England (Holland, 2005). Hobsbawm and Rudé (1969 p$xxiv$, $62-64$) observe, 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."

[73] Years 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.

[74] These 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.

Lyon.

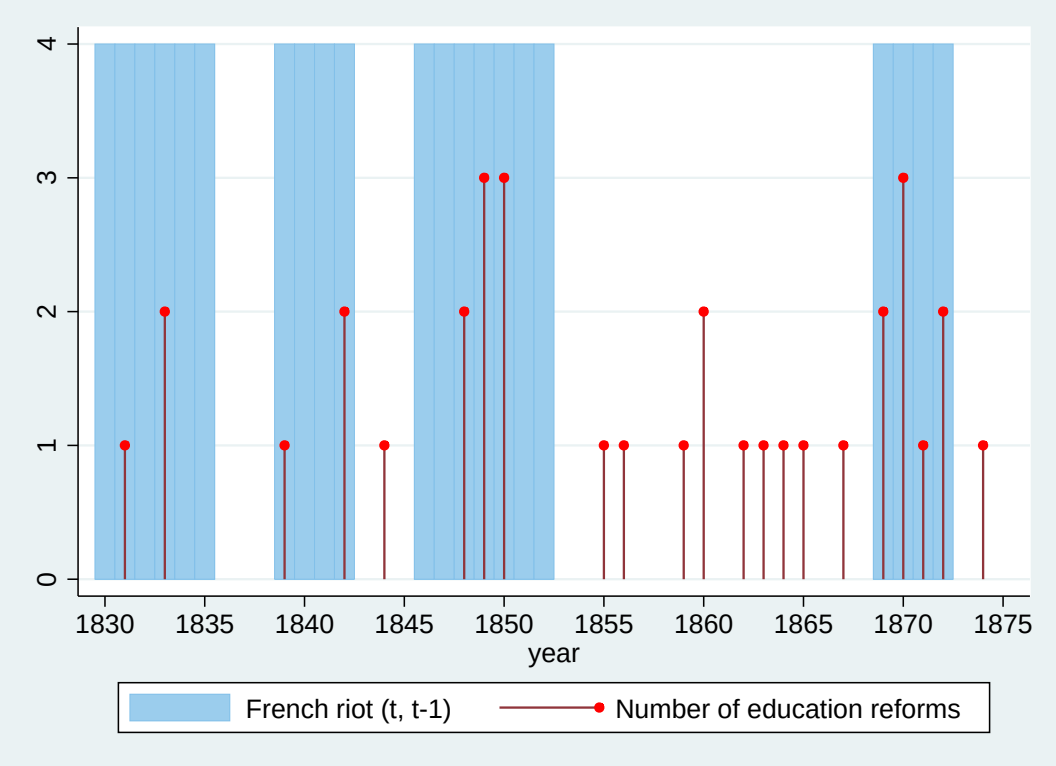

Figure 1: Years of major rioting in France and number of education reforms across Europe

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 only cause cited by more than 30 parishes was 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.[75] 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 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).[76]

Prediction 2 *The harshest types and the highest levels of homogenization are undertaken by non-democratic rulers under threat of democracy.*

Intense nation-building was not just a feature of 19th Century European elites. 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.[77] As in 19th Century Europe, public education was central to these efforts, in particular in forcing a Marxist-Leninist doctrine. Consistent with Prediction 2, 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

[75] Tilly (1998).

[76] Our 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 is an exception, and indeed Germany is responsible for the earliest welfare reforms.

[77] 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, the names of their original habitats changed and their own names removed from the list of Soviet peoples. In some cases these groups had attempted autonomy and were considered a threat to unity (Chechens and Crimean Tatars). Other cases were somewhat precautionary, but with a similar motive. The population of a block of land over which Turkey and Russia had fought over for more than a century was considered to have Turkish sympathies and its population was arguably deported (and repopulated with others more loyal) as a precaution to avoid future trouble in this area with Turkey (Gorenburg, 2006).

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.[78] 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.[79]

Prediction 2 says that the most extreme episodes of nation-building will be undertaken by non-democratic regimes under threat. The previous discussion illustrates particularly harsh policies under non-democratic regimes. However, we can obtain a more general picture by comparing policies across democracies and non-democracies in larger data sets. Since public education can be used to nation-build, Prediction 2 implies that we may see higher levels of government spending on education, more education reform, or higher levels of other education measures, in non-democratic regimes compared to democracies. Such a comparison has, in fact, been done for us. A number of studies compare education policies across democratic and non-democratic regimes and we here summarize recent empirical evidence. Aghion, Persson and Rouzet (2012), using annual data on 137 countries from $1830-2001$, find that autocracies have higher enrolment rates in primary education than democracies. The authors also determine the twenty year period for each country in which primary enrolment rose most sharply. In almost all countries, this period of education reform occurs before democratic transition. In fact, for only 2 out of 53 countries examined does democratic transition occur before the rise in primary enrolment. This is consistent with the evidence in Mulligan, Gil and Sala i Martin (2004) who 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

[78] Apart 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 churches in his diocese.

[79] From 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."

content of indoctrination, for instance only studying Marxist economics in communist dictatorships, not teaching minority languages or the history of ethnic or religious minorities.

Prediction 3 *Policies of divide and rule are more likely to be implemented in non-democratic regimes that are ruled by a colonizer or have low state capacity.*

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 policies and the implementation of divide and rule policies, 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 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.[80]

Zambia, a British colony from the 19th Century to independence in 1964, adheres to this pattern. Colonization was "a take-the-money-and-run affair" with education mainly provided by missionaries.[81] 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.[82]

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 indepen-

[80] Smith (2003) and Miguel (2004).

[81] Marten and Kula (2008).

[82] Marten 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.

dence, 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.[83] 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.

Alternative theories of nation-building

We briefly discuss other theories of nation-building. Our model focuses on internal threats, however external threats may also motivate nation-building. The threat of external war motivates governments to build unity and patriotism and thus the willingness of the population to fight the enemy. This follows a similar argument to Besley and Persson (2010) regarding the role of external war in encouraging state-building. This theory is examined by Aghion, Persson and Rouzet (2012) who establish a correlation between a "war risk" indicator and investment in public education. They argue that public education was, in part, an investment to increase the willingness and ability of citizens to defend the state.

Certainly external wars are relevant to nation-building. However war plays only a partial role in motivating governments to nation-build. Aghion, Persson and Rouzet (2012) find that the nature of the regime has a significant effect on education reforms and primary enrolment, independent of the threat of war. They find democratic reform is associated with lower enrolment. In Figure 1 we show that education reforms are particularly associated with periods of unrest. Indeed, in some of the case studies above, the external war incentive can be easily ruled out. For example, in Catalonia in 1939, repressive policies were particularly intense, yet at the same time Franco reduced the size of the army by three quarters, only increasing it again later as a result of the Second World War.[84] In Italy, the forming of Italian regiments made up of citizens from all parts of the country and stationing recruits far from home is argued to have actually been counter-productive to producing an efficient fighting force as it would have hindered rapid mobilization.[85]

[83] English became the "median of national communication," while also maintaining state level languages, as it gave no one language group a particular advantage.

[84] Young (1976) p236 – 241 and Bowen (2007).

[85] Duggan (2007), p289.

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.[86] In contrast, an industrial society based upon markets (as opposed to a stratified agrarian society with local markets) needs better means of communication. Universal schooling serves an economic purpose as well, necessary for the development of an industrial society.[87] In other words productivity would increase in an industrial society with more homogenization relative to an agrarian one.[88]

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.[89]

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 chooses a government that is ideal for himself and 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, a ruler who faces a high probability of overthrow may undertake the highest levels of homogenization. Indeed he may homogenize enough to ensure a single country where a democracy alone would instead break up. Fourth, contrary to a democracy, a dictator will always choose odious forms of homogenization that are particularly costly to minorities. Fifth, when the probability of democratic success decreases with the degree of homogeneity of the population, the dictator will strategically take it into account and homogenize more. 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. This

[86] Gellner (1983) writes that political units in agrarian societies "can de divided into two species: local self governing communities and large empires". Neither of these type of governments represent a modern nation state.

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

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

[89] As 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."

may be particularly likely to occur when the rulers are foreign colonizers. 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 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. Vouka (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. 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.

Aghion, P., T. Persson and D. Rouzet (2012)"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, and Politics." *American Economic Review*, 99, 1218 − 44.
Besley, T. and T. Persson (2010) "State Capacity, Conflict, and Development." *Econometrica*, 78, 1-34.
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 Unintened Effects of Language Prohibition in US Schools after World War I." Unpublished.
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.
Glaeser, E., G Ponzetto and A. Shleifer (2007) "Why does Democracy need Education?" *Journal of Economic Growth*, 12, 77-99.
Gorenburg, D. (2006) "Soviet Nationalities, Policy and Assimilation"in *Rebounding Identities: The Politics of Identity in Russia and Ukraine*, D. Arel, B. A. Ruble eds. Woodrow Wilson Center Press.
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

Proof of Proposition 1

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

For a single country and any government $j \in [1/4, 3/4]$, the level of homogenization chosen by majority rule satisfies $ga/4 = C(\lambda_j)$ 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), then the level of homogenization chosen by majority rule satisfies $ga(1/2 - j) = C'(\lambda_j)$, denote this by λ_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, 1/2. Let $\hat{l}_i$ denote this distance for individual $i \in [0, 1/2]$ who is indifferent between the government at some $j \in [0, 1/4)$ 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 an individual's value of locating the government at $j = 1/2$ versus some $j \in [0, 1/4)$ 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 $i = 3/4$ strictly prefers $j = 1/2$ and thus $\hat{\hat{l}}_i > 1/4$, 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

$$g - ga(1 - \lambda_{1/2}^m)\hat{l}_i + y - k - C(\lambda_{1/2}^m) = g - ga(1 - \lambda_j^m)(1/2 - j - \hat{l}_i) + y - k - C(\lambda_j^m) \quad (6)$$

$$g - ga(1 - \lambda_{1/2}^m)\hat{\hat{l}}_i + y - k - C(\lambda_{1/2}^m) = g - ga(1 - \lambda_j^m)(1/2 - j + \hat{\hat{l}}_i) + y - k - C(\lambda_j^m). \quad (7)$$

Expressions (6) and (7) 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 $\hat{l}_i + \hat{\hat{l}}_i > 1/2$. In the same way we can show that Country A and B will locate the government at their centers.□

To complete the proof it remains to determine the median voter in the choice to form a single country or split. Expression (2) for individual i at distance $l_i \in [0, 1/4]$ from the center of the population can be rewritten as

$$-[(1 - \lambda_{1/4}^m) + (1 - \lambda_{1/2}^m)]gal_i + (1 - \lambda_{1/4}^m)ga/4 + k - [C(\lambda_{1/2}^m) - C(\lambda_{1/4}^m)] \quad (8)$$

and for individual i at distance $l_i \in [1/4, 1/2]$

$$[(1-\lambda_{1/4}^m) - (1-\lambda_{1/2}^m)]gal_i - (1-\lambda_{1/4}^m)ga/4 + k - [C(\lambda_{1/2}^m) - C(\lambda_{1/4}^m)]. \tag{9}$$

Expression (8) is at a maximum when $l_i = 0$ and decreasing until $l_i = 1/4$; while expression (9) 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 (8) and (9) respectively and such that $l_i' + (0.5 - l_i'') = 1/4$. Thus l_i' and l_i'' have the median valuation of (2) and l_i' solves:

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

and hence

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

Proof of Proposition 3

Working backwards, we determine the choices that will be made by a democracy in period 2 following homogenization $\lambda_{j,1} \in [0, 1]$ in period 1.

For any $\lambda_{j,1} \in [0, 1]$, if in period 2 a democracy forms a single country, then the democracy locates $j = 1/2$ and undertakes homogenization $\lambda_{2,j} = \max\{0, \lambda_{1/2}^m - \lambda_{j,1}\}$. If the democracy locates $j = 1/2$ then for any individual i, optimal homogenization is

$$\underset{\lambda_{j,2} \in [0,1]}{\arg\max} \quad g - ga(1 - \lambda_{j,1} - \lambda_{j,2})d_{ij} + y - k - [C(\lambda_{j,1} + \lambda_{j,2}) - C(\lambda_{j,1})].$$

Individual i's optimal homogenization is weakly increasing in d_{ij} and all individuals have single-peaked preferences over homogenization, thus $d_{ij} = 1/4$ is the median distance and homogenization chosen by majority rule is $\lambda_{2,j} = \max\{0, \lambda_{1/2}^m - \lambda_{j,1}\}$. The result that a democracy will choose to locate the government at $j = 1/2$ follows by the same argument as Lemma 1, noting that for all i the utility from locating $j = 1/2$ weakly increases compared to that detailed in the proof of Lemma 1 while the utility from any $j \neq 1/2$ is the same.

For any $\lambda_{j,1} \in [0, 1]$, if in period 2 a democracy chooses to split, then the democracy locates $j = 1/4$ in country A and $j = 3/4$ in B and homogenizes by $\lambda_{1/4}^m, \lambda_{3/4}^m$. By lemma 1, in Country A, $j = 1/4$ beats all other locations $j \neq 1/2$ in a pairwise vote. Since the ruler undertakes homogenization at $j = 1/2$ we need to show that $j = 1/4$ necessarily beats $j = 1/2$. 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 a government at $j = 1/4$.[90] Then, since the population is symmetric, a majority in Country B must also prefer a single country with the government located at $j = 1/2$ to a split country with a government at $j = 3/4$.

It remains to determine when a democracy will form a single country or split. Following the method in Proposition 1, for any level of homogenization $\lambda^m_{1/2} > \lambda_{j,1} \geq 0$, the period 2 utility of individual i at distance $l_i \in [0, 1/4]$ from the center in a single country minus his utility in two countries is

$$[g-(1-\lambda^m_{1/2})gal_i+y-k-[C(\lambda^m_{1/2})-C(\lambda_{j,1})]]-[g-(1-\lambda^m_{1/4})ga(1/4-l_i)+y-2k-C(\lambda^m_{1/4})] \quad (10)$$

with median voter

$$l_i = \frac{(1-\lambda^m_{1/4})+(1-\lambda^m_{1/2})}{8(1-\lambda^m_{1/4})}. \quad (11)$$

Similarly for $\lambda_{j,1} \geq \lambda^m_{1/2}$, and $l_i \in [0, 1/4]$ this is

$$[g-(1-\lambda_{j,1})gal_i+y-k]-[g-(1-\lambda^m_{1/4})ga(1/4-l_i)+y-2k-C(\lambda^m_{1/4})] \quad (12)$$

with median voter given by

$$l_i = \frac{(1-\lambda^m_{1/4})+(1-\lambda_{j,1})}{8(1-\lambda^m_{1/2})}. \quad (13)$$

Expressions (12) and (10) evaluated for the respective median voters are equal at $\lambda_{j,1} = \lambda^m_{1/2}$ and both strictly increasing in $\lambda_{j,1}$. Further for $\lambda_{j,1} = 1$ every individual prefers a single country.

Proof of Proposition 4

Suppose the ruler forms a single country with the government at $j = 1/2$. How much homogenization does he undertake? Let us go through the different cases.

(i). When $\lambda^* = 0$ the ruler sets $\lambda_{j,1} = 0$. Any $\lambda_{j,1} > 0$ is costly and results in the same democratic outcome as $\lambda_{j,1} = 0$.

(ii). When $\lambda^* > 0$ the ruler sets either $\lambda_{j,1} = 0$ (any $\lambda^* > \lambda_{j,1} > 0$ is costly and results in the same democratic outcome as $\lambda_{j,1} = 0$) or $\lambda_{j,1} = \lambda^*$ (by the same argument).

(a) Thus when $\lambda^m_{1/2} > \lambda^* > 0$ the ruler sets $\lambda_{j,1} = \lambda^*$ if

$$[g+y-k-C(\lambda^*)]+p[g+y-k-[C(\lambda^m_{1/2})-C(\lambda^*)]]+(1-p)[g+y-k] \geq$$

[90] Given the same $\lambda_{j,1}$, homogenization $\lambda_{j,2}$ will be exactly the same in a single country and in country A with $j = 1/2$, since the median voter in either case has $d_{ij} = 0.25$.

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

and $\lambda_{j,1}=0$ otherwise. This can be rewritten

$$p[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]-p[C(\lambda^m_{1/2})-C(\lambda^*)]-C(\lambda^*)\geq 0. \tag{14}$$

By the optimality of $\lambda^m_{1/2}$ for $d_{ij}=1/4$, we have $(1-\lambda^m_{1/4})ga/4+C(\lambda^m_{1/4})>(1-\lambda^m_{1/2})ga/4+C(\lambda^m_{1/2})$ and from this we know (14) is increasing in p.

(b) Similarly, when $\lambda^*\geq\lambda^m_{1/2}$, the ruler sets $\lambda_{j,1}=\lambda^*$ if

$$p(1-\lambda^m_{1/4})ga/4+pk+pC(\lambda^m_{1/4})-C(\lambda^*)\geq 0 \tag{15}$$

and sets $\lambda_{j,1}=0$ otherwise. This is increasing in p.

It follows that if for any p the ruler sets $\lambda_{j,1}=\lambda^*$, then he does so for all p higher. Now for $p=1$, (14) is positive, by the same argument as above on the optimality of $\lambda^m_{1/2}$ for $d_{ij}=1/4$. For $p=1$, (15) is negative when

$$(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})<C(\lambda^*). \tag{16}$$

Observe (16) never holds whenever $\lambda^*\leq\lambda^m_{1/2}$, again by the same argument as above on the optimality of $\lambda^m_{1/2}$ for $d_{ij}=1/4$, and so (16) completely describes the conditions under which a ruler will not avoid secession for any $p\in[0,1]$. It remains to show in the following lemma that the ruler always chooses to form a single country.

Lemma 2 : The ruler will form a single country in period 1 and period 2

We show that for any $\lambda_{j,1}$ the ruler always does strictly better by forming a single country.

If the ruler splits the country in period 1 and democracy prevails in period 2, then if a democracy locates $j=1/2$ utility is

$$g-ga(1-\lambda_{j,1}-\lambda_{j,2})l_i+y-2k-[C(\lambda_{j,2}+\lambda_{j,1})-C(\lambda_{j,1})] \tag{17}$$

where $\lambda_{j,2}=\max\{0,\lambda^m_{1/2}-\lambda_{j,1}\}$. For $\lambda_{j,1}<\lambda^*$ from Proposition 4 we know for the median voter his utility from (17) is lower than from $j=1/4$ and a democracy will vote for $j=1/4$. Thus for any $\lambda_{j,1}<\lambda^*$, if the ruler forms a single country in period 1 his expected utility is

$$[g+y-k-C(\lambda_{j,1})]+p[g-ga(1-\lambda^m_{1/4})(1/4-l_i)+y-2k-C(\lambda^m_{1/4})]+(1-p)[g+y-k]. \tag{18}$$

and if the ruler splits the country in period 1 his utility is

$$[g+y-2k-C(\lambda_{j,1})]+p[g-ga(1-\lambda^m_{1/4})(1/4-l_i)+y-2k-C(\lambda^m_{1/4})]+(1-p)[g+y-2k] \tag{19}$$

For any $\lambda_{j,1} \geq \lambda^*$, if the ruler forms a single country in period 1 his expected utility is

$$[g + y - k - C(\lambda_{j,1})] + p[g + y - k - [C(\lambda_{j,2} + \lambda_{j,1}) - C(\lambda_{j,1})]] + (1 - p)[g + y - k]$$

where $\lambda_{j,2} = \max\{0, \lambda^m_{1/2} - \lambda_{j,1}\}$. It follows that if the ruler splits the population in period 1, his utility in period 2 if democracy prevails must be strictly lower and therefore his expected utility is strictly lower. Finally, observe that by forming a single country in period 2 the ruler attains his maximum possible utility $g + y - k$.□

Discussion of a ruler located anywhere

See the working paper version of the current paper for a slightly different model worked out for a ruler located anywhere. We here discuss how the motives of the ruler to homogenize change when the ruler is located at some $j' \neq 1/2$. The motives to homogenize that arise from an endogenous probability of overthrow are straightforward. If the probability of overthrow depends negatively on homogenization then this motive to homogenize remains, whatever the ruler's location. If the probability of overthrow depends positively on homogenization then "divide and rule" motives are present, whatever the ruler's location. Suppose the probability of overthrow is exogenous, how does the motive to homogenize change in this case? When the ruler is located at $j = 1/2$ he homogenizes to ensure a large state that reflects his preferences, in order to mitigate the effects of future democratization.

For a ruler located at $j' \neq 1/2$, there exists a level of homogenization above which a single country democracy would choose to locate the government at j' rather than $j = 1/2$. The intuition for this result is that if the ruler homogenizes enough to government j' (he builds roads to j', teaches the language of j', schooling is catered to j') he can make government j' less costly for the majority of the population than $j = 1/2$ and the population will choose a government at j' over $j = 1/2$. Similarly, if the ruler homogenizes enough he can ensure the population of Country A (supposing $j' \in [0, 1/2]$) would vote for a government at j' rather than 1/4. It follows that, for any j', enough homogenization can ensure a single country and the ruler's ideal government. Similarly to a ruler at $j = 1/2$, if such homogenization is not "too costly" then when p is high enough, a ruler will homogenize to ensure a single country and his ideal government.

Working out the details of the case of a ruler located anywhere is more complicated than $j = 1/2$ because homogenization can have multiple effects. One level of homogenization ensures a single population prefers j' to $j = 1/2$, another level of homogenization ensures Country A prefers j' to $j = 1/4$, and differing amounts of homogenization will determine whether the population prefers a single country or split. There are four different country size and government combinations possible rather than two.

The intuition behind the two key results remains. First, a ruler who faces a low probability of democracy continues to undertake no homogenization. Homogenization, when p is exogeneous, is only beneficial to the ruler because it can make the democratic outcome better for him. Therefore the higher the probability of the democratic outcome, p, the more willing he is to undertake homogenization. Second, a ruler facing a high probability of democracy may

undertake homogenization above and beyond a democracy and may ensure a single country whereas a democracy would otherwise choose to split. The ruler does so to ensure a large stable state that reflects his preferences. As discussed in the main body of the text, whether the ruler homogenizes and avoids secession above and beyond a democracy depends on two main factors: how badly the ruler does from democratization and the costs of homogenization. By how "badly" he does from democratization we mean how far the democratic government will be from his ideal and whether the population will break up. Changing the location of the ruler affects his motives to homogenize since it changes how far he is from the democratic government.

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 ruler faces negligible costs. This can be captured within our definition of an odious homogenization technology. Then, when the probably 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 is the median voter who, for the same level of homogenization, will face more substantial costs (whether using odious or non-odious technology), and therefore he will be less willing to undertake high homogenization. In other words, 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.

For online publication

Proof of Proposition 5

If the ruler forms a single country with $j = 1/2$ and undertakes homogenization $\lambda_{j,1}$, from the proof of Proposition 3 we know the value of $U_{i2,dem}$ and $U_{i2,ruler}$. We go through the different cases to determine the median value of $U_{i2,dem} - U_{i2,ruler}$ and show this is decreasing in $\lambda_{j,1}$.

Case 1: $\lambda^* = 0$.
For any homogenization $\lambda_{j,1} \leq \lambda^m_{1/2}$

$$U_{i2,dem} - U_{i2,ruler} = [g - ga(1-\lambda^m_{1/2})l_i + y - k - [C(\lambda^m_{1/2}) - C(\lambda_{j,1})]] - [g - ga(1-\lambda_{j,1})l_i + y - k]$$

where l_i denotes i's distance from the center. For any $\lambda_{j,1} \geq \lambda^m_{1/2}$ $U_{i2,dem} - U_{i2,ruler} = 0$. The median valuation of $U_{i2,dem} - U_{i2,ruler}$ occurs for $l_i = 1/4$ hence

$$\frac{\partial U_{i2,dem} - U_{i2,ruler}}{\partial \lambda_{j,1}} = C'(\lambda_{j,1}) - ga/4, \quad \text{for } \lambda_{j,1} \leq \lambda^m_{1/2}, \quad \frac{\partial U_{i2,dem} - U_{i2,ruler}}{\partial \lambda_{j,1}} = 0, \quad \text{for } \lambda_{j,1} \geq \lambda^m_{1/2}.$$

where $C'(\lambda_{j,1}) \leq ga/4$ for all $\lambda_{j,1} \leq \lambda^m_{1/2}$.

Case 2: $\lambda^* \geq \lambda^m_{1/2}$.
For any homogenization $\lambda_{j,1} < \lambda^*$

$$U_{i2,dem} - U_{i2,ruler} = [g - ga(1-\lambda^m_{1/4})(1/4 - l_i) + y - 2k - C(\lambda^m_{1/4})] - [g - (1-\lambda_{j,1})gal_i + y - k]. \quad (20)$$

for $l_i \in [0, 0.25]$. For any $\lambda_{j,1} \leq \lambda^m_{1/4}$, then $U_{i2,dem} - U_{i2,ruler}$ is increasing in $l_i \in [0, 1/2]$ and so $l_i = 1/4$ has the median valuation of $U_{i2,dem} - U_{i2,ruler}$. For any $\lambda_{j,1} > \lambda^m_{1/4}$, individual $l_i = 1/4 \frac{(1-\lambda^m_{1/4}) + (1-\lambda_{j,1})}{2(1-\lambda^m_{1/4})}$ has the median valuation of $U_{i2,dem} - U_{i2,ruler}$. Then

$$\frac{\partial U_{i2,dem} - U_{i2,ruler}}{\partial \lambda_{j,1}} = -ga/4 \text{ for } \lambda_{j,1} < \lambda^m_{1/4},$$

$$\frac{\partial U_{i2,dem} - U_{i2,ruler}}{\partial \lambda_{j,1}} = -\frac{(1-\lambda^m_{1/4}) + (1-\lambda_{j,1})}{2(1-\lambda^m_{1/4})} ga/2 \text{ for } \lambda_{j,1} \geq \lambda^m_{1/4}.$$

For any homogenization $\lambda_{j,1} \geq \lambda^*$ then $U_{i2,dem} - U_{i2,ruler} = 0$. Finally observe that as $\lambda_{1,j}$ increases towards λ^*, by definition of λ^*, (20) is positive and tends to zero for the median voter.

Case 3: $\lambda^m_{1/2} > \lambda^* > 0$.

For any homogenization $\lambda_{j,1} < \lambda^*$, then $U_{i2,dem} - U_{i2,ruler}$ is the same as Case 2. For any homogenization $\lambda^* \leq \lambda_{j,1} < \lambda^m_{1/2}$, then

$$U_{i2,dem} - U_{i2,ruler} = [g - ga(1-\lambda^m_{1/2})l_i + y - k - [C(\lambda^m_{1/2}) - C(\lambda_{j,1})]] - [g - ga(1-\lambda_{j,1})l_i + y - k], \quad (21)$$

and for any $\lambda_{j,1} \geq \lambda^m_{1/2} \geq \lambda^*$, then $U_{i2,dem} - U_{i2,ruler} = 0$. When $\lambda_{j,1} \geq \lambda^*$ this is the same as Case 1. Thus $U_{i2,dem} - U_{i2,ruler}$ is weakly decreasing in $\lambda_{j,1}$ for $\lambda_{j,1} < \lambda^*$ and for $\lambda_{j,1} \geq \lambda^*$. It remains to determine happens to $U_{i2,dem} - U_{i2,ruler}$ at λ^*. Subtract (20) from (21) at $\lambda_{j,1} = \lambda^*$ for $l_i \in [0, 1/4]$ to get

$$[g - ga(1-\lambda^m_{1/2})l_i + y - k - [C(\lambda^m_{1/2}) - C(\lambda^*)]] - [g - ga(1-\lambda^m_{1/4})(1/4 - l_i) + y - 2k - C(\lambda^m_{1/4})]. \quad (22)$$

By definition of λ^*, (22) is zero for the median voter $l_i = \frac{(1-\lambda^m_{1/2})+(1-\lambda^m_{1/4})}{2(1-\lambda^m_{1/4})}$. Thus, for this value of l_i, (20) and (21) are equal. However at λ^*, the median value of (21) occurs for $l_i = 1/4$ and the median value of (20) occurs for $l_i = \frac{(1-\lambda^*)+(1-\lambda^m_{1/4})}{2(1-\lambda^m_{1/2})}$. Thus for $l_i = 1/4$ compared to $l_i = \frac{(1-\lambda^m_{1/2})+(1-\lambda^m_{1/4})}{2(1-\lambda^m_{1/4})}$, expression (21) increases by

$$[(1-\lambda^*) - (1-\lambda^m_{1/2})] \frac{(1-\lambda^m_{1/4}) - (1-\lambda^m_{1/2})}{2(1-\lambda^m_{1/4})} ga/4$$

and for $l_i = \frac{(1-\lambda^*)+(1-\lambda^m_{1/4})}{2(1-\lambda^m_{1/2})}$, (20) increases by

$$[(1-\lambda^m_{1/4}) + (1-\lambda^*)] \frac{(1-\lambda^*) - (1-\lambda^m_{1/2})}{2(1-\lambda^m_{1/4})} ga/4.$$

Since (20) is strictly decreasing for all $\lambda_{j,1} \leq \lambda^*$ and (21) is strictly lower than (20) at λ^*, the result follows.

Proof of Proposition 6

Given $\lambda_{j,1}$, the probability of overthrow, $p(\lambda_{j,1}, v)$, is the same whether the ruler forms a single country or splits. Therefore, by Lemma 2, for any given $\lambda_{j,1}$ a ruler does strictly better by forming a single country in periods 1 and 2. We now go through the different cases and determine the ruler's optimal homogenization in period 1. Observe, only the probability of democracy changes, therefore in period 2 a democracy and ruler make the same choices as when p is exogenous.

Case 1: $\lambda^* = 0$.

The ruler's period 1 utility and period 2 expected utility is strictly higher if $\lambda_{j,1} = \lambda^m_{1/2}$ than for any $\lambda_{j,1} > \lambda^m_{1/2}$, since for $\lambda_{j,1} > \lambda^m_{1/2}$, the democratic outcome for the ruler is identical but the cost is higher. Then optimal homogenization is in the range $\lambda_{j,1} \leq \lambda^m_{1/2}$ and maximizes

$$[g + y - k - C(\lambda_{j,1})] + p(\lambda_{j,1}, v)[g + y - k - [C(\lambda^m_{1/2}) - C(\lambda_{j,1})]] + (1 - p(\lambda_{j,1}, v))[g + y - k] \quad (23)$$

The derivative with respect to $\lambda_{j,1}$ is

$$-\frac{\partial p(\lambda_{j,1}, v)}{\partial \lambda_{j,1}}[C(\lambda_{1/2}^m) - C(\lambda_{j,1})] - (1 - p(\lambda_{j,1}, v))C'(\lambda_{j,1}). \tag{24}$$

At $\lambda_{j,1} = 0$ (24) is strictly positive and at $\lambda_{j,1} = \lambda_{1/2}^m$ strictly negative. The derivative of (24) with respect to $\lambda_{j,1}$ is

$$-\frac{\partial^2 p(\lambda_{j,1}, v)}{\partial \lambda_{j,1}^2}[C(\lambda_{1/2}^m) - C(\lambda_{j,1})] + 2\frac{\partial p(\lambda_{j,1}, v)}{\partial \lambda_{j,1}}C'(\lambda_{j,1}) - (1 - p(\lambda_{j,1}, v))C''(\lambda_{j,1}).$$

All terms are strictly negative and thus there is a unique optimal level of homogenization, $\lambda_{j,1}$, which is strictly positive. This optimal level is strictly increasing in v since the derivative of (24) w.r.t v is $\frac{\partial p(\lambda_{j,1}, v)}{\partial v}C'(\lambda_{j,1})$.

Case 2: $\lambda^* \geq \lambda_{1/2}^m$.

For the ruler, optimal homogenization in the range $\lambda_{j,1} < \lambda^*$ maximizes

$$[g+y-k-C(\lambda_{j,1})]+p(\lambda_{j,1}, v)[g-(1-\lambda_{1/4}^m)ga0.25+y-2k-C(\lambda_{1/4}^m)]+(1-p(\lambda_{j,1}, v))[g+y-k] \tag{25}$$

The derivative with respect to $\lambda_{j,1}$ is

$$-\frac{\partial p(\lambda_{j,1}, v)}{\partial \lambda_{j,1}}[(1 - \lambda_{1/4}^m)ga0.25 + k + C(\lambda_{1/4}^m)] - C'(\lambda_{j,1}). \tag{26}$$

At $\lambda_{j,1} = 0$ this is strictly positive and is strictly decreasing in the range $\lambda_{j,1} < \lambda^*$. Thus there exists a unique optimal amount of homogenization in this range which is strictly positive. The ruler will never homogenize above λ^*, by the same argument as Case 1. Homogenization to λ^* is optimal for the ruler whenever

$$[g + y - k - C(\lambda^*)] + p(\lambda^*, v)[g + y - k] + (1 - p(\lambda^*, v))[g + y - k] \geq \tag{27}$$

$$[g+y-k-C(\lambda')]+p(\lambda', v)[g-(1-\lambda_{1/4}^m)ga0.25+y-2k-C(\lambda_{1/4}^m)]+(1-p(\lambda', v))[g+y-k].$$

where λ' maximizes (25). To see that optimal homogenization by the ruler is increasing in v, observe that optimal λ' does not change with v and that the derivative of the left hand side of (27) with respect to v is zero and the derivative of the right hand side is negative:

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

Case 3: $\lambda_{1/2}^m > \lambda^* > 0$.

Optimal homogenization by the ruler in the range $\lambda_{j,1} < \lambda^*$ is the same as Case 2. Optimal homogenization in the range $\lambda^m_{1/2} \geq \lambda_{j,1} \geq \lambda^*$ is the same as Case 1. The ruler never homogenizes above $\lambda^m_{1/2}$ in period 1.

To see that optimal homogenization by the ruler is weakly increasing in v, observe the derivative of (26) w.r.t v is zero and the derivative of (24) w.r.t v is $\frac{\partial p(\lambda_{j,1},v)}{\partial v} C'(\lambda_{j,1})$. It remains to show that if for any v

$$[g+y-k-C(\lambda'''',v)[g+y-k-[C(\lambda^m_{1/2})-C(\lambda'')]]+(1-p(\lambda'',v))[g+y-k]\geq$$

$$[g+y-k-C(\lambda'',v)[g-(1-\lambda^m_{1/4})ga/4+y-2k-C(\lambda^m_{1/4})]+(1-p(\lambda',v))[g+y-k]$$

where λ' and λ'' maximize respectively (25) in the range $\lambda_{j,1} < \lambda^*$ and (23) in the range $\lambda^m_{1/2} \geq \lambda_{j,1} \geq \lambda^*$, then this inequality holds for all v higher. To see this, take the derivative of either side of the inequality with respect to v. We find

$$-\frac{\partial p(\lambda'',v)}{\partial v}[C(\lambda^m_{1/2})-C(\lambda'')]+\frac{\partial \lambda''}{\partial v}[-\frac{\partial p(\lambda_{j,1},v)}{\partial \lambda_{j,1}}[C(\lambda^m_{1/2})-C(\lambda_{j,1})]-(1-p(\lambda_{j,1},v))C'(\lambda_{j,1})] \tag{28}$$

$$\geq -\frac{\partial p(\lambda',v)}{\partial v}[(1-\lambda^m_{1/4})ga0.25+k+C(\lambda^m_{1/4})],$$

where the left hand side of (28) is strictly greater than the right hand side since the second term is zero and $\frac{\partial p(\lambda'',v)}{\partial v} = \frac{\partial p(\lambda',v)}{\partial v}$.

Proof of Proposition 7

For any i in a democracy, preferences remain single peaked over $\lambda_{2,j} \in [-1,1]$, since reducing homogenization from zero both increases the cost of homogenization and the cost of distance. A democracy never votes for negative homogenization. For any $\lambda_{j,1} \geq 0$ then optimal homogenization under a democracy is the same as given in the proof of Proposition 3. If $\lambda_{j,1} < 0$ then utility in a single country in period 2 for $\lambda_{j,2} \geq 0$ is

$$g - ga(1-\lambda_{j,1}-\lambda_{j,2})d_{ij} + y - k - C(\lambda_{j,2})$$

The derivative with respect to $\lambda_{j,2}$ is

$$gad_{ij} - C'(\lambda_{j,2})$$

and optimal $\lambda_{j,2}$ for any i is increasing in d_{ij}. If a democracy forms a single country with $j = 1/2$, then median distance is $d_{ij} = 1/4$ and hence median homogenization is $\lambda_{j,2} = \lambda^m_{1/2}$. If a democracy splits, it locates $j = 1/4, 3/4$ with homogenization $\lambda^m_{1/4}, \lambda^m_{3/4}$, by the proof of Proposition 4.

We show there continues to exist a threshold λ^*, which is now in the range $[-1, 1)$. If the ruler forms a single country with $j = 1/2$ and $\lambda_{j,1} < 0$, then a period 2 democracy chooses to form a single country if

$$[g-(1-\lambda_{j,1}-\lambda^m_{1/2})gal_i+y-k-C(\lambda^m_{1/2})]-[g-(1-\lambda^m_{1/4})ga(1/4-l_i)+y-2k-C(\lambda^m_{1/4})]\geq 0$$

for the median voter

$$l_i=\frac{(1-\lambda^m_{1/4})+(1-\lambda_{j,1}-\lambda^m_{1/2})}{8(1-\lambda^m_{1/4})}.$$

The derivative of this expression evaluated for the median voter is

$$\frac{(1-\lambda^m_{1/4})+(1-\lambda_{j,1}-\lambda^m_{1/2})}{2(1-\lambda^m_{1/4})}ga/2.$$

Therefore if the population forms a single country for any $\lambda_{j,1} < 0$ the population will form a single country for all $\lambda_{j,1}$ higher and the result follows.

Suppose the ruler forms a single country with $j = 1/2$. Case by case, we determine the optimal homogenization for the ruler in period 1. In period 2 the ruler attains maximum utility $g + y - k$ if he forms a single country and undertakes no homogenization.

Case 1: $\lambda^* \geq \lambda^m_{1/2}$.

The ruler never homogenizes above λ^*, since the cost is higher but the democratic outcome for the ruler is the same. Expected utility for the ruler if he homogenizes by $\lambda_{j,1} < \lambda^*$ is

$$[g+y-k-C(\lambda_{j,1})]+p(\lambda_{j,1},v)[g-(1-\lambda^m_{1/4})ga/4+y-2k-C(\lambda^m_{1/4})]+(1-p(\lambda_{j,1},v))[g+y-k] \quad (29)$$

The derivative with respect to $\lambda_{j,1}$ is

$$-\frac{\partial p(\lambda_{j,1},v)}{\partial \lambda_{j,1}}[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]-C'(\lambda_{j,1}). \quad (30)$$

This is strictly negative for all $\lambda^* > \lambda_{j,1} \geq 0$ and strictly positive for $\lambda_{j,1} = -1$. Also (30) is strictly decreasing in $\lambda_{j,1}$ so there exists a unique optimal level of homogenization in this range, which is strictly negative. Denote this by $-\lambda'$. Then the optimal choice of the dictator is λ^* if the following is satisfied and is $-\lambda'$ if not:

$$[g+y-k-C(\lambda^*)]+p(\lambda^*,v)[g+y+k]+(1-p(\lambda^*,v))[g+y+k] \quad (31)$$

$$\geq [g+y-k-C(-\lambda'',v)[g-(1-\lambda^m_{1/4})ga/4+y-2k-C(\lambda^m_{1/4})]+(1-p(-\lambda',v))[g+y+k].$$

The derivative of (30) with respect to v is zero. The derivative of the left hand side of (31) with respect to v is zero. The derivative of the right hand side is

$$-\frac{\partial p(-\lambda',v)}{\partial v}[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]$$

Thus if for any v the ruler prefers λ^*, then he does so for all v higher.

Case 2: $\lambda^m_{1/2} > \lambda^* > 0$.

By the same argument as above, the ruler homogenizes to λ^* if the following inequality is satisfied or $-\lambda'$ which maximizes (29) if not:

$$[g+y-k-C(\lambda^*)]+p(\lambda^*,v)[g+y+k-[C(\lambda^m_{1/2})-C(\lambda^*)]]+(1-p(\lambda^*,v))[g+y+k] \quad (32)$$

$$\geq [g+y-k-C(\lambda')]+p(-\lambda',v)[g-(1-\lambda^m_{1/4})ga/4+y-2k-C(\lambda^m_{1/4})]+(1-p(-\lambda',v))[g+y+k]$$

The derivatives of the left hand side and right hand side of (32) are respectively

$$-\frac{\partial p(\lambda^*,v)}{\partial v}[C(\lambda^m_{1/2})-C(\lambda^*)] > -\frac{\partial p(-\lambda',v)}{\partial v}[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]$$

where the inequality holds since $\frac{\partial p(\lambda^*,v)}{\partial v} = \frac{\partial p(-\lambda',v)}{\partial v}$.

Case 3: $\lambda^* \leq 0$.

The elite never choose $\lambda_{j,1} > 0$, since the democratic outcome for the ruler is the same as when $\lambda_{j,1} = 0$, but the cost of homogenization and probability of overthrow are higher. Denote by $-\lambda'$ the value of $\lambda_{j,1} \in [\lambda^*, 0]$ that maximizes

$$[g+y-k-C(\lambda_{j,1})]+p(\lambda_{j,1},v)[g+y+k-C(\lambda^m_{1/2})]+(1-p(\lambda_{j,1},v))[g+y+k]$$

The derivative with respect to $\lambda_{j,1}$ is

$$-C'(\lambda_{j,1}) - \frac{\partial p(\lambda_{j,1},v)}{\partial \lambda_{j,1}} C(\lambda^m_{1/2})$$

where the first term is positive, the second negative and both terms are decreasing in $\lambda_{j,1}$. Thus there is a unique optimal level of homogenization in the range $\lambda_{j,1} \in [\lambda^*, 1]$. Denote by $-\lambda''$ the value of $\lambda_{j,1} \in [-1, \lambda^*)$ that maximizes

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

The derivative with respect to $\lambda_{j,1}$ is

$$-C'(\lambda_{j,1}) - \frac{\partial p(\lambda_{j,1},v)}{\partial \lambda_{j,1}}[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]$$

where the first term is positive, the second negative and both terms are decreasing in $\lambda_{j,1}$. Thus there is a unique $\lambda_{j,1} \in [-1, \lambda^*)$ that optimizes. In both cases the ruler implements divide and rule. He chooses $-\lambda'$ if the following inequality is satisfied and $-\lambda''$ if not

$$[g+y-k-C(-\lambda'',v)[g+y+k-C(\lambda^m_{1/2})]+(1-p(-\lambda',v))[g+y+k] \quad (33)$$

$\geq [g+y-k-C(-\lambda'')]+p(-\lambda'',v)[g-(1-\lambda^m_{1/4})ga/4+y-2k-C(\lambda^m_{1/4})]+(1-p(-\lambda'',v))[g+y-k]$

The derivative of the left hand side and right hand side of (33) are respectively

$$-\frac{\partial p(-\lambda',v)}{\partial v}C(\lambda^m_{1/2}) > -\frac{\partial p(-\lambda'',v)}{\partial v}[(1-\lambda^m_{1/4})ga/4+k+C(\lambda^m_{1/4})]$$

where the inequality holds since $\frac{\partial p(-\lambda',v)}{\partial v} = \frac{\partial p(-\lambda'',v)}{\partial v}$.

Lemma 3: the ruler forms a single country in period 1 and 2.
We fix the probability of democracy, p and extend Lemma 2 to show that for any $\lambda_{j,1} < 0$ a ruler does strictly better if he forms a single country in period 1 rather than split. For any $\lambda_{j,1} < 0$ a democracy always locates the government in the center of the country, therefore in a democratic Country A the government is always located at $j = 1/4$. The ruler's expected utility if he forms a single country and homogenizes by $\lambda_{j,1} < 0$ where $\lambda_{j,1} < \lambda^*$ is given by (18) and if the ruler splits by (19), which is strictly lower. The ruler's expected utility if he forms a single country and homogenizes by $\lambda_{j,1} < 0$ where $\lambda_{j,1} \geq \lambda^*$ is

$$[g+y-k-C(\lambda_{j,1})]+p[g+y-k-C(\lambda^m_{1/2})]+(1-p)[g+y-k] \tag{34}$$

and if the ruler splits is again given by (19). That (19) is strictly lower than (34) comes from observing that when democracy prevails in period 2, if $\lambda_{j,1} \geq \lambda^*$ then the median voter does weakly better by forming a single country, and therefore the ruler must do strictly better. As in Lemma 2, in period 2 the ruler attains maximum feasible utility from forming a single country.

Finally, given $\lambda_{j,1}$ if the probability of overthrow is endogenous it is $p(\lambda_{j,1}, v)$ which is identical whether the ruler forms a single country or splits and the above result holds in this case also.

Proof of Proposition 8

For clarity of exposition we here denote odious homogenization by μ_j and non-odious by λ_j. Without loss of generality we can write $M(\mu_j, d_{ij}) = \beta(\mu_j) + \alpha(\mu_j)d_{ij}$. Since total costs of homogenization by different technologies are equalized for $j = 1/2$

$$2\int_0^{0.5} [\beta(\mu_j) + \alpha(\mu_j)x]\,dx = C(\lambda_j),$$

where $\mu_j = \lambda_j$, $j = 1/2$. Hence $\beta(\mu_j)+\alpha(\mu_j)/4 = C(\lambda_j)$ and $\beta'(\mu_j)+\alpha'(\mu_j)/4 = C'(\lambda_j)$ for all $\mu_j = \lambda_j \in [0,1]$ and for all $j \in [0,1]$. The first order condition for the optimal level of odious homogenization for i can be written $gad_{ij} = \beta'(\mu_j) + \alpha'(\mu_j)d_{ij}$. Since $M(\mu_j, 0) = \beta(\mu_j)$, then $\beta'(\mu_j) > 0$, $\forall \mu_j > 0$ and since $M_{\mu_j}(\mu_j, d_{ij})$ is increasing in d_{ij}, $\alpha'(\mu_j) > 0$ $\forall \mu_j > 0$. Thus $ga > \alpha'(\mu_j) > 0$ so the optimal μ_j for an individual i is increasing in d_{ij} and preferences are single peaked over μ_j.

We examine the choice of homogenization for a single country democracy supposing no previous homogenization. If a strictly positive amount of homogenization has been undertaken in previous periods the same argument applies, and similarly for Country A or B. For $j = 1/2$ the median level of homogenization, denoted $\mu^m_{1/2}$, satisfies $ga/4 = \beta'(\mu^m_{1/2}) + \alpha'(\mu^m_{1/2})/4$. But since also $ga/4 = C'(\lambda^m_{1/2})$ then $C'(\lambda^m_{1/2}) = \beta'(\mu^m_{1/2}) + \alpha'(\mu^m_{1/2})/4$, and from above $\lambda^m_{1/2} = \mu^m_{1/2}$. Each individual evaluates the difference between their utility in the case of non-odious homogenization and their utility in the case of odious homogenization:

$$[g - (1 - \lambda^m_{1/2})gad_{ij} + y - k - C(\lambda^m_{1/2})] - [g - (1 - \mu^m_{1/2})gad_{ij} + y - k - M(\mu^m_{1/2}, d_{ij})]. \quad (35)$$

Since $\lambda^m_{1/2} = \mu^m_{1/2}$, expression (35) is increasing in d_{ij} and so the median voter when deciding between odious and non-odious homogenization is at $d_{ij} = 1/4$. He is indifferent between the two. The same argument applies if $j \in [1/4, 3/4]$.

For $j \in [0, 1/4)$, the median voter when choosing the level of odious or non-odious homogenization is $i = 1/2$ at $d_{ij} = 1/2 - j > 1/4$. Since $M_{\mu_j}(\mu_j, 0.25) = C'(\lambda_j)$ for $\mu_j = \lambda_j$, then for all $d_{ij} > 1/4$ $M_{\mu_j}(\mu_j, d_{ij}) > C'(\lambda_j)$ for $\mu_j = \lambda_j$, and so the median level of odious homogenization chosen will be lower than the median level of non-odious. Each individual evaluates the difference between utility in the case of non-odious homogenization and utility in the case of odious homogenization:

$$[g - (1 - \lambda_j)gad_{ij} + y - k - C(\lambda_j)] - [g - (1 - \mu_j)gad_{ij} + y - k - M(\mu_j, d_{ij})], \quad (36)$$

where μ_j and λ_j satisfy the relevant first order conditions for $d_{ij} = 1/2 - j$. Since $\mu_j < \lambda_j$, (36) is increasing in d_{ij} and the median valuation of (36) 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.

Wherever the government is located, a democracy chooses non-odious. Then, by previous results, a democracy locates the government at the center.

Proof of Proposition 9

Given odious homogenization μ_j by the ruler in period 1, the choices a democracy makes in period 2 are the same as if the ruler undertook non-odious homogenization of the same amount. Then by the same argument as for non-odious homogenization, the ruler always forms a single country with the government at $j = 1/2$. If $\lambda^* = 0$ the ruler sets $\mu_{j,1} \in [0, \lambda^m_{1/2}]$ to maximize

$$[g + y - k - M(\mu_{j,1}, 0)] + p[g + y - k - [C(\lambda^m_{1/2}) - C(\mu_{j,1})]] + (1 - p)[g + y - k]. \quad (37)$$

Optimal homogenization by the ruler may be positive. If $\lambda^* > 0$ and $\lambda^* \geq \lambda^m_{1/2}$ he sets $\mu_{j,1} = \lambda^*$ if

$$[g + y - k - M(\lambda^*, 0)] + p[g + y - k] + (1 - p)[g + y - k]$$
$$\geq [g + y - k] + p[g - (1 - \lambda^m_{1/4})ga/4 + y - 2k - C(\lambda^m_{1/4})] + (1 - p)[g + y - k]$$

rewritten

$$p(1 - \lambda^m_{1/4})ga/4 + pk + pC(\lambda^m_{1/4}) - M(\lambda^*, 0) \geq 0 \tag{38}$$

and sets $\mu_{j,1} = 0$ otherwise. This is increasing in p. If $\lambda^* > 0$ and $\lambda^* < \lambda^m_{1/2}$ the ruler sets $\mu_{j,1} \in [\lambda^*, \lambda^m_{1/2}]$ (where $\mu_{j,1} \in [\lambda^*, \lambda^m_{1/2}]$ maximizes the left hand side of (39)) if

$$[g + y - k - M(\mu_{j,1}, 0)] + p[g + y - k - [C(\lambda^m_{1/2}) - C(\mu_{j,1})]] + (1 - p)[g + y - k] \tag{39}$$
$$\geq [g + y - k] + p[g - (1 - \lambda^m_{1/4})ga/4 + y - 2k - C(\lambda^m_{1/4})] + (1 - p)[g + y - k]$$

rewritten

$$p[(1 - \lambda^m_{1/4})ga/4 + k + C(\lambda^m_{1/4})] - pC(\lambda^m_{1/2}) + pC(\mu_{j,1}) - M(\mu_{j,1}, 0) \geq 0 \tag{40}$$

and sets $\mu_{j,1} = 0$ otherwise. This is increasing in p since the term $\frac{\partial \mu_{j,1}}{\partial p}[p'C(\mu_{j,1}) - M_{\mu_{j,1}}(\mu_{j,1}, 0)] = 0$ from (37).

Second, show under what conditions avoiding secession is too costly even for $p = 1$. For $p = 1$ then (40) is positive. For $p = 1$ (38) is negative when

$$(1 - \lambda^m_{1/4})ga/4 + k + C(\lambda^m_{1/4}) < M(\lambda^*, 0). \tag{41}$$

By the optimality of $\lambda^m_{1/2}$ for $d_{ij} = 1/4$

$$(1 - \lambda^m_{1/4})ga/4 + C(\lambda^m_{1/4}) > (1 - \lambda^m_{1/2})ga/4 + C(\lambda^m_{1/2}) > M(\lambda^*, 0)$$

so (41) never holds whenever $\lambda^* \leq \lambda^m_{1/2}$.

A1

A population that would otherwise split may stay together with a technology of homogenization.

When there is no option to homogenize ($\lambda_j = 1$), the value of forming a single country versus splitting into two for the median voter is $-ga/4 + k$. When homogenization is possible, the value of forming a single country versus splitting into two for the median voter is

$$\left[g - (1 - \lambda^m_{1/2})gal_i + y - k - C(\lambda^m_{1/2})\right] - \left[g - (1 - \lambda^m_{1/4})ga(1/4 - l_i) + y - 2k - C(\lambda^m_{1/4})\right] \tag{42}$$

where $l_i = ((1-\lambda^m_{1/2}) + (1-\lambda^m_{1/4}))/8(1-\lambda^m_{1/4})$. To see that (42) is strictly higher than when there is no option to homogenize, observe that the expression

$$[g-(1-\lambda)gal_i + y - k - C(\lambda)] - \left[g-(1-\lambda^m_{1/4})ga(1/4-l_i) + y - 2k - C(\lambda^m_{1/4})\right]$$

when $\lambda = \lambda^m_{1/4}$ is

$$-(1-\lambda^m_{1/4})ga(2l_i - 1/4) + k$$

which is strictly higher than when no homogenization is allowed. Then differentiating this expression with respect to λ we get

$$\frac{(1-\lambda)}{(1-\lambda^m_{1/4})}ga/4 + ga/4 - C'(\lambda) \tag{43}$$

we see this is positive for all $\lambda \in [\lambda^m_{1/4}, \lambda^m_{1/2}]$.

A2

Example of an odious homogenization technology.

Without loss of generality the cost function will take the form $M(\mu_j, d_{ij}) = \beta(\mu_j) + \alpha(\mu_j)d_{ij}$. Let us take the particular function where $\beta(\mu_j) = C(\mu_j)/2$ for all μ_j. Intuitively this implies that the cost of this technology is divided into two parts. A cost that we can think of as the cost of using government apparatus (this could be enforcement for example), $C(\mu_j)/2$, which is divided equally among the population through taxes. The rest of the cost burden is shouldered in greater proportion by those who are homogenized by more. One way to think of this is as a personal cost, for example of being forced not to speak one's own language. Since we assume $\beta(\mu_j) = C(\mu_j)/2$ and from Proposition 8 we know $M(\mu_j, 1/4) = C(\mu_j)$, then we calculate that $\alpha(\mu_j) = 2C(\mu_j)$. Thus in this example $M(\mu_j, d_{ij}) = C(\mu_j)/2 + 2C(\mu_j)d_{ij}$. It is clear this satisfies all conditions: $M(0, d_{ij}) = 0$; $M_{\mu_j}(\mu_j, d_{ij}) = C'(\mu_j)/2 + 2C'(\mu_j)d_{ij} > 0$; it is then clear that $M_{\mu_j}(0, d_{ij}) = 0$, $\lim_{\mu_j \to 1} M_{\mu_j}(\mu_j, d_{ij}) = \infty$, $M_{\mu_j\mu_j}(\mu, d_{ij}) = C''(\mu_j)/2 + 2C''(\mu_j)d_{ij} > 0$. The marginal cost is higher for those who are homogenized by more since $M_{\mu_j}(\mu, d_{ij})$ is higher for higher d_{ij}. Finally it can be seen that for $j = 1/2$ the total cost sums to $C(\mu_j)$.

A3

Allowing for $\lambda_j \in [-1, 1]$ does not change any of the results.

From the proof of Proposition 7, for a static democracy, preferences over homogenization $\lambda_j \in [-1, 1]$ remain single peaked and ideal homogenization for any individual i does not change. For a static dictator, homogenization has no effect on his utility and he continues to choose zero homogenization.

In the proof of Proposition 7 we show that Proposition 3 extends to negative homogenization. To show that Proposition 4 continues to hold and a ruler will never choose negative homogenization, we show a ruler always does better by choosing $\lambda_{j,1} = 0$ than any $\lambda_{j,1} < 0$. Then since his total expected utility from any $\lambda_{j,1} \geq 0$ does not change, Proposition 4 does not change. Suppose in period 1 the ruler forms a single country with $j = 1/2$ and undertakes homogenization $\lambda_{j,1}$. A period 2 ruler will continue to form a single country with zero homogenization, whatever amount of homogenization is chosen in period 1. If $\lambda^* \leq 0$, the ruler's period 1 utility and period 2 expected utility from $\lambda_{j,1} = 0$ is

$$[g + y - k] + p[g + y - k - C(\lambda^m_{1/2})] + (1 - p)[g + y - k]$$

From $\lambda^* \leq \lambda_{j,1} < 0$ this is

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

which is strictly lower (and so is any $\lambda_{j,1} < \lambda^* \leq 0$). If $\lambda^* > 0$ then from Proposition 7, any $\lambda_{j,1} < 0$ results in a split so utility if $\lambda_{j,1} = 0$ is

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

and utility for any $\lambda_{j,1} < 0$ is

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

which is strictly lower. To complete the result, observe in Lemma 3 that for any $\lambda_{j,1} < 0$ the ruler does strictly better by forming a single country in period 1 rather than splitting. The argument when we introduce odious homogenization is the same.