

Nation-Building and Education*

Alberto Alesina
Harvard University and IGER Bocconi

Paola Giuliano
UCLA

Bryony Reich
Northwestern University

Revised: February 2020

Abstract

Democracies and dictatorships have different incentives when it comes to choosing how much and by what means to homogenize the population, i.e. “to build a nation”. We study and compare nation-building policies under the transition from dictatorship to democracy in a model where the type of government and the borders of the country are endogenous. We find that the threat of democratization provides the strongest incentive to homogenize. We focus upon a specific nation-building policy: the provision of mass primary education. We offer historical discussions of cases of nation-building across time and space, and provide correlations for a large sample of countries over the 1925-2014 period.

*We thank Nezhil Guner and five anonymous referees for comments that greatly improved the paper. For helpful feedback and comments, we also thank Tim Besley, Martin Cripps, Jeffrey Frieden, Oded Galor, Terri Kneeland, Mark Koyama, Alessandro Riboni, Enrico Spolaore as well as participants in many seminars and conferences. Kevin Bi, Elisabetta Campagna, Nicolas Dubost, Giulia Giupponi, Andrea Passalacqua, Mikhail Poyker, Sebastian Ottinger and Logan Spencer provided excellent research assistance.

1 Introduction

“There cannot be a firmly established political state unless there is a teaching body with definitely recognized principles. If the child is not taught from infancy that he ought to be a republican or a monarchist, a Catholic or a free-thinker, the state will not constitute a nation; it will rest on uncertain and shifting foundations; and it will be constantly exposed to disorder and change.” Napoleon I, 1805¹

From the French Revolution and throughout the 19th century, French rulers expressed the imperative “to form French citizens.”² Following the unification of Italy in 1860, Massimo d’Azeglio, a member of the Northern elite that led the unification process, remarked: “Italy has been made; now it remains to make Italians.” During the 19th and early 20th centuries, French and Italian rulers adopted a range of policies aiming to build commonality among the population and to form suitable “Frenchmen” and “Italians.” A critically important policy to this end was the introduction of state-controlled education, including compulsory elementary schooling. Other nation-building policies included the introduction of a “national language” in schools, religious services, and administration, and the introduction of compulsory military service, which often had the explicit aim of integrating and mixing individuals from different parts of the country.

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

The goal of this paper is to analyze nation-building through education in different political regimes. The terms state-building and nation-building are sometimes used interchangeably. However, state-building generally refers to the construction of bureaucratic institutions necessary for a functioning state, one able to collect revenues, provide public goods, fight wars, etc., while nation-building is the construction of a national identity, which in turn may also facilitate state-building. Thus, we define “nation-building” as a process leading to the formation of countries in which the citizens feel a sufficient amount of commonality of interests, goals, and preferences that they do not wish to separate from each other. We allow the degree of divergence of preferences amongst the population to be endogenous, in the sense that it can be affected by nation-building policies (like public education) chosen by the government. We explicitly model nation-building policies.

We consider different types of regimes (non-democratic and democratic) and their different incentives to nation-build. We begin with a completely secure non-democratic

¹Quote from Ramirez and Boli (1987).

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

ruling elite (the “ruler” for short) who is not threatened by any overthrow. The ruler extracts rents from his territories. He builds the type of government and adopts policies that match his preferences. He has no interest in (costly) nation-building. We show that a non-democratic ruler facing a probability of overthrow and the establishment of democracy has very different incentives to nation-build. If overthrow and democracy were to occur, the ruler would lose his rents, and the newly formed democracy may choose public goods and policies that differ from the preferences of the ruler or elite. In addition, when installed, a democratic regime may break apart the territories of the ruler in ways which would be costly for him. Homogenization and indoctrination allow a ruling elite to better maintain their preferred policies and a larger country if democracy prevails. More homogenization, if it reduces distaste towards the existing government, may also reduce the incentive of the population to overthrow the ruler in the first place. In more colorful terms: threatened rulers will indoctrinate people in order to teach them to “enjoy” the current regime and the current borders of the country and not break away from them.³

Well-functioning democracies also have reasons to promote the homogenization of their citizens. Within a country, individuals share the same public goods and policies, and therefore increasing commonality can increase welfare. Our model shows that nation-building is lowest in a “safe” dictatorship, while elites threatened by democratic revolutions undertake the highest levels of homogenization. Our main result finds that the threat of democratization provides strong incentives to implement homogenization policies. We show that this implication of the model is consistent with the data using historical examples from different continents and time periods, and econometric evidence on a large sample of countries.

Education facilitates nation-building in several ways. It can change individual preferences by indoctrination, that is, by convincing individuals who dislike the ruling government that it is not so bad after all. Cantoni et. al. (2017) show that a Chinese education reform, introduced with the explicit aim of shaping ideology, shifted the attitudes of students towards the ideological position of the government in aspects such as their view of free market economics and the political system. Schools, say in France or Scandinavia, emphasize the benefits of regulation and social welfare while in the UK and the US the merits of individualism are stressed more (Alesina and Glaeser, 2005). Mass education can also facilitate nation-building by teaching a common language. For example, in France and Italy the language of the elites became, via mass instruction, the national language. Teaching a common language can help individuals to better communicate with the government and access public services. You (2018) studies the effect of the Chinese reform in 1960 which enforced the use of Mandarin in all schools in China with the explicit goal of reducing diversity. Clots-Figueras and Masella (2013) show that compulsory Catalan language education encouraged Catalan identity.

While the social benefits and redistributive effects of public education are obvious and important, most state education systems were established before democratization, without any apparent social welfare motive. We contribute to a literature that examines the introduction of mass education by elites with goals other than social welfare in mind. Gellner (1983) (amongst others) argues that industrial societies, based upon markets,

³In this paper, we focus on internal factors that motivate governments to implement nation-building policies. Aghion, Jaravel, Persson, and Rouzet (2019) and Alesina, Reich, and Riboni (2017) study the importance of external motives for nation-building, namely the threat of external wars. Internal and external motives to nation-build may coexist as we will show below.

require better means of communication and that education was a response to this. In recent work, Hauk and Ortega (2015) provide a framework that examines the choice of elites to invest in education to increase mobility and productivity following an industrialization shock. War is also a motive to establish mass education, to build a better educated and more obedient army (Aghion, Jaravel, Persson, and Rouzet, 2019; Ramirez and Boli, 1987; Darden and Mylonas, 2016).⁴

Our framework shows that the transition to democracy is a core driver of mass education. Elites invested in mass education to homogenize nations, both in order to reduce the threat of overthrow and to maintain their interests under a future democracy.⁵

Our model can also account for “reverse homogenization” namely policies of “divide and rule.” While native elites may have incentives to spend parts of their rents to homogenize the population, colonizers do not. They do not care about building a nation for the long run since they are there simply to extract rents for the short or medium run. Therefore they may even engage in policies that pitch one ethnic group against another. These policies may prevent the formation of a coalition of native ethnicities against the invaders from Europe, precisely divide and rule.

Our paper is related to several strands of literature. The first is research which compares education policies across democratic and non-democratic regimes. Aghion, Jaravel, Persson, and Rouzet (2019), using annual data on 137 countries from 1830–2001, find that autocracies have higher enrollment rates in primary education than democracies. Consistent with this finding, Mulligan, Gil, and Sala-i-Martin (2004) examine cross-country data from 1960-1990 and find that there is no evidence that democracies spend more on public education than non-democratic regimes. Bursztyn (2016) finds that democracies spend less on public education than non-democracies for below-median income countries. Lott (1999) also examines education expenditure data from 99 countries in the period 1985-92 and finds that an increase in totalitarianism increases education spending, again with the strongest effects for lower income countries.

The second strand is the literature on endogenous borders by Alesina and Spolaore (1997, 2003). These authors take the diversity of preferences amongst individuals as given, whereas in our model the degree of divergence of preferences amongst the population is endogenous.⁶

The third strand studies policy reforms implemented by forward-looking elites when they are threatened with overthrow. Acemoglu and Robinson (2000), Aidt and Jensen (2014), and Aidt and Franck (2015) examine how a threat of revolution can prompt elites to democratize for strategic reasons. Acemoglu and Robinson (2008) argue that democratic transitions motivate elites to invest in institutions that allow them to maintain a higher degree of power under democracy and mitigate their economic losses. Besley, Persson, and Reynal-Querol (2016) present evidence that rulers facing a greater threat of loss of power invest in institutional reforms, namely improving executive constraints, to limit the ability of future regimes to act against their interests. Our model is related to both these points. First, we show that elites might nation-build to try to reduce the threat

⁴See 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, national identity, and state formation.

⁵Green (1990) and Hobsbawm (1990) discuss education as a way to protect the social order, but, as far as we are aware, in their work there is no explicit discussion linking this to democratic transitions and the motive of preserving policies that favor the elite.

⁶Bolton and Roland (1997) consider separatism due to income differences rather than differences in preferences.

of revolution itself. Second, our model shows that forward-looking elites threatened with democracy invest heavily in nation-building as a means to maintain their interests and their borders under a future democracy.

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

Finally, our paper is related to the literature on the need for education for the better functioning of institutions, as in Glaeser, Ponzetto, and Shleifer (2007) or Bourguignon and Verdier (2000). Papers by Gradstein and Justman (2002) and Ortega and Tangers (2008) examine schooling as a means to improve communication across groups and so increase growth.

This paper is organized as follows. Section 2 presents the model and Section 3 solves it to examine nation-building via education under different regimes. Section 4 presents some generalizations of the model. Section 5 describes several historical examples that speak to the relationship between mass education, nation-building, and the threat of democratization under non-democratic regimes. Section 6 presents correlations between mass education and the probability of a regime being overthrown for a large sample of 172 countries over the 1925-2014 period. The last section concludes.

2 A Model of Nation-Building

We consider a two-period model in which governments can choose to nation-build. In the first period, a ruler (also referred to as “the elite”) runs the country. In the second period, the population either becomes democratic or remains governed by the ruler. We first take the probability of democratization as exogenous, we later endogenize it.

Homogenization and distance

Assume a continuum of individuals of mass 1, with heterogeneous ideal points distributed uniformly on the segment $[0, 1]$. At time t individual i lives in a country with a government located at j . Individual i 's per period utility function at time t is given by

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

The first term $g(1 - a_t^j d_{ij})$, measures the value of the government to individual i . With the term “government” we refer to a set of public goods and policies provided by the authority, democratic or not. The term d_{ij} is the preference distance of individual i from government j . The value of the government to individual i falls with his distance from the government. We think of distance as the language, cultural, ideological, or preference differences between individual i and the public goods and policies provided by government j . The value a_t^j measures the cost of this distance.⁷ The parameter g is the

⁷For example, Laitin and Ramachandran (2016) show that an individual's language distance from the official language negatively affects his or her socioeconomic outcomes.

maximum utility an individual receives from the government when distance is zero. The remaining terms are income y , which is exogenously given, identical for everyone, and identical across time periods, and taxes in period t , r_t .

We model “homogenization” as a technology which uses state education to reduce the cost of distance from the government. Specifically, government j at time t implements a homogenization policy $\lambda_t^j \in [0, 1]$ such that

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

where a is the initial level of heterogeneity at time zero. Homogenization λ_t^j reduces the costs to individual i of facing policies and public goods j that are different from his ideal. From now on with the term “distance” we summarize any difference in preferences and with the term “homogenization,” a reduction in the costs of such a distance through education policies. Homogenization is durable: languages learned today are not forgotten tomorrow, preferences influenced by schooling today are persistent. The cost of the homogenization policy λ_t^j , for a country of mass s , is

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

where λ_{t-1}^j is the homogenization of this population by government j in the previous period. That is, homogenization by government j in the previous period persists such that $\lambda_t^j \geq \lambda_{t-1}^j$ and the cost of homogenization this period covers only any additional homogenization. We assume that homogenization policies influence the cost of distance from a government. Therefore homogenization by a different previous government is irrelevant if the location of the government changes.⁸

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

The cost of the homogenization policy is paid with period t taxes. We initially assume the cost of homogenization is split equally among the population of the country. In Section 4 we generalize this.⁹

Country Formation

In period 1, a non-democratic ruler is located at $1/2$. We model the ruler or elite as measure zero. Alternatively, we could represent the elite by a group of mass δ with their

⁸For example, teaching French reduces the cost of accessing government services and jobs provided in French but it does not necessarily affect the cost of accessing government services and jobs provided in German. An alternative way to model homogenization policies is that they work by influencing preferences directly. In our model, this would mean that homogenization shifts an individual’s ideal point and therefore the term d_{ij} . These two assumptions both work through the term $a_i^j d_{ij}$ and push this term in the same direction, and we think of them as interchangeable. The only difference between them is that from a modeling perspective a shift in ideal point towards government j will also change the distance of an individual’s ideal point vis a vis other government locations and this can complicate the analysis.

⁹In our model, income is exogenous. However, at least up to a point, diversity of skills, education, background, and culture may increase productivity. In this case a reduction in diversity would have costs and benefits. The latter are already modeled. The former would include not only the costs modeled above but also a reduction in productivity, therefore of income. Given that income enters linearly in the utility function and taxes are lump sum, we can treat a reduction in productivity as part of the tax.

ideal point at $1/2$ or in an interval around $1/2$. Such an extension would complicate notation and algebra with little advantage in terms of insight. In period 2, either the non-democratic regime survives, or democracy prevails. In a democracy, the population will choose whether to maintain the borders of the single country or to split into two equal-sized countries, A and B , comprising the intervals of ideal points $[0, 1/2]$ and $(1/2, 1]$ respectively. This split of the population assumes that preferences are perfectly correlated with geography.¹⁰ If there were no correlation between geography and preferences we might end up with countries that are not geographically connected. Alesina and Spolaore (2003) provide a discussion and justification of this assumption in the context of a model of country formation. We adopt the restriction of at most two equal-sized countries for simplicity.¹¹ We do not allow for unilateral secession, namely a situation in which, without any majority vote, a group of citizens form a third country. Each country has a single government located at some point within it, j . The borders and the location of the government can be altered by a democracy at the beginning of period 2 at no cost.

The cost of government in period t is k . Since k can be divided amongst all citizens in the country, this captures the benefits of a larger country.¹² When the population splits into two countries, each country is more homogeneous and the government of each country is closer to the median citizen, but the costs of government are higher in per capita terms. In a democracy, therefore, the voters face a trade-off between homogeneity and the costs of government. 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. Only a democracy in period 2 would have an incentive to separate. An elite would never split the country since they would have to provide two governments.

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

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

where for now we assume that taxes r_t are identical for all individuals in the country. The costs of government and costs of homogenization are split equally among the population. We relax this assumption in Section 4.

Decision-Making and Timing

In the initial period the elite runs the country, then in the second period a democracy may prevail. Utility for an individual i in period t , is denoted u_{it} and is given by (1). Utility for the ruler located at $i = 1/2$ in period t is denoted $U_{it} = u_{it} + R$, where u_{it} is given by (1) and $R \geq 0$ is rents received by the ruler. Under a non-democratic regime, the utility of the ruler or elite is equal to the same utility as a normal citizen located at $i = 1/2$, plus some exogenously given positive rents $R > 0$. Under a democracy, the ruler or elite does not collect rents and so $R = 0$ and the ruler or elite receives exactly the same utility as a normal citizen who is located at $i = 1/2$. In our model, rents are always

¹⁰See Campante, Do, and Guimaraes (2019) and Michalopoulos and Papaioannou (2014) for evidence consistent with this assumption.

¹¹Alesina 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.

¹²Alesina, Spolaore, and Wacziarg (2000) and Alesina and Spolaore (2003) investigate the sources of benefits of size, like the dimension of the market and diversity of inputs in productivity, or economies of scale in the provision of some public goods.

chosen at the maximum feasible level. The choice of rents by the ruler is relevant only when democratization is endogenous (Section 3.3) and we discuss endogenizing rents at that point in the paper.

Period 1

An elite rules the population and locates the government at $1/2$. The elite decides how much to invest in homogenization in period 1 to maximize expected utility $U_{i1} + E[U_{i2}]$, where $E[U_{i2}]$ takes into account the probability of democratization in period 2.

Period 2

If the elite remains in power in period 2, it leaves the government at $1/2$, and chooses homogenization to maximize utility U_{i2} .

If democracy arises, decisions are made by majority rule with the order of voting as follows:

- (1) the population decides whether to form a single country or split into two;
 - (2) the population of each country decides where to locate the government in that country;
 - (3) the population of each country decides the homogenization policy in that country.
- For tie-breaking, we assume that when indifferent between one country or two, the single country remains in place.

3 Solving the model

We solve the model backwards, starting with the decisions made by a democracy in period 2. We then consider the decisions of a ruler or elite in the different cases when democratization is exogenous and endogenous.

3.1 A Democracy

If democracy prevails in period 2, the population chooses whether to form a single country or split, where to locate the government, and how much to homogenize. All the decisions are made by majority rule and each individual i faces utility u_{i2} given by (1). Rents for the former elite are, of course, zero under democracy and they also face utility u_{i2} . The value of u_{i2} is influenced by the former elite's choice of homogenization in period 1. These choices made by a democracy in period 2 are solved in detail in the Appendix. Here we summarize the basic intuition.

For individual i the level of homogenization which equalizes the marginal cost and marginal benefit is given by:

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

The optimal level of homogenization for individual i depends upon the distance of individual i from the government and the cost of the homogenization technology. Since, for now, we assume that the cost of homogenization falls equally on those close to and far from the government, homogenization is a transfer from the center to the periphery, since the latter benefits more. In other words, a technology that reduces the distance

to the government may be especially beneficial to people with distant preferences. Of course, distant minorities may also resist homogenization attempts, distant minorities may be made to pay more for homogenization via a higher tax bill, or may be the target of more extensive homogenization or more painful repressive policies. In Section 4 we model unequal costs of homogenization.

Preferences over homogenization are single-peaked, thus a democracy homogenizes up to the point at which the marginal cost of homogenization equals the marginal benefit for the individual at median distance from the government. If homogenization by the ruler in period 1 exceeds this amount, then no additional homogenization will be undertaken by a democracy in period 2.

The “preference” interpretation of homogenization, literally speaking, implies that an individual “chooses” a policy that changes her preferences, knowing that after the change she would feel happier in the country in which she lives. This argument becomes more plausible if we think of a dynamic extension in which parents transmit values and educate their children in such a way which makes them fit better in the country in which they live by adopting certain social norms and types of behavior.¹³ This is not contradictory to strong attachment to cultural values which can be captured by very high costs of homogenization.

A democracy locates the government at the median ideal point in the population, namely the center of the country. Thus, in a single country, the government is located at $j = 1/2$. In Countries A and B, the government is located at $j = 1/4$ and $3/4$ respectively. These results are illustrated in Figure 1 below. We denote by $\lambda_{1/2}^m$ the optimal homogenization of the median voter for the homogenization decision in a single country. We denote by $\lambda_{1/4}^m$ the optimal homogenization of the median voter for the homogenization decision in a split country.

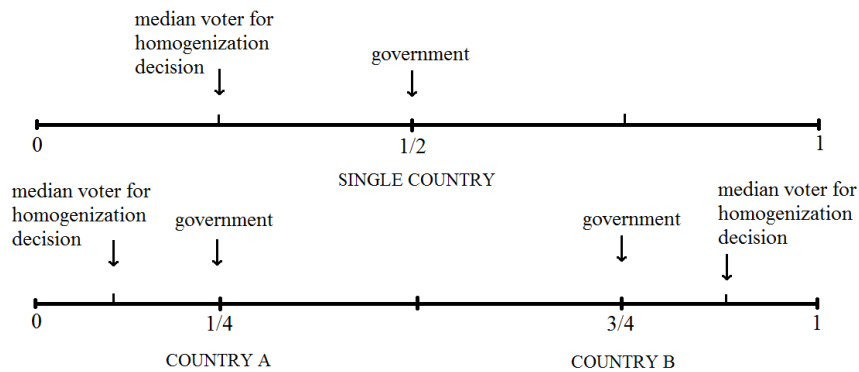


Figure 1: Homogenization and government location under a democracy, for a single country and a split.

The choice of whether to form a single country or split captures the trade-off between the benefits of a larger country and the costs of heterogeneity. In our model, however, a democracy also has the option of homogenizing. It is perfectly possible that with

¹³For models related to parents “choosing” values for children see Bisin and Verdier (2000). Algan, Mayer, and Thoenig (2013) discuss the costs of lack of assimilation in France. They document a substantial increase in salaries for children of Arabic families who signal assimilation by choosing French rather than Arabic first names.

($\lambda_t^j = 0$) a democracy would decide to split into two countries, but the option of choosing $\lambda_t^j \in [0, 1]$ would lead a democracy to homogenize somewhat and form a single country. This is democratic nation-building.

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

Lemma 1 *For a given g , a , k , and $C(\cdot)$, there exists a level of homogenization $\lambda^* \in [0, 1)$ such that:*

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

The proof is in the Appendix. Homogenization implemented by the ruler in period 1 changes citizens' relative payoffs from different types of government: it increases a citizen's period 2 utility from the ideal government of the ruler at $1/2$ relative to other government locations. It also reduces the costs of heterogeneity and so makes separation less attractive. Enough homogenization will, therefore, change the choices of a democracy in period 2 in the direction of the ruler.

3.2 Democratization

Suppose that with exogenous probability p , democracy arises in period 2. Given the choice of homogenization by the elite in period 1, a democracy behaves as described above in period 2. Here we consider the homogenization choices a ruler makes in period 1 and, if still in power, in period 2.

With probability $1 - p$ the elite maintains power in period 2. Since this is the final period, there is no threat of democracy, the elite maintains the same ideal government and undertakes no additional homogenization. The elite's utility, in this case, is $g + y - k + R$. With probability p , democracy prevails in period 2 and the elite's utility u_{it} is given by (1) where the country choice, location of the government, and homogenization choice in period 2 are made by majority rule.

In period 1 the elite is forward-looking. Define $\lambda_1^{1/2}$ as the homogenization level chosen in period 1 by the elite located at $1/2$. Suppose $\lambda_1^{1/2} < \lambda^*$. From Lemma 1, a democracy would choose to split and the elite is located at the border between the two countries, thus at the farthest point from the two governments. The elite's expected utility is

$$\left[g + y - k - C(\lambda_1^{1/2}) + R \right] + p \left[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m) \right] + (1-p) [g + y - k + R].$$

Suppose $\lambda_1^{1/2} \geq \lambda^*$. From Lemma 1, a democracy chooses a single country and the government is located at $1/2$, which is the elite's ideal point, although the elite loses its rents. The elite's expected utility if $\lambda_1^{1/2} \geq \lambda_{1/2}^m$ is

$$\left[g + y - k - C(\lambda_1^{1/2}) + R \right] + p [g + y - k] + (1 - p) [g + y - k + R]. \quad (2)$$

The elite faces a trade-off between paying a higher cost of homogenization in period 1, which is of no direct benefit to the elite in period 1, but which ensures the elite's ideal government and a single country in period 2, should democracy prevail. In other words, homogenization enables the elite to maintain public goods and policies under democracy that are liked by the elite; the only loss would be the rents. The more likely is the event of democratization, the more the elite is willing to invest in homogenization. When the probability of democracy is sufficiently high and homogenization is not too costly, then the ruler will nation-build to ensure that, under a democracy, the country does not split. This brings us to Proposition 1.

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

The next corollary builds on Proposition 1 and shows that secure rulers undertake less homogenization than would be undertaken by a democracy, while unsafe rulers homogenize more than a democracy would. Corollary 1 also highlights that homogenization by the ruler has a long-run impact on the homogeneity of a population even once that population becomes democratic. Take two identical populations which both become democratic in period 2. But suppose in period 1 one of these populations is governed by a ruler who faces a low probability of democracy and one is governed by a ruler who faces a high probability of democracy. The ruler who faces a high probability of democracy has enacted more homogenization and, after democratization, this population remains more homogenous and forms a larger country than the other. The implication is that today's democracies which followed a "smooth" path to democracy (where elites foresaw the advent of democracy) may be more homogeneous and bigger than they would be otherwise as a result of nation-building by those elites. A testable implication for future research.

In Corollary 1, we refer to "baseline" homogenization as the level of homogenization that would be undertaken by the democracy in period 2 if previous homogenization was zero.

Corollary 1 *For a given g , a , k , and $C(\cdot)$, there exists a threshold $\bar{p} \in (0, 1]$ such that:*

- (i). *for $p \leq \bar{p}$, the period 1 ruler chooses a strictly lower level of homogenization than baseline and, if democracy prevails in period 2, the democracy will increase homogenization to baseline;*
- (ii). *for $p > \bar{p}$, the period 1 ruler undertakes homogenization such that if democracy prevails in period 2 then homogenization in period 2 will be higher than baseline.*

When the probability of democracy is low, a "safe" ruler has little incentive to homogenize. A safe ruler has his ideal government, faces little threat of overthrow and break-up, and has no concern for general welfare, so he is largely unconcerned with the heterogeneity of the population in period 1 and expects the same in period 2. In contrast, a democracy homogenizes to improve the welfare of its citizens, particularly those at the periphery. Thus a democracy undertakes more nation-building than a relatively safe non-democratic regime.

When the probability of democracy is high, an “unsafe” ruler will homogenize in period 1 and a democracy can end up more homogenous than it otherwise would have been under the baseline. Under some parameters, a ruler will homogenize in period 1 to such an extent as to avoid secession and ensure his ideal government persists in period 2; whereas, without any homogenization by the ruler, a democracy in period 2 would choose less homogenization, split, and opt for governments representing preferences very different from the ruler’s.¹⁴ Thus an unstable non-democratic regime may over-invest in homogenization compared to a democracy in order to ensure the regime’s ideal government is preserved.

3.3 Endogenous Democratization

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

Suppose, as above, a democratic opportunity arrives at the beginning of period 2 with probability p . We now assume that if a democratic opportunity arises, the population can decide whether or not to overthrow the non-democratic regime and install a democracy. If the population chooses overthrow, then democracy prevails in period 2; if not, the ruler continues to hold power. The cost of overthrowing the ruler is $L > 0$. If a democratic regime is installed, then the utility attained by individual i in period 2 is denoted $u_{i2,dem}$, and if a non-democratic regime is in power in period 2, then the utility attained by individual i is denoted $u_{i2,ruler}$. Given the choices in period 1, the choices a democracy or a ruler will make in period 2 are known, and so the values of $u_{i2,dem}$ and $u_{i2,ruler}$ are known at the beginning of period 2. Individual i prefers overthrow if

$$u_{i2,dem} - u_{i2,ruler} - L \geq 0. \quad (3)$$

An overthrow occurs if a majority prefers it to the status quo, thus the median voter is critical. Consider the median value of (3) as a measure of opposition to the ruler’s regime.

Proposition 2 *The opposition to the ruler’s regime is decreasing in homogenization by the ruler in period 1. For a given g , a , k , L , and $C(\cdot)$, there exists a threshold, $\bar{\lambda} \in [0, 1)$, such that if the ruler homogenizes to $\bar{\lambda}$ or above in period 1, the population will choose not to overthrow the ruler.*

Proposition 3 describes the choices of a ruler.

Proposition 3

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

¹⁴There are two situations under which $\bar{p} = 1$ in Corollary 1, implying a ruler undertakes less homogenization than a democracy whatever the probability of overthrow. These situations are straightforward to interpret. These occur: 1. when homogenization is extremely costly and the ruler cannot preserve his ideal government without a very large cost, and 2. when the ruler’s ideal policies are preserved anyway with very little or no homogenization.

The proof is in the Appendix. Two forces now generate the positive relationship between homogenization and threat of democracy. In addition to the one discussed above, now homogenization also reduces the probability of a democratization and the loss of rents, R . It follows that the higher the rents available to the ruler, the more homogenization he will undertake to try protect those rents. In other words, rulers who extract more rents are also more willing to homogenize using possibly unpleasant means.

Corollary 2 *In period 1, the ruler undertakes a level of homogenization which is (weakly) increasing in the size of rents, R .*

3.3.1 Endogenous Rents

If rents could be manipulated by the ruler, he might also want to reduce them strategically to lower opposition to the regime. Rents, in other words, could be a further tool, in addition to nation-building policies, to try to avoid overthrow. Let us briefly consider what happens if rents are endogenous. Suppose the elite can choose rents R in the range $[0, \bar{R}]$, and suppose the cost to a normal citizen of this rent extraction is $E(R)$, which is increasing in the rents extracted R . So the utility of a normal citizen i at time t if the elite are in power and choose to extract rents R is $u_{it} - E(R)$, where u_{it} is given by (1). As above, with probability p , a revolution opportunity occurs at the beginning of period 2. If a revolution opportunity occurs, the population decides whether or not to overthrow the regime and install a democracy. An individual prefers to overthrow if

$$u_{i2,dem} - u_{i2,ruler} - L \geq 0,$$

where overthrow occurs if it is preferred by a majority. The ruler can lower rents R to increase the utility of normal citizens under the ruler in period 2, $u_{i2,ruler}$, and thereby reduce opposition to the regime. However, for this to work as a way to avoid overthrow, the ruler needs some way to commit to low rents in period 2. If the ruler promises low rents but has no way to commit to them then, once the population chooses not to overthrow the ruler, the ruler can go ahead and extract the maximum rents in period 2. The idea that the promise of future redistribution to avoid overthrow suffers from a credibility problem is made in Acemoglu and Robinson (2000).

We assume exogenous rents because once we allow for endogenous rents in our model, lowering rents is not credible and therefore not a tool to avoid overthrow. In contrast, because homogenization is persistent, nation-building to avoid democratization does not suffer a commitment problem.

3.4 Divide and Rule

Our framework uncovers two different motives to homogenize when faced with a threat of democratization. One motive is to reduce the threat of democracy. The other is to build a more homogeneous nation that reflects the ruler's preferences, so that if democracy prevails the population will choose to maintain the status quo. Notice that the relevance

of each motive depends on the “type” of non-democratic regime. The motive to homogenize to maintain the status quo after democratization only applies if the ruler or elite expects to stay in the country after democratization. It does not apply if a ruler or elite expects to be kicked out or physically eliminated if democracy prevails. For example, harsh dictators might expect to be kicked out or captured and punished/killed, in which case this incentive to homogenize is not relevant. In contrast, in 19th century Europe, elites largely foresaw the advent of democracy which progressed over the course of that century. European elites were in most part not eliminated and remained part of the new democracies. Thus European elites faced an incentive to homogenize to try to preserve policies that favored their position in society. The incentives of native elites which expect to remain in the country and elites which know they will leave are different.

We now consider the possibility that homogenization policies make democratization more likely. In some cases, revolutions are considered to be more likely when a population is more homogeneous since a more homogenous population can communicate better and this might make collective action easier. By the same argument, policies that increase diversity and its costs can hinder collective action, implying rulers may face an incentive to increase heterogeneity. Indeed the term “divide and rule” is coined to capture precisely this effect. Thus, in this subsection, when introducing into the framework the idea that more homogeneity can make democratization more likely, we also give rulers the option to increase heterogeneity within the population. That is, we allow for both positive and negative homogenization to specifically model the notion of “divide and rule”.

We allow for positive and negative homogenization policies, $\lambda_t^j \in [-1, 1]$, in any period and by any regime. We need to update the assumption on costs to allow for negative homogenization (we state this updated assumption in the proof of Proposition 4 in the Appendix). We now assume that the probability of a democratic opportunity is given by $v(p, \lambda_1^{1/2})$, which depends on both $p \in [0, 1]$ (some exogenous factors affecting the likelihood of democratization), and the homogenization undertaken by the ruler in period 1, denoted $\lambda_1^{1/2}$. The function $v : [0, 1] \times [-1, 1] \rightarrow (0, 1)$ is strictly increasing in p , strictly increasing and convex in $\lambda_1^{1/2}$, and twice continuously differentiable. That is, a higher exogenous threat of a revolution opportunity and higher homogeneity both increase the probability of a revolution opportunity occurring. If a revolution opportunity occurs, we assume democracy arises in period 2. Convexity in $\lambda_1^{1/2}$ ensures a unique optimal homogenization policy. Otherwise, the framework is exactly as detailed so far. We make one simplifying technical assumption, that a democracy always locates the government at the center of the country. Under a sufficient condition, detailed in the Appendix, which implies that the marginal effect of p on the revolution opportunity is not too sensitive to homogenization, we obtain Proposition 4.

Proposition 4 *For a given $g, a, k, C(\cdot)$, and $v(\cdot, \cdot)$, there exists a threshold \hat{p} such that in period 1:*

- (i). *if $p \leq \hat{p}$, the ruler undertakes strictly negative homogenization (divide and rule policies);*
- (ii). *if $p > \hat{p}$, the ruler undertakes positive homogenization.*

A democracy never chooses negative homogenization. However, a period 1 ruler may choose to undertake strictly negative homogenization, that is to divide and rule. A ruler

has an incentive to increase the costs of diversity (to divide and rule) only when homogenization increases the probability of collective action. The proof is in the Appendix.

The ruler faces two conflicting forces. On the one hand, if he implements the divide and rule policy he makes collective action more difficult and reduces the probability of a revolution. On the other hand, if he implements the divide and rule policy and democracy prevails, then the country may split and the new government may not reflect the interests of the elite. Homogenization to ensure the status quo post-democratization is costly and therefore when conditions make democracy unlikely, p low, the incentive to divide and rule dominates. When conditions favor democracy, p high, the probability of democracy is high enough that the elite wants to put in place safeguards should democracy occur and so the incentive to homogenize dominates.¹⁵

As noted above, the elite's motivation to homogenize in order to maintain certain policies and borders under democracy does not apply if the elite expects to leave if overthrown, which is the case for instance of colonizers. Thus colonizers have lower incentives to homogenize than native dictators and, if homogenization increases the ability of the population to engage in collective action and therefore increases the probability of democratization, it may even be the case that colonizers have incentives to follow policies of "divide and rule."

Higher rents R increase the incentive of the ruler or elite to avoid overthrow and maintain his rents, and so, in this case, higher rents reduce homogenization and increase divide and rule policies.

Corollary 3 *In period 1 the ruler undertakes a level of homogenization which is (weakly) decreasing in the size of rents, R .*

This is in contrast to Corollary 2 where we showed that when homogenization reduces the willingness of the population to overthrow the regime, then higher rents increase homogenization. Thus the effect of higher rents could be to increase or decrease homogenization depending on whether the dominant effect of homogenization is to reduce or increase the chance of overthrow.

4 Extensions

4.1 The Elite's Location

In this section we consider an elite located at any point on the unit interval, $e \in [0, 1]$. The model is as described above, for the case of an exogenous probability of democratization, p . Adding endogenous democratization would make the analysis unnecessarily complicated and it is obvious that the forces from endogenizing democratization push in the same direction wherever the ruler is located. The details of the analysis are in the Appendix.

We make a technical simplifying assumption: if democracy prevails, we restrict a democratic country to a choice of two locations for the government, either at the median location in the democratic country or at the original location of the elite. Our restriction implies the following. If a democracy chooses to remain a single country, then the population chooses between locating the government at $1/2$ or at the original location of the elite. If a democracy chooses to split, suppose the original location of the elite was

¹⁵Since we limit assumptions on the function $v(\cdot, \cdot)$, we can also have the degenerate cases in Proposition 4 where homogenization is negative for all $p \in [0, 1]$ or homogenization is positive for all $p \in [0, 1]$.

between 0 and 1/2, then Country A chooses between locating the government at 1/4 or at the original location of the elite. Country B chooses to locate the government at 3/4. The other way around if the elite was originally located between 1/2 and 1.

Under these assumptions our key results do not change for an elite located anywhere along the unit interval, $e \in [0, 1]$. This is described in Proposition 5.

Proposition 5

- (i). *There exists a level of homogenization, denoted $\lambda_e^* \in [0, 1)$, such that if the ruler homogenizes by at least λ_e^* in period 1, a democracy in period 2 would locate the government at the elite's location, e .*
- (ii). *In period 1, the ruler undertakes a level of homogenization which is weakly increasing in the probability of democratization, p .*

If the elite homogenizes enough, the ruler can ensure that his ideal government persists should democracy arise. A higher threat of democracy induces the ruler or elite to homogenize to a greater extent for exactly the same reason as before. He wants to mitigate the democratic outcome and homogenization is one way to do that.

How does the location of the elite affect homogenization? As an example, consider an elite located in $e \in [0, 1/4]$. The homogenization that is required to ensure that the elite's ideal government persists under democracy increases when the location of the elite moves closer to the periphery (both if the democracy splits or forms a single country). Thus more peripheral elites may undertake more homogenization.

4.2 Different Methods of Homogenization

In this section we introduce a second type of homogenization technology which imposes higher costs on those further from the government. This can capture an unequal tax burden, but can also be interpreted more broadly, as methods of homogenization which result in higher personal costs for more distant minorities. We allow the government to choose between these two different types of homogenization, which we term equal-cost and unequal-cost homogenization. We return to the case of an exogenous probability of democratization, p , and the case of the ruler located in the middle of the country at 1/2.

The costs for individual i of homogenization to level λ_t^j using the unequal-cost technology is $M(\lambda_t^j, d_{ij}) - M(\lambda_{t-1}^j, d_{ij})$, where $M(\lambda_t^j, d_{ij})$ is strictly increasing, strictly convex and twice continuously differentiable as λ_t^j increases from 0 to 1, with $M(0, d_{ij}) = 0$, $M_{\lambda_t^j}(0, d_{ij}) = 0$ and $\lim_{\lambda_t^j \rightarrow 1} M_{\lambda_t^j}(\lambda_t^j, d_{ij}) = \infty$. In contrast to equal-cost technologies, $M(\lambda_t^j, d_{ij})$, is linearly increasing in d_{ij} , the distance of the individual from the government; that is, the cost of homogenization is higher for those who are homogenized by more. We also assume the marginal cost of homogenization is increasing in distance from the government. Thus for each individual i in a population of size s we can write i 's tax burden as $r_t = k/s + [M(\lambda_t^j, d_{ij}) - M(\lambda_{t-1}^j, d_{ij})]$. As mentioned, we can also consider these idiosyncratic costs as personal costs rather than increased taxes directly.

To make comparisons between these two technologies, we impose that the total cost of homogenizing a country to λ_t^j (when $\lambda_{t-1}^j = 0$) has to be the same under both technologies and they differ only in the allocation of costs. That is,

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

The timing now also incorporates the choice over homogenization technologies. A ruler chooses first the homogenization technology and then the amount of homogenization. A democracy chooses first borders, second the location of the government, third the homogenization technology, and finally the amount of homogenization.

Proposition 6 *When both an unequal and equal-cost technology are available, in period 1 a ruler strictly prefers to homogenize using the unequal-cost technology, while in period 2 a democracy weakly prefers to homogenize using the equal-cost technology.*

Unequal-cost technologies are cheaper for the ruler; with the latter the ruler can shift the burden of homogenization onto more distant individuals. In contrast, for a government located at the center of a country, a democracy is indifferent between the two homogenization technologies.¹⁶ This is because choices in a democracy are made by the median voter and so they better represent the preferences of the wider population, some of whom are harshly affected by unequal-cost homogenization. The result that a democracy is exactly indifferent relies on both the linearity in distance of the cost function for the unequal-cost technology and on the fact that the population is distributed uniformly. Allowing for any type of distribution of costs, as well as any distribution of the population, would make the problem intractable. Our modeling device is meant to capture the fact that, in general, a non-democratic regime has more latitude in the allocation of costs, while a democracy considers (to a greater degree, at least) the views of the whole population. In that sense, democracy can limit harsh homogenization policies.

Proposition 7 extends and compares our results to the case where both an unequal and equal-cost technology are available. We assume that when indifferent, the equal-cost technology is chosen. The qualitative results remain, but the means of homogenization undertaken by the ruler is harsher and the level of homogenization is higher.

Proposition 7

- (i). *In period 1 the ruler uses the unequal-cost technology to undertake an amount of homogenization which is (weakly) increasing in p .*
- (ii). *The amount of homogenization undertaken by the ruler is weakly higher, and for some parameter values strictly higher, than when only the equal-cost technology is available.*

Proposition 7 implies, again, that a higher probability of democratization induces the ruler to homogenize by more. Second, greater latitude in homogenization technologies will not only induce the ruler to choose more unequal technologies, but will also induce him to homogenize to a greater extent, because he can use technologies which place more of the costs on minorities. One final point of note. Unequal-cost technologies place

¹⁶Proposition 6 states that a democracy weakly prefers homogenization via the equal-cost technology. This is because for a government not located at the center, a democracy sometimes strictly prefers the equal-cost technology.

high costs on those further from the center, lowering the level of homogenization that is desired by those at the periphery. While distant minorities may have the most to gain from reducing the cost of distance to the government, they may not necessarily welcome homogenization from the center and instead may prefer a split.

5 Historical Examples

In this section we present a range of case studies examining a number of autocratic regimes at various levels of threat of democratization and spanning different geographical regions and periods of time. We also focus on colonial regimes where the ruler is not a native. Overall, the literature we reviewed from political history and comparative education appears to be consistent with our model. Threatened autocratic regimes of all kinds expanded and extended their education programs, in large part, with nation-building ends in mind. 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. In the next section, we present more systematic evidence on a large sample of 172 countries.

The main prediction from our model is that compulsory education came before democratization, because it occurred in response to the rise of that very threat to power. The time variation in the historical analysis broadly aligns with this prediction. With the exception of a few attempts to offer universal education, the phenomenon of the state taking full responsibility for the education of all youth did not arise until the mid-to-late 19th century. Regimes facing the risk of democratization during this period of time, like France, Japan, Prussia, and Britain, instituted compulsory primary education relatively early, while those facing the same threat later, like Turkey, did so much later. In a very comprehensive analysis, Paglayan (2018) provides evidence consistent with this view by showing that compulsory education preceded democratization by an average of half a century in a sample of European and Latin American countries. She shows that twenty years before democratization, the world average primary enrollment rate was already sixty percent. It is clear, then, that states made proactive efforts to ensure that the bulk of their population was educated on the state's terms before voters demanded it. Below we provide a more systematic analysis of different historical cases. For simplicity, we group the main historical experiences by geographical region. We study colonies separately.

Europe

Public policies to educate the population were implemented in force during the 19th and early 20th centuries. During the 19th century, European countries moved from little to no government-supported schooling (and low enrollment rates) to centralized and compulsory primary schooling. This was a significant shift in government policy over a short period of time, made all the more interesting because in many cases it occurred decades before other welfare policies and was generally unpopular with the masses.

France The French Revolution was a turning point for threats to traditional elites. Although the Ancien Régime was a very centralized state, there was little homogenization of the population before the French Revolution (Tilly, 1975) and little, if any, effort to foster more widely a nation of French-speakers. Weber (1979) writes that the French

Crown showed “little concern with the linguistic conquest of the regions under its administration” and that “diversity had not bothered earlier centuries very much [...]”. In fact, the ruling elites made a point of distinguishing themselves from the masses, using language as a barrier (Gellner, 1983). Primary schooling was predominantly provided by the church. (Katznelson and Weir, 1985). Weber (1979) notes “the Revolution had brought with it the concept of national unity as an integral and integrating ideal at all levels.” Schooling was one way to homogenize and, after the Revolution, schools became a key concern of elites. The Constitution of 1791 called for the establishment of free public instruction for all. A major role for schooling was to enforce a national language. The Convention (the legislative assembly from September 1792 to October 1795) decreed that in the Republic children should learn to “speak, read and write in the French language” and that “instruction should take place only in French.” The Jacobins insisted “The unity of the Republic demands the unity of speech.”¹⁷

Weber (1979) notes that “linguistic diversity had been irrelevant to administrative unity. But it became significant when it was perceived as a threat to political - that is, ideological unity.” The first serious attempt to implement mass schooling was made in 1833 following a period of major rebellion (the “July Revolution”, 1830–32). In France, as elsewhere in Europe, the emergence of state intervention in schooling was in no way a concession to the demands of the population; state-provided schooling was, at least into the last quarter of the 19th century, largely unpopular (Katznelson and Weir, 1985; Weber, 1979). What was perhaps the most intense period of schooling reform followed the establishment of the Third Republic in 1870. Hobsbawm (1990) describes this period as one in which the inevitability of a shift of power to the wider population became clear. Schooling was regarded as a key tool in moving the values and way of life of the population towards those of the elite. Weber (1979) highlights the chasm between the way of life and culture of the urban elite and that of the rural masses throughout much of the 19th century. He writes of the perceived need after the Revolution to integrate this part of the population and to make it “French”: “the unassimilated rural masses had to be integrated into the dominant culture as they had been integrated into an administrative entity.” Weber notes “the village school, compulsory and free, has been credited with the ultimate acculturation process that made the French people French - finally civilized them, as many nineteenth-century educators liked to say.”

Policies of homogenization were also motivated by concerns of secession, as highlighted by the case of Brittany. A report on the Breton departments in the 1880s noted that “Brittany, which was not willingly joined to France, which never wholeheartedly accepted its annexation, which still protests” had still to be merged into the nation. The report urged the use of education to “Frenchify Brittany as promptly as possible [...] integrate western Brittany with the rest of France,” and that only schooling could “truly unify the peninsula with the rest of France and complete the historical annexation always ready to dissolve.” 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 Northern France, yet it finally became thoroughly French” (Tilly, 1975). Ensuring French was spoken was considered a vital component in integrating the French population and avoiding secessionist threats. Indeed, the use of languages other than French was viewed as a particular threat to the

¹⁷Both quotations from Weber (1979).

stability of the French state: in 1891, the Minister of the Interior argued that preaching in local dialects “may endanger French unity.” In 1890, a ministerial decree banned religious instruction in Flemish and in 1902, the government banned Breton language sermons.

Italy Italian unification was completed by Northern elites in the 1860s, with virtually no involvement of local populations. Southern regions saw reunification more as a conquest from the elites of the North. At the time of unification, Italy included a diverse population speaking a range of very different languages and dialects. At best, 10% of the population spoke what would become Italian. This was a time of increasing pressure for more democracy (the largest proportion of adult males were enfranchised in Italy in 1912). The governing elite considered homogenization vital to ensure internal stability. Duggan (2007) documents that “during the 1860s the government had embarked on extensive discussions about what form of Italian should be adopted as the national language.” He writes “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.” Holding the country together and avoiding a break up was a major concern and goal of the rulers.

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

England Public education first appeared in minimal form in 1833, following three years of widespread rioting in rural England and the Great Reform Act of 1832. With further political reforms in the 1860s, the “full democratization of the political realm was seen as inevitable” (Ramirez and Boli, 1987). Green (1990) writes that the “Education Act of 1870, which established a quasi-national system, was a result, as much as anything, of the desire to control the political effects of the extension of the franchise in 1867 to the skilled working class.” The connection between democratization and the introduction of mass education can be read directly from the English political debate of the time. The desire to protect the status quo, and the idea of using education to do so, is explicitly stated by those implementing the reforms. British politician Robert Lowe (later Home Secretary and Chancellor of the Exchequer), in an address in 1867, highlighted the urgency for education reforms following the 1867 Reform Act as a means to protect the status quo: “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.”¹⁸ In 1870, when W.E. Forster put forward the bill for his education act in Parliament, his speech argued that if people were given the vote it was necessary to educate them in order to maintain the current system: “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 a popular government [...] now that we have given [the people] political power we must not wait any longer to give them education.”¹⁹

Prussia Frederick II implemented changes to the school system with the “General School Regulations for the Rural Schools” in 1763. The law introduced a common curriculum and was broadly viewed as a way to promote loyalty to the king (Paglayan, 2018). These changes were introduced as a response to peasant revolts that plagued Prussia in 1740 and 1750. Peasants revolted to protest an increase in the days they were required to work, a change demanded by landlords as a reaction to an increase in grain prices. The first attempt of Frederik II to stop peasants’ discontent was the introduction of an agrarian reform. After that response failed, a change in the school system was implemented to re-establish loyalty to the king. Paglayan (2018) notes that historians generally view Prussian autocratic rulers as conceiving compulsory schooling to impose social control and to indoctrinate submission to Prussian rulers.

Russian Empire Alexander III (1881-1894) tried to create a Russian identity. Russian language and culture were imposed over a linguistically diverse population. Russian was made the compulsory language of instruction not only in all state schools in the Baltic provinces (in 1887) but also in private schools. The Armenian church lost control of all its schools in 1896 and part of its wealth was confiscated to pay for the costs of the new schools. Finland, which up until the 1890s was considered as a state distinct from Russia and only sharing the same ruler, was forcibly transformed into a province of the Russian empire by Alexander III with the imposition, among other things, of Russian as a subject in Finnish schools. This process of Russification, illustrated in these examples, was imposed throughout the whole empire.

Alexander III started the process of Russification as a reaction to numerous uprisings. The different nationalities that comprised the empire started to rebel during the 1860s, first with the Polish uprising of 1863, and later with revolutionary activities in universities during the 1870s that eventually lead to the assassination of Alexander II in 1881. Shortly before his death, Alexander II had considered reforms to introduce some elected positions into government. However, Alexander III’s position was that autarchy and nationalism were the right way to prevent the empire from collapsing.²⁰

For all these cases, one alternative explanation is that rioters demanded public education and the latter was a concession under duress on the part of the rulers. As noted previously, state-run mandatory schooling was unpopular and opposed by peasantry for much of the 19th century in France. In England, violent and non-violent protests spread

¹⁸Quote from Marcham (1973). The 1867 act enfranchised a part of the male urban working-class population.

¹⁹Quote from Young and Handcock (1964). Bandiera et. al (2018) highlight this side of nation-building in the context of the United States. Americans introduced compulsory education, in large part, to civilize and instill common civic and other values in migrants, in order to influence their participation in American life.

²⁰From Andersen (2016) and Wachtel et al. (2020).

across the country in the first years of the 1830s. The Royal Commission into the Poor Laws in 1834, which was set up in part in response to this unrest, asked the following question: “Can you give the commissioners any information respecting the causes and consequences of the agricultural riots and burning of 1830 and 1831?” In England, 526 parishes responded. The only causes cited by more than 30 parishes were labor concerns (unemployment, wages, and mechanization of jobs that previously provided employment), subsidies for the poor (poor law) and beer shops (where it is believed many of the protests were organized). Not a single response considered demand for education or anything related to education as a cause of the unrest (Holland, 2005). Similarly, Tilly (1998) provides a detailed study of episodes of collective disturbances in France 1830-1860 with information on the objective of the group involved in the disturbance. Education is not mentioned. In Prussia, peasants did not demand anything related to education after the increase in grain prices.

If education in the 19th century was provided with a nation-building motive and not as a welfare policy, then we should expect differences in the implementation of education policies compared to other welfare policies, such as social security or health care, especially since direct redistributive concerns were closer to population demands than education. Indeed, there are stark differences in the timing of education reforms 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 European countries had compulsory universal education by the time welfare reforms were introduced and in some countries it was highly developed (e.g. France and Germany). Welfare reform then tended to follow franchise extension. In contrast, education reforms preceded it. We will test this hypothesis more formally in Section 6.1.2.²¹

Latin America

Paglayan (2018) shows that, like Europe, Latin American states introduced compulsory mass primary education before universal male suffrage and democratization. Below we discuss the expansion of primary education in Chile and Argentina, our evidence is drawn from Paglayan (2017) and Paglayan (2018).

Chile Paglayan (2017) examines education provision in Chile in the years following the passing of the 1860 “General Law of Education.” This law transferred control of primary education to the Chilean central government. Immediately before the passing of this law, in 1859, there were mass rebellions in the north of the country. Paglayan (2017) examines quantitative data on the subsequent expansion of primary education by the government of Chile and shows that the number of new primary schools established in a province and enrollment in primary schools by province was positively related to the extent of rebellion in that province. As with the introduction of education in Europe, there was no apparent demand for education by the masses and no apparent demand by rebels for education. Political discourse at the time is consistent with a nation-building motive behind education reforms and highlights the power of education as a means to indoctrinate and prevent future rebellion. The Amunategui brothers (politicians heavily

²¹This is consistent with the historical discussion in Acemoglu and Robinson (2000) on the extension of the franchise. They suggest that in many cases redistributive concessions were not credible before franchise extension (Germany, which introduced the first European insurance systems, being an exception).

involved in influencing education at this time) wrote that “children generally acquire in school habits of order, of submission” and “the best way to prevent future revolutions is to educate children” (Paglayan, 2017).

Argentina Paglayan (2017) argues that a similar pattern is evident in Argentina. The Argentinean government faced a series of civil wars from 1814 to 1880, which were then followed by the “Law of Common Education” in 1884 to introduce government regulation, funding, and mandated primary education. It was believed that the ability of schooling to instill certain moral values would help to reduce future opposition to the government.

Asia

Japan The Meiji Restoration in 1868 ousted the last shogun and returned power to the Emperor. The Meiji Restoration marked both a turning point in Japanese politics, as well as the introduction of significant nation-building policies. The Meiji regime faced a multitude of threats. Takayama (1988) describes the Meiji regime as “confronted with the double threat of Western civilization and internal disintegration.” By the 1870s, the new regime met significant opposition from within Japan, including rebellion from samurai and peasant uprisings. A civil rights movement generated pressure towards “wider participation” and the “creation of a constitutional government,” and the government responded by promising a constitution by 1890.²² It is in this context of threats to the regime, that the Japanese government started to nation-build.

Compulsory education was a way to impose both “state Shinto” and a unified Japanese language. The Imperial Rescript on Education in 1890 was the core education reform that reinforced these principles.²³ Takayama (1988) describes state Shinto as “essentially the newly devised religion of Japanese nationalism at the time of the Meiji Restoration in 1868,” with the emperor cult as its central element. The idea was that the imperial system would be used “as an ideological weapon for controlling the entire Japanese population” (Takayama, 1988). The Imperial Rescript on Education pushed state Shinto and loyalty to the emperor through schools. In the 1890s, the head of the Ministry of Education’s Compulsory Education Agency put forward a process to deal with unified speech: first to “artificially refine” a version of Japanese, which “would then be delivered to the public through the new compulsory education system,” seen as the most effective means of its dissemination (Shimoda, 2010). Shimoda (2010) describes the preoccupation within the Meiji regime to create a unified speech in Japan and writes that “just as a French state previously indifferent to linguistic inconsistency” was replaced by one concerned with unity, “a newly national Imperial Japan took a similar turn.” Education reforms in Japan occurred well before the introduction of social welfare and state provision of health services, which did not begin to develop until the 1920s (Odaka, 2002).

Siam-Thailand Nation-building policies in Siam-Thailand were implemented by Wachirawut, who reigned from 1910-1925. He regarded himself as the country’s first nationalist and pushed an official nationalism through the education system. Among other measures (like providing an official re-writing of history), Wachirawut introduced compulsory state-controlled primary education and free government schools, granted education to girls, and

²²Encyclopaedia Britannica (2018).

²³Lawson et al. (2019).

brought all private schools under the control of the state (as a part of the Private School Act). These measures were part of a broader process of Thai-ification of all Chinese schools.²⁴

Wachirawut's official nationalism was a response to a strong threat of revolution and increasing challenge from the middle class. The modernization that the reign had attained under his father, had eventually created a new generation of intellectuals (both Thai and Chinese) and a new middle class, that began criticizing the absolutist state and demanding a constitutional democracy. Just a few months before he became king, he also had to suppress a big general strike organized by Chinese merchants and workers. Similar types of revolts in China, the Wuchang Uprising, organized by a heterogeneous assortment of groups including merchants, lead to the overthrow of the Qing dynasty and eventually to the establishment of the Republic of China. Chinese people started then to be seen as pioneers of a popular republicanism, that constituted a severe threat to the dynastic principle.

Middle East and North Africa

Turkey Like the Ancien Régime in France, the 19th century rulers of the Ottoman Empire were unconcerned by the diversity of their population. Pluralism of language, religion, and culture was tolerated. One illustration of this is the "millet system" under which a religious community was allowed a large degree of autonomy in aspects such as law and education. As the Ottoman Empire lost more territory from external pressure and through secession, particularly towards the end of the 19th century, homogenization policies were introduced.

Both nation-building policies and the arrival of democracy came later in Turkey than in Europe. Homogenizing practices were implemented in force after the founding of the Turkish Republic in 1923 by Kemal Atatürk. The focus was to create a "Turkish" nation. Aslan (2007) writes that "the spread of a common and standardized national language had been one of the main concerns of the state elite from the beginning of the Republic." Large changes in education policies played a big role in nation-building under Atatürk. Secularization reforms abolished religious education. In 1924 the "Law on the Unification of Education" brought education under the control of the state and instituted a common curriculum. Over the following years, education reforms included a mandatory five hours of Turkish teaching, teachers had to have Turkish as their native language, and later all children had to attend a Turkish primary school.²⁵ Zürcher (2010) writes that this Turkish identity was "imposed gradually on the population through a process of nation-building in which, as in similar processes the world over, historiography and linguistics played a key role, as did suppression of alternative or sub-identities." Zürcher (2010) estimates that "the adoption of Turkish nationalism led to the forced assimilation of the 30 percent or so of the population which did not have Turkish as its mother tongue."

Colonial rulers

Colonizers are different from domestic rulers because if overthrown they can leave the country and go home where they maintain a "high status" as part of the elite. Thus colo-

²⁴ Andersen (2016), Murray (2014), Skinner (1957), Suryadinata (1997).

²⁵ Aslan (2007)

nizers face lower incentives to homogenize to preserve the status quo and avoid break-up after democratization. Consistent with our model, there is little evidence of an expansion of compulsory education by colonizers.

The British Caribbean In Britain, the Education Act of 1870 established a system of local education authorities, essentially expanding the reach of the government to the education of children. In 1880 education was made compulsory for children between the ages of 5 and 10. By the early 1890s, 82% of children aged 5 to 10 attended school.

Education in the British territories in the Caribbean during the same period looked vastly different. Compulsory education came much later, for example, in 1915 in British Honduras and in 1921 in Trinidad (Lewis, 2000). Yet, even when education was made compulsory, it was largely on paper only. In 1931, a commission ordered by the British government to assess educational systems in the British Caribbean stated that “compulsion is applied only in a few islands, and even in these not effectively” (Gordon, 1964). The report showed that “the total percentage of the population enrolled in Government and aided primary schools is 16.3%,” while attendance of those enrolled is 66%, with most in the infant classes of primary school. In general, the report highlighted the “backwardness of primary education” and the lack of adequate financial provision.

Even though compulsory schooling was established in 1915 in British Honduras, Hitchen (2000) describes “the British neglect of education in the colony” and education as “a primitive affair up to 1931, with only limited state involvement.” Instead, education was more of a religious affair and was “organized randomly by the various denominations.” Lewis (2000) writes that “like other schools in the British Empire, education was a missionary effort” and describes how, in 1923, rather than the British government, it was religious organizations who “took it upon themselves to establish a clear educational policy”. Lewis (2000) summarizes the education system as “based on the British colonial model”, which was “a model that did not want to educate the colonized.”

Zambia Colonizers of Africa did not make any effort to build cohesive nation-states (Easterly and Levine, 1997; Herbst, 2000; Alesina, Easterly, and Matuszeski, 2011; Michalopoulos and Papaioannou, 2016). They viewed broad-based education as an unnecessary expense and a threat. Levels of primary education enrollment in many cases even declined during the period of colonial occupation (Benavot and Riddle, 1988). In Zambia, a British colony from the 19th century to its independence in 1964, colonization was a “take-the-money-and-run affair” with education mainly provided by missionaries. Colonization exacerbated differences among the Zambian population (Marten and Kula (2008) on language; Phiri (2006) on regional divisions). On independence, a multitude of languages were spoken, with English existing as the main language of commerce and administration. Kenneth Kuanda, the first president of Zambia, claimed that although nationalism had led to independence, national identity in Zambia was completely lacking. Phiri (2006) writes that “Zambia’s experience in the first eight years of independence is a typical example of how most newly independent African countries grappled with the need to create a sense of national identity.”

Uganda Before colonization, education in Uganda was carried out privately within the various tribes. The beginning of missionary activity in Uganda, in the late 19th century, marked the introduction of western education as a means of conversion to Christianity.

Education under colonization was almost entirely left in the hand of missionaries. Only in the 1950s did the colonial government announce its aim of integrating education in Uganda, organizing education on a racial basis (with separate schools for Europeans, Africans, and Asians). However, the content and scope of primary and secondary education under British rule was geared toward building a small local public bureaucracy.²⁶

Kenya The British wanted to make Kenya the center of East African development (Barkan, 1994). Their divide and rule policy was implemented even before Kenya was colonized, when the British played off different tribes against each other: while some tribes managed to ally against the British, others were used by the British to conquer much of the Kenyan territory. After the conquest of Kenya by the British was completed, the divide and rule policy involved exacerbating ethnic conflicts within Kenyan borders. One way this was done was to prohibit any form of inter-ethnic cooperation, for example, by prohibiting settlement in between neighboring ethnic entities (Weber, 2009). Similar to other African colonies, education in Kenya was left in the hands of missionaries. Only in the first years of the 20th century, did the British start to take an interest in education. Their aim was exclusively to expand the colony’s economy. The Commission of 1919 opposed the spread of a common language, Swahili, the lingua franca, and instead allowed the use of any vernacular language for the preliminary stages of education (Urch, 1971). This policy stands in stark contrast to the previous examples of domestic rulers who used primary education to create a common language. The permanent advisory committee, set up in 1923 to supervise education in African colonies, explicitly stated that education was to be adapted to local conditions “to conserve all sound elements in local tradition and social organization, while functioning as an instrument of progress and evolution” (Urch, 1971).

India India had a similar experience. The British had done little, if anything, to homogenize a diverse population, even using specific policies of divide and rule (Christopher, 1988). For most of the 19th century, the East India Company was legally responsible for funding education, however, all substantial efforts were dismissed as “premature or impractical” (Mondal, 2017). Colonial administrators neglected primary education, especially for the indigenous and non-English speakers (Ojha, 1966). A commission on education in 1882 made no recommendation for compulsory or free education. Instead, Indian leaders tried to step in and in 1906, the Indian National Congress declared education “the birthright of the people of India” (Mondal, 2017). In the same year, the Maharaja of Baroda made education compulsory and free for all young boys and girls in his state.²⁷ The British government responded to these developments by appointing a Commission, which concluded that the time for compulsory education had not yet come, because of “backwardness among the large masses” and the risk that it might “cause endless friction between the Government and the people” (Ojha, 1966).²⁸

²⁶See US Department of Health and Welfare (1964).

²⁷Mondal (2017)

²⁸Laitin (1989) provides an interesting discussion on the consequences of lack of nation-building under colonial powers for the case of India. Laitin (1989) describes a contemporary multi-linguistic state: Hindi and English are the de facto common language across states, Indians also learn their state official language (often primary education is taught in the language of the state), and some Indians, in addition, have a mother tongue which is neither Hindi, English, nor their state language. This is not a result of a lack of desire by elites for a single language: Laitin (1989) discusses attempts by Congress to introduce a single unified language, and attempts by State governments to do so within states. Our framework suggests

Other colonizers The Italian colonization of Ethiopia is another case of divide and rule. Under Italian rule, many schools were closed and those that stayed open were intended to teach fascist values and loyalty to Italy. Instead of Ethiopia’s national languages, colonial education officials recommended the teaching of lessons in local administrative languages. This was not undertaken with any educational rationale in mind, but rather with the express purpose of sowing disunity in the Ethiopian population (Bishaw and Lasser, 2012).

Education in French and Belgian colonies was even more neglected than in British colonies. Lewis (2000) points to estimates that in 1940 the enrollment level in French colonies was about 50 years behind that of British colonies. This should be compared to the French education policy in France itself which was arguably widespread and effective from the 1870s. Rates of enrollment in British colonies were two or three times higher than French colonies, and as much as five times higher than Belgian colonies (Benavot and Riddle, 1988). These differences may reflect a lower willingness to depart in the event of their political overthrow, due to the relatively more commercial nature of their colonial business interests. In terms of our model, this points to the idea that the type of colonial power most likely to nation-build through education is one that does not see itself as having an easy choice of abandoning its investments and place in the colonized society. Examples like Australia, Canada, and perhaps Brazil and South Korea, in which the colonial power viewed the colony as something like a fully-fledged part of the larger nation, would therefore not be expected to fall under a divide and rule strategy.

6 Econometric evidence

6.1 Data and Specification

6.1.1 Sources and Variable Definitions

Education. We use an unbalanced panel with ten year averages data on primary educational enrollment per capita for 172 countries between 1925 and 2014. Our measure of imputed reform is a dummy indicating if enrollment grew by more than 20 percent over the previous 10 year period. For robustness, we also report results using a binary variable set equal to one if enrollment grew by more than 10 percent over the previous 10 year period. We collapse the data into 10-year averages so as to minimize measurement error. The variable on primary enrollment is defined according to the UNESCO criteria and expressed per 10,000 inhabitants. The data come from the CNTS Data Archive of Banks and Wilson (2018). For a sample of 14 European countries, we also use a dummy indicating whether and when an education reform was adopted. The data is from Flora (1983). Flora (1983) defines educational reforms as any new law which extended compulsory education, lowered the cost of education (by abolishing school fees or providing free primary education), or increased the number of schools (by making it compulsory for each municipality to set up at least one primary school). We look at the relationship between our measure of imputed reform and the measure of legal reform provided by Flora (1983) in a panel regression where the left hand side is our measure of imputed reforms and the

democracies (like India after independence), given they have to take into account the wishes of the wider population to a greater degree, are more restricted in the homogenization they can undertake. A country that moves from colonization to democracy misses the opportunity for intense nation-building by elites, implying it may be more heterogenous today (or even break-up).

controls include population growth, country and period fixed effects. The coefficient on the measure by Flora is 0.8 and significant at the 10% level.

Political Regimes. The autocracy variable is constructed from the polity2 variable taken from the Polity IV database (Marshall, Gurr, and Jaggers, 2016). This variable ranges from -10 to 10, where a higher score means that the country is more democratic. The variable is based on information on constraints on the executive, the openness and competitiveness of the executive recruitment, and the competitiveness of political participation. We define autocracy when the polity2 variable is lower than zero.²⁹

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

- *major government crises:* document any rapidly developing situation that threatens to bring the downfall of the present regime;
- *revolutions:* documents any illegal or forced change in the top government elite, any attempt at such a change, or any successful or unsuccessful armed rebellion whose aim is independence from the central government;
- *weighted conflict average (WCI):* a weighted average of all the conflict indicators contained in the dataset.³⁰

In some of our specifications we measure the “safety” of a ruler by using the variable *durable* from the Polity 4 database. This variable measures the number of (cumulative) years since the last substantive change in authority characteristics (defined as a three-point change in the polity2 score). Various scholars argued that the current stability of the regime is the best predictor of future regime stability (Clemens and Cook (1999), Gates, Hegre, Jones, and Strand (2006), Poyker (2018)).

Control variables. We control for population growth to limit the possibility that our measure could be affected by demographic population shifts. Our measure of educational reform is based upon enrollment per capita, rather than enrollment per school-age child. Additional controls included in some of our specifications are whether the country has been involved in an external war in the previous 10 years (the variable is taken from the

²⁹Democracy indices are subject to considerable measurement error. Acemoglu, Naidu, Restrepo, and Robinson (2019), following Papaioannou and Siourounis (2008), improve on the measurement of democracy. Their measure is, however, available only after 1960 and, as a result, is not useful for our research question. Like Aghion, Jaravel, Persson, and Rouzet (2019), in our paper, we use the measure of autocracy based on polity2 because it is the only one available for a long period of time.

³⁰The CNTS dataset contains various measures of domestic conflict. In addition to the ones mentioned above it also contains the following variables. *Assassinations:* records the occurrence of any politically motivated murder or attempted murder of a high government official or politician. *General strikes:* lists strikes of 1,000 or more industrial or service workers that involve more than one employer and that are aimed at national government policies or authority. *Guerrilla warfare:* gives information about armed activities, sabotage, or bombings carried out by independent bands of citizens or irregular forces and aimed at the overthrow of the present regime. *Purges:* identifies any systematic elimination by jailing or execution of political opposition within the ranks of the regime or the opposition. *Riots:* records the occurrence of any violent demonstration or clash of more than 100 citizens involving the use of physical force. *Anti-government demonstrations:* records any peaceful public gathering of at least 100 people for the primary purpose of displaying or voicing their opposition to government policies or authority, excluding demonstration of a distinctly anti-foreign nature. We do not consider any of these as part of our analysis as some of them are not related to the probability of the regime being overthrown (riots, anti-government demonstrations, and general strikes). Guerrilla warfare could also be relevant but it does not refer to a desire of regime overthrow from the general population, whereas assassination refers to the assassination of any high government official and not only to the assassination of the ruler.

Correlates of War database), fiscal capacity, measured as revenue and expenditure over GDP (taken from the CNTS dataset), GDP per capita (taken from Madison), a measure of trade (the proportion of world trade represented by the country) and a measure of urbanization (measured as the population of cities with more than 100,000), both taken from the CNTS dataset.

Dataset on leaders. We use the *Archigos* dataset (Goemans, Chiozza, and Gleditsch, 2009) to test some of the predictions of the model. This dataset contains information on leaders for 188 countries from 1875 to 2004. This dataset constitutes an advantage to the standard Polity 4 dataset because it allows us to identify political changes in autocracies not apparent in data that consider only the democratic nature of institutions.

Descriptive statistics for all our variables are provided in Table 1. In our sample, 26 percent of countries experienced an increase in the per capita enrollment in primary schooling larger than 20 percent, compared to the previous decade; 35 percent experienced at least an increase equal to 10 percent. When we look at legal education reforms, for the sample limited to European countries, almost 50 percent of them implemented a legal reform in education. In the overall sample, around half of the time period was under autocratic regimes.³¹ Figure 1 describes the fraction of countries implementing educational reforms in our sample, by continent. As is apparent from the figure, the results are not driven by a specific continent, and there is sufficient heterogeneity to identify the effects.

6.1.2 Empirical specification

Our baseline regression equation is expressed as:

$$\begin{aligned} \text{educational reform}_{it} = & \alpha_0 + \alpha_1 \text{autocracy}_{i,t-1} + \alpha_2 \text{threat to regime}_{i,t-1} + \\ & + \alpha_3 \text{autocracy}_{i,t-1} \cdot \text{threat to regime}_{i,t-1} + \alpha_4 X_{i,t-1} + \delta_i + \gamma_t + \epsilon_{it}. \end{aligned} \quad (4)$$

where *educational reform*_{*i,t*} is a dummy indicating whether educational enrollment has increased by more than 10 or 20 percent in the last 10 years. Our coefficient of interest is α_3 , which indicates that more unstable autocracies are likely to implement education reforms. All our specifications include country (δ_i) and year (γ_t) fixed effects and population growth. We also test the robustness of our results to a larger set of controls, $X_{i,t-1}$, including the level of development, trade, urbanization, fiscal capacity and whether the country was involved in a war in the previous 10 years. The standard errors are clustered at the country level.

Table 2 shows the results for our baseline estimation. The first three columns show the results when the education reform is defined as an increase in primary enrollment higher than 10 percent from the previous 10 years, whereas Columns 4-6 report the results with the 20 percent threshold. Columns 7-9 use the definition of reform constructed by Flora (1983) and it is limited to a sample of 14 European countries. Our coefficient of interest, α_3 , is always positive and significant, indicating that the threat to the regime is associated with an increase in educational enrollment.

³¹This fraction is much smaller in the European sample used for the legal reform measure, where only 4 percent was under autocratic regimes in the period 1925-2014.

Alternative explanations Aghion, Jaravel, Persson, and Rouzet (2019) show that the threat of war is associated with increased primary education enrollment (considered as a measure of nation-building), but that the threat of war may only be relevant when countries are sufficiently democratic. This result is consistent with our model: a dictator can “force” armies to fight by fear, but in a more democratic regime it may be more difficult to do so and teaching nationalism may be more compelling and necessary (see also Alesina, Reich, and Riboni (2017)).

We add to our specification a variable indicating whether the country was involved in an external war in the previous 10 years, and an interaction term with the fraction of years spent under autocratic regimes (Table 3, columns 1-3). Consistent with Aghion, Jaravel, Persson, and Rouzet (2019) we find that education reforms respond more positively to military threats in democracies, however the interaction term between threat to democracy and the presence of autocratic regimes remains significant and of similar magnitude. We see our argument about nation-building for fear of democratization and splitting of countries, and state-building for fear of aggression, as complementary.

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

A third prominent theory is that industrialization prompted governments to undertake significant nation-building. Gellner (1983) argues that an industrial society, based upon broad markets needs better means of communication than an agrarian society.³² We use GDP per capita and a measure of urbanization as proxies of industrialization. None of the two variables alter our main findings (columns 7-12). This is also consistent with several scholars who question the timing of this theory. Smith (1998) and Green (1990) argue that education reforms were implemented country by country in a way that is inconsistent with industrialization acting as a major driver. 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 provide children the appropriate technical skills.

Another phenomenon linked to the modernization of society relies on trade and globalization. Openness to trade should have a positive effect on education in industrial countries where the complementarity between education and technology was important (Collins, 1971). Controlling for trade does not alter our results (Columns 13-15).

Autocracies could start with a lower initial level of public education provision, therefore when a threat to democracy is looming, they react by providing more public education to

³²See 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.

appease the masses. Various papers have compared education policies across democratic and non-democratic regimes. Aghion, Jaravel, Persson, and Rouzet (2019) using annual data on 137 countries from 1830-2001, find that autocracies have higher enrolment rates in primary education than democracies. Consistent with this, Mulligan, Gil, and Sala-i-Martin (2004) examine cross-country data from 1960-1990 and find that there is no evidence that democracies spend more on public education than non-democratic regimes. Looking at the same dataset, Bursztyn (2016) finds that democracies spend less on public education than non-democracies for below-median income countries. We also explore more directly this hypothesis in our dataset. We do not find any systematic difference in education between autocratic and democratic countries (Appendix, Table A1)

Another possibility is that the introduction of education provision just had a redistributive and equal opportunities motive, rather than a nation-building motive. If autocratic regimes provided education for redistributive reasons, they should have also implemented other types of welfare reforms of similar nature. We use data from Flora (1983) on three types of reforms having redistributive nature: health insurance, pension insurance and unemployment insurance reforms. We do not find that the threat to democracy was in general associated with other types of redistributive reforms to compensate the masses (Table 4). This is consistent with 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, finding either no effect of totalitarianism or a negative effect.

The threat of revolution could also be related to the extension of the voting franchise in Europe. Acemoglu and Robinson (2000), Aidt and Jensen (2014) and Aidt and Franck (2015) show theoretically and empirically that the extension of the franchise by elites can be a reaction to a strong threat of revolution as a way to avoid that revolution. Our analysis relies on the idea that not all crises, unrest, or revolutions necessarily result in an immediate extension of the franchise (which is also consistent with the literature above). In our data we have many incidents of government crises, unrest, and revolutions that increased the perceived threat to the elite but did not result in immediate franchise extension and so gave elites time to implement nation-building policies in-line with our theory. We show this in Table A2, where we use our sample and the data provided by Aidt and Jensen (2014) on the extension of the franchise.

6.1.3 Additional implications of our model

One of the implications of Proposition 2 is that, if the ruler homogenizes, this will reduce opposition to the regime and the population may not choose to overthrow the ruler. We try to look at this possibility by using a dataset on leaders by Goemans, Chiozza, and Gleditsch (2009). This dataset allows us to calculate for each leader and each country the number of years in power. The implication of Proposition 2 is that autocratic leaders who implemented nation-building policies are more likely to stay in power. Table 5 shows the results of a regression where the left hand side is the number of years in power starting from the decade of their first regime, whereas the controls (in addition to the level of GDP, country and decade fixed effects) are the increase in the level of primary education during the reign, a dummy for whether the leader was autocratic, and an interaction between these two variables. We find nation-building policies helped the leader to stay

in power.

6.1.4 Robustness checks

Extension of the time period of analysis The measures we have of threat to democracy are only available after 1925. To test the robustness of our results to a longer time period, we use a variable on the durability of the regime (see data section for details). The idea behind it is that the current durability of the regime is the best proxy for its future durability. Our theory predicts that unsafe rulers are the ones more likely to implement nation-building policies. Table 6 therefore controls for a measure of durability of the regime and its interaction with autocracies. The prediction is that durable (safe) rulers should be less likely to implement nation-building reforms. We find that this is indeed the case, showing that our results are not only valid but even stronger when we extend our period of analysis. In this table we also control for all the other alternative explanations analyzed in Table 3.

Sample selection Our model also makes predictions on the possibility that a country separates if it is not homogenous enough. We do not test this prediction in this paper since it has been object of study in the “size of nations” literature (Alesina and Spolaore, 1997, 2003). In our context, however, it has direct implications for the selection of countries in our sample: if countries can potentially break apart, the sample of countries and their endogenous choices with regard to homogenization will be a result of this selection process. Table A3 shows the results restricting the sample to countries that never split. The results stay the same. Figure A1 indicates that the results are not driven by some specific continents. We also rule out this possibility more formally, by controlling for continent specific linear trends (Tables A4 and A5).

Endogeneity In our main specification all the explanatory variables are lagged, therefore limiting the possibility of reverse causality. To further limit endogeneity concerns, we also attempt an instrumental variable strategy using system generalized method of moments. Lags of all endogenous variables are used as instruments for all not strictly exogenous variables, including autocracy and measures of political unrest. In all the different specifications used, the Hansen test and the second order correlation tests indicate that we cannot reject the validity of the moment conditions assumed for the estimation. The results are reported in Table A6.

Education as a continuous outcome and dynamic panel specifications To improve our estimate of the impact of the instability of the regime on education, we also estimate a regression in levels for primary education, which will help us to better estimate the dynamics of the educational variable (see Acemoglu, Naidu, Restrepo, and Robinson (2019)). Table A7 presents the results of this estimation controlling for different numbers of lags. The first column of the table controls for a single lag of primary enrollment. As expected there is a sizable amount of persistence in annual levels of education. For our purposes, the interaction between the threat to democracy and autocracy is positive and significant.³³ Column 2 adds a second lag to this specification, and shows that both

³³The coefficient is always positive and significant, with the exception of revolutions, for which we lose power, due to an increase in the standard errors.

lags are highly statistically significant and point to richer dynamics, with the first lag being positive and greater than one and the second negative. Columns 3 and 4 include three/four lags of education and column 5 includes up to ten lags and shows that these lags are not significant.

7 Conclusion

We have studied when and how governments nation-build. Safe dictators who do not fear any revolt do not have an incentive to homogenize; they simply extract rents. Instead non-democratic regimes homogenize when threatened by pro-democracy revolts which may also split the country, turn the former elite into a minority not in control of policies in any of the new countries, and facing a loss of rents. Mass primary education is a central policy instrument used to homogenize populations, and so we studied this policy empirically. We reviewed many historical examples, covering different geographical areas and periods of time, which have several features in common. First, government implementation of mass primary education occurs under non-democratic regimes, before the largest extensions of the franchise. Second, the timing of expansions of mass primary education is linked to threats of democratization. Third, expansions of mass primary education occurred with nation-building goals in mind.

In contrast, colonizers are not native and have no interest in remaining in the country if a democratic revolution is successful. Thus they have no such incentive to homogenize, and in fact may have an interest in policies of divide and rule, encouraging different ethnicities to fight and building upon ancient animosities amongst ethnic groups. This lack of nation-building policies by former colonizers is one of the reasons why former colonies have had difficulties in their transition to independence, due to internal divisions and no communal sense of a nation.

We then analyzed cross-country data on mass primary schooling and educational reforms. Our statistical evidence is consistent with the implications of our model. Using evidence on a large sample of countries and covering more than 150 years of data, we confirm that the threat of democracy is an important driver of nation-building education policies. We successfully compare our explanation for education policies against several alternative hypotheses, including state capacity, external wars, and industrialization. We also rule out the possibility that education provision had a redistributive motive rather than a nation-building one. Consistent with our model, we also show that leaders who implemented educational reforms tend to stay in power longer.

References

- ACEMOGLU, D., S. NAIDU, P. RESTREPO, AND J. A. ROBINSON (2019): “Democracy Does Cause Growth,” *Journal of Political Economy*, 127(1), 47–100.
- ACEMOGLU, D., AND J. A. ROBINSON (2000): “Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective,” *The Quarterly Journal of Economics*, 115(4), 1167–1199.

- (2008): “Persistence of Power, Elites, and Institutions,” *American Economic Review*, 98(1), 267–293.
- AGHION, P., X. JARAVEL, T. PERSSON, AND D. ROUZET (2019): “Education and Military Rivalry,” *Journal of the European Economic Association*, 17(2), 376412.
- AIDT, T. S., AND R. FRANCK (2015): “Democratization Under the Threat of Revolution: Evidence From the Great Reform Act of 1832,” *Econometrica*, 83, 505–547.
- AIDT, T. S., AND P. S. JENSEN (2014): “Workers of the world, unite! Franchise extensions and the threat of revolution in Europe, 18201938,” *European Economic Review*, 72(C), 52–75.
- ALESINA, A., W. EASTERLY, AND J. MATUSZESKI (2011): “Artificial States,” *Journal of the European Economic Association*, 9(2), 246–277.
- ALESINA, A., AND E. GLAESER (2005): *Fighting Poverty in the US and Europe: A World of Difference*. Oxford University Press, Oxford UK.
- ALESINA, A., B. REICH, AND A. RIBONI (2017): “Nation-Building, Nationalism and Wars,” *NBER working paper No. 23435*.
- ALESINA, A., AND E. SPOLAORE (1997): “On the Number and Size of Nations,” *The Quarterly Journal of Economics*, 112(4), 1027–1056.
- (2003): *The Size of Nations*. MIT Press, Cambridge MA.
- ALESINA, A., E. SPOLAORE, AND R. WACZIARG (2000): “Economic Integration and Political Disintegration,” *American Economic Review*, 90(5), 1276–1296.
- ALGAN, Y., T. MAYER, AND M. THOENIG (2013): “The Economic Incentives of Cultural Transmission: Spatial : Spatial Evidence from Naming Patterns across France,” Working Papers 2013-25, CEPII research center.
- ANDERSEN, B. (2016): *Imagined Communities. Reflections on the Origins and Spread of Nationalism*. Verso Books, Revised Edition.
- ASLAN, S. (2007): “Citizen, Speak Turkish!: A Nation in the Making,” *Nationalism and Ethnic Politics*, 13(2), 245–272.
- BANDIERA, O., M. MOHNEN, I. RASUL, AND M. VIARENGO (2018): “Nation-Building Through Compulsory Schooling During the Age of Mass Migration,” *Economic Journal*, 129(617), 62–109.
- BANKS, A., AND K. WILSON (2018): “Cross-National Time-Series Data Archive,” .
- BARKAN, J. D. (1994): “Divergence and Convergence in Kenya and Tanzania: Pressures for Reform,” in *Beyond Capitalism vs. Socialism in Kenya and Tanzania*, ed. by J. D. Barkan. Nairobi: East African Educational Publishers.

- BENAÏOT, A., AND P. RIDDLE (1988): “The Expansion of Primary Education, 1870-1940: Trends and Issues,” *Sociology of Education*, 61(3), 191–210.
- BESLEY, T., AND T. PERSSON (2009): “The Origins of State Capacity: Property Rights, Taxation, and Politics,” *American Economic Review*, 99(4), 1218–1244.
- (2010): “State Capacity, Conflict, and Development,” *Econometrica*, 78(1), 1–34.
- BESLEY, T., T. PERSSON, AND M. REYNAL-QUEROL (2016): “Resilient Leaders and Institutional Reform: Theory and Evidence,” *Economica*, 83(332), 584–623.
- BISHAW, A., AND J. LASSER (2012): “Education in Ethiopia: Past, Present and Future Prospects,” *Africa Nebula*, 5, 53–69.
- BISIN, A., AND T. VERDIER (2000): “‘Beyond the Melting Pot’: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits,” *The Quarterly Journal of Economics*, 115(3), 955–988.
- BOLTON, P., AND G. ROLAND (1997): “The Breakup of Nations: A Political Economy Analysis,” *The Quarterly Journal of Economics*, 112(4), 1057–1090.
- BOURGUIGNON, F., AND T. VERDIER (2000): “Oligarchy, Democracy, Inequality and Growth,” *Journal of Development Economics*, 62(2), 285–313.
- BOWLES, S. (1998): “Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions,” *Journal of Economic Literature*, 36(1), 75–111.
- BURSZTYN, L. (2016): “Poverty And The Political Economy Of Public Education Spending: Evidence From Brazil,” *Journal of the European Economic Association*, 14(5), 1101–1128.
- CAMPANTE, F. R., Q.-A. DO, AND B. GUIMARAES (2019): “Capital Cities, Conflict, and Misgovernance,” *American Economic Journal: Applied Economics*, 11(3), 298–337.
- CANTONI, D., Y. CHEN, D. Y. YANG, N. YUCHTMAN, AND Y. J. ZHANG (2017): “Curriculum and Ideology,” *Journal of Political Economy*, 125(2), 338–392.
- CHRISTOPHER, A. J. (1988): “‘Divide and Rule’: The Impress of British Separation Policies,” *Area*, 20, 233–240.
- CLEMENS, E., AND J. COOK (1999): “Politics and Institutionalism: Explaining Durability and Change,” *Annual Review of Sociology*, 25(1), 441–466.
- CLOTS-FIGUERAS, I., AND P. MASELLA (2013): “Education, Language and Identity,” *Economic Journal*, 123(570), 332–357.
- COLLINS, R. (1971): “Functional and Conflict Theories of Educational Stratification,” *American Sociological Review*, 36(6), 1002–1019.

- DARDEN, K., AND H. MYLONAS (2016): “Threats to Territorial Integrity, National Mass Schooling, and Linguistic Commonality,” *Comparative Political Studies*, 49(11), 14461479.
- DUGGAN, C. (2007): *The Force of Destiny: A History of Italy since 1796*. Penguin Books, London.
- EASTERLY, W., AND R. LEVINE (1997): “Africa’s Growth Tragedy: Policies and Ethnic Divisions,” *The Quarterly Journal of Economics*, 112(4), 1203–1250.
- ENCYCLOPAEDIA BRITANNICA (2018): *Meiji Restoration*. Encyclopaedia Britannica, inc.
- FLORA, P. (1983): *State, Economy, and Society in Western Europe 1815 - 1975. A Data Handbook. Vol. 1. The Growth of Mass Democracies and Welfare States*. Campus Verlag, Frankfurt am Main.
- GATES, S., H. HEGRE, M. JONES, AND H. STRAND (2006): “Institutional Inconsistency and Political Instability: Polity Duration, 1800-2000,” *American Journal of Political Science*, 50(4), 893–908.
- GELLNER, E. (1983): *Nations and Nationalism*. Cornell University Press, Ithaca.
- GLAESER, E., G. PONZETTO, AND A. SHLEIFER (2007): “Why does Democracy Need Education?,” *Journal of Economic Growth*, 12(2), 77–99.
- GOEMANS, H., G. CHIOZZA, AND K. S. GLEDITSCH (2009): “Introducing Archigos: A Dataset of Political Leaders,” *Journal of Peace Research*, 46(2), 269–283.
- GORDON, S. C. (1964): “Documents which have guided Educational Policy in the West Indies, No. 8: Report of the Commissioners Mayhew and Marriott on Secondary and Primary Education in Trinidad, Barbados, Leeward Islands and Windward Islands. 1931-1932,” *Caribbean Quarterly*, 10(4), 3–32.
- GRADSTEIN, M., AND M. JUSTMAN (2002): “Education, Social Cohesion, and Economic Growth,” *American Economic Review*, 92(4), 1192–1204.
- GREEN, A. (1990): *Education and State Formation in England, France and the USA*. Palgrave Macmillan, UK.
- HAUK, E., AND J. ORTEGA (2015): “Schooling, Nation Building, and Industrialization: a Gellnerian Approach,” *Barcelona GSE Working Paper n 825*.
- HERBST, J. (2000): *States and Power in Africa: Comparative Lessons in Authority and Control*. Princeton University Press, Princeton, NJ.
- HITCHEN, P. (2000): “State and Church in British Honduran Education, 1931-1939: a British Colonial Perspective,” *History of Education*, 29(3), 195–211.

- HOBBSAWM, E. (1990): *Nations and Nationalism since 1780: Programme, Myth, Reality*. Cambridge University Press, Cambridge UK.
- HOLLAND, M. (2005): “The Captain Swing Project,” in *Swing Unmasked: The Agricultural Riots of 1830 to 1832 and Their Wider Implications*, ed. by M. Holland. FACHRS publications, UK.
- KATZNELSON, I., AND M. WEIR (1985): *Schooling for All: Class, Race, and the Decline of the Democratic Ideal*. Basic Books, New York.
- LAITIN, D. (1989): “Language Policy and Political Strategy in India,” *Policy Sci*, 22, 415-436.
- LAITIN, D. (2007): *Nation, States, and Violence*. Oxford University Press, Oxford UK.
- LAITIN, D., AND R. RAMACHANDRAN (2016): “Language Policy and Human Development,” *American Political Science Review*, 110(3), 457–480.
- LAWSON, R. F., ET AL. (2019): *Education*. Encyclopaedia Britannica, inc.
- LEWIS, K. (2000): “Colonial Education: A History of Education in Belize,” *Paper prepared for the Annual Meeting of the American Educational Research Association, New Orleans, USA, 24-28 Apr., 2000*.
- LOTT, JR., J. R. (1987): “Why Is Education Publicly Provided? A Critical Survey,” *Cato Journal*, 7(2), 475–501.
- (1999): “Public Schooling, Indoctrination, and Totalitarianism,” *Journal of Political Economy*, 107(6), 127–157.
- MARCHAM, A. (1973): “Educating our Masters: Political Parties and Elementary Education 1867 to 1870,” *British Journal of Educational Studies*, 21(2), 180–191.
- MARSHALL, M. G., T. GURR, AND K. JAGGERS (2016): “Polity IV Project: Political Regime Characteristics and Transitions, 1800-2015 Dataset Users’ Manual,” *Centre for Systemic Peace*.
- MARTEN, L., AND N. KULA (2008): “Zambia: One Zambia, One Nation, Many Languages,” in *Language and National Identity in Africa*, ed. by A. Simpson. Oxford University Press, Oxford.
- MICHALOPOULOS, S., AND E. PAPAIOANNOU (2014): “National Institutions and Subnational Development in Africa,” *The Quarterly Journal of Economics*, 129(1), 151–213.
- (2016): “The Long-Run Effects of the Scramble for Africa,” *American Economic Review*, 106(7), 1802–1848.
- MONDAL, A. (2017): “Free and Compulsory Primary Education in India Under the British Raj: A Tale of an Unfulfilled Dream,” *Sage Open*, 7(3), 1–12.

- MULLIGAN, C. B., R. GIL, AND X. SALA-I-MARTIN (2004): “Do Democracies Have Different Public Policies than Nondemocracies?,” *Journal of Economic Perspectives*, 18(1), 51–74.
- MURRAY, T. (2014): *Schooling in the ASEAN Region: Primary and Secondary Education in Indonesia, Malaysia, the Philippines, Singapore and Thailand*. Pergamon Press.
- ODAKA, K. (2002): “The Evolution of Social Policy in Japan,” *World Bank Institute Working Paper No. 33053*.
- OJHA, P. (1966): *Progress of Compulsory Education in India*. Delhi: Universal Book and Stationery Co.
- ORTEGA, J., AND T. TANGERS (2008): “Unilingual Versus Bilingual Education: A Political Economy Analysis,” *Journal of the European Economic Association*, 6(5), 1078–1108.
- PAGLAYAN, A. S. (2017): “Civil War, State Consolidation, and the Spread of Mass Education,” *Working paper*.
- (2018): “Democracy and Educational Expansion: Evidence from 200 Years,” *Working paper*.
- PAPAIIOANNOU, E., AND G. SIOUROUNIS (2008): “Democratisation and Growth,” *Economic Journal*, 118(532), 1520–1551.
- PHIRI, B. J. (2006): *A Political History of Zambia: From the Colonial Period to the 3rd Republic*. Africa World Press, Inc.
- POYKER, M. (2018): “Regime Stability and the Persistence of Traditional Practices,” *Columbia University, mimeo*.
- RAMIREZ, F., AND J. BOLI (1987): “The Political Construction of Mass Schooling: European Origins and Worldwide Institutionalization,” *Sociology of Education*, 60(1), 2–17.
- SHIMODA, H. (2010): “Tongues-Tied: The Making of a National Language and the Discovery of Dialects in Meiji Japan,” *The American Historical Review*, 115(3), 714–731.
- SKINNER, G. W. (1957): “Chinese Assimilation and Thai Politics,” *The Journal of Asian Studies*, 16(2), 237–250.
- SMITH, A. (1998): *Nationalism and Modernism: A Critical Survey of Recent Theories of Nations and Nationalism*. Routledge, London.
- SURYADINATA, L. (1997): *Ethnic Chinese as Southeast Asians*. Singapore: Institute of Southeast Asian Studies.

- TAKAYAMA, K. (1988): “Revitalization Movement of Modern Japanese Civil Religion,” *Sociological Analysis*, 48(4), 328–341.
- TILLY, C. (1975): “Reflections on the History of European State-Making,” in *The Formation of National States in Western Europe*, ed. by C. Tilly. Princeton University Press, Princeton NJ.
- (1998): *Disturbances in France, 1830-1860 and 1930-1960: Intensive Sample*. Inter-University Consortium for Political and Social Research, Ann Arbor, MI. doi:10.3886/ICPSR00051.v1.
- URCH, G. (1971): “Education and Colonialism in Kenya,” *History of Education Quarterly*, 11(3), 249–264.
- US DEPARTMENT OF HEALTH, E., AND WELFARE (1964): “Education in Uganda,” .
- WACHTEL, A. B., ET AL. (2020): *Russia*. Encyclopaedia Britannica, inc.
- WEBER, A. (2009): “The Causes of Politicization of Ethnicity: A Comparative Case Study of Kenya and Tanzania,” *CIS Working Paper No. 47*.
- WEBER, E. (1979): *Peasants into Frenchmen: Modernization of Rural France, 1870-1914*. Chatto and Windus, London.
- YOU, Y. (2018): “Unified Language, Labor and Ideology,” *Mimeo*.
- YOUNG, G., AND W. HANDCOCK (1964): *English Historical Documents: Vol. 9, 1833-1874*. Eyre and Spottiswoode, London.
- ZÜRCHER, E. J. (2010): *The Young Turk Legacy and Nation Building: From the Ottoman Empire to Atatürk’s Turkey*. I.B. Taurus, London.

Table 1
Descriptive statistics

| | Obs. | Mean | St. dev. | | Obs. | Mean | St. dev. |
|--------------------------------------|------|----------|----------|--|------|-------|----------|
| Educational reform: 10% threshold | 977 | 0.350 | 0.477 | Legal reform (Flora, 1983) | 110 | 0.464 | 0.501 |
| Educational reform: 20% threshold | 977 | 0.255 | 0.436 | Autocracy | 110 | 0.038 | 0.164 |
| Autocracy | 977 | 0.514 | 0.461 | Government crises | 110 | 0.441 | 0.622 |
| Government crises | 977 | 0.224 | 0.358 | Revolutions | 110 | 0.049 | 0.214 |
| Revolutions | 977 | 0.214 | 0.353 | All internal conflicts | 110 | 0.839 | 1.631 |
| All internal conflicts | 977 | 1.080 | 1.451 | Autocracy*government crises | 110 | 0.025 | 0.195 |
| Durability | 1387 | 0.201 | 0.245 | Autocracy*revolutions | 110 | 0.004 | 0.033 |
| Autocracy*government crises | 977 | 0.077 | 0.202 | Autocracy*all internal conflicts | 110 | 0.079 | 0.540 |
| Autocracy*revolutions | 977 | 0.127 | 0.277 | Population growth | 110 | 0.054 | 0.039 |
| Autocracy*all internal conflicts | 977 | 0.534 | 1.050 | Legal reforms (Flora, 1983) | 211 | 0.521 | 0.501 |
| Autocracy*durability | 1387 | 0.102 | 0.168 | Durability | 211 | 0.289 | 0.271 |
| Autocracy*international war | 977 | 0.087 | 0.271 | Autocracy*durability | 211 | 0.078 | 0.154 |
| Population growth | 977 | 0.187 | 0.152 | Autocracy | 211 | 0.341 | 0.460 |
| International war | 977 | 0.179 | 0.384 | Population growth | 211 | 0.063 | 0.087 |
| Log(revenue_expend) | 781 | 10.026 | 1.952 | Health insurance reform (Flora, 1983) | 110 | 0.036 | 0.188 |
| Log(GDP per capita) | 781 | 7.922 | 1.008 | Pension insurance reform (Flora, 1983) | 110 | 0.055 | 0.228 |
| Urbanization | 949 | 0.629 | 2.339 | Unempl. insurance reform (Flora, 1983) | 110 | 0.064 | 0.245 |
| Trade | 962 | 0.930 | 2.268 | Franchise extension | 175 | 0.006 | 0.076 |
| Primary school enrollment per capita | 2798 | 1202.949 | 447.186 | Leader's number of years in power | 1002 | 5.388 | 6.406 |

Table 2
Educational reforms and threat to democracy

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

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 10% (20%) from the previous 10 years (columns 1-3, and columns 4-6). Legal reform is a dummy if the country experienced at least one legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All explanatory variables are lagged. Standard errors are clustered at the country level.

Table 3
Educational reforms and threat to democracy, controlling for alternative explanations

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
|---|-----------------------------------|----------------------------|---------------------------|-------------------------|----------------------------|---------------------------|-------------------------|---------------------------|---------------------------|-------------------------|---------------------------|--------------------------|-------------------------|----------------------------|---------------------------|
| | Educational reform: 20% threshold | | | | | | | | | | | | | | |
| Autocracy*gov. crises | 0.171* (0.097) | | | 0.134 (0.107) | | | 0.143 (0.100) | | | 0.137 (0.101) | | | 0.159 (0.099) | | |
| Autocracy*revolutions | | 0.235*** (0.075) | | | 0.227*** (0.083) | | | 0.184** (0.076) | | | 0.208** (0.095) | | | 0.204*** (0.076) | |
| Autocracy*all internal conflicts | | | 0.057** (0.023) | | | 0.051** (0.023) | | | 0.049** (0.022) | | | 0.043* (0.023) | | | 0.047** (0.023) |
| Autocracy | 0.010 (0.066) | 0.005 (0.058) | -0.012 (0.065) | 0.006 (0.073) | -0.006 (0.063) | -0.017 (0.072) | -0.038 (0.063) | -0.031 (0.055) | -0.054 (0.064) | -0.001 (0.066) | -0.017 (0.064) | -0.015 (0.066) | -0.014 (0.066) | -0.028 (0.065) | -0.028 (0.065) |
| Pop. growth | 0.027 (0.169) | -0.007 (0.168) | 0.012 (0.167) | -0.084 (0.155) | -0.111 (0.154) | -0.098 (0.154) | -0.076 (0.144) | -0.095 (0.146) | -0.084 (0.142) | -0.029 (0.166) | -0.029 (0.165) | -0.039 (0.165) | -0.109 (0.146) | -0.113 (0.145) | -0.113 (0.145) |
| Gov. crises | -0.120** (0.054) | | | -0.096* (0.058) | | | -0.122** (0.057) | | | -0.120** (0.053) | | | -0.136** (0.054) | | |
| Revolutions | | -0.006 (0.057) | | | -0.018 (0.068) | | | -0.004 (0.058) | | | -0.009 (0.060) | | | -0.016 (0.057) | |
| All internal conflicts | | | -0.014 (0.016) | | | -0.008 (0.017) | | | -0.015 (0.016) | | | -0.016 (0.015) | | | -0.017 (0.016) |
| International war | 0.020 (0.053) | 0.012 (0.054) | 0.013 (0.054) | | | | | | | | | | | | |
| Autocracy*international war | -0.135* (0.081) | -0.139* (0.079) | -0.142* (0.080) | | | | | | | | | | | | |
| Log(revenue) | | | | 0.060* (0.034) | 0.065** (0.033) | 0.068** (0.033) | | | | | | | | | |
| Log(per capita GDP) | | | | | | | 0.084* (0.050) | 0.097* (0.049) | 0.095* (0.049) | | | | | | |
| Urbanization | | | | | | | | | | -0.013 (0.009) | -0.012 (0.008) | -0.012 (0.008) | | | |
| Trade | | | | | | | | | | | | | -0.005 (0.016) | -0.006 (0.018) | -0.007 (0.017) |
| Number of countries | 172 | 172 | 172 | 158 | 158 | 158 | 156 | 156 | 156 | 171 | 171 | 171 | 171 | 171 | 171 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 977 | 977 | 977 | 861 | 861 | 861 | 871 | 871 | 871 | 949 | 949 | 949 | 962 | 962 | 962 |
| R-squared | 0.410 | 0.418 | 0.412 | 0.445 | 0.452 | 0.448 | 0.460 | 0.464 | 0.461 | 0.420 | 0.423 | 0.419 | 0.434 | 0.437 | 0.433 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 20% from the previous 10 years. All explanatory variables are lagged. Standard errors are clustered at the country level.

Table 4
Other redistributive reforms and threat to democracy

| | (1) | (2) | (3) | (7) | (8) | (9) | (10) | (11) | (12) |
|---|------------------|----------------|----------------|-------------------|----------------|----------------|------------------------|----------------|----------------|
| | Health insurance | | | Pension insurance | | | Unemployment insurance | | |
| Autocracy*gov. crises | -0.069 | | | -0.066 | | | -0.018 | | |
| | (0.158) | | | (0.176) | | | (0.082) | | |
| Autocracy*revolutions | | -0.191 | | | -0.510 | | | -0.137 | |
| | | (0.575) | | | (0.984) | | | (0.328) | |
| Autocracy*all internal conflicts | | | -0.024 | | | -0.059 | | | -0.040 |
| | | | (0.058) | | | (0.085) | | | (0.044) |
| Autocracy | 0.283 | 0.243 | 0.279 | 0.012 | 0.019 | 0.120 | -0.209 | -0.180 | -0.101 |
| | (0.353) | (0.298) | (0.361) | (0.241) | (0.209) | (0.284) | (0.180) | (0.129) | (0.154) |
| Pop. growth | -1.309 | -1.369 | -1.325 | -1.115 | -1.099 | -0.999 | 0.476 | 0.577 | 0.534 |
| | (0.803) | (0.888) | (0.850) | (0.665) | (0.736) | (0.738) | (0.835) | (0.845) | (0.840) |
| Gov. crises | 0.010 | | | -0.007 | | | 0.061 | | |
| | (0.043) | | | (0.056) | | | (0.057) | | |
| Revolutions | | -0.032 | | | 0.011 | | | 0.320 | |
| | | (0.050) | | | (0.152) | | | (0.203) | |
| All internal conflicts | | | -0.001 | | | 0.024 | | | 0.045 |
| | | | (0.009) | | | (0.039) | | | (0.035) |
| Number of countries | 14 | 14 | 14 | 14 | 14 | 14 | 14 | 14 | 14 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 110 | 110 | 110 | 110 | 110 | 110 | 110 | 110 | 110 |
| R-squared | 0.281 | 0.281 | 0.281 | 0.177 | 0.178 | 0.192 | 0.166 | 0.202 | 0.192 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Health, pension and unemployment insurance are dummies indicating if the country implemented an health, pension or unemployment insurance reform. The dates come from Flora (1983). All explanatory variables are lagged. Standard errors are clustered at the country level.

Table 5
Leaders' number of years in power and increase in education during their reign

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
|--|--------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| | Number of years in power | | | | | |
| Autocracy*(growth in primary education) | 13.810** | 14.483** | 14.834** | 13.471** | 13.692** | 13.959** |
| | (6.716) | (6.721) | (6.043) | (6.353) | (6.717) | (6.668) |
| Growth in primary education | -8.296* | -8.752** | -7.137* | -6.743 | -8.254* | -8.359* |
| | (4.311) | (4.305) | (3.995) | (4.101) | (4.309) | (4.266) |
| Autocracy | 1.594** | 1.539** | 1.714*** | 1.878*** | 1.559** | 1.556** |
| | (0.686) | (0.682) | (0.591) | (0.569) | (0.681) | (0.687) |
| International war | | 4.158 | | | | |
| | | (3.954) | | | | |
| Autocracy*(international war) | | 0.204*** | | | | |
| | | (0.007) | | | | |
| Pop. Growth | -13.441 | -14.175 | -27.418 | -32.935 | -11.224 | -11.320 |
| | (15.862) | (15.879) | (17.384) | (20.136) | (15.748) | (15.630) |
| Log(revenue) | | | 0.031*** | | | |
| | | | (0.004) | | | |
| Log(per capita GDP) | | | | 4.295*** | | |
| | | | | (1.244) | | |
| Urbanization | | | | | 0.043 | |
| | | | | | (0.044) | |
| Trade | | | | | | 0.021** |
| | | | | | | (0.010) |
| Number of countries | 106 | 106 | 106 | 106 | 106 | 106 |
| Country fixed effects | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes |
| Observations | 1,002 | 1,002 | 971 | 971 | 1,002 | 1,002 |
| R-squared | 0.497 | 0.502 | 0.584 | 0.597 | 0.498 | 0.500 |

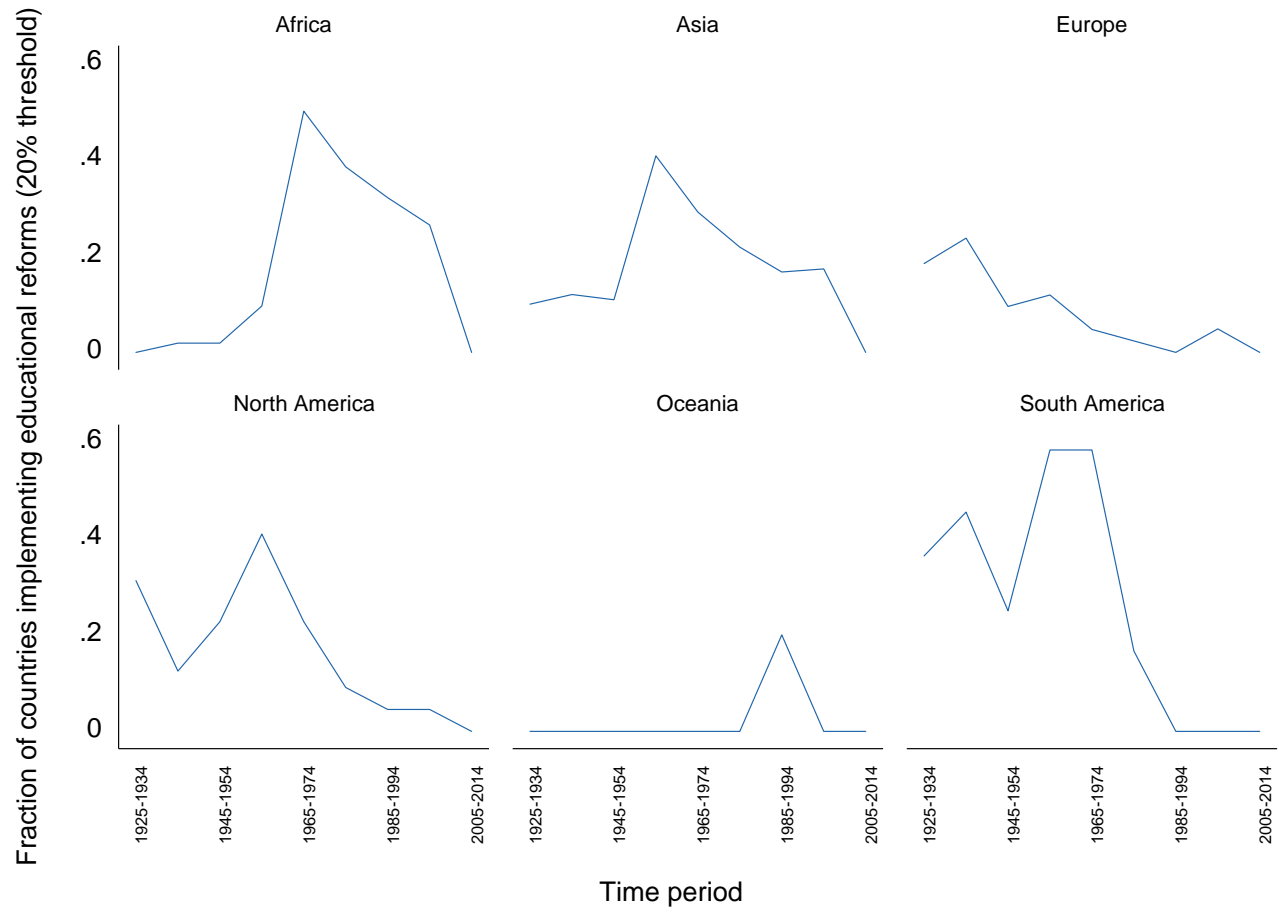
Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1875-2004 period. Standard errors are clustered at the country level

Table 6
Educational reforms and regime durability

| | (1) | (2) | (3) | (4) | (5) | (6) |
|--------------------------------------|-----------------------------------|-----------------|------------------|------------------|----------------|------------------|
| | Educational reform: 20% threshold | | | | | |
| Autocracy*(regime durability) | -0.300** | -0.302** | -0.508*** | -0.552*** | -0.174 | -0.726*** |
| | (0.149) | (0.149) | (0.133) | (0.157) | (0.121) | (0.165) |
| Regime durability | 0.030 | 0.032 | 0.217** | 0.207** | 0.072 | 0.440*** |
| | (0.108) | (0.113) | (0.098) | (0.096) | (0.058) | (0.106) |
| Autocracy | 0.084 | 0.092* | 0.071 | 0.049 | 0.061 | 0.070 |
| | (0.053) | (0.053) | (0.055) | (0.054) | (0.050) | (0.067) |
| Pop. growth | 0.155 | 0.150 | 0.026 | 0.015 | -0.029 | -0.230* |
| | (0.103) | (0.105) | (0.105) | (0.114) | (0.085) | (0.125) |
| International war | | -0.005 | | | | |
| | | (0.051) | | | | |
| Autocracy*international war | | -0.054 | | | | |
| | | (0.070) | | | | |
| Log(revenue) | | | 0.009 | | | |
| | | | (0.025) | | | |
| Log(per capita GDP) | | | | 0.066 | | |
| | | | | (0.042) | | |
| Urbanization | | | | | -0.007 | |
| | | | | | (0.011) | |
| Trade | | | | | | 0.002 |
| | | | | | | (0.012) |
| Number of countries | 183 | 183 | 162 | 156 | 175 | 171 |
| Country fixed effects | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes |
| Observations | 1,387 | 1,387 | 1,131 | 1,100 | 1,308 | 974 |
| R-squared | 0.348 | 0.349 | 0.373 | 0.401 | 0.335 | 0.420 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1825-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 20% from the previous 10 years. All explanatory variables are lagged. Standard errors are clustered at the country level.

Figure 1
Fraction of countries implementing structural reforms, by continent



Appendix

Analysis of the Choices made by a Period 2 Democracy.

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

We first show a democracy will locate the government at the center of the country.

First examine the case where the ruler undertakes no homogenization in period 1, that is $\lambda_1^{1/2} = 0$. Where would a period 2 democracy choose to locate the government in a single country?

For any government location $j \in [1/4, 3/4]$, the median voter is at distance $d_{ij} = 1/4$, and so the level of homogenization chosen by majority rule, denoted λ_j^m , is the same for all $j \in [1/4, 3/4]$; it satisfies $ga/4 = C'(\lambda_j^m)$. Since the level of homogenization is the same for any $j \in [1/4, 3/4]$ and since at least half of the population are closer to $j = 1/2$ than any $j \in [1/4, 3/4]$, locating the government at $j = 1/2$ beats all other $j \in [1/4, 3/4]$ in a pairwise vote.

For any government location $j \in [0, 1/4)$, the level of homogenization chosen by majority rule satisfies $ga(1/2 - j) = C'(\lambda_j^m)$, which is strictly higher than $\lambda_{1/2}^m$ above. 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 $1/2$, the center of the population. Let \hat{l}_i denote this distance for the individual $i \in [0, 1/2]$ who is indifferent between a government at some fixed $j \in [0, 1/4)$ and a government at the center $j = 1/2$. Similarly denote by $\hat{\hat{l}}_i$ the distance of the individual that satisfies the same condition on the interval $i \in [1/2, 1]$. If all $i \in [1/2, 1]$ prefer $j = 1/2$ then set $\hat{\hat{l}}_i = 1/2$. First note that the relative utility from $j = 1/2$ versus some fixed $j \in [0, 1/4)$ is strictly decreasing in l_i . It follows that \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. Individual $i = 3/4$ strictly prefers a government at $1/2$ with homogenization $\lambda_{1/2}^m$ to a government at $j \in [0, 1/4)$ with homogenization λ_j^m since the government at $1/2$ is closer and $\lambda_{1/2}^m$ is the optimal homogenization in that case for $i = 3/4$ and therefore a government at $j \in [0, 1/4)$ can only be strictly worse. It follows that necessarily $\hat{\hat{l}}_i > 1/4$, and 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

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

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

Expressions (1) and (2) 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 $\hat{l}_i + \hat{\hat{l}}_i > 1/2$. The argument for why a government located at $j = 1/2$ beats all $j \in (3/4, 1]$ is symmetric.

Second, examine the case where the ruler undertakes positive homogenization in period 1, that is $\lambda_1^{1/2} \in (0, 1]$. Where would a period 2 democracy choose to locate the government in a single country? The ruler implements some positive amount of homogenization in period 1 and the cost is incurred in period 1. In a period 2 democracy, each individual's utility from a government at $j = 1/2$ is strictly higher than in the case above where the ruler implements zero homogenization in period 1. In contrast, the utility from a government at any $j \neq 1/2$ is the same as above. From above we know that a majority of the population prefer a government at $j = 1/2$ in period 2 when the ruler undertakes no homogenization in period 1. It follows that a majority must prefer a government at $j = 1/2$ when the ruler undertakes positive homogenization.

Finally we examine the case where a ruler undertakes homogenization $\lambda_1^{1/2} \in [0, 1]$ in period 1 and a period 2 democracy chooses to split. A democracy that splits will choose to locate the governments at $1/4$ in Country A and $3/4$ in Country B. By the same argument as a single country, $j = 1/4$ beats all other locations $j \neq 1/2$ in a pairwise vote. Additionally, we need to show that $j = 1/4$ necessarily beats $j = 1/2$, even when the ruler has implemented positive homogenization. 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 the government $j = 1/2$. By symmetry, a majority in Country B must also prefer a single country with $j = 1/2$ to a split and so a democracy forms a single country. \square

We show when a democracy will form a single country or split.

Suppose the ruler implements homogenization $\lambda_1^{1/2} \in [0, \lambda_{1/2}^m)$ in period 1. Consider a period 2 democracy. Expression (3) gives the period 2 utility of individual i at distance $l_i \in [0, 1/4]$ from the center if a single country is formed minus his utility from a split:

$$\left[g - (1 - \lambda_{1/2}^m)gal_i + y - k - C(\lambda_{1/2}^m) + C(\lambda_1^{1/2}) \right] - \left[g - (1 - \lambda_{1/4}^m)ga(1/4 - l_i) + y - 2k - C(\lambda_{1/4}^m) \right] \quad (3)$$

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

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

Expression (3) is decreasing in l_i and expression (4) is increasing in l_i (since $\lambda_{1/4}^m < \lambda_{1/2}^m$), with both expressions equal at $l_i = 1/4$. Thus there exist uniquely two different individuals, $l'_i \in [0, 1/4]$ and $l''_i \in [1/4, 1/2]$, with the same value of (3) and (4) respectively and such that $l'_i + (0.5 - l''_i) = 1/4$. These two individuals at distance l'_i, l''_i have the median valuation of a single country versus a split. The value of l'_i solves

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

and so

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

Suppose instead the ruler implements homogenization $\lambda_1^{1/2} \geq \lambda_{1/2}^m$ in period 1. We can find the two individuals who have the median valuation of a single country versus a split in the same way as above. The utility of individual i at distance $l_i \in [0, 1/4]$ from the center in a single country versus a split is

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

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

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

By the same argument as above, there exist uniquely two different individuals, $l'_i \in [0, 1/4]$ and $l''_i \in [1/4, 1/2]$, with the same value of (5) and (6) respectively, such that $l'_i + (0.5 - l''_i) = 1/4$. Solving for l'_i and l''_i we find

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

Proof of Lemma 1

Suppose the ruler homogenizes by $\lambda_1^{1/2}$ in period 1. Expressions (3) and (5) evaluated at l'_i give the median voter's valuation of forming a single country versus splitting in period 2. Expressions (3) and (5) are strictly increasing in $\lambda_1^{1/2}$ and are equal at $\lambda_1^{1/2} = \lambda_{1/2}^m$. At $\lambda_1^{1/2} = 1$, expression (5) is positive. Therefore there exists some level of homogenization undertaken by the ruler in period 1, denoted λ^* , above which a period 2 democracy will form a single country and below which it will split. If (3) and (5) are positive for all $\lambda_1^{1/2} \geq 0$, then $\lambda^* = 0$ and a democracy will always form a single country. Otherwise $1 > \lambda^* > 0$ and is the unique value of $\lambda_1^{1/2}$ at which (3) or (5) evaluated at l'_i is equal to zero. \square

Proof of Proposition 1

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

The ruler's expected utility in period 1 if the ruler homogenizes by $\lambda_1^{1/2} < \lambda^*$ is

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

The ruler's expected utility in period 1 if the ruler homogenizes by $\lambda_1^{1/2} \geq \lambda^*$ is

$$\left[g + y - k - C(\lambda_1^{1/2}) + R \right] + p \left[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})] \right] + (1-p) [g + y - k + R] \quad (8)$$

if $\lambda_1^{1/2} < \lambda_{1/2}^m$, and his utility is

$$\left[g + y - k - C(\lambda_1^{1/2}) + R \right] + p [g + y - k] + (1-p) [g + y - k + R] \quad (9)$$

if $\lambda_1^{1/2} \geq \lambda_{1/2}^m$.

The ruler chooses either $\lambda_1^{1/2} = 0$ or $\lambda_1^{1/2} = \lambda^*$, since his expected utility from any other level of homogenization is strictly lower than at least one of $\lambda_1^{1/2} = 0$ or $\lambda_1^{1/2} = \lambda^*$. When $p = 0$ expected utility is maximized at $\lambda_1^{1/2} = 0$. Since $\lambda_{1/2}^m$ is the optimal homogenization for an individual at distance $d_{ij} = 1/4$ from the government j then we have the following inequality $(1 - \lambda_{1/4}^m)ga/4 + C(\lambda_{1/4}^m) > (1 - \lambda_{1/2}^m)ga/4 + C(\lambda_{1/2}^m)$. From this inequality we know (7) is decreasing in p at a faster rate than (8) and (9). Therefore if λ^* is chosen by the ruler for any p it is chosen for all p higher.

Case 2: $\lambda^* = 0$. In this case a period 2 democracy will always form a single country and the ruler undertakes no homogenization for all p .

□

Proof of Corollary 1

From the proof of Proposition 1 above, if $\lambda_{1/2}^m > \lambda^* > 0$, then the ruler homogenizes to λ^* if

$$\begin{aligned} & [g + y - k - C(\lambda^*) + R] + p [g + y - k - [C(\lambda_{1/2}^m) - C(\lambda^*)]] + (1-p) [g + y - k + R] \\ & > [g + y - k + R] + p [g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1-p) [g + y - k + R] \end{aligned}$$

rewritten

$$p \left[(1 - \lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m) - C(\lambda_{1/2}^m) \right] > (1-p)C(\lambda^*).$$

From the proof of Proposition 1, we know this inequality is satisfied for $p = 1$ and so there exists a $\bar{p} < 1$ such that for all $p > \bar{p}$ the ruler chooses λ^* . If the ruler undertakes λ^* then a democracy forms a single country and the population is homogenized to $\lambda_{1/2}^m > \lambda_{1/4}^m$. If $p \leq \bar{p}$ the ruler undertakes zero homogenization in period 1 and a democracy in period 2 splits and the population is homogenized to $\lambda_{1/4}^m$.

From the proof of Proposition 1, if $\lambda^* \geq \lambda_{1/2}^m$, then the ruler homogenizes to λ^* if

$$\begin{aligned} & [g + y - k - C(\lambda^*) + R] + p [g + y - k] + (1 - p) [g + y - k + R] \\ & > [g + y - k + R] + p [g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p) [g + y - k + R] \end{aligned}$$

rewritten

$$p [(1 - \lambda_{1/4}^m)ga/4 + k + C(\lambda_{1/4}^m)] > C(\lambda^*).$$

From the proof of Proposition 1, we know that there exists a some \bar{p} such that for all $p > \bar{p}$ the ruler chooses λ^* . If the ruler undertakes λ^* then a democracy forms a single country and the population is homogenized to $\lambda^* > \lambda_{1/4}^m$. However, if the required homogenization λ^* is too costly then it could be that $\bar{p} = 1$, such that the ruler never homogenizes to ensure a single country and his ideal government because it is too expensive. If $p \leq \bar{p}$ the ruler undertakes zero homogenization in period 1 and a democracy in period 2 splits and the population is homogenized to $\lambda_{1/4}^m$.

If $\lambda^* = 0$, the ruler does not undertake homogenization because a single country and his ideal government will be chosen by a democracy anyway.

Proof of Proposition 2 and 3

Let $\lambda_1^{1/2}$ denote homogenization by the ruler in period 1. We first show that $u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2})$ is decreasing in $\lambda_1^{1/2}$.

Case 1: The ruler homogenizes by $\lambda_1^{1/2} < \lambda^*$.

Utility under a democracy in period 2 relative to utility under a ruler in period 2 for an individual i at distance l_i from the center is

$$u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2}) = [g - (1 - \lambda_{1/4}^m)ga|0.25 - l_i| + y - 2k - C(\lambda_{1/4}^m)] - [g - (1 - \lambda_1^{1/2})gal_i + y - k] \quad (10)$$

For $\lambda_1^{1/2} \leq \lambda_{1/4}^m$, (10) is increasing in distance l_i and so the median valuation occurs for the individual at distance $l_i = 1/4$. For $\lambda_1^{1/2} > \lambda_{1/4}^m$ (following the same process as in the analysis of a democracy above), we find the median valuation of (10) occurs for the individuals at distances

$$l_i = \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda_1^{1/2})}{8(1 - \lambda_{1/4}^m)} \quad \text{and} \quad l_i = 0.25 + \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda_1^{1/2})}{8(1 - \lambda_{1/4}^m)}.$$

Plugging into (10) the value of l_i for the median voter and taking the derivative of (10) with respect to $\lambda_1^{1/2}$, we find that the median value of $u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2})$ is decreasing in $\lambda_1^{1/2}$ at rate $-0.25ga$ when $\lambda_1^{1/2} \leq \lambda_{1/4}^m$ and decreasing in $\lambda_1^{1/2}$ at rate $-2ga \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda_1^{1/2})}{8(1 - \lambda_{1/4}^m)}$ when $\lambda_1^{1/2} > \lambda_{1/4}^m$.

Case 2: The ruler homogenizes by $\lambda_1^{1/2} \geq \lambda^*$.

If $\lambda_1^{1/2} < \lambda_{1/2}^m$ then

$$u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2}) = [g - (1 - \lambda_{1/2}^m)gal_i + y - k - [C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})]] - [g - (1 - \lambda_1^{1/2})gal_i + y - k] \quad (11)$$

The median valuation of (11) occurs for individual i at distance $l_i = 1/4$. The median value of expression (11) is decreasing in $\lambda_1^{1/2}$ at rate $C'(\lambda_1^{1/2}) - 0.25ga$, which we know is negative since $\lambda_1^{1/2} < \lambda_{1/2}^m$ and by definition $C'(\lambda_{1/2}^m) = 0.25ga$.

If $\lambda_1^{1/2} \geq \lambda_{1/2}^m$ then

$$u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2}) = [g - (1 - \lambda_1^{1/2})gal_i + y - k] - [g - (1 - \lambda_{1/2}^m)gal_i + y - k] \quad (12)$$

This expression is zero for all l_i and does not change in $\lambda_1^{1/2}$.

We have shown that the median values of (10), (11) and (12) are weakly decreasing in $\lambda_1^{1/2}$ in the relevant range. It remains to show that as $\lambda_1^{1/2}$ approaches λ^* , the median value of (10) is weakly higher than the median value of (11) (respectively (12)) at λ^* . We know from above that for the individual at distance

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

the value of (10) at λ^* is equal to the value of (11) at λ^* . However at λ^* the median value of (11) occurs for the individual at distance $l_i = 1/4$ and the median value of (10) occurs for the individual at distance $l_i = \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda^*)}{8(1 - \lambda_{1/4}^m)}$. For an individual at distance $l_i = \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda^*)}{8(1 - \lambda_{1/4}^m)}$ relative to $l_i = \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda_{1/2}^m)}{8(1 - \lambda_{1/4}^m)}$ the value of (10) increases by

$$0.25ga[(1 - \lambda_{1/4}^m) + (1 - \lambda^*)] \frac{(1 - \lambda^*) - (1 - \lambda_{1/2}^m)}{2(1 - \lambda_{1/4}^m)}. \quad (13)$$

For an individual at distance $l_i = 0.25$ relative to $l_i = \frac{(1 - \lambda_{1/4}^m) + (1 - \lambda_{1/2}^m)}{8(1 - \lambda_{1/4}^m)}$ the value of (11) increases by

$$0.25ga[(1 - \lambda^*) - (1 - \lambda_{1/2}^m)] \frac{(1 - \lambda_{1/4}^m) - (1 - \lambda_{1/2}^m)}{2(1 - \lambda_{1/4}^m)} \quad (14)$$

Expression (13) is higher than expression (14) and the result follows. Finally we show that as $\lambda_1^{1/2}$ approaches λ^* , the median value of (10) is weakly higher than the median value of (12) at λ^* . This is immediate since (10) approaches zero at λ^* and (12) is zero at λ^* .

At $\lambda_1^{1/2} = \max\{\lambda^*, \lambda_{1/2}^m\} < 1$, the median value of $u_{i2,dem}(\lambda_1^{1/2}) - u_{i2,ruler}(\lambda_1^{1/2})$ is zero and the population will choose not to overthrow the ruler. Let $\bar{\lambda}$ denote the lowest value

of $\lambda_1^{1/2}$ at which the population will choose not to overthrow the ruler.

We next examine the ruler's optimal homogenization in period 1.

Any level of homogenization other than 0, $\bar{\lambda}$ or λ^* is not optimal for the ruler. Further, if $\bar{\lambda} \leq \lambda^*$ then only homogenization 0 or $\bar{\lambda}$ can be optimal.

Case 1: $\lambda^* = 0$. The ruler chooses 0 or $\bar{\lambda}$.

The ruler's expected utility if he undertakes zero homogenization is

$$[g + y - k + R] + p[g + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k + R] \quad (15)$$

and if he undertakes homogenization $\bar{\lambda}$ is

$$[g + y - k - C(\bar{\lambda}) + R] + [g + y - k + R]. \quad (16)$$

When $p = 0$, (15) is strictly higher than (16). Expression (15) is decreasing in p while (16) does not change in p , therefore if $\bar{\lambda}$ gives higher expected utility for any p then $\bar{\lambda}$ gives higher expected utility for all p higher. Let \bar{p} denote the value of p at which (16) is equal to (15). Since (15) is strictly increasing in R at rate $2 - p$ and (16) at rate 2, then \bar{p} is falling in R .

Case 2: $\lambda^* > 0$. The ruler potentially chooses between 0, $\bar{\lambda}$ or λ^* .

The ruler's expected utility if he undertakes zero homogenization is

$$[g + y - k + R] + p[g - ga(1 - \lambda_{1/4}^m)0.25 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k + R] \quad (17)$$

The ruler's expected utility if he undertakes homogenization $\lambda^* < \lambda_{1/2}^m$ is

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

The ruler's expected utility if he undertakes homogenization $\lambda^* \geq \lambda_{1/2}^m$ is

$$[g + y - k - C(\lambda^*) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \quad (19)$$

The ruler's expected utility if he undertakes homogenization $\bar{\lambda}$ is

$$[g + y - k - C(\bar{\lambda}) + R] + [g + y - k + R] \quad (20)$$

When $p = 0$, expression (17) is strictly higher than (18),(19),(20). The ruler's expected utility from homogenization $\lambda^* < \lambda_{1/2}^m$ minus expected utility from homogenization zero is given by subtracting (18) from (17) to get

$$-C(\lambda^*) - p[C(\lambda_{1/2}^m) - C(\lambda^*)] + p[ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m)].$$

This expression is increasing in p . To see this use the fact that for the individual at distance $l_i = 1/4$ from the center the optimal homogenization is $\lambda_{1/2}^m$, and so we have that $ga(1 - \lambda_{1/4}^m)0.25 + C(\lambda_{1/4}^m) > ga(1 - \lambda_{1/2}^m)1/4 + C(\lambda_{1/2}^m)$ and so $ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m) > C(\lambda_{1/2}^m)$. The ruler's expected utility from homogenization $\lambda^* \geq \lambda_{1/2}^m$ minus expected utility from homogenization zero is given by subtracting (19) from (17) and is straightforward to see that it is increasing in p . Thus if expected utility from homogenization λ^* is higher than homogenization zero for any p then it is higher for all p higher.

The ruler's expected utility from homogenization $\bar{\lambda}$, given by (20), does not change in p . Given that (17), (18), (19) are decreasing in p , if $\bar{\lambda}$ gives higher expected utility than zero (respectively λ^*) at any p , then it gives higher expected utility than zero (respectively λ^*) for all p higher.

The above statements put together show that optimal homogenization chosen by the ruler is increasing in p .

The expected utility of homogenization zero and homogenization λ^* are increasing in R at rate $2 - p$. The expected utility of $\bar{\lambda}$ is increasing in R at rate 2. Therefore the value of p at which expected utility from 0 is equal to expected utility from λ^* does not change in R but the value of p at which expected utility from 0 or λ^* is equal to expected utility from $\bar{\lambda}$ is decreasing in R .

Proof of Proposition 4

We have to update the costs to describe the costs of negative homogenization. We update the function $C(\cdot)$ for $\lambda_t^j \in [-1, 1]$ such that

$$C(\lambda_t^j) = C(|\lambda_t^j|)$$

That is, the costs of negative and positive homogenization are symmetric around zero. If $\lambda_{t-1}^j \geq 0$ and $\lambda_t^j \geq \lambda_{t-1}^j$ or if $\lambda_{t-1}^j \leq 0$ and $\lambda_t^j \leq \lambda_{t-1}^j$, then the cost of homogenization λ_t^j following homogenization λ_{t-1}^j is as before

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

However if homogenization this period reverses the homogenization of the previous period, that is if $\lambda_{t-1}^j > 0$ and $\lambda_t^j < \lambda_{t-1}^j$ or if $\lambda_{t-1}^j < 0$ and $\lambda_t^j > \lambda_{t-1}^j$, then the cost is

$$sC(|\lambda_t^j - \lambda_{t-1}^j|).$$

We write it like this to deal appropriately with the case where for example $\lambda_t^j > 0$ but $\lambda_{t-1}^j < 0$. In the analysis of the Divide and Rule case we assume that the government in a democracy is always located at the center of the democratic country, rather than deriving it.¹

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

First examine the level of homogenization that will be chosen by a period 2 democracy. If a democracy splits then a democracy chooses $\lambda_{1/4}^m$ in Country A and $\lambda_{3/4}^m$ in Country B as before. If a democracy forms a single country and $\lambda_1^{1/2} \geq 0$ in period 1, then homogenization chosen by a democracy is unchanged from the results above. If a democracy forms a single country and $\lambda_1^{1/2} < 0$, a democracy will never choose to make the population even less homogenous. Suppose the democracy chooses homogenization $\lambda_2^{1/2}$, then utility of an individual at distance d_{ij} from the center is

$$g - ga(1 - \lambda_2^{1/2})d_{ij} + y - k - C(|\lambda_2^{1/2} - \lambda_1^{1/2}|).$$

Regardless of whether $\lambda_2^{1/2} < 0$ or $\lambda_2^{1/2} \geq 0$, the derivative w.r.t $\lambda_2^{1/2}$ is

$$gad_{ij} = C'(\lambda_2^{1/2} - \lambda_1^{1/2}).$$

The optimal value of $\lambda_2^{1/2} - \lambda_1^{1/2}$ is higher for a higher d_{ij} and so the median voter is at distance $d_{ij} = 1/4$. Thus following $\lambda_1^{1/2} < 0$, a democracy that forms a single country will choose homogenization $\lambda_2^{1/2} = \lambda_1^{1/2} + \lambda_{1/2}^m$.

Next we note that Lemma 1 extends to this specification of the model. That is, there exists a $\lambda^* \in [-1, 1]$ such that if the ruler homogenizes above this level a period 2 democracy forms a single country and below it splits. The proof is a replication of the proof of Lemma 1 and we leave it out.

The final step of this proof is to examine what level of homogenization a ruler will choose in period 1.

Case 1: $\lambda^* \geq 0$

The ruler will not homogenize above λ^* . This follows the same argument as Proposition 1 and by noting that in this case homogenization has a further negative effect on the utility of the ruler because it increases the probability of democracy, v . The expected utility of the ruler if he homogenizes by some $\lambda_1^{1/2}$ less than λ^* is

$$[g+y-k-C(\lambda_1^{1/2})+R]+v(p, \lambda_1^{1/2})[g-ga(1-\lambda_{1/4}^m)0.25+y-2k-C(\lambda_{1/4}^m)]+(1-v(p, \lambda_1^{1/2}))[g+y-k+R]$$

The derivative with respect to $\lambda_1^{1/2}$ is

$$-C'(\lambda_1^{1/2}) - \frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2}} [ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m) + R]. \quad (21)$$

For $\lambda_1^{1/2} \geq 0$, since $\frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2}} > 0$ and $C'(\lambda_1^{1/2}) > 0$, (21) is negative. As $\lambda_1^{1/2} \rightarrow -1$, (21) is positive since $C'(\lambda_1^{1/2})$ is negative and large in absolute value. Finally we note that

(21) is strictly decreasing in $\lambda_1^{1/2}$ for all $\lambda_1^{1/2} \in [-1, 1]$, since $C''' > 0$ and $\frac{\partial v(p, \lambda_1^{1/2})}{\partial^2 \lambda_1^{1/2}} \geq 0$. Thus there is a unique optimal level of homogenization for the ruler in the range $[-1, \lambda^*]$, denote it by $-\lambda^\circ$, which is strictly negative.

The expected utility of the ruler from homogenizing by $-\lambda^\circ$ is decreasing in p at rate

$$-\frac{\partial v(p, -\lambda^\circ)}{\partial p} [ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m) + R]$$

If $\lambda^* \geq \lambda_{1/2}^m$ then the ruler's expected utility from homogenizing to λ^* is

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

which is decreasing in p at rate

$$-\frac{\partial v(p, \lambda^*)}{\partial p} R$$

If $\lambda^* < \lambda_{1/2}^m$ then the ruler's expected utility from homogenizing to λ^* is

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

which is decreasing in p at rate

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

We know $ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m) > C(\lambda_{1/2}^m) - C(\lambda^*)$ and so provided $\frac{\partial v(p, \lambda_1^{1/2})}{\partial p}$ does not increase too fast in $\lambda_1^{1/2}$, the ruler's utility from homogenization to λ^* is decreasing in p at a lower rate than his utility from optimal homogenization below λ^* . So if the ruler chooses λ^* at any p he does so for all p higher. Below that he chooses strictly negative homogenization.

Homogenization chosen by the ruler is decreasing in R . First, note that the value of $-\lambda^\circ$ is decreasing in R . Since an increase in R lowers the second term in (21), and a fall in $\lambda_1^{1/2}$ increases the first term and reduces the term $\frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2}}$, the value of $\lambda_1^{1/2}$ at which (21) is equal to zero is decreasing in R . Second, note that the lowest value of p at which the ruler chooses λ^* is increasing in R , since the expected utility from $-\lambda^\circ$ is increasing in R at rate $2 - v(p, -\lambda^\circ)$ and the expected utility from λ^* is increasing in R at strictly lower rate of $2 - v(p, \lambda^*)$.

Case 2: $\lambda^* < 0$

If $\lambda^* < 0$ it is never optimal for the ruler to undertake positive homogenization. The ruler's expected utility from homogenizing to some $\lambda_1^{1/2} \in [\lambda^*, 0]$ is

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

The derivative with respect to $\lambda_1^{1/2}$ is

$$-C'(\lambda_1^{1/2}) - \frac{\partial v(p, \lambda_1^{1/2})}{\partial \lambda_1^{1/2}} [C(\lambda_{1/2}^m) + R]. \quad (22)$$

As above, (22) is negative at $\lambda_1^{1/2} = 0$ and is decreasing in $\lambda_1^{1/2}$. There is a unique optimal level of homogenization for the ruler in the range $[\lambda^*, 1]$ which is strictly negative. We denote it by $-\lambda^{oo}$. The ruler's utility from homogenization $-\lambda^{oo}$ is decreasing in p at rate

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

The ruler's expected utility from homogenizing below λ^* is

$$[g+y-k-C(\lambda_1^{1/2})+R]+v(p, \lambda_1^{1/2})[g-ga(1-\lambda_{1/4}^m)0.25+y-2k-C(\lambda_{1/4}^m)]+(1-v(p, \lambda_1^{1/2}))[g+y-k+R]$$

As above, there is also some unique optimal level of homogenization for the ruler in the range $[-1, \lambda^*)$, denote this by $-\lambda^o$. The ruler's utility from homogenization $-\lambda^o$ is decreasing in p at rate

$$-\frac{\partial v(p, -\lambda^o)}{\partial p} [ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m) + R]$$

We know $C(\lambda_{1/2}^m) < ga(1 - \lambda_{1/4}^m)0.25 + k + C(\lambda_{1/4}^m)$ and so the same argument as above applies.

Homogenization chosen by the ruler in this case is also decreasing in R . As above, the value of $-\lambda^o$ is decreasing in R , and so is the value of $-\lambda^{oo}$. The expected utility from homogenization by $-\lambda^o$ is increasing at rate $2 - v(p, -\lambda^o)$ and the expected utility from homogenization by $-\lambda^{oo}$ is increasing at a lower rate of $2 - v(p, -\lambda^{oo})$, and so the lowest value of p at which the ruler chooses $-\lambda^{oo}$ is increasing in R .

Proof of Proposition 5

The ruler is located at some $e \in [0, 1]$ and faces an exogenous probability of democracy. We make one simplifying technical assumption: when a democratic country chooses the location of the government, the democracy chooses by majority rule between the ruler's location e or the location of the median individual in the country.

In the following analysis we first suppose the ruler is located at $e < 1/2$, locates the government at $j = e$, and homogenizes by λ_1^e in period 1 (the case $e > 1/2$ is symmetric). We examine the subsequent choice of a period 2 democracy of whether to form a single country or split, where to locate the government, and how much homogenization to undertake. We then show that the ruler's optimal choice of homogenization is increasing in p .

Choices of a Democracy in Period 2

Working backwards, we first examine the choice of homogenization and location of the government in a period 2 democracy.

If a Democracy Forms a Single Country

If a democracy forms a single country in period 2 the government will be located at either $1/2$ or e . If the government is located at $1/2$, the median voter will choose homogenization $\lambda_{1/2}^m$. If the democratic government is located at $e \in [1/4, 1/2)$, the optimal level of homogenization for the median voter, denoted λ_e^m , satisfies $ga/4 = C'(\lambda_e^m)$, since the median voter is at distance $d_{ij} = 1/4$. For $e \in [0, 1/4)$ the optimal level of homogenization for the median voter satisfies $ga(1/2 - e) = C'(\lambda_e^m)$. A democracy increases homogenization to λ_e^m if $\lambda_1^e < \lambda_e^m$ otherwise the democracy does not homogenize further.

We next examine whether a democracy that forms a single country in period 2 would choose to locate the government at $1/2$ or e given homogenization by the ruler of λ_1^e . We show that there exists a level of homogenization λ_e^* such that, for all $\lambda_1^e \geq \lambda_e^*$ a democracy that forms a single country will choose to locate the government at e .

Case 1: $e \in [1/4, 1/2)$

If the ruler homogenizes by $\lambda_1^e < \lambda_{1/2}^m$ will a democracy locate the government at e or $1/2$? Consider $i \in [1/2, 1]$ at distance l_i from $1/2$ and distance $l_i + (1/2 - e)$ from e . Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_{1/2}^m)(l_i + (1/2 - e)) + y - k - C(\lambda_{1/2}^m) + C(\lambda_1^e)] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)].$$

This expression is positive for all i if

$$ga(1 - \lambda_{1/2}^m)(1/2 - e) < C(\lambda_1^e) \quad (23)$$

in which case all $i \in [1/2, 1]$ prefer a government at e . Otherwise it is negative for all i and a government at $1/2$ is chosen. For e approaching $1/2$ the left hand side of (23) is positive but approaching zero and so expression (23) will hold for some level of homogenization $0 < \lambda_1^e < \lambda_{1/2}^m$ and all λ_1^e higher. The minimum value of λ_1^e needed to satisfy (23) is increasing as e falls. For some e low enough it may not be satisfied for any $\lambda_1^e \leq \lambda_{1/2}^m$.

If the ruler homogenizes by $\lambda_1^e \geq \lambda_{1/2}^m$ will the democracy locate the government at e or $1/2$? Consider $i \in [1/2, 1]$ at distance l_i from $1/2$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

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

Rewritten

$$ga(\lambda_1^e - \lambda_{1/2}^m)l_i - ga(1 - \lambda_1^e)(1/2 - e) + C(\lambda_{1/2}^m) \quad (24)$$

This is increasing in l_i . Consider $i \in [e, 1/2)$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_1^e)((1/2 - e) - l_i) + y - k] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)].$$

Rewritten

$$ga(1 - \lambda_{1/2}^m)l_i + ga(1 - \lambda_1^e)l_i - ga(1 - \lambda_1^e)(1/2 - e) + C(\lambda_{1/2}^m) \quad (25)$$

This is increasing in l_i .

Consider $i \in [0, e)$. Since all $i \in [0, e]$ are closer to the government at e than a government at $1/2$ and since the cost of homogenization λ_1^e (which is at least as large as $\lambda_{1/2}^m$) has already been paid for by the ruler, a government at e is preferred to $1/2$ by all $i \in [0, e]$. We will therefore look for an individual i who is indifferent between e and $1/2$ in the range $i \in [e, 1/2)$ and $i \in [1/2, 1]$. We set (24) equal to zero and solve for l'_i to find the distance of the individual $i \in [1/2, 1]$ who is indifferent, where if $l'_i > 1/2$ all prefer $1/2$ and if $l'_i < 0$ all prefer e . Notice that l'_i is decreasing in λ_1^e and is negative for $\lambda_1^e = 1$. Set (25) equal to zero and solve for l''_i to find $i \in [e, 1/2]$ who is indifferent, where if $l''_i > 1/2 - e$ all prefer $1/2$ and if $l''_i < 0$ all prefer e . Notice that l''_i is decreasing in λ_1^e and is negative for $\lambda_1^e = 1$. Write

$$l'_i + l''_i = [ga(1 - \lambda_1^e)(1/2 - e) - C(\lambda_{1/2}^m)] \left[\frac{1}{ga(\lambda_1^e - \lambda_{1/2}^m)} + \frac{1}{ga(1 - \lambda_{1/2}^m) + ga(1 - \lambda_1^e)} \right].$$

For any $e \in [1/4, 1/2)$ there exists a value of $\lambda_1^e < 1$ such that $l_i + l'_i \leq 1/2$ (with relevant truncation) and $l'_i + l''_i$ is decreasing in λ_1^e . The lowest value of λ_1^e required to ensure the population choose a government at e is increasing as e falls.

In summary, there exists a value of $\lambda_1^e < 1$ at which a democracy will choose to locate the government at e and will do so for all homogenization higher. We denote this threshold level of homogenization by λ_e^* . The value of λ_e^* is decreasing in e .

Case 2: $e \in [0, 1/4]$

If the ruler homogenizes by $\lambda_1^e < \lambda_e^m$ will a democracy locate the government at e or $1/2$? Consider $i \in [1/2, 1]$ at distance l_i from $1/2$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_e^m)(l_i + (1/2 - e)) + y - k - C(\lambda_e^m) + C(\lambda_1^e)] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)]$$

Rewritten

$$ga(\lambda_e^m - \lambda_{1/2}^m)l_i - ga(1 - \lambda_e^m)(1/2 - e) - C(\lambda_e^m) + C(\lambda_1^e) + C(\lambda_{1/2}^m) \quad (26)$$

This is increasing in l_i . Consider $i \in [e, 1/2)$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_e^m)((1/2 - e) - l_i) + y - k - C(\lambda_e^m) + C(\lambda_1^e)] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)] \\ ga(1 - \lambda_e^m)l_i + ga(1 - \lambda_{1/2}^m)l_i - ga(1 - \lambda_e^m)(1/2 - e) - C(\lambda_e^m) + C(\lambda_1^e) + C(\lambda_{1/2}^m) \quad (27)$$

This is increasing in l_i . Consider $i \in [0, e)$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_e^m)(l_i - (1/2 - e)) + y - k - C(\lambda_e^m) + C(\lambda_1^e)] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)]$$

This is increasing in l_i since $\lambda_{1/2}^m < \lambda_e^m$.

If an individual in the range $i \in [0, e]$ is indifferent between a government at e and $1/2$ then the individual in the range $i \in [0, 1/2]$ at the same distance from the center strictly prefers a government at $1/2$ since their utility from a government at $1/2$ is the same but their utility from a government at e is lower. Then more than half the population prefer a government $1/2$ to e . We therefore look for an individual i who is indifferent between e and $1/2$ in the range $i \in (e, 1/2]$ and $i \in [1/2, 1]$. Set (26) equal to zero and solve for l'_i to find the distance of the individual $i \in [1/2, 1]$ who is indifferent between a government at $1/2$ and e

$$l'_i = \frac{ga(1 - \lambda_e^m)(1/2 - e) + C(\lambda_e^m) - C(\lambda_1^e) - C(\lambda_{1/2}^m)}{ga(\lambda_e^m - \lambda_{1/2}^m)},$$

where if $l'_i > 1/2$ all prefer $1/2$ and if $l'_i < 0$ all prefer e . Notice that l'_i is decreasing in λ_1^e . Set (27) equal to zero and solve for l''_i to find the distance of the individual $i \in (e, 1/2]$ who is indifferent between a government at $1/2$ and a government at e

$$l''_i = \frac{ga(1 - \lambda_e^m)(1/2 - e) + C(\lambda_e^m) - C(\lambda_1^e) - C(\lambda_{1/2}^m)}{ga(1 - \lambda_e^m) + ga(1 - \lambda_{1/2}^m)}$$

where if $l''_i > 1/2 - e$ all prefer $1/2$ and if $l''_i < 0$ all prefer e . Notice that l''_i is decreasing in λ_1^e . A government at e is preferred if $l'_i + l''_i \leq 1/2$ (with the relevant truncation). We know from above $l'_i + l''_i$ is decreasing in λ_1^e and so if there exists a value of $\lambda_1^e < \lambda_e^m$ at which $l'_i + l''_i = 1/2$ then a democracy will choose to locate the government at e for that level of homogenization by the ruler and all homogenization $\lambda_1^e < \lambda_e^m$ higher. As e decreases a higher value of λ_1^e is required to satisfy $l'_i + l''_i = 1/2$. If there does not exist a value of $\lambda_1^e < \lambda_e^m$ such that $l'_i + l''_i = 1/2$ then we consider $\lambda_1^e \geq \lambda_e^m$.

If the ruler homogenizes by $\lambda_1^e \geq \lambda_e^m$ will a democracy locate the government at e or $1/2$? Consider $i \in [1/2, 1]$ at distance l_i from $1/2$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_1^e)(l_i + (1/2 - e)) + y - k] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)].$$

Rewritten

$$ga(\lambda_1^e - \lambda_{1/2}^m)l_i - ga(1 - \lambda_1^e)(1/2 - e) + C(\lambda_{1/2}^m). \quad (28)$$

This is increasing in l_i . Consider $i \in [e, 1/2)$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

$$[g - ga(1 - \lambda_1^e)((1/2 - e) - l_i) + y - k] - [g - ga(1 - \lambda_{1/2}^m)l_i + y - k - C(\lambda_{1/2}^m)].$$

Rewritten

$$ga(1 - \lambda_1^e)l_i + ga(1 - \lambda_{1/2}^m)l_i - ga(1 - \lambda_1^e)(1/2 - e) + C(\lambda_{1/2}^m). \quad (29)$$

This is increasing in l_i . Consider $i \in [0, e)$. Individual i 's utility from a government at e minus his utility from a government at $1/2$ is

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

This is increasing in l_i .

By the same argument as above, if an individual $i \leq e$ is indifferent between a government at $1/2$ and e then a majority prefer a government at $1/2$. We will therefore look for an individual i who is indifferent between e and $1/2$ in the range $i \in (e, 1/2]$ and $i \in [1/2, 1]$. Set (28) equal to zero and solve for l'_i to find the distance of the individual $i \in [1/2, 1]$ who is indifferent

$$l'_i = \frac{ga(1 - \lambda_1^e)(1/2 - e) - C(\lambda_{1/2}^m)}{ga(\lambda_1^e - \lambda_{1/2}^m)}.$$

where if $l'_i > 1/2$ all prefer $1/2$ and if $l'_i < 0$ all prefer e . Notice that l'_i is decreasing in λ_1^e and is negative for $\lambda_1^e = 1$. Set (29) equal to zero and solve for l''_i to find the distance of the individual $i \in (e, 1/2]$ who is indifferent

$$l''_i = \frac{ga(1 - \lambda_1^e)(1/2 - e) - C(\lambda_{1/2}^m)}{ga(1 - \lambda_1^e) + ga(1 - \lambda_{1/2}^m)}$$

where if $l''_i > 1/2 - e$ all prefer $1/2$ and if $l''_i < 0$ all prefer e . Notice that l''_i is decreasing in λ_1^e and is negative for $\lambda_1^e = 1$. Then $l'_i + l''_i$ (with the relevant truncation) is decreasing in λ_1^e and there exists $\lambda_1^e < 1$ such that $l'_i + l''_i = 1/2$. Notice also that the value of λ_1^e required for $l'_i + l''_i \leq 1/2$ is increasing as e decreases.

In summary, there exists a value of $\lambda_1^e < 1$ at which a democracy will choose to locate the government at e and will do so for all homogenization higher. We denote this threshold level of homogenization by λ_e^* . The value of λ_e^* is decreasing in e .

If a Democracy Splits

If a democracy splits then the ruler is located at e in Country A. The population of Country A on the interval $[0, 1/2]$ chooses between a government at the ruler's location

e and a government at the median individual in Country A, $i = 1/4$. If the democracy locates the government at $1/4$ they will homogenize by $\lambda_{1/4}^m$. If the democracy locates the government at e , denote the optimal level of homogenization for the median voter by λ_e^{mA} . If the ruler homogenizes by $\lambda_1^e < \lambda_e^{mA}$ in period 1 and the democracy chose to locate the government at e then a democracy will increase homogenization to λ_e^{mA} , otherwise a democracy will not undertake additional homogenization. The analysis here is analogous to above where a population on the interval $[0, 1]$ is choosing between a government at $e \neq 1/2$ and a government at the median individual, $i = 1/2$. We denote by λ_{eA}^* the level of homogenization at and above which Country A would choose to locate the government at e . In Country B the government will be located at $i = 3/4$.

The Ruler's Optimal Homogenization is Increasing in p

We have shown that there exists a threshold level of homogenization, denoted λ_e^* , such that a single country democracy would locate the government at e , and a threshold level of homogenization, denoted λ_{eA}^* , such that a Country A democracy would locate the government at e . We next examine whether a democracy would form a single country or split given the choice of homogenization by the ruler, λ_1^e . Finally we show that this choice of homogenization by the ruler is increasing in p .

We work through three possibilities: $\lambda_e^* = \lambda_{eA}^*$, $\lambda_e^* < \lambda_{eA}^*$, and $\lambda_e^* > \lambda_{eA}^*$.

Suppose $\lambda_e^* = \lambda_{eA}^*$

If the ruler homogenizes to $\lambda_e^* = \lambda_{eA}^*$ or higher then both a democracy that forms a single country and Country A in a democracy that splits would locate the government at e . Everyone in Country A does strictly better by forming a single country and so a single country will be formed with the government at e . The ruler's expected utility from homogenization $\lambda_1^e \geq \lambda_e^*$ if $\lambda_1^e \geq \lambda_e^m$ is

$$[g + y - k - C(\lambda_1^e) + R] + p[g + y - k] + (1 - p)[g + y - k + R]$$

and if $\lambda_1^e < \lambda_e^m$ is

$$[g + y - k - C(\lambda_1^e) + R] + p[g + y - k - C(\lambda_e^m) + C(\lambda_1^e)] + (1 - p)[g + y - k + R].$$

Expected utility is decreasing in homogenization in this range and so for the ruler $\lambda_e^* = \lambda_{eA}^*$ strictly dominates all homogenization higher.

If the ruler homogenizes to $\lambda_1^e < \lambda_e^*$ then if $\lambda^* = 0$ a democracy forms a single country with the government located at $1/2$. The ruler's expected utility is

$$[g + y - k - C(\lambda_1^e) + R] + p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k + R].$$

If $\lambda^* > 0$ then a democracy splits and the government in Country A is located $1/4$. The ruler's expected utility is

$$[g + y - k - C(\lambda_1^e) + R] + p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k + R]$$

Zero homogenization strictly dominates all other homogenization in this range.

The ruler will choose either homogenization 0 or λ_e^* . We show that if the ruler chooses λ_e^* at some p he does so for all p higher.

Case 1: $\lambda^* = 0$

Compare the ruler's expected utility from homogenization $\lambda_e^* = \lambda_{eA}^*$ versus homogenization 0. If $\lambda_e^* \geq \lambda_e^m$ then expected utility from homogenization λ_e^* minus expected utility from zero homogenization is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + p[ga(1 - \lambda_{1/2}^m)(1/2 - e) + C(\lambda_{1/2}^m)] \end{aligned}$$

This is increasing in p . If $\lambda_e^* < \lambda_e^m$

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k - [C(\lambda_e^m) - C(\lambda_e^*)]] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) - p[C(\lambda_e^m) - C(\lambda_e^*)] + p[ga(1 - \lambda_{1/2}^m)(1/2 - e) + C(\lambda_{1/2}^m)] \end{aligned}$$

This is either increasing in p or if it is decreasing in p the whole expression is negative and so the ruler never chooses λ_e^* for any p .

Case 2: $\lambda^* > 0$

Compare the ruler's expected utility from homogenization $\lambda_e^* = \lambda_{eA}^*$ versus homogenization 0. If $\lambda_e^* \geq \lambda_e^m$ then expected utility from λ_e^* minus zero homogenization is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + p[ga(1 - \lambda_{1/4}^m)|1/4 - e| + k + C(\lambda_{1/4}^m)] \end{aligned}$$

This is increasing in p . If $\lambda_e^* < \lambda_e^m$ then

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k - [C(\lambda_e^m) - C(\lambda_e^*)]] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + p[g + y - k - [C(\lambda_e^m) - C(\lambda_e^*)]] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] \end{aligned}$$

This is either increasing in p or if it is decreasing in p the whole expression is negative and so the ruler never chooses λ_e^* for any p .

Suppose $\lambda_e^* < \lambda_{eA}^*$

If the ruler homogenizes by less than λ_e^* then if $\lambda^* = 0$ a democracy will form a single country with the government located at $1/2$ and if $\lambda^* > 0$ then a democracy will split with the governments located at $1/4$ and $1/2$. As above, homogenization zero strictly dominates all other levels of homogenization in this range. If the ruler homogenizes to λ_{eA}^* or higher then, as above, a democracy would form a single country with the government located at e and homogenization λ_{eA}^* dominates all homogenization higher. If the ruler homogenizes in the range $[\lambda_e^*, \lambda_{eA}^*)$ then if a democracy forms a single country it will locate the govt at e , and if a democracy splits, the government in Country A will be located at $1/4$. If the ruler homogenizes in the range $\lambda_1^e \in [\lambda_e^*, \lambda_{eA}^*)$ then utility for an individual $i \in [0, 1/2]$ in a democracy at distance l_i from the center if homogenization $\lambda_1^e \geq \lambda_e^m$ is undertaken and a single country is formed is

$$g - ga(1 - \lambda_1^e)|(1/2 - e) - l_i| + y - k$$

and if the population splits

$$g - ga(1 - \lambda_{1/4}^m)|1/4 - l_i| + y - k - C(\lambda_{1/4}^m).$$

For all $i \in [0, 1/2]$ utility from a single country is increasing in homogenization by the ruler while utility from a split does not change. Similarly if the ruler homogenizes in the range $\lambda_1^e \in [\lambda_e^*, \lambda_{eA}^*)$ with $\lambda_1^e < \lambda_e^m$ and $i \in [1/2, 1]$. Thus if a democracy would choose to form a single country for some $\lambda_1^e \in [\lambda_e^*, \lambda_{eA}^*)$ then a democracy would form a single country for all homogenization higher. Let λ_e^{1*} denote the lowest value of homogenization such that a democracy would choose to form a single country, where we know $\lambda_e^{1*} \in [\lambda_e^*, \lambda_{eA}^*]$. If $\lambda_e^{1*} \neq \lambda_e^*$ then for $\lambda_1^e \in [\lambda_e^*, \lambda_e^{1*})$ a democracy would split and locate the governments at $1/4$ and $3/4$, and homogenization λ_e^* strictly dominates all higher homogenization in this range.

Therefore the ruler may choose one of three levels of homogenization: zero homogenization, λ_e^* if $\lambda_e^* \neq \lambda_e^{1*}$ (where a democracy would choose to split with the governments at $1/4$ and $3/4$), or λ_e^{1*} (where a democracy would choose to form a single country with the government at e).

Case 1: $\lambda^* = 0$

The ruler might choose one of three levels of homogenization: zero homogenization (where a democracy would choose to form a single country with the government at $1/2$), λ_e^* if $\lambda_e^* \neq \lambda_e^{1*}$ (where a democracy would choose to split with the governments at $1/4$ and $3/4$), and λ_e^{1*} .

From above, we know that if λ_e^{1*} gives the ruler higher expected utility than zero for

some p then it does so for all p higher. Next we want to show that if λ_e^* gives the ruler higher expected utility than zero for some p then it does so for all p higher. The ruler's expected utility from homogenization λ_e^* minus homogenization zero is

$$\begin{aligned} & [g+y-k-C(\lambda_e^*)+R]+p[g-ga(1-\lambda_{1/4}^m)|1/4-e|+y-2k-C(\lambda_{1/4}^m)]+(1-p)[g+y-k+R] \\ & -[g+y-k+R]-p[g-ga(1-\lambda_{1/2}^m)(1/2-e)+y-k-C(\lambda_{1/2}^m)]-(1-p)[g+y-k+R] \\ & = -C(\lambda_e^*)-p[ga(1-\lambda_{1/4}^m)|1/4-e|+k+C(\lambda_{1/4}^m)]+p[ga(1-\lambda_{1/2}^m)(1/2-e)+C(\lambda_{1/2}^m)] \end{aligned}$$

This is either increasing in p or if not then it is negative for all p . Finally we want to show that if λ_e^{1*} gives the ruler higher expected utility than λ_e^* for some p then it does so for all p higher. The ruler's expected utility from homogenization λ_e^{1*} minus homogenization λ_e^* is

$$\begin{aligned} & [g+y-k-C(\lambda_e^{1*})+R]+p[g+y-k]+(1-p)[g+y-k+R] \\ & -[g+y-k-C(\lambda_e^*)+R]-p[g-ga(1-\lambda_{1/4}^m)|1/4-e|+y-2k-C(\lambda_{1/4}^m)]-(1-p)[g+y-k+R] \\ & = -C(\lambda_e^{1*})+C(\lambda_e^*)+p[ga(1-\lambda_{1/4}^m)|1/4-e|+k+C(\lambda_{1/4}^m)] \end{aligned}$$

This is increasing in p .

Case 2: $\lambda^* > 0$

The ruler might choose one of two levels of homogenization: zero homogenization (where a democracy would split with the governments at 1/4 and 3/4), or λ_e^{1*} . As above, if λ_e^{1*} gives the ruler higher expected utility than zero for some p then it does so for all p higher.

Suppose $\lambda_e^* > \lambda_{eA}^*$.

As above, if the ruler homogenizes to λ_e^* or higher then a democracy will form a single country with the government at e , and λ_e^* strictly dominates all higher levels of homogenization. If the ruler homogenizes less than λ_{eA}^* , then if $\lambda^* = 0$ a democracy forms a single country with the government at 1/2 an if $\lambda^* > 0$ then a democracy splits with the governments at 1/4 and 3/4. Homogenization zero strictly dominates all homogenization in the range $(0, \lambda_{eA}^*)$.

For homogenization in the range $[\lambda_{eA}^*, \lambda_e^*)$, if a democracy forms a single country the government will be located at 1/2 and if a democracy splits it will be located at e in Country A and 3/4 in Country B. Utility from a single country is not changing in homogenization by the ruler. Utility from a split for an individual in Country A, when the ruler homogenizes by $\lambda_1^e \in [\lambda_{eA}^*, \lambda_e^*)$, where $\lambda_1^e \geq \lambda_e^{mA}$, is

$$g-ga(1-\lambda_1^e)|(1/2-e)-l_i|+y-k$$

This is increasing in homogenization by the ruler and similarly if $\lambda_1^e < \lambda_e^{mA}$. Utility from a split for an individual in Country B is not changing in homogenization by the ruler. Therefore if a democracy would choose to split for some level of homogenization in the range

$[\lambda_{eA}^*, \lambda_e^*)$ they do so for all homogenization higher in this range. Let $\lambda_e^{2*} \in [\lambda_{eA}^*, \lambda_e^*)$ denote the lowest value of homogenization in this range for which a democracy would choose to split, if it exists. This dominates all homogenization higher in the range $(\lambda_e^{2*}, \lambda_e^*)$. Then for the range $[\lambda_{eA}^*, \lambda_e^{2*})$ a democracy will locate the government at $1/2$. If $\lambda^* = 0$ then homogenization in the range $[\lambda_{eA}^*, \lambda_e^{2*})$ is strictly dominated by zero homogenization. If $\lambda^* > 0$ then we can only say that homogenization λ_{eA}^* strictly dominates any other homogenization in the range $[\lambda_{eA}^*, \lambda_e^{2*})$. Notice that if λ_e^{2*} exists then a democracy chooses to locate the government at e for this homogenization and all homogenization higher (although a democracy only forms a single country if also homogenization is above λ_e^*). Otherwise λ_e^* and all homogenization higher results in a democracy choosing to form a single country and locate the government at e .

Therefore the ruler may choose one of four homogenization levels: 0 , λ_{eA}^* , λ_e^{2*} , or λ_e^* . We show that, for each pair of these four homogenization levels, if the higher level of homogenization gives the ruler higher expected utility than the lower level of homogenization for some p then it does so for all p higher, and so homogenization chosen by the ruler in period 1 is increasing in p .

Case 1: $\lambda^* = 0$

The ruler chooses between: homogenization zero (a democracy would form a single country and locates the government at $1/2$), homogenization $\lambda_e^{2*} \in [\lambda_{eA}^*, \lambda_e^*)$ (a democracy would split and Country A locate the government at e) if such a value of homogenization exists, and λ_e^* .

From above we know that if the ruler has higher expected utility from λ_e^* than 0 at any p , then this is true for all p higher. We next show that if λ_e^{2*} gives the ruler higher expected utility than zero for any p it gives higher expected utility for all p higher. Compare the ruler's expected utility from homogenization $\lambda_e^{2*} \geq \lambda_e^{mA}$, supposing it exists, with his expected utility from homogenization zero:

$$\begin{aligned} & [g + y - k - C(\lambda_e^{2*}) + R] + p[g + y - 2k] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^{2*}) - pk + p[ga(1 - \lambda_{1/2}^m)(1/2 - e) + C(\lambda_{1/2}^m)] \end{aligned}$$

Either this is increasing in p , or, if it is decreasing in p then the ruler always has higher expected utility from 0 than λ_e^{2*} . If instead $\lambda_e^{2*} < \lambda_e^{mA}$ then the ruler's expected utility

from homogenization λ_e^{2*} minus homogenization zero is

$$\begin{aligned} & [g + y - k - C(\lambda_e^{2*}) + R] + p[g + y - 2k - [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^{2*}) - pk - p[C(\lambda_e^{mA}) - C(\lambda_e^{2*})] + pga(1 - \lambda_{1/2}^m)(1/2 - e) + pC(\lambda_{1/2}^m) \end{aligned}$$

Either this is increasing in p , or, if it is decreasing in p then the ruler always has higher expected utility from 0 than λ_e^{2*} .

Finally we show that if λ_e^* gives higher expected utility than λ_e^{2*} for any p it gives higher expected utility for all p higher. The ruler's expected utility from homogenization $\lambda_e^* \geq \lambda_e^m$ versus $\lambda_e^{2*} \geq \lambda_e^{mA}$ is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \\ & - [g + y - k - C(\lambda_e^{2*}) + R] - p[g + y - 2k] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + C(\lambda_e^{2*}) + pk \end{aligned}$$

This is increasing in p . The ruler's expected utility from homogenization $\lambda_e^* \geq \lambda_e^m$ versus $\lambda_e^{2*} < \lambda_e^{mA}$ is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \\ & - [g + y - k - C(\lambda_e^{2*}) + R] - p[g + y - 2k - [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + C(\lambda_e^{2*}) + p[k + [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] \end{aligned}$$

This is increasing in p . The ruler's expected utility from homogenization $\lambda_e^* < \lambda_e^m$ versus $\lambda_e^{2*} \geq \lambda_e^{mA}$ is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k - [C(\lambda_e^m) - C(\lambda_e^*)]] + (1 - p)[g + y - k + R] \\ & - [g + y - k - C(\lambda_e^{2*}) + R] - p[g + y - 2k] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) - p[C(\lambda_e^m) - C(\lambda_e^*)] + C(\lambda_e^{2*}) + pk \end{aligned}$$

Either this is increasing in p , or, if it is decreasing in p then the ruler has higher expected utility from λ_e^{2*} than λ_e^* for all p . The ruler's expected utility from homogenization $\lambda_e^{2*} < \lambda_e^{mA}$ versus $\lambda_e^* < \lambda_e^m$ is

$$\begin{aligned} & [g + y - k - C(\lambda_e^*) + R] + p[g + y - k - [C(\lambda_e^m) - C(\lambda_e^*)]] + (1 - p)[g + y - k + R] \\ & - [g + y - k - C(\lambda_e^{2*}) + R] - p[g + y - 2k - [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] - (1 - p)[g + y - k + R] \\ = & -C(\lambda_e^*) + C(\lambda_e^{2*}) - p[C(\lambda_e^m) - C(\lambda_e^*)] + p[k + [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] \end{aligned}$$

Either this is increasing in p , or, if it is decreasing in p then the ruler has higher expected utility from λ_e^{2*} than λ_e^* for all p .

Case 2: $\lambda^* > 0$

The ruler may choose one of four homogenization levels: 0 (a democracy would split and locate the governments at 1/4 and 3/4), λ_{eA}^* (a democracy would form a single country and locate the government at 1/2), λ_e^{2*} (a democracy would split and Country A would locate the government at e), or λ_e^* .

From above we know that if the ruler has higher expected utility from λ_e^* than 0 at any p . We next show that if λ_e^{2*} gives the ruler high expected utility than zero for any p it gives higher expected utility for all p higher. The ruler's expected utility from homogenization $\lambda_e^{2*} \geq \lambda_e^{mA}$ and homogenization zero is

$$\begin{aligned} & [g + y - k - C(\lambda_e^{2*}) + R] + p[g + y - 2k] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] - (1 - p)[g + y - k + R] \\ & = -C(\lambda_e^{2*}) + p[ga(1 - \lambda_{1/4}^m)|1/4 - e| + C(\lambda_{1/4}^m)] \end{aligned}$$

This is increasing in p . The ruler's expected utility from homogenization $\lambda_e^{2*} < \lambda_e^{mA}$ and homogenization zero is

$$\begin{aligned} & [g + y - k - C(\lambda_e^{2*}) + R] + p[g + y - 2k - [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] + (1 - p)[g + y - k + R] \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] - (1 - p)[g + y - k + R] \\ & = -C(\lambda_e^{2*}) - p[C(\lambda_e^{mA}) - C(\lambda_e^{2*})] + pga(1 - \lambda_{1/4}^m)|1/4 - e| + pC(\lambda_{1/4}^m) \end{aligned}$$

Either this is increasing in p , or, if it is decreasing in p then the ruler has higher expected utility from 0 than λ_e^{2*} for all p . Thus if the ruler has higher expected utility from λ_e^{2*} than 0 at any p this is true for all p higher.

We next show that if λ_{eA}^* gives the ruler higher expected utility than zero for any p it gives higher expected utility for all p higher. The ruler's expected utility from homogenization λ_{eA}^* versus homogenization zero is

$$\begin{aligned} & [g + y - k - C(\lambda_{eA}^*) + R] + p[g - ga(1 - \lambda_{1/2}^m)(1/2 - e) + y - k - C(\lambda_{1/2}^m)] + (1 - p)[g + y - k + R]. \\ & - [g + y - k + R] - p[g - ga(1 - \lambda_{1/4}^m)|1/4 - e| + y - 2k - C(\lambda_{1/4}^m)] - (1 - p)[g + y - k + R]. \\ & = -C(\lambda_{eA}^*) - p[ga(1 - \lambda_{1/2}^m)(1/2 - e) + C(\lambda_{1/2}^m)] + p[ga(1 - \lambda_{1/4}^m)|1/4 - e| + k + C(\lambda_{1/4}^m)] \end{aligned}$$

This is either increasing in p or, if it is decreasing in p then the ruler has higher expected utility from 0 than λ_{eA}^* for all p .

We next show that if λ_e^{2*} gives the ruler higher expected utility than λ_{eA}^* for any p it gives higher expected utility for all p higher. Compare the ruler's expected utility from homogenization $\lambda_e^{2*} \geq \lambda_e^{mA}$ with homogenization λ_{eA}^*

$$[g + y - k - C(\lambda_e^{2*}) + R] + p[g + y - 2k] + (1 - p)[g + y - k + R].$$

$$\begin{aligned}
& -[g+y-k-C(\lambda_{eA}^*)+R]-p[g-ga(1-\lambda_{1/2}^m)(1/2-e)+y-k-C(\lambda_{1/2}^m)]-(1-p)[g+y-k+R]. \\
& = -C(\lambda_e^{2*}) - pk + C(\lambda_{eA}^*) + p[ga(1-\lambda_{1/2}^m)(1/2-e) + C(\lambda_{1/2}^m)]
\end{aligned}$$

This is either increasing in p or, if not, the ruler has higher expected utility from λ_{eA}^* than λ_e^{2*} for all p . Compare the ruler's expected utility from homogenization $\lambda_e^{2*} < \lambda_{eA}^*$ versus λ_{eA}^*

$$\begin{aligned}
& [g+y-k-C(\lambda_e^{2*})+R] + p[g+y-2k - [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] + (1-p)[g+y-k+R]. \\
& -[g+y-k-C(\lambda_{eA}^*)+R]-p[g-ga(1-\lambda_{1/2}^m)(1/2-e)+y-k-C(\lambda_{1/2}^m)]-(1-p)[g+y-k+R]. \\
& = -C(\lambda_e^{2*}) - p[k + [C(\lambda_e^{mA}) - C(\lambda_e^{2*})]] + C(\lambda_{eA}^*) + p[ga(1-\lambda_{1/2}^m)(1/2-e) + C(\lambda_{1/2}^m)]
\end{aligned}$$

This is either increasing in p or, if not, the ruler has higher expected utility from λ_{eA}^* than λ_e^{2*} for all p . Then if λ_e^{2*} gives the ruler higher expected utility than λ_{eA}^* for any p , this is true for all p higher.

We showed above that if λ_e^* gives high expected utility than λ_e^{2*} for any p it gives higher expected utility for all p higher. Finally we show that if λ_e^* gives high expected utility than λ_{eA}^* for any p it gives higher expected utility for all p higher. Compare the ruler's expected utility from homogenization $\lambda_e^* \geq \lambda_e^m$ with homogenization λ_{eA}^*

$$\begin{aligned}
& [g+y-k-C(\lambda_e^*)+R] + p[g+y-k] + (1-p)[g+y-k+R] \\
& -[g+y-k-C(\lambda_{eA}^*)+R]-p[g-ga(1-\lambda_{1/2}^m)(1/2-e)+y-k-C(\lambda_{1/2}^m)]-(1-p)[g+y-k+R]. \\
& = -C(\lambda_e^1) + C(\lambda_{eA}^*) + p[ga(1-\lambda_{1/2}^m)(1/2-e) + C(\lambda_{1/2}^m)]
\end{aligned}$$

This is increasing in p . Compare the ruler's expected utility from homogenization $\lambda_e^* < \lambda_{eA}^*$ with homogenization λ_{eA}^*

$$\begin{aligned}
& [g+y-k-C(\lambda_e^*)+R] + p[g+y-k - [C(\lambda_e^m) - C(\lambda_e^*)]] + (1-p)[g+y-k+R]. \\
& -[g+y-k-C(\lambda_{eA}^*)+R]-p[g-ga(1-\lambda_{1/2}^m)(1/2-e)+y-k-C(\lambda_{1/2}^m)]-(1-p)[g+y-k+R]. \\
& = -C(\lambda_e^1) - p[C(\lambda_e) - C(\lambda_e^1)] + C(\lambda_{eA}^*) + p[ga(1-\lambda_{1/2}^m)(1/2-e) + C(\lambda_{1/2}^m)]
\end{aligned}$$

This is either increasing in p or, if not, the ruler has higher expected utility from λ_{eA}^* than λ_e^* for all p .

Proof of Proposition 6

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

$$gad_{ij} = \beta'(\lambda_t^j) + \alpha'(\lambda_t^j)d_{ij}. \quad (30)$$

By the above, since $\beta'(\lambda_t^j) > 0$ then for λ_t^j that satisfies (30) it must be that $\alpha'(\lambda_t^j) < ga$ and so the value of λ_t^j that satisfies (30) is increasing in d_{ij} . Preferences are single peaked over λ_t^j .

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

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

$$M(\lambda_t^j, 1/4) = \beta(\lambda_t^j) + \alpha(\lambda_t^j)/4 = C(\lambda_t^j), \quad \beta'(\lambda_t^j) + \alpha'(\lambda_t^j)/4 = C'(\lambda_t^j), \quad \forall \lambda_t^j \in [0, 1]. \quad (31)$$

From (31), for $d_{ij} = 1/4$ the levels of homogenization that satisfy (30) and $gad_{ij} = C'(\lambda_t^j)$ are equal; for $d_{ij} < 1/4$ the level of homogenization that satisfies (30) is strictly higher; and for $d_{ij} > 1/4$ it is strictly lower than the level of homogenization that satisfies $gad_{ij} = C'(\lambda_t^j)$. A ruler can always homogenize more cheaply using unequal-cost homogenization technologies and so he will always choose the unequal-cost technology.

Suppose a democracy forms a single country with the government at some $j \in [0, 1]$ and the ruler has undertaken no homogenization in period 1, $\lambda_1^{1/2} = 0$. For $j \in [1/4, 3/4]$, the median voter for homogenization is $d_{ij} = 1/4$, thus the levels of homogenization chosen by majority rule under unequal and equal-cost technologies are the same. For $j \in [0, 1/4)$, the median voter for homogenization is $i = 1/2$ with $d_{ij} = 1/2 - j > 1/4$, thus unequal homogenization chosen by majority rule will be lower than equal-cost homogenization. Each individual evaluates the difference between their utility in the case of equal-cost homogenization and their utility in the case of unequal homogenization:

$$[g - (1 - \lambda_{eq})gad_{ij} + y - k - C(\lambda_{eq})] - [g - (1 - \lambda_{uneq})gad_{ij} + y - k - M(\lambda_{uneq}, d_{ij})], \quad (32)$$

where λ_{uneq} (respectively λ_{eq}) is the level of unequal (respectively equal-cost) homogenization chosen by majority rule. For $j \in [1/4, 3/4]$, then $\lambda_{uneq} = \lambda_{eq}$, and so expression (32) is increasing in d_{ij} . The median voter when deciding between unequal and equal-cost homogenization is at $d_{ij} = 1/4$, and he is indifferent between the two technologies. For $j \in [0, 1/4)$, $\lambda_{uneq} < \lambda_{eq}$, expression (32) is increasing in d_{ij} and the median valuation of (32) is at $d_{ij} = (1/2 - j)$. The median voter for the choice of homogenization technology is $d_{ij} = (1/2 - j)$ and he is also the median voter for the level of homogenization. He must prefer equal-cost homogenization since for any level of unequal homogenization he can homogenize to the same level instead using equal-cost technologies and do strictly better.

The same argument applies for any positive level of homogenization by the ruler in period 1, $\lambda_1^{1/2} \in (0, 1]$. Notice that only the level of homogenization in period 1, $\lambda_1^{1/2}$, matters for the choice of a democracy in period 2, and not what technology was used. The total cost of homogenization λ_t^j in period t following homogenization λ_{t-1}^j is

$$\int_{i \in \text{country}} M(\lambda_t^j, d_{ij}) di - \int_{i \in \text{country}} M(\lambda_{t-1}^j, d_{ij}) di,$$

where the cost to an individual at distance d_{ij} is

$$M(\lambda_t^j, d_{ij}) - M(\lambda_{t-1}^j, d_{ij}).$$

Period 2 utility for individual i in a single country with homogenization λ_t^j is

$$g - ga(1 - \lambda_t^j)d_{ij} + y - k - [M(\lambda_t^j, d_{ij}) - M(\lambda_1^{1/2}, d_{ij})].$$

The optimal level of period 2 homogenization for this individual satisfies

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

The median individual for homogenization when the government is located at $j \in [1/4, 3/4]$ is $d_{ij} = 1/4$. If $\lambda_1^{1/2} \geq \lambda_{1/2}^m$ then both for equal and unequal-cost technologies zero homogenization is chosen and so the median voter is indifferent between either technology. If $\lambda_1^{1/2} < \lambda_{1/2}^m$, the median voter will choose homogenization $\lambda_{eq} = \lambda_{uneq} = \lambda_{1/2}^m$. Each individual evaluates the difference between their utility in the case of equal-cost homogenization and their utility in the case of unequal homogenization:

$$[g - (1 - \lambda_{eq})gad_{ij} + y - k - [C(\lambda_{eq}) - C(\lambda_1^{1/2})]] - [g - (1 - \lambda_{uneq})gad_{ij} + y - k - [M(\lambda_{uneq}, d_{ij}) - M(\lambda_1^{1/2}, d_{ij})]],$$

which is increasing in d_{ij} . Thus the median voter over choice of technology is $d_{ij} = 1/4$ and, since $\lambda_{eq} = \lambda_{uneq}$ and he faces the same costs, he is indifferent between the two. The median individual for homogenization when the government is located at $j \in [0, 1/4]$ is $d_{ij} = (1/2 - j)$. We are in the same situation as above because homogenization by the ruler occurs at location $1/2$.

An analogous argument applies for Country A and B. \square

Proof of Proposition 7

The analysis of what a democracy will do in period 2 given homogenization by the ruler in period 1 to $\lambda_1^{1/2}$ is the exactly the same as above, with the value of λ^* unchanged. The ruler however will now use unequal-cost homogenization technologies in period 1. The ruler's expected utility if he homogenizes to $\lambda_1^{1/2} < \lambda^*$ is

$$[g + y - k - M(\lambda_1^{1/2}, 0) + R] + p[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k + R].$$

Notice that zero homogenization dominates all other homogenization levels $\lambda_1^{1/2} < \lambda^*$. The ruler's expected utility if he homogenizes to $\lambda_1^{1/2} \geq \lambda^*$ is

$$[g + y - k - M(\lambda_1^{1/2}, 0) + R] + p[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})]] + (1 - p)[g + y - k + R].$$

if $\lambda_1^{1/2} < \lambda_{1/2}^m$, and is

$$[g + y - k - M(\lambda_1^{1/2}, 0) + R] + p[g + y - k] + (1 - p)[g + y - k + R].$$

if $\lambda_1^{1/2} \geq \lambda_{1/2}^m$.

Suppose $\lambda^* \geq \lambda_{1/2}^m$, then the ruler chooses λ^* if

$$\begin{aligned} & [g + y - k - M(\lambda^*, 0) + R] + p[g + y - k] + (1 - p)[g + y - k + R] \\ & > [g + y - k + R] + p[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k + R] \end{aligned}$$

and zero homogenization otherwise. For a given value of λ^* , the left hand side is strictly higher when the ruler uses unequal-cost compared to equal cost technologies, and so the ruler will choose λ^* at a lower value of p than for equal-cost technologies.

Suppose $\lambda^* < \lambda_{1/2}^m$, then the optimal value of $\lambda_1^{1/2} \geq \lambda^*$ satisfies

$$M_\lambda(\lambda_1^{1/2}, 0) = pC_\lambda(\lambda_1^{1/2})$$

which is increasing in p and if $p = 1$ the optimal value is $\lambda_1^{1/2} = \lambda_{1/2}^m$. The ruler will choose to homogenize to this optimal level of $\lambda_1^{1/2} \geq \lambda^*$ if at this optimal level of $\lambda_1^{1/2}$ the following holds

$$\begin{aligned} & [g + y - k - M(\lambda_1^{1/2}, 0) + R] + p[g + y - k - [C(\lambda_{1/2}^m) - C(\lambda_1^{1/2})]] + (1 - p)[g + y - k + R] \\ & > [g + y - k + R] + p[g - (1 - \lambda_{1/4}^m)ga/4 + y - 2k - C(\lambda_{1/4}^m)] + (1 - p)[g + y - k + R]. \end{aligned}$$

Otherwise the ruler will choose zero homogenization. At $\lambda_1^{1/2} = \lambda^*$ the value of the left hand side of this expression is strictly higher for unequal-cost homogenization than equal cost homogenization. Additionally, the value of left hand side is weakly higher still for the optimal value of $\lambda_1^{1/2} \geq \lambda^*$. Thus the ruler will choose to homogenize to some $\lambda_1^{1/2} \geq \lambda^*$ at a lower level of p than for equal-cost technologies. \square

Table A1
Educational reforms and political regimes

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|------------------------|-----------------------------------|----------------|----------------|-----------------------------------|----------------|----------------|-------------------------------------|----------------|----------------|
| | Educational reform: 10% threshold | | | Educational reform: 20% threshold | | | Legal reform, based on Flora (1983) | | |
| Autocracy | -0.048 | -0.034 | -0.037 | 0.027 | 0.040 | 0.037 | -0.115 | -0.089 | -0.086 |
| | (0.055) | (0.054) | (0.055) | (0.061) | (0.059) | (0.061) | (0.303) | (0.274) | (0.305) |
| Gov. crises | -0.092* | | | -0.059 | | | 0.055 | | |
| | (0.052) | | | (0.047) | | | (0.082) | | |
| Revolutions | | 0.079 | | | 0.120** | | | 0.296 | |
| | | (0.052) | | | (0.053) | | | (0.185) | |
| All internal conflicts | | | 0.002 | | | 0.012 | | | 0.041 |
| | | | (0.016) | | | (0.015) | | | (0.038) |
| Pop. growth | 0.317 | 0.303 | 0.322 | 0.029 | 0.004 | 0.031 | 0.432 | 0.527 | 0.430 |
| | (0.210) | (0.211) | (0.209) | (0.165) | (0.163) | (0.163) | (0.965) | (0.931) | (0.908) |
| Number of countries | 172 | 172 | 172 | 172 | 172 | 172 | 14 | 14 | 14 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 977 | 977 | 977 | 977 | 977 | 977 | 110 | 110 | 110 |
| R-squared | 0.450 | 0.450 | 0.448 | 0.404 | 0.408 | 0.404 | 0.626 | 0.634 | 0.632 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 10% (20%) from the previous 10 years (columns 1-3, and columns 4-6). Legal reform is a dummy if the country experienced at least one legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All explanatory variables are lagged. Standard errors are clustered at the country level

Table A2
The extension of the franchise and threat to democracy

| | (1) | (2) | (3) |
|---|---------------------|----------------|----------------|
| | Franchise extension | | |
| Autocracy*gov. crises | 0.003 | | |
| | (0.004) | | |
| Autocracy*revolutions | | -0.001 | |
| | | (0.005) | |
| Autocracy*all internal conflicts | | | 0.000 |
| | | | (0.001) |
| Autocracy | -0.005 | -0.004 | -0.004 |
| | (0.005) | (0.004) | (0.004) |
| Pop. growth | -0.003 | -0.000 | -0.002 |
| | (0.005) | (0.003) | (0.004) |
| Gov. crises | -0.001 | | |
| | (0.001) | | |
| Revolutions | | -0.008 | |
| | | (0.007) | |
| All internal conflicts | | | -0.001 |
| | | | (0.001) |
| Number of countries | 21 | 21 | 21 |
| Country fixed effects | yes | yes | yes |
| Year fixed effects | yes | yes | yes |
| Observations | 175 | 175 | 175 |
| R-squared | 0.168 | 0.202 | 0.204 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Data on the extension of the franchise comes from Aidt and Jensen (2014). All explanatory variables are lagged. Standard errors are clustered at the country level.

Table A3
Educational reforms and threat to democracy, restricting the samples to countries that never split

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|-----------------------------------|----------------|----------------|-----------------------------------|-----------------|----------------|-------------------------------------|----------------|----------------|
| | Educational reform: 10% threshold | | | Educational reform: 20% threshold | | | Legal reform, based on Flora (1983) | | |
| Autocracy*gov. crises | 0.138 | | | 0.115 | | | 0.592** | | |
| | (0.089) | | | (0.102) | | | (0.228) | | |
| Autocracy*revolutions | | 0.144* | | | 0.208*** | | | 2.057* | |
| | | (0.078) | | | (0.076) | | | (0.956) | |
| Autocracy*all internal conflicts | | | 0.048** | | | 0.044* | | | 0.202* |
| | | | (0.022) | | | (0.024) | | | (0.099) |
| Autocracy | -0.061 | -0.050 | -0.080 | 0.018 | 0.007 | -0.003 | -0.659 | -0.379 | -0.614 |
| | (0.059) | (0.057) | (0.061) | (0.068) | (0.059) | (0.067) | (0.393) | (0.374) | (0.479) |
| Pop. growth | 0.319 | 0.303 | 0.310 | 0.037 | 0.005 | 0.026 | -0.210 | 0.316 | -0.075 |
| | (0.216) | (0.220) | (0.217) | (0.166) | (0.165) | (0.164) | (1.083) | (0.937) | (0.962) |
| Gov. crises | -0.135** | | | -0.084 | | | -0.043 | | |
| | (0.062) | | | (0.062) | | | (0.080) | | |
| Revolutions | | 0.004 | | | 0.007 | | | 0.410 | |
| | | (0.066) | | | (0.062) | | | (0.290) | |
| All internal conflicts | | | -0.022 | | | -0.009 | | | 0.013 |
| | | | (0.018) | | | (0.018) | | | (0.054) |
| Number of countries | 154 | 154 | 154 | 154 | 154 | 154 | 11 | 11 | 11 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 896 | 896 | 896 | 896 | 896 | 896 | 96 | 96 | 96 |
| R-squared | 0.448 | 0.449 | 0.448 | 0.404 | 0.413 | 0.407 | 0.608 | 0.605 | 0.609 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 10% (20%) from the previous 10 years (columns 1-3, and columns 4-6). Legal reform is a dummy if the country experienced at least one legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All explanatory variables are lagged. Standard errors are clustered at the country level.

Table A4
Educational reforms and threat to democracy, controlling for continent specific linear trends

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
|----------------------------------|-----------------------------------|----------------------------|---------------------------|-------------------------|---------------------------|--------------------|-------------------------|--------------------------|-------------------|-------------------------|---------------------------|----------------------|-------------------------|---------------------------|-------------------|
| | Educational reform: 20% threshold | | | | | | | | | | | | | | |
| Autocracy*gov. crises | 0.163* (0.098) | | | 0.113 (0.105) | | | 0.105 (0.102) | | | 0.137 (0.101) | | | 0.147 (0.099) | | |
| Autocracy*revolutions | | 0.215*** (0.076) | | | 0.184** (0.093) | | | 0.129* (0.076) | | | 0.203** (0.094) | | | 0.168** (0.075) | |
| Autocracy*all internal conflicts | | | 0.052** (0.022) | | 0.043* (0.022) | | | 0.037* (0.021) | | | 0.042* (0.022) | | | 0.039* (0.022) | |
| Autocracy | 0.009 (0.064) | 0.008 (0.056) | -0.008 (0.062) | 0.027 (0.070) | 0.019 (0.061) | 0.009 (0.068) | -0.015 (0.060) | -0.005 (0.050) | -0.024 (0.058) | 0.002 (0.064) | | -0.011 (0.062) | -0.006 (0.063) | | -0.013 (0.062) |
| Pop. growth | 0.076 (0.171) | 0.048 (0.174) | 0.069 (0.172) | -0.075 (0.165) | -0.094 (0.167) | -0.082 (0.164) | -0.045 (0.154) | -0.064 (0.159) | -0.048 (0.154) | 0.055 (0.169) | | 0.051 (0.169) | -0.052 (0.151) | | -0.051 (0.152) |
| Gov. crises | -0.100* (0.058) | | | -0.074 (0.062) | | | -0.083 (0.060) | | | -0.095* (0.057) | | | -0.108* (0.058) | | |
| Revolutions | | 0.015 (0.057) | | | 0.008 (0.070) | | | 0.036 (0.058) | | | 0.011 (0.058) | | | 0.020 (0.057) | |
| All internal conflicts | | | -0.006 (0.016) | | 0.001 (0.017) | | | | -0.002 (0.016) | | | -0.008 (0.015) | | | -0.005 (0.016) |
| International war | 0.020 (0.057) | 0.014 (0.057) | 0.014 (0.057) | | | | | | | | | | | | |
| Autocracy*international war | -0.128 (0.083) | -0.133 (0.081) | -0.135 (0.082) | | | | | | | | | | | | |
| Log(revenue) | | | | 0.066* (0.036) | 0.070* (0.036) | 0.073** (0.035) | | | | | | | | | |
| Log(per capita GDP) | | | | | | | 0.064 (0.056) | 0.073 (0.056) | 0.074 (0.056) | | | | | | |
| Urbanization | | | | | | | | | | -134.660 (91.018) | -125.540 (84.307) | -121.537 (82.727) | | | |
| Trade | | | | | | | | | | | | | 0.637 (16.805) | 1.123 (17.564) | 0.966 (17.229) |
| Number of countries | 172 | 172 | 172 | 158 | 158 | 158 | 156 | 156 | 156 | 171 | 171 | 171 | 171 | 171 | 171 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Continent-time linear trends | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 977 | 977 | 977 | 861 | 861 | 861 | 871 | 871 | 871 | 949 | 949 | 949 | 962 | 962 | 962 |
| R-squared | 0.415 | 0.424 | 0.419 | 0.464 | 0.471 | 0.469 | 0.484 | 0.489 | 0.486 | 0.432 | 0.438 | 0.433 | 0.448 | 0.453 | 0.449 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 20% from the previous 10 years. All explanatory variables are lagged. Standard errors are clustered at the country level.

Table A5
Educational reforms and regime durability, controlling for continent specific linear trends

| | (1) | (2) | (3) | (4) | (5) | (6) |
|------------------------------------|----------------|----------------|------------------|------------------|----------------|------------------|
| Educational reform: 20% threshold | | | | | | |
| Autocracy*regime durability | -0.292* | -0.299* | -0.441*** | -0.476*** | -0.293* | -0.758*** |
| | (0.170) | (0.169) | (0.144) | (0.168) | (0.164) | (0.199) |
| Autocracy | 0.065 | 0.073 | 0.062 | 0.049 | 0.057 | 0.106 |
| | (0.053) | (0.053) | (0.056) | (0.054) | (0.054) | (0.066) |
| Pop. growth | 0.120 | 0.114 | 0.001 | 0.037 | 0.071 | -0.045 |
| | (0.105) | (0.106) | (0.114) | (0.119) | (0.103) | (0.145) |
| Regime durability | 0.032 | 0.037 | 0.139 | 0.171 | 0.078 | 0.396*** |
| | (0.131) | (0.135) | (0.120) | (0.115) | (0.118) | (0.138) |
| International war | | -0.012 | | | | |
| | | (0.048) | | | | |
| Autocracy*international war | | -0.050 | | | | |
| | | (0.067) | | | | |
| Log(revenue) | | | 0.030 | | | |
| | | | (0.028) | | | |
| Log(per capita GDP) | | | | 0.056 | | |
| | | | | (0.048) | | |
| Urbanization | | | | | -218.516 | |
| | | | | | (144.590) | |
| Trade | | | | | | 7.110 |
| | | | | | | (14.486) |
| Number of country | 183 | 183 | 162 | 156 | 175 | 171 |
| Country fixed effects | yes | yes | yes | yes | yes | yes |
| Continent-time linear trends | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes |
| Observations | 1,387 | 1,387 | 1,131 | 1,100 | 1,308 | 974 |
| R-squared | 0.353 | 0.354 | 0.377 | 0.407 | 0.368 | 0.458 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1825-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 20% from the previous 10 years. All explanatory variables are lagged. Standard errors are clustered at the country level.

Table A6
Educational reforms and threat to democracy, SGMM specification

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|-----------------------------------|----------------|----------------|-----------------------------------|-----------------|----------------|-------------------------------------|-----------------|-----------------|
| | Educational reform: 10% threshold | | | Educational reform: 20% threshold | | | Legal reform, based on Flora (1983) | | |
| Autocracy*gov. crises | 0.198* | | | 0.190* | | | 0.470*** | | |
| | (0.102) | | | (0.113) | | | (0.172) | | |
| Autocracy*revolutions | | 0.152** | | | 0.233*** | | | 1.884*** | |
| | | (0.070) | | | (0.062) | | | (0.613) | |
| Autocracy*all internal conflicts | | | 0.053** | | | 0.047** | | | 0.172*** |
| | | | (0.021) | | | (0.023) | | | (0.064) |
| Autocracy | 0.018 | 0.001 | -0.050 | 0.009 | 0.010 | -0.026 | -0.488* | -0.315 | -0.528** |
| | (0.058) | (0.052) | (0.061) | (0.059) | (0.053) | (0.060) | (0.256) | (0.213) | (0.269) |
| Pop. growth | 0.566*** | 0.530*** | 0.602*** | 0.357*** | 0.286** | 0.374*** | -0.720 | -0.433 | -0.552 |
| | (0.179) | (0.178) | (0.179) | (0.132) | (0.129) | (0.131) | (0.894) | (0.867) | (0.810) |
| Gov. crises | -0.093 | | | -0.118* | | | 0.023 | | |
| | (0.057) | | | (0.060) | | | (0.048) | | |
| Revolutions | | 0.050 | | | 0.017 | | | 0.273** | |
| | | (0.063) | | | (0.050) | | | (0.117) | |
| All internal conflicts | | | -0.017 | | | -0.009 | | | 0.026 |
| | | | (0.016) | | | (0.014) | | | (0.021) |
| Number of country | 172 | 172 | 172 | 172 | 172 | 172 | 14 | 14 | 14 |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Ar(2) test | 0.482 | 0.489 | 0.46 | 0.81 | 0.916 | 0.902 | 0.414 | 0.496 | 0.501 |
| P-value Hansen test | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Observations | 977 | 977 | 977 | 977 | 977 | 977 | 110 | 110 | 110 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Observations are 10 year country averages, for the 1925-2014 period. Educational reform is a dummy if primary per capita school enrollment increased more than 10% (20%) from the previous 10 years (columns 1-3, and columns 4-6). Legal reform is a dummy if the country experienced at least one legal reform during the 10 year period (the definition and timing of legal reforms come from Flora (1983)). All explanatory variables are lagged. Standard errors are clustered at the country level

Table A7
Primary school enrollment per capita and threat to democracy

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
|---|--------------------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | Primary school enrollment per capita | | | | | | | | | | | | | | |
| Autocracy*gov. crises | 8.364*** | 5.183** | 5.212* | 5.039* | 4.978* | | | | | | | | | | |
| | (3.124) | (2.615) | (2.706) | (2.725) | (2.790) | | | | | | | | | | |
| Autocracy*revolutions | | | | | | 3.858 | 1.588 | 1.908 | 2.102 | 2.149 | | | | | |
| | | | | | | (3.987) | (3.255) | (3.308) | (3.402) | (3.482) | | | | | |
| Autocracy*all internal conflicts | | | | | | | | | | | 1.827** | 1.551** | 1.537* | 1.536* | 1.602** |
| | | | | | | | | | | | (0.905) | (0.769) | (0.787) | (0.795) | (0.780) |
| Autocracy | -3.671 | -2.698 | -2.780 | -2.884 | -2.496 | -1.940 | -1.674 | -1.880 | -2.099 | -1.765 | -3.924 | -3.820 | -3.938 | -4.131 | -3.892 |
| | (5.762) | (4.772) | (4.996) | (5.208) | (5.181) | (5.812) | (4.810) | (5.035) | (5.217) | (5.180) | (5.836) | (4.812) | (5.078) | (5.300) | (5.270) |
| Pop. growth | -291.355 | -381.496 | -387.862 | -387.680 | -381.869 | -293.121 | -384.719 | -390.843 | -390.643 | -385.126 | -292.145 | -383.540 | -389.864 | -389.521 | -384.014 |
| | (297.626) | (429.866) | (436.184) | (440.539) | (437.487) | (298.584) | (431.707) | (438.042) | (442.279) | (439.270) | (297.806) | (430.062) | (436.465) | (440.710) | (437.680) |
| Gov. crises | -3.195* | -0.122 | 0.236 | 0.318 | 0.249 | | | | | | | | | | |
| | (1.924) | (1.617) | (1.686) | (1.675) | (1.699) | | | | | | | | | | |
| Revolutions | | | | | | -1.402 | 1.253 | 0.991 | 0.922 | 1.002 | | | | | |
| | | | | | | (3.230) | (2.287) | (2.318) | (2.387) | (2.639) | | | | | |
| All internal conflicts | | | | | | | | | | | -0.316 | -0.174 | -0.151 | -0.146 | -0.190 |
| | | | | | | | | | | | (0.642) | (0.516) | (0.517) | (0.515) | (0.523) |
| Primary school enrollment p.c. (first lag) | 1.000*** | 1.393*** | 1.378*** | 1.374*** | 1.374*** | 1.001*** | 1.394*** | 1.379*** | 1.375*** | 1.375*** | 1.001*** | 1.393*** | 1.379*** | 1.374*** | 1.374*** |
| | (0.008) | (0.047) | (0.050) | (0.051) | (0.050) | (0.008) | (0.048) | (0.051) | (0.052) | (0.050) | (0.008) | (0.048) | (0.051) | (0.052) | (0.050) |
| Primary school enrollment p.c. (second lag) | | -0.407*** | -0.384*** | -0.382*** | -0.383*** | | -0.408*** | -0.384*** | -0.382*** | -0.382*** | | -0.407*** | -0.384*** | -0.382*** | -0.383*** |
| | | (0.047) | (0.067) | (0.065) | (0.062) | | (0.048) | (0.067) | (0.065) | (0.063) | | (0.048) | (0.067) | (0.065) | (0.063) |
| Primary school enrollment p.c. (third lag) | | | -0.008 | -0.009 | -0.007 | | | -0.008 | -0.010 | -0.009 | | | -0.007 | -0.009 | -0.008 |
| | | | (0.032) | (0.042) | (0.041) | | | (0.032) | (0.042) | (0.042) | | | (0.032) | (0.042) | (0.041) |
| Primary school enrollment p.c. (fourth lag) | | | | 0.004 | 0.004 | | | | 0.005 | 0.005 | | | | 0.005 | 0.004 |
| | | | | (0.026) | (0.031) | | | | (0.026) | (0.032) | | | | (0.026) | (0.032) |
| Primary school enrollment p.c. (fifth lag) | | | | | -0.048 | | | | | -0.048 | | | | | -0.047 |
| | | | | | (0.035) | | | | | (0.035) | | | | | (0.035) |
| Primary school enrollment p.c. (sixth lag) | | | | | | 0.085** | | | | 0.085** | | | | | 0.085** |
| | | | | | | (0.036) | | | | (0.037) | | | | | (0.037) |
| Primary school enrollment p.c. (seventh lag) | | | | | | | -0.074 | | | -0.076 | | | | | -0.075 |
| | | | | | | | (0.052) | | | (0.052) | | | | | (0.052) |
| Primary school enrollment p.c. (eighth lag) | | | | | | 0.014 | | | | 0.013 | | | | | 0.013 |
| | | | | | | (0.033) | | | | (0.033) | | | | | (0.033) |
| Primary school enrollment p.c. (ninth lag) | | | | | | 0.011 | | | | 0.013 | | | | | 0.012 |
| | | | | | | (0.034) | | | | (0.034) | | | | | (0.034) |
| Primary school enrollment p.c. (tenth lag) | | | | | | 0.017 | | | | 0.017 | | | | | 0.017 |
| | | | | | | (0.025) | | | | (0.025) | | | | | (0.025) |
| Country fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Year fixed effects | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes | yes |
| Observations | 2,893 | 2,817 | 2,758 | 2,702 | 2,694 | 2,893 | 2,817 | 2,758 | 2,702 | 2,694 | 2,893 | 2,817 | 2,758 | 2,702 | 2,694 |
| R-squared | 0.989 | 0.991 | 0.991 | 0.991 | 0.991 | 0.989 | 0.991 | 0.991 | 0.991 | 0.991 | 0.989 | 0.991 | 0.991 | 0.991 | 0.991 |

Notes: ***, ** and * indicate significance at the 1, 5 and 10% level. Standard errors are clustered at the country level.