Democratic Values and Institutions[*]

Timothy Besley   Torsten Persson
LSE and CIFAR   IIES and CIFAR

February 2017

Abstract

This paper attempts to bridge political science/sociology research that sees democratic values as a crucial pre-requisite for democratic institutions with economics research that sees equilibrium institutions as the product of strategic decisions by key groups in society. We propose a framework for analyzing the two-way interplay between democratic values and democratic institutions. Some citizens hold values that make them protest to preserve strong executive constraints or open executive recruitment, and the share of such citizens evolves over time. Our model suggests a natural, but important, complementarity between values and institutions so that the dynamics of values tend to reinforce institutional choices. This approach can help explain variation in democratic values and the history of political institutions across countries and over time, giving insight into country-level heterogeneity. It can also be used to consider how foreign intervention may influence values and institutions.

"(I)f a political system is not characterized by a value system allowing the peaceful "play" of power ... there can be no stable democracy." Lipset (1959, p. 71)

"During the nineteenth century most Western societies extended voting rights, ... these political reforms can be viewed as strategic decisions by the political elite to prevent widespread social unrest and revolution." Acemoglu and Robinson (2000, p. 1167).

1   Introduction

A body of research by economists and economic historians gives a central role to political institutions in explaining cross-country differences in economic growth and development.[1] Apart from being intrinsically interesting, understanding the origins and prospects for reform of institutions is therefore instrumentally important. In this paper, we pick up the thread from another body of research by political scientists and sociologists where democratic values play a key rule in supporting democratic institutions. Incorporating slowly evolving democratic values generates new insights into institutional dynamics.

Looking across the world and over history, the heterogeneity in democratic experiences is striking. Some polities have made secure transitions into open elections of leaders who face constraints on their executive authority – the two hallmarks of democratic institutions. In such cases, these arrangements go largely unchallenged and the rulers and policies chosen through these institutions are accepted with little or no protest. At the other end of the spectrum, we observe a set of states which have never secured democracy: elections are at best a fig leaf to give dictators and single-party regimes a veneer of legitimacy, while incumbent powers remain unchecked by legal and parliamentary institutions. A third group occupies a middle ground where progress towards democracy or autocracy is punctuated by protests and institutional reversals, while occasionally a country joins one of the stable groups.[2]

While these broad patterns are well-known, another aspect of heterogeneity has received less attention. Unbundling democracy into openness of executive recruitment (particularly how far elections are used) and strong constraints on the executive, reveals that reforms across these two dimensions are often not synchronized. Moreover, at least in the modern era, countries tend to achieve openness in executive recruitment earlier than they build strong executive constraints.

We model the drivers of democratic development in a way that encompasses both aspects of country heterogeneity. By giving a role to the dynamics of democratic values, our suggested approach supplements the recent economics literature which focuses on institutional change as the product of strategic investments, including costly decisions to fight by those who would gain or lose from institutional reversals.[3] The model that we develop gives neither institutions nor values an upper hand in the process of democratic change – rather, the two evolve jointly and interdependently. It

---

[1]Interest in these issues goes back, at least, to the seminal work of North (1983).

[2]Fact 1 and 2 in Section 4 below gives a graphical interpretation of these patterns.

[3]Section 2 below reviews some of the key ideas in these existing literatures.

accommodates both commonalities and differences between the forces that shape open executive recruitment and strong executive constraints, helping to understand a richer set of institutional experiences.

We use the model to interpret broad patterns in the data in ways that are consistent with a number of seemingly separate ideas and findings in the existing literature. For example, our model suggests a mechanism behind a long-lived effect of historical institutions, like the colonial-origins hypothesis of Acemoglu, Johnson, and Robinson (2001). It also suggests why cumulated values – like social or democratic capital – may consolidate change, as in Putnam (1993) and Persson and Tabellini (2009). Even though incumbents are free to make reforms in each period, political institutions become sticky in equilibrium due to slow-moving democratic values. Our model provides a new perspective on the resource curse as a source of political disruption (Mehlum, Moene, and Torvik 2006). It also gives a theoretical underpinning to the notion of critical junctures in history and subsequent institutional paths, as emphasized by Acemoglu and Robinson (2012). Moreover, the theory allows for different types of reforms of political institutions. Our dynamic equilibria thus entail "defensive" reforms – a ruling elite voluntarily relinquishing political control given the expected costs of trying to hold on to it (Acemoglu and Robinson 2000, 2006) – as well as classic "offensive" reforms, when citizens oust incumbent elites from power to bring about institutional change (Marx and Engels 1848, Kuran 1995). Finally, it also gives a framework for understanding foreign influence and its persistent effects.

The paper is organized as follows. The next section gives a short selective overview of the cultural and strategic approaches to political institutions. Section 3 provides some background facts about the dispersion of democratic institutions and values over countries and time. Section 4 sketches a canonical modeling approach to the interplay between democratic institutions and democratic values. Section 5 puts this approach to work: first we show how to model executive constraints and executive recruitments separately, then we extend the framework to sequential choices of both institutions. In Section 6, we return to the data and explain how the model help us understand the patterns discussed in Section 3. Section 7 sketches how the theory can easily be extended to think about the effects of active or passive influence by foreign powers on the paths of values and institutions, and presents some correlations in line with the model. Section 8 concludes. Four Appendixes collect proofs and some additional material.

2   Background

The idea of a cultural basis for democracy goes back, at least, to Aristotle. But the *locus classicus* is Montesquieu (1748), which spells out how factors like geography and climate interact with different cultures to produce "spirits" that shape the working of alternative political institutions, including the separation of powers. In modern political science, Lipset (1959) and Almond and Verba (1963) stand out in their study how political culture is vital for democracy. One strand of this literature

focuses on the importance of education in supporting civic values. These ideas have been taken up, with a global focus on measurable attitudes, in the work of Inglehart (1997) and Inglehart and Welzel (2005). The latter argues that pro-democratic attitudes can serve to measure the demand for democratic change.

In line with our model's mechanism, Welzel (2007) argues that emancipating mass attitudes measured in the World Values Survey demonstrate citizens' willingness to struggle for democracy.[4] We use such values to enumerate the facts in Section 3 and use our model to illuminate them in Section 5. Also related is Gorodnichenko and Roland (2015), who emphasize values in their theoretical and empirical analysis of why individualistic rather than collectivistic cultures are more likely to underpin democratizations.

Although suspicious of cultural explanations as potentially circular, Moore (1966) recognizes that if culture matters it does so in a dynamic fashion. He notes that "to take values as the starting point of sociological explanation makes it very difficult to understand the obvious fact that values change in response to circumstance" (p. 487). In the same vein, our approach focuses on changing, rather than static, democratic values.

Almond and Verba (1963, p. 367) discuss the importance of the socialization process that shapes civic culture, which "includes training in many social institutions – family, peer group, school, work place, as well as in the political system itself". Crucial in our approach is a model of cultural evolution, which is inspired by earlier research in anthropology beginning with Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985).

In economics, research on culture and individual behavior has increased in the past ten to fifteen years; a body of work is surveyed and discussed in Bisin and Verdier (2011). We model cultural change through the dynamics of preferences or values (rather than dynamics of behavior or beliefs) of a specific group. In that sense, we follow the lead of Güth and Yaari (1992). Unlike the earlier economics literature, we focus on the values that govern political, rather than economic or social, behavior.

Some aspects of our modelling approach parallels the theory of franchise extension, particularly that in Acemoglu and Robinson (2000, 2006), who also offer insightful case-study evidence. These authors emphasize the quest for political instability, due to the threat of revolution. The franchise is used as a commitment device by the elite to guarantee the masses more favorable treatment. Aidt and Jensen (2014) find some econometric evidence supporting this view. The strategic reform approach typically has no state variable while, in our model, democratic values play this role.

Our model is also related to Lizzeri and Persico (2004) where a ruling group voluntarily extends the franchise to shift spending from targeted transfers to broad-based programs. This resembles a classic argument – first made by Rokkan (1970), then extended by Boix (1999) – that fears of electoral losses explain the move from plurality to proportional representation to protect the center-right from

---

[4]See also Welzel, Inglehart and Kruse (2015).

a labor electoral landslide in some countries in early 20th-century Europe where landed and industrial elites had not forged their interests.

Our paper also relates to a range of contributions which have modeled the dynamics of political rights. Weingast (1997) develops a game-theoretic model in which the rights of citizens emerge as a self-enforcing equilibrium. Lagunoff (2001) develops a dynamic game between two groups, in which greater political turnover leads to greater constitutional support for civil liberties. Congleton (2007) discusses forces that promoted parliamentary oversight of royal power, focusing on instability due to shocks to the monarch's preferences.

By emphasizing constraints on the executive (in addition to the franchise), our research relates to Acemoglu, Robinson and Torvik (2011) who propose a model of endogenous checks and balances stressing how these change the ability of special interests to influence policy. Building on the ideas in Besley and Persson (2011), Besley, Persson, and Reynal-Querol (2016) analyze theoretically and empirically how the resilience of incumbent leaders may rub off on their motives to introduce stronger executive constraints. One key idea in these papers is the need to contrast the motives to create open elections with those that constrain incumbents. This is also at the heart of Mukand and Rodrik (2015) who draw a distinction between electoral and liberal democracy.

Some authors maintain that human capital and education, rather than institutions, is the main driver of development – see, e.g., Glaeser et al (2004). They emphasize two related dimensions of education which are relevant to this paper. It is often the educated (like university students) who protest, and education can transform values. In our paper, concerned citizens are the conduit of democratic values and, although we emphasize parental transmission, the education system may also be important.

Closest to our approach is the important unpublished paper by Ticchi, Verdier and Vindigni (2013), who also model the interaction between value formation through socialization and reforms of political institutions, when democratic values shape citizens' willingness to defend democracy. As well as differences in the modelling, they consider only a single democratic dimension and place less emphasis of matching theory and facts. On the other hand, they do consider the role of education explicitly.

Persson and Tabellini (2009) study interactions between political reforms and economic growth, when people are more prone to support democracy at higher levels of democratic capital, which is passively and exogenously accumulated depending on the incidence of democracy in a country's past history or its nearby geography.

The mechanism that we propose for sustaining institutional change builds on the idea that values create a credible threat to protest against institutions that produce unfair outcomes relative to a reference point. As such, it is related to Passarelli and Tabellini (2016) who consider how values underpin the willingness to protest in the wake of policies which citizens regard as unfair.

Finally, our paper relates to a few recent studies of the two-way interaction between belief (rather

than value) formation and formal institutions. One example is Benabou (2008) who studies interactions between the size of government and people's beliefs (ideologies) about the relative efficiency of state vs. market solutions. Another is the recent paper by Levy and Razin (2016) who analyze interactions between segregation in private and public school choices and polarization of beliefs about the relative merits of these types of schools.

3   Facts on Institutions and Values

In this section, we focus on two sets of facts obtained from the Polity IV (PIV) and the World Values Survey (WVS) datasets. These facts tee up the aspects of country-level experience that our model will focus on.

Institutions   We highlight two important, but inherently different, aspects of democracy: constraints on executive power – i.e., the presence of checks and balances – and open recruitment of the executive – i.e., the presence of open elections. In each case, we define a dummy variable for each country and year. For the first, PIV measures the executive constraints on a seven-point scale, and our binary measure of *strong executive constraints* gives a 1 to a country and year with the highest score on this scale, which is reserved for regimes where a legislature or the courts can and does refute the executive's proposals.[5] For the second, we define a dummy variable from the four-point scale in PIV, and set *open executive recruitment* at 1 for a country and year with the highest score on this scale, which is reserved for regimes where the top executive is directly elected or appointed.[6]

As background, Figure 1 plots the time paths of these two measures. The left-hand panel holds the sample constant, taking the 50 countries that appear in the PIV data in 1875. The right-hand panel displays all countries in the data in each year – i.e., it includes countries that enter the data in the post-war period after becoming independent at the end of colonialism.

To explore the relationship between these two aspects of democracy, we create four distinct regime types depending on openness of recruitment and executive constraints. In Figure 2, we track the distribution of countries across this four-way split by decade from 1900 to 2011 for the fixed sample of 50 countries. The proportion of countries which are both open *and* strong has increased over time, while the fraction of open and weak countries has remained fairly stable. The strong but closed regime type disappears from the sample early in the 20th century, and there has also been a decline in the closed and weak type. Explaining what underpins Figure 2 reinforces the need for a theory that considers separate dimensions of institutional change rather than looking at democracy in a one-dimensional way.

---

[5]The coders designate this as a case where *"(a)ccountability groups have effective authority equal to or greater than the executive in most areas of activity." (Polity IV, Coding Manual)*

[6]These variables are respectively, the "xconst" variable (measured on a one to seven scale) and "xropen" (measured on a one to four scale) from PolityIV. Both of these are components of the aggreate Polity2 democracy score.

Our main focus in this paper is on the heterogeneity of country histories as expressed in:

Facts about Institutions (i) *Histories of reforms towards strong executive constraints are heterogenous across countries and can be classified into three broad forms: permanent transitions into strong or weak executive constraints, or churning between the two, with the churning group being the most prevalent one.* (ii) *Reforms of executive recruitment display a similar pattern, with the same three broad forms.* (iii) *Most countries tend to introduce open executive recruitment before acquiring strong executive constraints.*

Table 1 illustrates fact (i) for the 50 countries with PIV data from 1875. It classifies each country according to its history for executive constraints. The left-most column shows how long-standing transitions into permanent autocracy tend to be. Similarly, the top of the right-most column shows a striking longevity of democratic traditions in countries with democratic institutions from the outset (or from 1800). Our model will produce permanent transitions in both directions. That said, transitions to democracy have taken place much more recently in countries at the bottom of the right-most column, with the exception of Costa Rica and Sweden. The middle column is the largest part of the sample, with transitions in both directions. For some countries, however, this was due to strong external influence – e.g., German occupation in World War II. We return to such episodes in Section 7.

Table 2 illustrates fact (ii) for the openness of executive recruitment using the same sample of countries. Comparing it to Table 1, we see a similar structure with the same three broad forms, although few countries have been permanently closed in terms of executive recruitment.

Table 3 provides a window on fact (iii); it gives the year when open recruitment and strong executive constraints was first introduced for the same set of countries. As we know from Tables 1 and 2, this is only a partial picture, as many countries have gone back and forth. Table 3 illustrates the fact that the most common historical pattern is for open executive recruitment to precede strong executive constraints. The exceptions to this general pattern – Belgium, Ethiopia, Japan, Netherlands, Norway, Portugal, Spain, and the UK – all started out with an entrenched aristocratic system and a strong monarch whose powers were gradually curtailed by courts and legislatures. The model presented in Section 5.3 offers insights into when mixed regimes are likely, even if there is a common driver for both dimensions of democracy.

Values We consider heterogeneity in democratic values using individual data on attitudes in waves 5 and 6 of the WVS. We focus on a question which asks people to rate the importance of democracy on a ten-point scale. In line with our theoretical approach below – that emphasizes the share of people with strong democratic values – we adopt a bivariate measure where a citizen has (strong) democratic values if she gives democracy a grade above 8. The global mean of this dummy variable is about 0.7. This yields:

Facts about Values (i) *Support for democracy varies not only across individuals, but also across countries.* (ii) *Values are systematically correlated with political histories, with the strongest (weakest) support among those that have experienced long (short) histories of democracy.*

To illustrate these two facts, the left-hand panel in Figure 3 plots each country's average (deviation from the sample mean) against the fraction of years it has had high openness as well as strong executive constraints, showing a positive relation between the two.[7] The right-hand panel shows a similar relation, when we use residual support for democracy, holding constant each individual's gender, education, age and income.[8]

Figure 4 shows fact (ii) more clearly: the systematic relation between democratic values and political histories. It relies on individual data in the three groups of countries used in Tables 1-3, given that they appear in WVS waves 5 and 6. As the figure shows, the average support for democracy is strongest among citizens in countries with long-standing democratic traditions, weakest among those with long-standing autocratic traditions, and in between among those with mixed histories where the grouping of countries follows Table 1.

4   A Canonical Model

Our core framework is centred around a conflict of interest between an incumbent group and some subordinated groups over democratic institutions. In each period, an incumbent leader chooses whether or not to install a democracy without being able to commit future leaders. Incumbents may follow their immediate interests and resist democratic institutions at the price of confronting protests by some citizens. The *only* dynamic element in the model is the proportion of individuals who protest because they have democratic values, a proportion which evolves along the equilibrium path.

Groups and conflicts of interest   The framework has $M + 1$ groups of equal size, normalized to unity. These groups are labelled $m = I, 1, ..., M$, where group I denotes the incumbent elite. Institutions $D_t \in \{0, 1\}$ – where $D_t = 1$ denotes democracy and $D_t = 0$ autocracy – directly affect the material payoffs to these groups as does the realization of random variable $x_t \in [\underline{x}, \overline{x}]$, with distribution function is $H(\cdot)$. Group m's expected payoff with realization x and institution D is:

$$u^m(x, D),$$

---

[7]Specifically, we compute the number of years for which each country has had openess in recruitment and strong constraints then take the sum of these two variables divided by 2.

[8]We estimate a linear probability model at the individual level with the dummy for democratic support on the left-hand side and also including on the right-hand side controls for gender, ten dummies for income groups, three for education groups, and three age bands. To construct the figure, we average the residuals at the country level.

which we assume is (weakly) increasing in x. The core conflict of interest is captured in the following two assumptions:

$$u^I(x, 0) - u^I(x, 1) > 0 \text{ and increasing in } x \text{ for all } x \in [\underline{x}, \overline{x}]$$

and

$$u^m(x, 0) - u^m(x, 1) < 0 \text{ for all } x \in [\underline{x}, \overline{x}] \text{ and } m = 1, ...M.$$

Thus the incumbent elite and the subordinated groups have different institutional preferences: the incumbent prefers autocracy over democracy, whereas everyone else has the opposite preference. The next two subsections give concrete alternative interpretations to D_t and x_t as well as the conflict of interest.

Types, democratic values, and losses from injustice   Individuals are of two types. A fraction $1 - \mu_t$ are standard economic agents: they are *passive*, never protest with date-t utility is

$$u^m(x_t, D_t).$$

The remaining fraction, μ_t are *concerned* citizens and constitute a prospective civil-society to support democracy. These agents are assumed to be equally distributed across all of the groups in society including the incumbent group.[9]

Concerned citizens differ from passive citizens in two key ways. First, they have democratic values which we capture by suposing that they care about social justice in addition to their material payoffs and second, they are willing to join protests against the government if $D_t = 0$, i.e., if the government does not grant democratic rights.

To capture the first, their payoffs are now

$$u^m(x_t, D_t) + S(r_g, r_l, D_t, x_t)$$

where

$$S(r_g, r_l, D, x) = \sum_{m=1}^{M} \left[\min\left\{ u^m(x, D) - u^m(x, r_l), 0 \right\} + \max\left\{ u^m(x, r_g) - u^m(x, D), 0 \right\} \right] \quad (1)$$

reflects their view of the society-wide benefits or cost from a particular political institution relative to a reference point for gains r_g and a reference point for losses r_l.[10]   We think of these reference

---

[9]For simplicity we assume that this applies also to the ruling group. This assumption also seem to have support in history where some elite members have frequently seen injustice in institutional arrangements even if those favor their own group.

[10]The key role of reference dependence in our framework is the implication that citizens are aggrieved when they

points as defining their values and it is natural to choose $r_g = 0$ and $r_l = 1$ so that gains are measured relative to the worst institution and losses relative to the best. This approach enables us to combine in a simple way that (i) the concerned experience a loss from not having the (democratic) institution they prefer and (ii) democracy provides a reference point for the loss.[11] This way of capturing injustice is, of course, quite specific. But the idea of reference-dependent preferences is well-established following the seminal work of Kahneman and Tversky (1979) and a range of psychological studies.[12] The formulation also makes democratic values distinct from standard preferences.[13] The payoffs of concerned citizens depend on their sense of injustice based on *societal* gains/losses. Values thus embody concerned citizens' views about the kind of society that they wish to live in.

Concerned citizens are assumed *always* to join in any protest to protect democracy, whenever they have the opportunity of doing so. Thus they are behavioral rather than strategic. Thus, the size of group of concerned citizens determines the magnitude of resistance that the government faces if $D_t = 0$. This captures the idea by Weber (1922) of value rationality (*wertrational*). An action is value-rational when it is determined by a conscious belief in the value for its own sake – due to some ethical, aesthetic, religious, or other form of behavior – independently of its prospects of success.[14]

Fighting by the incumbent The incumbent can respond to anticipated protests by spending resources to fight them.[15] Let f_t denote the forces recruited for fighting in period t, with wf_t being the cost of hiring this group. In each period, concerned citizens have an opportunity to join protests against the incumbent with exogenous constant probability ρ. We denote a protest opportunity by $r = 1$ and no opportunity by $r = 0$. If a protest occurs in period t and a fraction ϕ_t of individuals participate, the probability of unseating the incumbent group is $p(\phi_t, f_t)$. This probability is increasing in ϕ, decreasing and convex in f, and satisfies:

$$p_\phi - \frac{p_f p_{\phi f}}{p_{ff}} \geq 0 \text{ for all } \phi \in [0, 1] \text{ and } f, \text{ and } \lim_{\phi \to 0} p_f(\phi, f) \to 0 \text{ for all } f \tag{2}$$

---

lack constitutional rights. This emerges naturally when a reference point is used to value outcomes.

[11]Gains and losses here are necessarily equal and opposite given the discrete nature of the institutional choice. Below, we study the case of two dimensional institutional choice where some institutional configurations can lead to simultaneous gains and losses.

[12]There are many applications of reference-dependent preferences to concrete phenomena as discussed, for example, in Kahneman et al (1991). For a more recent theoretical treatment of reference-dependent preferences, see Koszegi and Rabin (2006).

[13]Here, we assume that gains and losses are valued symmetrically. However, there is ample evidence for loss aversion and our framework could introduce this straightforwardly by supposing that concerned citizens weigh losses more than gains. The core results below would still go through under that alternative assumption.

[14]Weber (1922) contrasts this with instrumental rationality which he terms *zweckrational*. This more standard model could be pursued in our framework as long as the *size* of the protesting group is increasing in μ. Thus, the model could be extended following the approach of Persson and Tabellini (2009) to model protests as a global game where the concerned receive a payoff from participating. Then, ϕ_t and μ_t may diverge. However, as long as there are strategic complementarities of the kind proposed by Persson and Tabellini, the equilibrium fraction of the concerned who protest will increase in μ. The qualitative properties of this extended model will be similar to what we have here, but the strategic choice of participation will tend to magnify the effects.

[15]We do not allow the incumbent to buy off protesters, although this would lead to similar trade-offs.

The first condition guarantees that $p(\cdot)$ increases in ϕ, even when the leadership optimally fights a protest. The second implies that the incumbent puts no resources into fighting if nobody is protesting.

Population structure and socialization We consider a model with successive generations, which overlap only in so far as parents endow their children with values, as modeled in Besley (2015). Children have two parents and – to keep the population balanced – all pairs have two children. We also assume that all marriage matching is random.[16]

Children are socialized into having democratic values. For simplicity, we model socialization as resulting from a form of osmosis rather than strategic behavior by parents.[17] Two parents of the same type simply pass along the values associated with their common type. However, children whose parents have different types get their type depending on the expected utilities of being concerned with democratic values rather than passive. Let $\Delta(\mu)$ be the expected utility difference between these types – their relative fitness – when the proportion concerned in the population is μ. Moreover, let $\eta \in (-\infty, \infty)$ be a couple-specific idiosyncratic negative shock to this utility difference. Then, a child with mixed parentage becomes concerned with democratic values, if and only if $\eta \leq \Delta$.

We assume that η has a symmetric single-peaked distribution with c.d.f. G and p.d.f. g. This implies that a mixed-marriage child holds democratic values with probability $G(\Delta(\mu))$ at utility difference $\Delta(\mu)$. By the law of large numbers, this is also the proportion among those with mixed parentage. By definition, $G(\cdot)$ is monotonically increasing, and by symmetry $G(0) = 1/2$.

Evolution of democratic values Using the notation above, we can write the evolution of democratic values as:

$$\mu_{t+1} = \mu_t^2 + 2\mu_t (1 - \mu_t) G(\Delta(\mu_t)).$$

A fraction μ_t^2 of matches occur between concerned citizens, whose children all obtain democratic values, while $2\mu_t (1 - \mu_t)$ is the fraction of mixed households, whose children obtain democratic values with probability $G(\Delta(\mu_t))$. Manipulating this expression, we get the change in the proportion of citizens with democratic values:

$$\mu_{t+1} - \mu_t = 2\mu_t (1 - \mu_t) [G(\Delta(\mu_t)) - 1/2]. \tag{3}$$

Clearly, μ goes up (down) whenever relative fitness $\Delta(\mu)$ is positive (negative). This is the one and only dynamic equation in the model.[18]

---

[16] For the results to go through, we only require that there is at least some element of random matching. With full assortative matching, there would be no socialization as all offspring would have parents of the same type.

[17] This makes the model simpler and does not fundamentally affect the insights compared to the strategic socialization model of Bisin and Verdier (2001). Appendix B shows that this is the case.

[18] There are a variety of approaches which yield a dynamical process analogous to (3). For example, something similar would follow from replicator dynamics with different behavioral types – see Appendix C.

The evolution of democratic values is grounded in the payoffs different types receive; the link between $\Delta(\mu_t)$ and the dynamics of values drives the model over time and this in turn depends on the relative fitness of being concerned rather than passive. Hence, the sign of $\Delta_\mu(\mu)$ is crucial to the equilibrium dynamics. This formulation assumes that socialization is based on on μ_t rather than μ_{t+1}. Thus, parents ask themselves about the expected relative fitness of holding democratic values (under a veil of ignorance about state variable x) for a member of their own generation.[19]

Timing　The timing within a generation has four steps:

1. A leader in generation t is selected from incumbent group I and x_t is realized.

2. This leader chooses D_t and f_t.

3. Under democracy $D_t = 1$, the payoffs are $u^m(x_t, 1)$.

3´. Under autocracy $D_t = 0$, the concerned get an opportunity to protest with probability ρ. If the protest is unsuccessful, payoffs are $u^m(x_t, 0)$. If it is successful, the incumbent is unseated and $u^m(x_t, 1)$.

4. Payoffs are realized, a new generation is born and socialized by their parents before they die. If the incumbent group is not unseated, it stays until period $t+1$. If it is unseated, then a new incumbent is drawn at random.

Before studying socialization at step 4, we first solve the rest of the generational equilibrium backwards.

Protests and payoffs – step 3　Given that all concerned citizens protest, the fraction of protesters is equal to μ_t and the probability of unseating the incumbent becomes $p(\mu_t, f_t)$. Now consider what happens when $D_t = 0$. The expected payoff to the incumbent leader with his preferred institution is

$$\widetilde{V}(x_t, \mu_t, f_t) = [1 - rp(\mu_t, f_t)] u^I(x_t, 0) + rp(\mu_t, f_t) u^I(x_t, 1) - wf_t \ ,$$

which takes into account whether or not a protest occurs, $r = 1$, or $r = 0$, as well as f, the intensity with which the leader fights at step 2.

With democracy $D_t = 1$, we can write the leader's payoff as

$$\widetilde{U}(x_t, f_t) = u^I(x_t, 1) - wf_t, \tag{4}$$

Moreover, Appendix B shows that we get a similar structure in a model with strategic socialization along the lines of Bisin and Verdier (2001).

[19]It would be straightforward – and lead to the same qualitative results – to assume instead that parents look to the future and based their socialization on μ_{t+1}.

which takes into account the fact that no protest occurs in this case.

Choice of f – **step 2**   Since fighting is costly, $f_t = 0$ if $D_t = 1$. We can therefore write the equilibrium expected payoff to democracy as

$$U(x_t) = \text{Max}_f \widetilde{U}(x_t, f) = u^I(x_t, 1).$$

Under autocracy $D_t = 0$, the leader faces a trade-off as – even though it is costly – more fighting f decreases the probability of a successful protest. The equilibrium expected payoff to $D_t = 0$ is

$$V(\mu_t, x_t,) = \text{Max}_f E_r\{\widetilde{V}(x_t, \mu_t, f_t)\}. \tag{5}$$

Denote the optimal choice of fighting by $f^*(x, \mu)$. The envelope theorem implies that $V(\cdot)$ increases in x and decreases in μ. We assume that for all $x \in [\underline{x}, \overline{x}]$

$$\lim_{\mu \to 1} \left[w f^*(x, \mu) - [1 - \rho p(\mu_t, f^*(x, \mu))] \times [u^I(x, 0) - u^I(x, 1)] \right] > 0. \tag{6}$$

By (6) and (2), $dp(\mu, f^*(\mu, x)) / d\mu > 0$ – i.e., a larger share of concerned raises the probability of a successful protest even though the incumbent is fighting optimally.

Choice of institutions – **step 2**   To choose D_t at step 2, the incumbent leader compares $V(x_t, \mu_t)$ with $U(x_t)$, given the realization of x_t, and the share of concerned μ_t. The equilibrium choice of D_t satisfies:[20]

Proposition 1 *There exist* $\{\mu^L, \mu^H\}$ *with* $\mu^L < \mu^H$ *such that for*

1. $\mu \leq \mu^L$, $D(\mu, x) = 0$ *for all* $x \in [\underline{x}, \overline{x}]$;

2. $\mu \geq \mu^H$, $D(\mu, x) = 1$ *for all* $x \in [\underline{x}, \overline{x}]$ *and*

3. $\mu \in [\mu^L, \mu^H]$ *there exists* $\widehat{x}(\mu) \in [\underline{x}, \overline{x}]$ *such that* $D(\mu, x) = 0$ *if and only* $x \geq \widehat{x}(\mu)$.

The result makes intuitive sense. With weak democratic values (low μ), protesters are unlikely to win and the incumbent leader can safely choose autocracy $D_t = 0$ while spending little on fighting. When democratic values are strong, the leader is likely to lose any protest and fighting is costly so citizens get democracy. In these extreme cases, this holds independently of x_t. However, in an intermediate region, the institutional choice depends on the realization of x_t; with high x, the leader stays with autocracy with low x he picks democracy.

---

[20]The proof of this and all subsequent results appears in the Appendix.

Evolution of values – step 5 The evolution of democratic values is governed by the relative fitness of being concerned rather than passive as determined by expected utilities perceived at date t. The material payoffs of passive and concerned citizens are the same and hence cancel out. Hence all that matters is (1), i.e., the society-wide component of utility for concerned citizens. To calculate this, let $\gamma(x) = \sum_{m=1}^{M} [u^m(x,1) - u^m(x,0)]$ be the expected gain when $D_t = 1$. Then

$$S(x, D_t) = \begin{cases} \gamma(x) & \text{if } D_t = 1 \\ -\gamma(x) & \text{if } D_t = 0. \end{cases} \tag{7}$$

Finally, let

$$\lambda(\mu, x) = [1 - \rho p(\mu, f^*(\mu, x))] \gamma(x) \tag{8}$$

be the expected loss when $D_t = 0$, taking into consideration the probability of a successful protest.

Cultural dynamics We know from (3) that $\mu_{t+1} - \mu_t$ is positive (negative) when $\Delta(\mu_t)$ is positive (negative). Using (8) and Proposition 1, and recalling that x has c.d.f. H, we can write the expression for $\Delta(\mu_t)$ as:

$$\Delta(\mu) = \begin{cases} \int_{\underline{x}}^{\overline{x}} \gamma(x)\, dH(x) & \mu \geq \mu^H \\ \int_{\underline{x}}^{\widehat{x}(\mu)} \gamma(x)\, dH(x) - \int_{\widehat{x}(\mu)}^{\overline{x}} \lambda(\mu, x)\, dH(x) & \mu \in [\mu^L, \mu^H] \\ -\int_{\underline{x}}^{\overline{x}} \lambda(\mu, x)\, dH(x) & \mu \leq \mu^L. \end{cases} \tag{9}$$

There are three cases in (9). In the top row, democratic values have evolved to a point where incumbents always choose democracy $D_t = 1$ and there are no protests. The concerned have an intrinsic gain from this institution, so their share is growing over time. In the bottom row, the incumbent group get its preferred autocracy $D_t = 0$ for any realization of x and the few concerned individuals feel a perpetual sense of injustice, which gives them an intrinsic loss. As a result, democratic values are shrinking. In the middle row, the realization of x matters for the incumbent's institutional choice. Recalling Proposition 1 and (7), a gain only occurs if $D_t = 1$ which requires $x \leq \widehat{x}(\mu)$. Otherwise, incumbent leaders choose $D_t = 0$, which leads to losses as defined in (8). Democratic values are growing or shrinking depending on whether expected gains outweigh expected losses, which in turn depends on the extent to which x is expected to fall short of threshold $\widehat{x}(\mu)$ (according to distribution H).

From (8), the loss from being a concerned citizen is higher when x is high and the probability of unseating the incumbent in a protest is low which will be the case when μ is low, since $p(\mu, f^*(\mu, y))$ is then close to zero. At the other extreme, the loss is low when the incumbent is almost certain to lose a rebellion as $p(\mu, f^*(\mu, y))$ is close to one which occurs when μ is high. It is also straightforward

to check that $\Delta_\mu (\mu) \geq 0$ for all $\mu \in [0, 1]$ after observing that

$$\frac{\partial \widehat{x} (\mu)}{\partial \mu} = \frac{f^* (\mu, \widehat{x} (\mu, q))}{\frac{\partial p(\mu, f^* (\mu, \widehat{x}(\mu,q)))}{\partial \mu} \widehat{x} (\mu, q)} > 0.$$

Thus the model uncovers a complementarity between having more democratic values and the expected utility of holding such values, which shapes the cultural dynamics.

Steady states and equilibrium inertia The possible steady states are described in the following result:

Proposition 2 *There exists a critical value $\hat{\mu}$ defined by*

$$\int_{\underline{x}}^{\widehat{x}(\hat{\mu})} \gamma (x) \, dH (x) = \int_{\widehat{x}(\hat{\mu})}^{\overline{x}} \lambda (x, \hat{\mu}) \, dH (x) .$$

Whenever $\mu_0 \geq \hat{\mu}$, the polity converges to $\mu = 1$. However, for $\mu < \hat{\mu}$, the polity converges to $\mu = 0$.

To see why this is true, note that $\Delta (0) < 0$ and $\Delta (1) > 0$. Because $\Delta (\mu)$ is (weakly) monotonically increasing, there must exist a unique level $\hat{\mu}$ such that $\Delta (\hat{\mu}) = 0$. Moreover, this interior point is unstable, meaning that the dynamics described in (3) will converge slowly to either of two extremes.

This convergence of democratic values is associated with a specific path of democratic institutions. Once democratic values on an upward path have reached region $\mu \geq \mu^H$, democracy becomes permanently chosen. Equally, once democratic values on a downward path have reached the region where $\mu \leq \mu^L$, autocracy becomes the permanent choice. In the middle region for democratic values, however, we see reforms in both directions depending on the realization of shocks x_t.

Democratic institutions are naturally sticky in this framework even without assuming that commitment is possible to that each incumbents is free to reform in any period. This inertia reflects slow-moving democratic values which feed back to the motives for democratic reform. There is also a feedback effect in the other direction, i.e. from democratic institutions to democratic values. As we make clear below, it this property of the model which helps us to understand the patterns in the data which we discussed in Section 3.

5   The Model at Work

In this section, we apply the canonical model to the specific institutions discussed in Section 3. The two first subsections deal separately with executive constraints and executive recruitment, while the third deals jointly with both institutions.

5.1 Executive Constraints

Let $X_t \in \{0,1\}$ denote weak and strong executive constraints in place of generic variable D_t in the canonical model. The role of the political process is to distribute a revenue y_t in each period. Revenues come from (labor) taxation at rate τ and the value of natural resources n_t

$$y_t = [\tau w + n_t],$$

where n_t is now the specific stochastic variable taking the place of x_t in the canonical model (with c.d.f. $K(n)$). Under weak executive constraints, $X_t = 0$, the incumbent leader just freely picks any allocation. But if executive constraints are strong, $X_t = 1$, the representatives of the citizens must approve the allocation.

Legislative bargaining To model strong constraints, we assume a simple two-stage legislative bargaining. At stage 1, the leader proposes a set of shares $S_t = \{s_{i,t}, .., s_{M,t}\}$. Then the legislature votes with majority rule. If a majority accepts the proposal, it is implemented. If not, another group is picked at random to make a stage-2 proposal.

It is natural to think about the vote on the stage-1 allocation as a vote of confidence on the incumbent's proposal. But who gets to make a new proposal at stage 2 if that vote fails? We assume that this part of the institution is parametrically given. Specifically, incumbent group I is picked again with zero probability, while every other group is picked with probability $1/M$. The new agenda setter proposes a new allocation. If this is not accepted, a (breakdown) default outcome gives nothing to everybody.[21]

Payoffs If executive constraints are weak $(X_t = 0)$, the incumbent group just grabs the whole pie, so that:

$$u^I(n_t, 0) = \tau w + n_t \quad \text{and} \quad u^m(n_t, 0) = 0 \text{ for } m = 1, ...M.$$

If constraints are strong $(X_t = 1)$, it is straightforward to solve for the legislative-bargaining outcome. At stage 2, any group will accept any small amount to avoid the bad default allocation; so can set this amount to zero. The expected payoff for groups $1, .., M+1$, from rejecting the stage-1 proposal is thus $1/M$. With this continuation value, the stage-1 leader from group i needs only offer $s_{m,t} = 1/M$ to $M/2$ groups to get his proposal approved. The share of y_t captured by the leader's group, $s_{i,t}$, is therefore:

$$u^I(n_t, 1) = \frac{1}{2}[\tau w + n_t] - wf_t.$$

---

[21] This default allocation replaces the usaual assumption in the Baron-Ferejohn (1989) framework of infinite-horizon bargaining with a shrinking pie, but gives similar results. In the previous version of the paper, we assumed that the incumbent is recognized in the second round with exogenous probability q. The simplicifaction that $q = 0$ yields similar results, but saves on notational complexity.

For citizens in groups $m = 1, .., M$, material payoffs are:

$$u^m(n_t, 1) = \begin{cases} \frac{1}{M}[\tau w + n_t] & \text{with probability } 1/2 \\ 0 & \text{with probability } 1/2. \end{cases}$$

Using this, optimal fighting by the incumbent solves:

$$-\rho p_f(\mu, f^*(\mu, n))\frac{[\tau w + n]}{2} = w,$$

assuming an interior solution.

Equilibrium executive constraints The choice of executive constraints hinges on $\tau + n_t/w$. All else equal, positive resource shocks makes it less likely that strong executive constraints are chosen, as they enhance the incentive to maintain power. Following Proposition 1, μ can fall in three ranges: constraints are consolidated ($X_t = 1$) above μ_X^H, are not implemented at all ($X_t = 0$) for $\mu < \mu_X^L$, and dependent on the realization of n_t for $\mu \in [\mu_X^L, \mu_X^H]$.

If $\mu \in [\mu_X^L, \mu_X^H]$, define $\hat{n}(\mu)$ from:

$$\left[\frac{1}{2} - \rho p(\mu, f^*(\mu; \hat{n}(\mu)))\right][\tau w + \hat{n}(\mu)] = w f^*(\mu; \hat{n}(\mu)).$$

For $n \geq \hat{n}(\mu)$, executive constraints are weak ($X_t = 0$).

Equilibrium dynamics To apply Proposition 2, note that the gains and losses for concerned citizens are now

$$\gamma(n) = \frac{1}{2}[\tau w + n] \quad \text{and} \quad \lambda(\mu, n) = [1 - \rho p(\mu, f)]\frac{1}{2}[\tau w + n].$$

We can then construct $\Delta(\mu)$ according to equation (9). Moreover, it is easy to verify that $\Delta_\mu(\mu) \geq 0$. Applying Proposition 2, a society with an initially low value of μ will never evolve high enough democratic values (share of concerned citizens) to support strong executive constraints.

The critical level of democratic values, $\hat{\mu}_X$, from Proposition 2 is implicitly defined by

$$\int_{\hat{n}(\hat{\mu}_X)}^{\bar{n}} (\tau w + n)\, dK(n) = \int_{\underline{n}}^{\hat{n}(\hat{\mu}_X)} [1 - \rho p(\hat{\mu}_X, f^*(\hat{\mu}; n))](\tau w + n)\, dK(n). \tag{10}$$

Assuming an initial distribution of μ across countries, we will observe some converging monotonically to strong democratic values and strong executive constraints, while others will converge to weak democratic values and weak executive constraints. For those countries starting out in the mid region for democratic values, we may see institutional reversals depending on the realization of stochastic

natural resource rents.

5.2 Executive Recruitment

The canonical model also straightforwardly applies to the openness of executive recruitment. The analog to institutions D_t is now denoted by $O_t \in \{0, 1\}$ where $O_t = 0$ is closed executive recruitment and $O_t = 1$ is open. In the latter case, the incumbent group participates in an open contest for power, like an election among representatives of all groups. As with executive constraints, the dynamic evolution of μ helps shape and smooth out the time path of executive-recruitment institutions.

In an open contest, an incumbent's probability of securing power is denoted by $z_t \in [\underline{z}, \overline{z}] \subset (0, 1)$. This is drawn from a distribution with c.d.f. $Q(z_t)$. Now z_t plays the role of x_t in the canonical model. If $O_t = 1$ and $z_t = 1/(M+1)$ executive recruitment becomes a fair lottery over groups, while if $z_t = \overline{z}$ the incumbent group is maximally popular. With $O_t = 0$, we assume that whichever group holds power at the end of $t-1$ remains in power in t.

Payoffs with exogenous executive constraints We now suppose that strong executive constraints are in place with exogenous and time-invariant probability $X \in [0, 1]$ (which we endogenize in the next subsection). Let $u^i(O, z)$ be the payoff for the incumbent with executive recruitment O. Further, let $\bar{y} = \int_{\underline{x}}^{\overline{x}} [\tau w + n] dH(n)$ be expected public revenue at the time when O is chosen, which is before X is determined. Now we have

$$
\begin{aligned}
\tilde{u}(z) &= \bar{y} \left\{ z \left[1 - \frac{X}{2} \right] + (1-z) X \frac{1}{2M} \right\} \\
&= a + bz
\end{aligned}
\tag{11}
$$

where $a = \bar{y} \frac{X}{2M}$ and $b = \bar{y} \left[1 - \frac{X(M+1)}{2M} \right] > 0$. We can interpret b as the value of being the incumbent which is decreasing in X, but does not depend on n except through the value of public revenue. Thus, $\tilde{u}(\cdot)$ is an (affine) increasing function of z. We can then define payoffs $u^I(z, 1) = \tilde{u}(z)$ and $u^I(z, 0) = \tilde{u}(z)(1 - \rho p(\mu, f)), u^m(z, 1) = \tilde{u}\left(\frac{(1-z)}{M} \right)$, and $u^m(z, 0) = \tilde{u}(0)$. Optimal fighting now solves:

$$
-\rho p_f(\mu, f^*(\mu; b)) b = w.
$$

Given our assumptions, the probability of a successful protest if the incumbent chooses $O_t = 0$ is $p(\mu, f^*(\mu; b))$, which is decreasing in b.[22]

Equilibrium executive recruitment We can now apply Proposition 1; democratic values, μ, together with incumbent popularity, z, determine whether the incumbent picks open recruitment.

---

[22] Unlike executive constraints, this does not depend on the realization of the "shock", i.e. z in this case.

For $\mu \geq \mu_O^H$, we get $O_t = 1$ and for $\mu \leq \mu_O^L$, we get $O_t = 0$, for all levels of popularity. With democratic values between these thresholds, executive recruitment depends on the realization of z.[23]

For the case when $\mu \in [\mu_O^L, \mu_O^H]$, we can define threshold $\hat{z}(\mu; b)$, below which realizations of z lead to closed executive recruitment:

$$\{[1 - \rho p(\mu, f^*(\mu; b))] - \hat{z}(\mu; b)\} b = w f^*(\mu; b).$$

Observe that $\hat{z}(\mu; b) < [1 - \rho p(\mu, f^*(\mu; b))]$, i.e., an incumbent only picks closed executive recruitment when he has a strictly lower probability of staying in power under open recruitment than if he chooses to fight the protest by concerned citizens if he picks closed recruitment. Note also that $\hat{z}(\mu; b)$ is increasing in b and decreasing in μ – i.e., an incumbent group is less likely to take a bet on relinquishing power when holding power is more valuable or when there are fewer citizens in the population.

Equilibrium dynamics   The gains and losses for concerned citizens are now given by

$$\gamma(z) = (1 - z) b \text{ and } \lambda(\mu, z) = [1 - \rho p(\mu, f^*(\mu; b))](1 - z) b.$$

The apparatus of the canonical model applies, and $\Delta_\mu(\mu) \geq 0$ as required for Proposition 2.[24]   The critical level of democratic values $\hat{\mu}_O$ from Proposition 2 is now implicitly defined by

$$\int_{\hat{z}(\hat{\mu}_O; b)}^{\bar{z}} (1 - z) dQ(z) = [1 - \rho p(\hat{\mu}_O, f^*(\hat{\mu}; b))] \int_{\underline{z}}^{\hat{z}(\hat{\mu}_O; b)} (1 - z) dQ(z) \tag{12}$$

This depends on the expected value of holding power, b, only through its effect on the critical threshold $\hat{z}(\mu; b)$ and its impact on fighting. By analogy with Subsection 5.1, the polity converges either to fully democratic values and open executive recruitment, or to absent democratic values and closed executive recruitment, when initial democratic values are above or below the critical level $\hat{\mu}_O$.

5.3   Executive Constraints *and* Recruitment

We now extend the model to include both dimensions of democratic institutions. As we argued in Section 3, an essential – but mostly ignored – aspect of democratic history is that the openness of executive recruitment and the strength of executive constraints do not move perfectly together. Therefore, we study a process with a two-dimensional sequential choice.[25]

Each period begins with an interim leader from generation t from the ruling group at $t - 1$.

---

[23]Thresholds μ_X^L, μ_X^H in the previous subsection are generally different from thresholds μ_O^L, μ_O^H, as the former depend on the distribution of n while the latter depend on the distribution of z.

[24]To see this recall that $\hat{z}(\mu; b)$ is decreasing in μ and $p(\mu, f^*(\mu, b))$ is increasing in μ.

[25]Details are in Appendix D.

The choice of executive recruitment follows the model in the previous subsection, subsequent to the popularity shock z_t. This institutional decision is taken at the same time as committing resources to fighting protests which is now denoted by F_t.

Following this – and with an incumbent leader in office – the revenue shock, n_t, is realized and the choice of executive constraints follows with resources committed to fighting over executive constraints given by f_t. Once X_t is determined, payoffs are realized, with a new generation being born and socialized by their parents.

In this sequential approach, each shock is thus realized in connection with the institutional choice to which it applies (z to O, and n to X). As before, we assume that a generation cannot commit the next generation to any institutional arrangements. Neither do we assume any within-period commitment: each one of the two institutional dimensions has to be incentive compatible.

Payoffs   One key difference in this sequential model is that the function $u^i(O, z)$ – which played a key role in the last subsection – now depends on μ since the decision over executive constraints has yet to be made when openness is chosen. However, the sequential structure implies that the analysis of executive constraints is identical to the single-institution case; we can therefore summarize this choice using a function $X(\mu, n)$ which incorporates the institution-specific thresholds μ_X^L, μ_X^H. Given this function, we can focus on the choice of executive recruitment.

First, define

$$\tilde{u}(\mu, z) = a(\mu) + zb(\mu), \tag{13}$$

where $a(\mu) = \frac{1}{2M} \int_{\underline{n}}^{\overline{n}} [X(\mu, n) + [1 - X(\mu, n)] \rho p(\mu, f^*(\mu, n))] [\tau w + n] dK(n)$ and $b(\mu) = \int_{\underline{n}}^{\overline{n}} X(\mu, n) \frac{1}{2} [\tau w$ $[1 - X(\mu, n)] V(\mu, n) dK(n) - a(\mu)$. This function is the expected utility of an interim leader with survival probability z as she contemplates the choice of executive recruitment summarized in $X(\mu, n)$. Just as in the previous subsection, this function does not depend on n other than through expected revenue, which reflects the assumption that n is realized in connection with the choice of X. As for the components of this function, $a(\mu)$ is the utility of a group with no chance of being in power, while $b(\mu)$ is the expected gain from holding office. Comparing (13) and (11), it should be clear that a and b are now functions of μ since μ affects the probability of strong executive constraints being chosen.

With these preliminaries, we can specify the incumbent's payoffs from the choice of open recruitment as $u^I(\mu, z, O)$ with $u^I(\mu, z, 0) = \tilde{u}(\mu, z)$ and $u^I(\mu, z, 1) = \tilde{u}(\mu, 1 - \rho p(\mu, F))$. The material payoffs $u^m(\mu, z, O)$ of the other groups as a function of openness are now $u^m(\mu, z, 1) = \tilde{u}\left(\mu, \frac{1-z}{M}\right)$ and $u^m(\mu, z, 0) = \tilde{u}(\mu, 0)$, which are all linear in z.

Equilibrium choices Optimal fighting to quell protests against a lack of open executive recruitment in the sequential model solves

$$-\rho p_F \left(\mu, F^* \left(\mu; b \left(\mu \right) \right) \right) b \left(\mu \right) = w.$$

Our assumptions imply that $b \left(\mu \right)$ is decreasing, i.e., the gain from holding office is lower when the prospect of strong executive constraints is greater.[26] As before, higher democratic values μ therefore reduce fighting by the incumbent.

The logic of Subsection 5.2 can now be used. In particular, when μ is in the range $\mu \in \left[\mu_O^L, \mu_O^H \right]$ we can define a critical threshold $\hat{z} \left(\mu; b \right)$ from the equilibrium condition:

$$\left(\left[1 - \rho p \left(\mu, F^* \left(\mu; b \left(\mu \right) \right) \right) \right] - \hat{z} \left(\mu; b \left(\mu \right) \right) \right) b \left(\mu \right) = w F^* \left(\mu; b \left(\mu \right) \right).$$

It is immediate that $\hat{z} \left(\mu; b \left(\mu \right) \right) < \left[1 - \rho p \left(\mu, F^* \left(\mu; b \left(\mu \right) \right) \right) \right]$, since an incumbent will only choose closed executive recruitment when he has a strictly lower probability of staying in power with open recruitment than he would by choosing closed recruitment and fighting protestors. Note also that $\hat{z} \left(\mu; b \left(\mu \right) \right)$ is decreasing in μ. This represents an institutional complementarity between executive constraints and openness – a higher μ lowers the value of holding office for any group of citizens.

Mixed regimes and their drivers It is clear that the central insights derived earlier on the drivers of executive recruitment and executive constraints hold up in this two-dimensional setting. Studying the sequential model helps us to understand the mixed regimes and the timing of their reforms and therefore giving some insight into the facts tabled in Section 3. The most interesting mixed regime to pinpoint – given fact (iii) on institutions – is where executive recruitment is open and executive constraints are weak. Our model predicts that such mixed regimes come about due to a specific interplay between shocks z and n and institution-specific thresholds μ_O^L, μ_O^H and μ_X^L, μ_X^H.

A specific result on mixed regimes is (see Appendix D for details):

Corollary *All else equal, a polity is more likely to have a mixed regime with open (closed) executive recruitment and weak (strong) executive constraints, if:*

1. *the incumbent group is expected to be more (less) popular – a first-order stochastically dominating outward (inward) shift in z.*

2. *expected resource rents are higher (lower) – a first-order stochastically dominating outward (inward) shift in n.*

---

[26] To see this, observe that $V \left(\mu, \hat{n} \left(\mu \right) \right) = \frac{1}{2} \hat{n} \left(\mu \right)$ if there is an interior value of n above which weak executive constraints are chosen and $V \left(\mu, n \right)$ is decreasing in μ for all n such that the incumbent group chooses to fight to maintain weak executive constraints.

This corollary follows because higher (lower) resource rents make strong executive constraints less (more) likely to be chosen, while a more (less) popular incumbent make open executive recruitment more (less) likely to be chosen. We return to these mixed regimes when we discuss the data in the next section.

Cultural dynamics The exact institutional path that a polity will follow as μ changes is not clear a priori. However, democratic values are a common driver and underpin the observed positive correlation between institutional dimensions found in the data. Moreover, the theory identifies a clear institutional complementarity since strong executive constraints makes fighting over open recruitment less attractive.

To study the dynamics of democratic values in a model with two dimensions of institutions, we invoke two reference points – one for gains and one for losses. For gains, the reference point is all-out autocracy $X_t = O_t = 0$ and for losses, it is all-out democracy $X_t = O_t = 1$. This implies that, for intermediate cases with mixed regimes, there can be simultaneous gains and losses so that the *net* gain or loss shapes the cultural dynamics.

Appendix D analyzes these gains and losses in detail and shows that if $\underline{z} \geq 1/(M+1)$, then $\Delta(\mu)$ is still increasing, implying that there is a complementarity similar to the one that drives the value dynamics in Proposition 2.[27] As democratic values change over time, the polity thus converges either to fully democratic values and all-out democracy, or to absent democratic values and all-out autocracy. Which case obtains depends on whether initial democratic values are above or below a critical level $\hat{\mu}$.

6 Interpreting the Evidence

In this section, we discuss how the model(s) in Section 5 can help us interpret the facts in Section 3 on democratic institutions and values. It also motivates some additional empirical observations as well as allowing us to interpret existing empirical regularities uncovered in previous research.

The specific models of executive constraints and recruitment suggest that shocks to resources available to government (n) and to executive survival prospects (z) should affect institutional choices for a given starting point (μ_0). Moreover, the dynamics of democratic values (μ) give some smoothness to the institutional choices, except in a middle region for democratic values. Finally, the model suggests the possibility of multiple steady states.

---

[27]The additional assumption is needed because subordinate citizens might prefer weak executive constraints, conditional on open executive recruitment, as this gives them a chance of sharing in the "plunder" if they are selected into office. If $M = 1$, the single subordinate group will be in power for sure if the incumbent loses. But this is unlikely for large M, which (given the survival probability of the leader) tends to rule out this case and guarantee the complementarity that delivers the core result. Even if the condition $\underline{z} \geq 1/(M+1)$ were not to hold, there is always a critical value of μ above which $X_t = O_t = 1$ is consolidated and another (lower) critical value of μ below which $X_t = O_t = 0$ is consolidated.

6.1 Institutional Dynamics

Tables 1 and 2 documented three groups of country histories, namely permanent transitions into democracy or autocracy, and flip-flopping between the two. These three groups correspond pretty closely to the combined prediction from Propositions 1 and 2, namely an upper (lower) region for democratic values where democracy (autocracy) becomes an absorbing state, and an intermediate range where we may observe reforms in both directions. The model emphasizes that responses to shocks are likely to be heterogeneous depending on the value of μ.

Observing strong executive constraints and/or open executive recruitment does indicate that a polity is on a path towards consolidation. But the model suggests why this path need not be deterministic. Even a country with democratic values μ below its cutoff level $\hat{\mu}$ could have fleeting episodes with strong executive constraints, as with Ethiopia and Liberia in Table 1, along a downward trajectory – e.g., due to negative *temporary* resource shocks to public revenue n. Looking at Table 2, temporary shocks to leader survival chances z seem less likely to generate a permanent downward spiral. This is perhaps not surprising given that leaders have finite lives and orderly transitions within groups are hazardous. Only the most entrenched of monarchies are likely to be viable as permanently closed.

Critical junctures The model emphasizes that *permanent* shocks can matter too. In fact, our model underpins the important idea championed by Acemoglu and Robinson (2012) that there are crucial moments – so-called critical junctures – in national political history which have long-run consequences for better or worse. To see this, recall the critical and country-specific threshold value $\hat{\mu}$ for the long-run dynamics in Proposition 2, which applies to both executive recruitment and executive constraints. Countries with levels of μ just above and below their specific $\hat{\mu}$ can have radically different trajectories. Moreover, a permanent shock at such a juncture, say a shift of the whole distribution for resource revenue n,can flip a country to the opposite side of $\hat{\mu}$. In this case, the heterogeneous effects of shocks are particularly dramatic. As Proposition 2 shows, such a shift would have major long-run consequences for its democratic values and institutions. As we discuss further below, this insight may be particularly relevant in the case of natural resource discoveries or in creating a legacy from colonial domination.

Values (μ) as a force for consolidation According to the model, a lasting reform which puts strong executive constraints in place leads to a gradual building of democratic values, which helps to consolidate democratic reforms. This fits the slow-moving history of England where Parliament gained power over the executive from the late 17th century, at a time when elites still had firm control over Parliamentary representation. The monarch gradually became more constrained, while the ruling elite (the aristocracy) remained securely in control until the 19th century reform acts. But by then democratic values may have been entrenched enough that the risk of reversion was low.

22

At the other end of the spectrum, strong parliamentary powers early on in a country's history before democratic values have been built may be harder to sustain. For instance, some of the post-colonial regimes in Africa – in Nigeria, Sudan, Somalia, and Uganda – started out with strong executive constraints with substantial powers to the legislature. But these were repealed within a decade, perhaps because a lack of democratic values made it hard to muster protests against these reform reversals. Colonialism may also have played a role in eroding democratic values as we discuss below.

Roots of μ If democratic values are so important for institutional reform, then understanding why countries would have different starting points for μ becomes a central issue. In the next section, we emphasize the importance of foreign influence. However, some differences are more home-grown. The European enlightenment transformed thinking on representative government. The ideas of John Locke (1690), Charles de Montesquieu (1748), and Thomas Paine (1776), which were considered radical in their day, influenced the US Founding Fathers, and challenged ruling elites in other parts of the world. As emphasized in Siedentop (2014), Christian teaching also played a key role in evolving European democratic thought. On top of this, universities and other institutions of learning became more prominent along with higher education and exposure to liberal democratic ideas.

Even though democratic values were initially scant throughout the world, these new ideas may awakened democratic values in some other locations (such as new nations in Latin America). Our model suggests that once the democratic genie is out of the bottle and the threshold $\hat{\mu}$ is crossed, democratic reform will eventually take off.

A less novel force, but perhaps equally important in history, is the role of economic development. In our model, this occurs via higher wages w, which raise the opportunity cost of fighting, and also (for given tax rates) reduce the share of income from natural resources.[28] As such changes lower the critical threshold $\hat{\mu}$, sustained and wage-increasing growth could also trigger a path towards democratic reform. There is also a complementarity between development and democratic values, as emphasized by modernization theorists like Lipset (1959). However, in our model values coevolve with development and may be subject to a critical take-off point.

Executive constraints and n Proposition 2 suggests that countries with high spoils from office may choose weak executive constraints. This regime could be quite stable (for given μ) if governments are willing to quash protests, and have the capacity to do so, as in the stable but undemocratic oil kingdoms of the Gulf. Equally, high-income countries where resource rents are small tend to spawn strong executive constraints (for given μ). These are examples of high average n.

Structural change has created many economies less dependent on rents from land and primary

---

[28]It would be interesting, following Besley and Persson (2009), to make fiscal capacity endogenous, in this case the determination of τ.

commodities. Our model predicts that such change is conducive to building stronger executive constraints. This reminds of the standard view that resource dependence fosters violence, especially absent good governance (Mehlum, Moene, and Torvik 2006). However, the mechanism is novel. All else equal, citizens in resource-rich countries face higher prospective losses and are thus more resentful of bad governance. This reduces the relative fitness of concerned types, while passive citizens accept their fate more readily. While concerned citizens will protest against the infringements on democracy, they will often be unsuccessful, partly due to repression by incumbent leaders to protect the resources captured by his group. According to our model, the state-dependent correlation of violence and bad governance may reflect a joint dependence of bad institutions and violence on resources, rather than bad institutions mapping resource dependence into violence.

Executive recruitment and z Slow-changing networks which regulate access to power – such as aristocratic families and social ties keeping an elite in power – are an important part of the political landscape in many countries. Proposition 1 suggests a role for shocks to these networks that permanently alter the security of leaders. A shift in the ruling group's grip on power is like a leftward shift in the distribution of z. We would expect this to create a move towards open executive recruitment, but the timing will vary by country, depending on the nature of the elites and their social ties (as well as the value of μ at the time).

Thus, the stable oil-rich kingdoms of the Gulf illustrate one kind of experience of perpetually low openness in executive recruitment. The model explains this in terms of these countries having high average z, perpetuated by strong political family networks. It is no surprise that the nascent reform movements in these countries following the Arab Spring were met with brutal repression.

The historical move towards more open executive recruitment in Western Europe since 1800 can be thought of as reflecting decaying aristocratic power networks.[29] Our model suggests that (above some threshold), this would be reinforced by increasing values which increased the cost of repression for those elites that wished to stay in office. The purely strategic view of Acemoglu and Robinson (2006) sees the franchise reforms as reflecting shifting interests, but would attribute the secular change experienced across a range of countries to the rising importance of the working class rather than to emerging values. Obviously, these two forces are mutually reinforcing in our framework.

Temporary shocks to the popularity of leaders matter most for countries with democratic values in the range $\mu \in [\mu_O^L, \mu_O^H]$. Strong charisma or deep unpopularity of leaders may change the likelihood z that an incumbent group can sustain its power in an election. Such shocks may also influence long-run political development if they arrive at a juncture where μ is close to $\hat{\mu}_O$. Our model predicts that particularly entrenched and/or popular leaders may open up executive recriuitment, since they do not fear the consequences of open contests. Thus, a stable one-party system could actually encourage

---

[29]The French revolution may also have emboldened protest movements. But understanding this in our model would require an extension where the level of protest is endogenous given μ.

openness since this will not appear as a major threat.

Mixed regimes   Even though our model identifies democratic values μ as a common factor, institutional choices over the two dimensions of democracy need not be synchronized. The model (see the Corollary in Subsection 5.3) suggests separate drivers of each dimension, n and z, which will affect the dynamics of institutions when μ is in a middle range. The model leads us to expect some churning of institutions until they become consolidated in either direction. This churning is indicative of the reversing reforms seen in Tables 1 and 2 (as well as indirectly in Figures 3 and 4).

While there is no set pattern of institutional transitions, strong executive constraints more often tend to emerge after open executive recruitment (see Table 3). However, a number of old monarchies (mostly in Western Europe) introduced strong executive constraints before going to open executive recruitment. The Corollary in Subsection 5.3 suggests that this could have reflected a combination of weakly entrenched/unpopular rulers and meagre government resources (low values of z and n).

Table 3 also shows how new nations in Africa and Asia (and Latin America before them) introduced open executive recruitment, but stuck with weak executive constraints. According to the model, this might have reflected popular and/or charismatic leaders from ruling groups – often tied to liberation movements – along with considerable rents to distribute from plentiful natural resources (high values of z and n). Even if executive constraints lag behind, the growth in democratic values when μ begins above $\hat{\mu}$ will, in the end, bring about strong executive constraints. This has indeed been the pattern across a range of countries, as we see in Table 3.

6.2   Values

Our model focuses on the evolution of democratic values. Even though the data does not allow us to observe the long-run evolution of these values, we can still relate contemporary values to long-lived features of countries.

Trapped in autocracy?   Our model provides an underpinning for fact (ii) on democratic values in Section 3, namely that people in countries which have never or rarely transited into democratic institutions have low levels of democratic support. Weak democratic values can thus lead to an "autocracy trap". A case in point is Russia, with a very short history of democratic institutions and low democratic values according to WVS. Up to 1990, the Soviet state effectively repressed any nascent democratic reforms. In our model, this would weaken democratic values over time. Weak values in turn would undermine a democratic reform – such as that attempted by Boris Yeltsin – giving it little chance of becoming permanent. Instead, we would predict a reversal to weak executive constraints, especially in the wake of rising resource rents.

According to the model, changing Russia's trajectory would require changing the fundamentals or a favorable exogenous shock to values μ. For example, the entrenched autocratic regime could be

weakened by a lower willingness or capacity to repress, which would raise the prospective influence of citizens holding democratic values. Both could reflect a state collapse with y becoming persistently low. That said, one would also need a very high value of b, the value of incumbency in the model, for any chance of a switch into a democratic regime, once μ has fallen to a low value. Interestingly, Russia has maintained (semi) open executive recruitment, which makes sense when the incumbent leader is popular and has little to fear from elections. But a less popular successor to Putin could be tempted to shift towards a more closed form of executive recruitment.

A foundation for democratic capital   Some countries – including Sweden, UK and the US – have made once-and-for-all stable transitions to strong executive constraints and open recruitment. Our model would interpret these transitions as passing the critical threshold $\max\left\{\mu_X^H, \mu_O^H\right\}$ above which fully democratic institutions are sustained. This is consistent too with the observation in fact (ii) on values, namely that democratic support is strong in such countries. We documented this pattern in Figure 6 with the highest democratic values being present for countries with consolidated strong executive constraints and open executive recruitment (in the right-most columns of Tables 1 and 2).

Persson and Tabellini (2009) interpreted the data on institutional persistence in terms of "democratic capital". Our model suggests a theoretical foundation for this idea in terms of evolving democratic values as represented by μ. As these values also manifest themselves in longer experience with democratic institutions, past institutional patterns may not actually cause persistence but reflect an omitted driver of democratic values. Moreover, our model is one where causality is running in both directions, requiring care in interpreting the correlation between democratic values and institutions.

Values and resource rents   A distinctive aspect of our theory is that values and institutions coevolve. The model predicts that natural-resource rich economies require high democratic values to enter a path towards strong executive constraints. It also says that – absent such reforms – high resource rents will gradually erode democratic values. Moreover, the multiple steady states of our theory suggest that the timing of resource discoveries may be critical in a country's history depending on the value of μ at the time. A case in point is Norway, which already had an established history of strong executive constraints prior to its major gas and oil discovery. However, the oil-dependent Gulf states have no experience with strong executive constraints.

We used the World Bank website to code a list of countries with oil rents (revenues less production costs) exceeding 10% of GDP in 2006. Of the 27 such countries in our data only two – Trinidad and Tobago and Timor-Leste – had strong executive constraints in 2006. In the same year, however, 14 had open executive recruitment. This is consistent with our theory since high average n undermines executive constraints more than openness. However, our model also suggests that, due to the lack of

strong executive constraints, natural resources undermine democratic values. This makes it harder to sustain open elections and if values fall below threshold $\hat{\mu}$, chances are low of a reform towards all-out democracy.

To investigate the link to values further, we create a dummy for the 13 of the 27 countries with oil rents above 10% of GDP in 2006 which also appear in the WVS. These are Algeria, Azerbaijan, Ecuador, Iran, Iraq, Kazahkstan, Kuwait, Libya, Nigeria, Qatar, Russia, Trinidad and Tobago, and Yemen. Around 55% of citizens in these oil rich countries give a score of 9 or 10 when asked to express their support for democracy, compared to around 61% for citizens in other countries in the sample.

To check whether this raw difference in support for democracy is due to omitted personal or country characteristics, columns (1) and (2) in Table 4 look at the correlation between natural resources and democratic support in WVS waves 5 and 6. They report estimates from a linear probability model where the dependent variable is our dummy for strong democratic support (standard errors clustered at the country level). We control for gender, ten dummies for income, three dummies for education, and three age bands. We also hold constant a wave dummy and three country-level variables: human capital from Barro-Lee, income per capita from the Penn World Tables, and the history of strong executive constraints from PIV.

For the countries in Table 1 with WVS data, column (1) shows a significant negative correlation between oil-rent dependence and support for democracy. Specifically, living in a resource-dependent country is associated with a 0.17 (more than a quarter of the sample mean) lower probability of supporting democracy. Column (2) shows that the correlation is weaker for all countries in PIV, although we cannot reject the correlation being the same in the two columns. While it is well-known that oil-rich countries are less likely to have strong executive constraints (which we hold constant), it is less well-known that they also exhibit low levels of democratic support *ceteris paribus*. Our theory links these correlations together through the mechanism of the model.

7   Foreign Influence

Our framework can be extended to consider international interdependencies and how these influence institutions and values. While such issues are often discussed, we are not aware of any existing framework to model them. We distinguish between two forms of foreign influence: direct and indirect.

7.1   Direct Influence

To study this, we revert to the single-institution $D_t \in \{0, 1\}$ as in the canonical model of Section 4. We first consider colonialism and/or military occupation. Here, a foreign power is in fact the incumbent and decides over both D_t and f_t. We then consider foreign military influence, such as

the Soviet Union dominating Eastern Europe during the Cold War, or Western powers supplying weapons to Arab-state rulers. Here, D_t is chosen by a domestic incumbent, but the foreign power can put resources into fighting protest which may lower the probability of a successful protest.

Colonialism Colonial powers directly established governance rules over a polity. In most cases, this involved weak executive constraints and closed executive recruitment. However, UK colonies such as Australia, Canada, New Zealand and South Africa were given elements of democratic institutions under colonialism. Acemoglu, Johnson and Robinson (2001) emphasized this heterogeneity in treatment using the distinction between extractive and inclusive institutions, which can be interpreted as different values of D_t in our model. In line with our model, colonial powers often faced citizen resistance from subjugated populations – as the Kenyan Mau-Mau uprising or the Indian Mutiny – which required fighting to preserve the status quo. A colonial regime with sufficient powers was therefore able to enforce institutions ($D_t = 0$) with sufficient investment in fighting, so that short-term shocks to x_t did not lead to pro-democratic reforms.

The model allows us to think about the longer-term impact of colonialism and hence cast light on the empirical finding in Acemoglu, Johnson and Robinson (2001) that has received widespread attention. By maintaining $D_t = 0$ (extractive institutions) or $D_t = 1$ (inclusive institutions), colonialism may have a permanent effect on post-colonial democratic institutions by inhibiting or promoting the emergence of democratic values.[30] Countries whose values were repressed to keep them below $\hat{\mu}$ would potentially face long-run effects of colonialism beyond any initial efforts to bring in democratic reforms after independence. Thus, the model is consistent with the historic abandonment of parliamentary democracy is large parts of the post-colonial world. Subsection 7.3 provides some further evidence on this channel.

Occupation The same basic idea applies to the subjugation of a country following an external conflict. The experience of Europe during World War II, or the influence of the USSR during the cold war, are interesting examples in history of repressed democratic institutions. On the other side of the cold war, we have the Japanese and West German examples, where victorious Western allies imposed democratic institutions. The fact that the US would not tolerate a descent back into autocracy would provide a boost to democratic values according to our framework. Our model suggests, however, that such intervention may be a hazardous exercise. While political institutions can be imposed to some degree, democratic values cannot easily be controlled and this can create an unstable situation until values and institutions are aligned. Like colonialism, occupations can potentially have long-run effects through the evolution of democratic values. Subsection 7.3 shows some correlations in line with such an effect.

---

[30]In fact, Acemoglu, Johnson and Robinson use strong executive constraints as their dependent variable in the post-colonial era.

Military support   Turning to influence without direct rule, a foreign power may also give military support f^R with the aim of strengthening a domestic ruling elite's capacity to fight rebellions. This could be modeled as the probability of successful protest becoming $p\left(\mu, \theta f^R + f\right)$, where a parameter value $\theta > 1$ would capture a more efficient fighting technology by a foreign power. In our model, this might affect D_t by lowering the cost of choosing $D_t = 0$. Examples include the puppet regime in Vichy France, or Vidkun Quisling's leadership in Norway during World War II. Such military support may also affect the time path of democratic values, potentially with long-term consequences depending on the length of the regime in question.

Yet another form of direct influence could involve a foreign power changing the distribution of shocks which a country faces or their consequences. Aid building the economy can raise w, with an indirect effect on institutions. The same would be true of foreign responses to economic and political shocks. For example, foreign governments are sometimes accused of intervening to enhance the survival prospects of specific leaders.

7.2   Indirect Influence

Other forms of foreign influence can be thought of as acting directly on μ. In principle, the canonical model could be extended to allow values to be influenced by events in other countries, creating cross-border spillovers which may reinforce the patterns in the data that we have already discussed.

Migration   One example to illustrate this idea is to consider the impact of emigration which brings democratic or non-democratic values to a new location. Proposition 2 says that such emigration will mostly have a temporary effect by shifting the value of μ in the recipient country. But it could also have a long-run effect if it is on a sufficient scale to push democratic values above or below $\hat{\mu}$ – e.g., a large enough influx of concerned citizens could help foster a shift from autocracy to democracy. In this spirit, Bandiera et al (2016) discuss an example of interplay between immigrant values and institutions in 19th-century US states.

Migration may also be endogenous to democratic institutions. Thus, concerned citizens in an autocracy may chose to leave rather than await a challenge the ruling elite – to paraphrase Hirschman (1970), they use "exit" rather than "voice". This would decrease democratic values in autocracies and increase them in democracies which would speed up the dynamics compared to a single-polity model.

One could also imagine other external influences on μ such as that stressed in the empirical work by Spilimbergo (2009) who argues that foreign education may change democratic values enough to help promote democratic reforms once students return home.

Cultural spillovers   Another interesting mechanism of indirect influence is cross-border cultural influence. Persson and Tabellini (2009) and Acemoglu et al (2015) exploit the fact that democracy in

one country is systematically related to democracy in its surroundings. Our model can explain this by supposing that direct social interactions or media influence works across national borders. Our framework could incorporate such interactions between two countries A and B, by assuming that relative payoff fitness, entering into country-A democratic values, follow

$$\tilde{\Delta}\left(\mu_A, \mu_B\right) = \left(1 - \beta\right) \Delta^A\left(\mu_A\right) + \beta \Delta^B\left(\mu_B\right),$$

with spillover parameter β. If $\Delta^J\left(\mu_J\right)$ is increasing in both countries, this yields a two-way evolution of democratic values with cross-country complementarities: democratic values in one location reinforce those in other locations.

Cultural spillovers relate to discussions in the international-relations literature about "soft power", an idea pushed by Nye (2004):

> "(a) country may obtain the outcomes it wants in world politics because other countries – admiring its values, emulating its example, aspiring to its level of prosperity and openness – want to follow it. ... This soft power – getting others to want the outcomes that you want – co-opts people rather than coerces them."

Using media and imagery may increase parameter β. This provides a window on the large funding of foreign-news networks and other kinds of cultural influence by governments around the world. Our approach suggests that such activities may reinforce a direction of on-going change and potentially lead to quicker adoption of a stable form of institutions. But if they operate at critical junctures around $\hat{\mu}$, they could have a more fundamental effect on long-run trajectories.

7.3   Correlations of Values and Foreign Influence

In light of the discussion above, the second part of Table 4 correlates our measure of individual-level values with variables representing foreign influence.

Soviet post-WWII influence   One interesting episode is the long-term postwar influence of the USSR. Some independent countries – like the Baltic ones – were absorbed into the USSR, while others became satellite countries in Eastern Europe, such as Bulgaria, Hungary, Poland, and Romania. We would expect Soviet influence not just to have suppressed democratic institutions, but to have weakened democratic values. As in the first part of Table 4, we use waves 5 and 6 of the WVS microdata for different samples. Again we estimate a linear probability model, where strong individual democratic values is the left-hand side variable, and individual and country-level controls appear on the right-hand side together with the dummy for Soviet postwar influence. The countries in the WVS which we code as being Soviet influenced are Armenia, Azerbaijan, Bulgaria, Belarus, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Moldova, Poland, Russia, Ukraine and Uzbekistan. The

proportion of the population who grade their support for democracy as 9 or 10 in the Soviet influenced countries is 54% compared to 61% in other countries in the WVS.

Table 4 shows a significant negative correlation between Soviet influence and support for democracy.[31] The coefficient is particularly large in column (3) for the sample of countries in Table 1, but it continues to be significant in column (4) for a larger sample. Moreover, it holds up in column (5), when we also include the natural-resources dummy variable. The negative correlation suggests that long-term Soviet influence did erode democratic values, as our model predicts it should.[32]

German WWII occupation A second example is German World War II occupation. Some countries in the middle category of Table 1 had achieved strong executive constraints prior to the war and their only reversal of the democratic path is due to German influence or occupation. These countries are Austria, Belgium, Denmark, Netherlands and Norway. Through the lens of our model, their prewar strong executive constraints suggest that their μ was relatively high. We thus expect that they may have resumed their "natural state" of strong constraints once German influence ended, which is exactly what happened after the war. All else equal, the occupied countries should thus have stronger democratic values today than the other churning countries in Table 1.

To explore this, we create a dummy for countries that dropped strong executive constraints only due to World War II. We then regress the individual indicator for high democratic values on this dummy, also including the individual and country control variables as outline above. The estimate in column (6) of Table 4 is only for the churning sample in Table 1. Citizens in countries that reversed only due to German occupation have higher democratic vales. The result holds up in column (7) when we widen the sample to all countries with mixed histories in PIV.

This finding is consistent with our model if nations acquiring strong executive constraints prior to German occupation did so because of strong democratic values. If these values persisted through a relatively short interruption, the model would also predict the immediate return to strong constraints. This differentiates the occupied subset of countries from others which have experienced reversals. These are plausibly in the range $\mu \in \left[\mu_X^L, \mu_X^H\right]$ where values may still be building, or were reaching a consolidation point after the countries that suffered Nazi occupation.[33]

Settler mortality Finally, we consider the possibility that colonial institutions leave a long-run legacy for democratic values. As mentioned in Section 7.1, our approach suggests that values may

---

[31] As in other columns, they estimates are for a linear probability model where the left-hand side variable is our dummy for a score of 9 or 10 of democratic support (standard errors clustered at the country level). We control for a wave dummy, gender, ten dummies for income, three dummies for education, and three age bands plus three country-level variables: human capital from Barro-Lee, income per capita from the Penn World Tables, and the history of strong executive constraints from PIV.

[32] These findings are consistent with Neundorf (2010) who studies satisfaction with democracy in Eastern Europe between 1990 and 2003 and finds that respondents with a "socialist upbringing" tend to show less satisfaction.

[33] Although we have couched this discussion in terms of executive constraints, similar considerations should apply to openness.

help drive the finding by Acemoglu, Johnson and Robinson (2001) that colonial regimes – when instrumented by settler mortality – persistently affect institutions. We exploit the intersection between their data and waves 5 and 6 of the WVS (only 12 countries). Specifically, we include settler mortality in an individual-level regression for democratic values, along with the earlier battery of individual-level and country-level controls. The result in column (8) of Table 4 shows a negative and significant relationship between democratic values and settler mortality. Because we control for institutional history, this correlation suggests a persistent effect consistent with the key idea of Acemoglu et al (2001). The twist in this story supplied by our model is that persistence may arise (partly) via endogenous democratic values.

8   Concluding Comments

We have developed a model of the two-way interaction between values and institutions. The single state variable in this model is the proportion of citizens who hold strong enough democratic values that they are willing to defend democratic institutions. These concerned citizens resent nondemocratic institutions which in turn helps to propagate democratic values. Over time, this leads to a complementarity between values and institutions. Institutional change becomes a gradual process, not because incumbents can commit future incumbents, but because the future incumbents pay close attention to democratic values which themselves change gradually over time.

We argue that this model can cast light on the heterogeneity in country experience with institutional change. A key part of this heterogeneity is to highlight two dimensions of institutions: executive recruitment and executive constraints. Even if democratic values are a common driver of both, we are able to explain why their evolution need not be synchronized, as is the case in the data.

Our paper builds a bridge between the cultural and strategic approaches to institutional change. In our framework, it is natural to see democratization as a process where the formation of democratic values and experience with democratic institutions reinforce each other. These joint dynamics help us to better understand both persistence and changes in political institutions across countries and time.

The paper suggests a wider agenda. The model has a number of empirical implications which could be explored beyond simple correlations. Another interesting issue concerns the within-country inter-generational distribution of democratic values, conditional on the country's political history. It is also important to build more specific models of the process whereby democratic values change.

So far, there has been relatively little research on the co-determination of values and institutional rules. Models like the one developed in this paper could be deployed to study a range of related phenomena, such as the joint dynamics of organizational cultures and organizational forms.

References

[1] Acemoglu, Daron, Simon Johnson, and James Robinson [2001], "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, 1369-1401.

[2] Acemoglu Daron, Suresh Naidu, Pasqual Restrepo, and James Robinson [2015], "Democracy Does Cause Growth," Mimeo, MIT.

[3] Acemoglu, Daron and James Robinson [2000], "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective, " *Quarterly Journal of Economics* 115, 1167-1199.

[4] Acemoglu Daron and James Robinson [2006], *Economic Origins of Dictatorship and Democracy*, Cambridge, MA: Cambridge University Press.

[5] Acemoglu Daron and James Robinson [2012], *Why Nations Fail*, New York, NY: Crown Publishers.

[6] Acemoglu, Daron, James Robinson, and Ragnar Torvik [2011], "Why Do Voters Dismantle Checks and Balances?," mimeo, MIT.

[7] Aidt, Toke and Peter Jensen [2014], "Workers of the World, Unite! Franchise Extensions and the Threat of Revolution in Europe, 1820-1938," *European Economic Review* 72, 52-75.

[8] Almond, Gabriel and Sidney Verba [1963], *The Civic Culture: Political Attitudes and Democracy in Five Nations*, New York: Sage Publications.

[9] Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo [2016], "Nation-Building through Compulsory Schooling During the Age of Mass Migration," mimeo, LSE.

[10] Baron, David and John Ferejohn [1989], "Bargaining in Legislatures", *American Political Science Review* 83, 1181-1206.

[11] Benabou, Roland [2008], "Ideologies", Joseph Schumpeter Lecture, *Journal of the European Economic Association* 6, 321-352.

[12] Besley, Timothy [2015], "Aspirations and the Political Economy of Inequality," Hicks Lecture.

[13] Besley, Timothy and Torsten Persson [2009], "The Origins of State Capacity: Property Rights, Taxation and Politics", *American Economic Review*, 99, 1218-1244.

[14] Besley Timothy and Torsten Persson [2011], *Pillars of Prosperity: The Political Economics of Development Clusters*, Princeton, NJ: Princeton University Press.

[15] Besley Timothy, Torsten Persson, and Marta Reynal-Querol [2016], "Resilient Leaders and Institutional Reform: Theory and Evidence," *Economica* 83, 584-523.

[16] Bisin, Alberto and Thierry Verdier [2001], "The Economics of Cultural Transmission and the Dynamics of Preferences," *Journal of Economic Theory*, 97, 298–319.

[17] Bisin, Alberto and Thierry Verdier [2011], "The Economics of Cultural Transmission and Socialization," Chapter 9 in Benhabib, Jess (ed.), *Handbook of Social Economics*, Volume 1A, Elsevier.

[18] Boix, Carles [1999], "Setting the Rules of the Game: The Choice of Electoral Systems in Advanced Democracies," *American Political Science Review* 93, 609-624.

[19] Boyd, Robert and Peter Richerson [1985], *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.

[20] Cavalli-Sforza, Luigi L. and Marcus Feldman [1981], *Cultural Transmission and Evolution*, Princeton, NJ: Princeton University Press.

[21] Congleton, Roger D. [2007], "From Royal to Parliamentary Rule Without Revolution: The Economics of Constitutional Exchange within Divided Government," *European Journal of Political Economy* 23, 261-284.

[22] Diamond, Larry [1992], "Economic Development and Democracy Reconsidered," in *Reexamining Democracy: Essays in Honor of Seymour Martin Lipset*, Marks, Gary and Larry Diamond (eds.), Newbury Park, CA: Sage.

[23] Ellison, Glenn and Drew Fudenberg [1993], "Rules of Thumb for Social Learning," *Journal of Political Economy* 101, 612-643.

[24] Glaeser, Edward L, Rafael LaPorta, Florencio López-de-Silanes, and Andrei Shleifer [2004], "Do Institutions Cause Growth?." *Journal of Economic Growth* 9, 271-303.

[25] Gorodnichenko, Yuriy and Gerard Roland [2015], "Culture, Institutions, and Democratization," mimeo, UC Berkeley.

[26] Güth, Werner and Menahem E. Yaari [1992], "Explaining Reciprocal Behavior in Simple Strategic Games: An Evolutionary Approach," in Witt, Ulrich (ed.) *Explaining Process and Change: Approaches to Evolutionary Economics*, Ann Arbor, MI: University of Michigan Press.

[27] Hirschman, Albert [1970], *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States,* Cambridge, MA: Harvard University Press.

[28] Inglehart, Ronald [1997], *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*, Princeton, NJ: Princeton University Press.

[29] Inglehart, Ronald and Christian Welzel [2005], *Modernization, Cultural Change, and Democracy: The Human Development Sequence*, Cambridge, UK: Cambridge University Press.

[30] Kahneman, Daniel and Amos Tversky [1979], "Prospect Theory: An Analysis of Decision under Risk," *Econometrica* 47, 263–291.

[31] Kahneman, Daniel, Jack L. Knetsch, and Richard H. Thaler [1991], "The Endowment Effect, Loss Aversion, and Status Quo Bias: Anomalies," *Journal of Economic Perspectives*, 5, 193-206.

[32] Koszegi, Botond and Matthew Rabin [2006], "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121, 1133-1165.

[33] Kuran, Timur [1995], *Private Truths, Public Lies: The Social Consequences of Preference Falsification,* Cambridge, MA: Harvard University Press.

[34] Lagunoff, Roger [2001], "A Theory of Constitutional Standards and Civil Liberty," *Review of Economic Studies* 68, 109-132.

[35] Levy, Gilat and Ronny Razin [2016], "The Echo Chamber Effect: Segregation, Polarisation, and Discrimination," mimeo, LSE.

[36] Lipset, Seymour Martin [1959], "Some Social Requisites of Democracy: Economic Development and Political Legitimacy," *American Political Science Review* 53, 69-105.

[37] Lipset, Seymour Martin [1960], *Political Man*, London: Heinemann.

[38] Lizzeri, Alessandro, and Nicola Persico [2004], "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's Age of Reform," *Quarterly Journal of Economics* 119, 707-765.

[39] Locke, John [1988, 1690], *Two Treatises on Government*, Cambridge Texts in the History of Political Thought.

[40] Marshall, Monty G. and Keith Jaggers [2005], "POLITY IV Project: Political Regime Characteristics and Transitions, 1800-2004," <www.cidcm.umd.edu/polity>.

[41] Marx, Karl and Friedrich Engels [2004, 1848], *Manifesto of the Communist Party*, Marxist Internet Archive, <www.marxists.org/archive/marx/works/1848/communist-manifesto/index.htm>.

[42] Mehlum, Halvor, Karl Moene and Ragnar Torvik [2006], "Institutions and the Resource Curse," *Economic Journal* 116, 1–20.

[43] Montesquieu, Charles de [1989, 1748], *The Spirit of the Laws*, Cambridge, UK: Cambridge University Press.

[44] Moore, Barrington [1966], *Social Origins of Dictatorships and Democracy: Lord, Peasant in the Making of the Modern World*, Boston: Beacon Press.

[45] Mukand, Sharun and Dani Rodrik [2015], "The Political Economy of Liberal Democracy," NBER Working Paper No. 21540.

[46] Neundorf, Anja [2010], "Democracy in Transition: A Micro Perspective on System Change in Post-Socialist Societies," *Journal of Politics* 72, 1096-1108.

[47] North, Douglass [1983], *Structure and Change in Economic History*, W.W. Norton and Company (new edition).

[48] Nye, Joseph [2004], *Soft Power: The Means to Success in World Politics,* New York: Public Affairs.

[49] O'Donnell, Guillermo and Philippe Schmitter [1986], *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies*, Baltimore: Johns Hopkins University Press.

[50] Paine, Thomas [1986, 1776], *Common Sense*, New York: Penguin Classics

[51] Passarelli, Francesco and Guido Tabellini [2016], "Emotions and Political Unrest," forthcoming in *Journal of Political Economy.*

[52] Persson, Torsten and Guido Tabellini [2008], "Political Regimes and Economic Growth," in Helpman, Elhanan (ed.), *Institutions and Economic Performance*, Cambridge, MA: Harvard University Press.

[53] Persson, Torsten and Guido Tabellini [2009], "Democratic Capital: The Nexus of Economic and Political Change," *American Economic Journal Macroeconomics* 1, 88-126.

[54] Przeworski, Adam, Michael E. Alvarez, Jose Antonio Cheibub, and Fernando Limongi [2000], *Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990*, Cambridge: Cambridge University Press.

[55] Putnam, Robert [1993], *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton, NJ: Princeton University Press.

[56] Rokkan, Stein [1970], *Citizens, Elections, Politics. Approaches to the Comparative Study of the Processes of Development*, New York, NY: MaKay.

[57] Siedentop, Larry [2014], *Inventing the Individual: The Origins of Western Liberalism*, Allen Lane.

[58] Spilimbergo, Antonio [2009], "Democracy and Foreign Education," *American Economic Review* 99, 528-543.

[59] Ticchi, Davide, Thierry Verdier, and Andrea Vindigni [2013], "Democracy, Dictatorship and the Cultural Transmission of Political Values," IZA Discussion Paper, No. 7441.

[60] Weber, Max, [1968, 1922], *Economy and Society*, Berkeley: Univ. Calif. Press

[61] Weingast, Barry, [1997], "The Political Foundations of Democracy and the Rule of Law," *American Political Science Review*, 91(2), 245-263.

[62] Welzel, Christian [2007], "Are Levels of Democracy Affected by Mass Attitudes? Testing Attainment and Sustainment Effects on Democracy," *International Political Science Review*, 28, 397-424.

[63] Welzel, Christian, Ronald Inglehart and Stefan Kruse [2015], "Pitfalls in the Study of Democratization: Testing the Emancipatory Theory of Democracy," *British Journal of Political Science*, forthcoming.

Figure 1
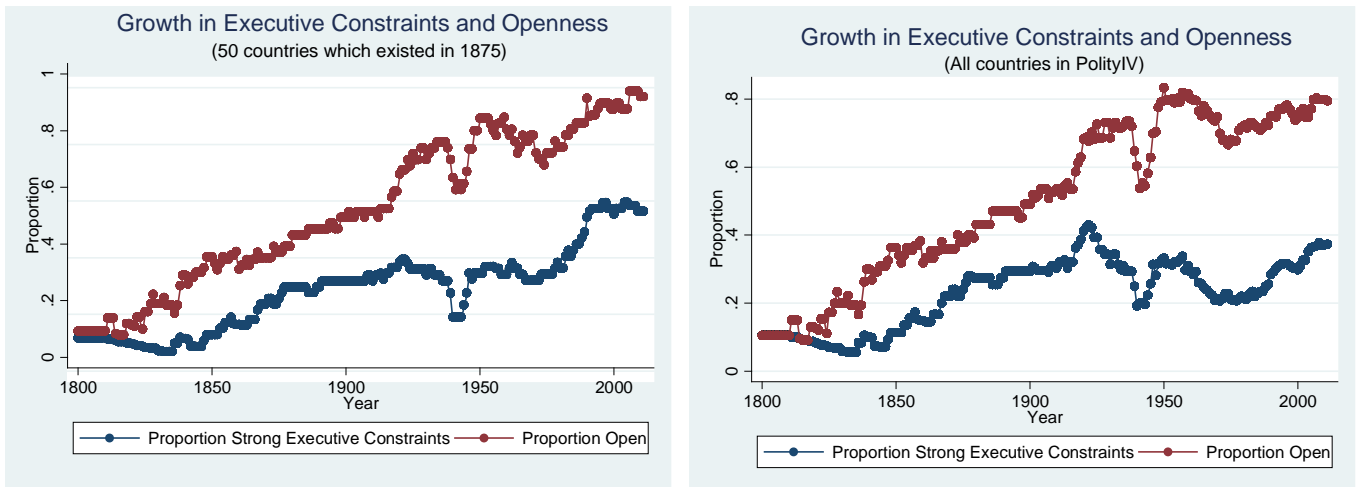
Notes: The data come from the Polity IV website http://www.systemicpeace.org/polityproject.html. For executive constraints, we use *xconst* which is measured by a dummy variable which is equal to 1 if the value of this variable in a given country-year is equal to 7; for openness, we *xropen* which is measured by a dummy variable which is equal to 1 in a given country-year is equal to 4. The left-hand panel is for all countries which are in the Polity IV dataset in 1875 and the right-hand panel for all countries in the data in a given year.

Figure 2
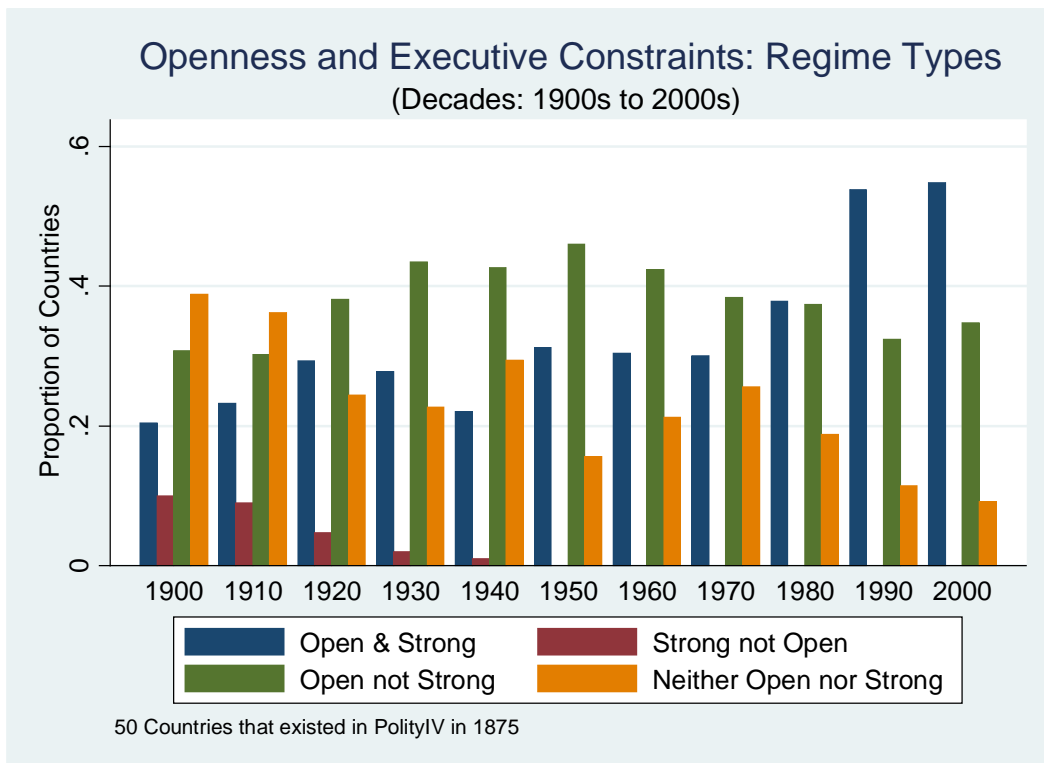
Notes: The data come from the Polity IV website http://www.systemicpeace.org/polityproject.html. For executive constraints, we use *xconst* which is measured by a dummy variable which is equal to 1 if the value of this variable in a given country-year is equal to the highest value of 7 and 0 otherwise; for executive constraints, we us *xropen* which is measured by a dummy variable which is equal to 1 in a given country-year is equal to the highest value of 4 and 0 otherwise. The figure shows the proportions of the four regime types defined by four possible combinations of these dummy variables by decade since 1900, among all countries which are in the Polity IV dataset in 1875.

Figure 3
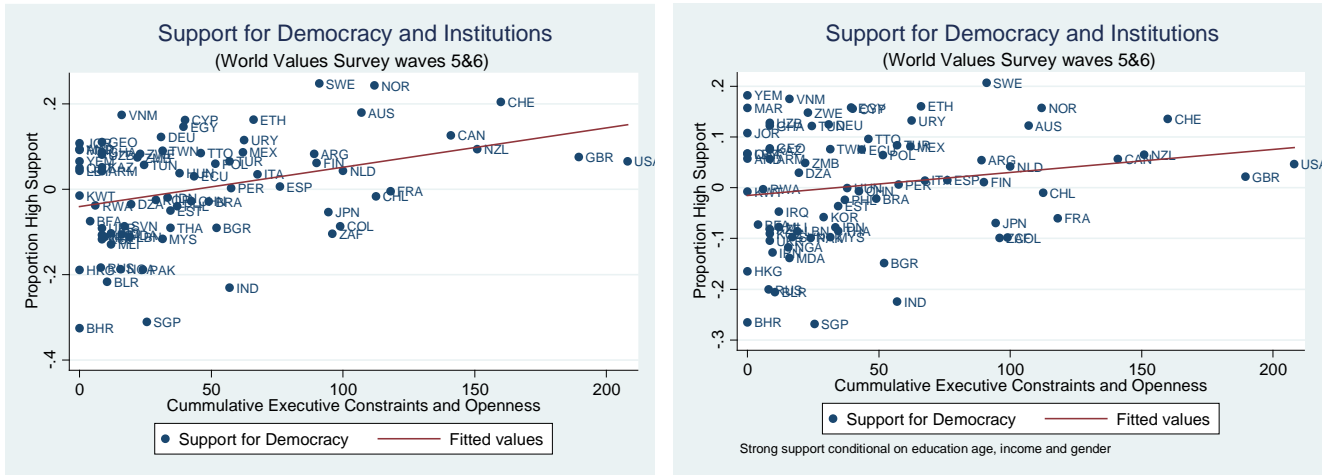
Notes: The data on institutions come from the Polity IV website http://www.systemicpeace.org/polityproject.html . For executive constraints, we use *xconst* which is measured by a dummy variable which is equal to 1 if the value of this variable in a given country-year is equal to the highest value of 7; for openness, we use *xropen* which is measured by a dummy variable which is equal to 1 if the value in a given country year is equal to the highest value of 4. The horizontal axis displays the number of years for which a country has strong executive constraints plus the number of years for which it has open executive recruitment divided by two. Support for democracy is an individual dummy variable from the World Values Survey http://www.worldvaluessurvey.org/wvs.jsp waves 5 and 6 which equals 1 if the individual expresses Support for Democracy (on a 10 point Scale) at 9 or 10. The vertical axis gives the average value of the dummy variable for each country across both waves.

Figure 4
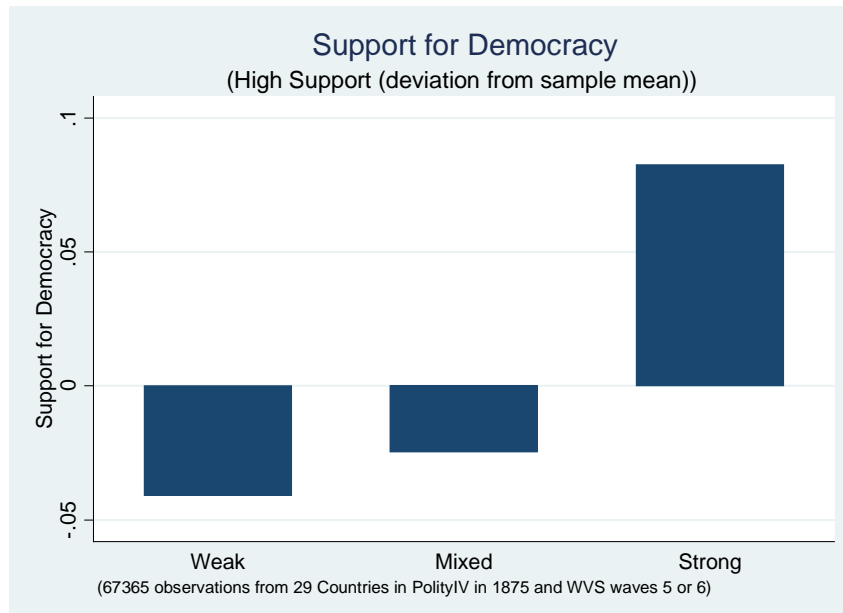
Notes: Support for democracy is a dummy variable which equals one if the individual expresses Support for Democracy (on a 10 point Scale) at 9 or 10. This is averaged by country across the two waves less the mean value across the whole sample (8.58). The country categories (weak, mixed, strong) follow Table 1 for the countries in the World Values Survey http://www.worldvaluessurvey.org/wvs.jsp (waves 5 and 6).

Table 1: Classification of Countries by Executive Constraints

Weak	Mixed	Strong
Always Weak Executive Constraints	Multiple Changes (Number of Upward Changes, Number of Downward Changes)	Always Strong Executive Constraints
Afghanistan China*† El Salvador Guatemala Honduras Iran*† Mexico*† Morocco*† Oman Nepal Russia*† Venezuela	Argentina*† (1,1) Austria (2,1) Belgium (3,2) Bolivia (1,1) Brazil*† (1,1) Chile*† (2,1) Colombia*† (2,2) Denmark (2,1) Dominican Republic (1,1) Ecuador† (2,2) France* (2,2) Greece (4,3) Haiti (1,1) Japan*† (2,1) Netherlands*† (2,1) Norway* (2,1) Peru*† (2,1) Portugal (4,3) Paraguay (2,1) Serbia* (1,1) Spain*† (4,3) Thailand*† (1,1) Turkey*† (3,2)	Canada*† New Zealand† Switzerland* United Kingdom*† United States*†
Permanent Switch to Weak Executive Constraints (Year of Switch		Permanent Switch to Strong Executive Constraints (Year of Switch)
Ethiopia*† (1930) Liberia (1884)		Costa Rica (1875) Germany*† (1990) Hungary* (1990) Italy* (1948) Nicaragua (1995) Romania (2004) Sweden*† (1917) Uruguay* (1985)

Notes: Sample is 50 countries which appear in the PolityIV data base as independent countries in 1875. The data set covers the period 1800 to 2011 and Table 3 displays when each country first entered the data. Data for Germany are for unified Germany; West Germany had strong executive constraints from 1950 onwards. A * denotes a country in wave 5 and a † denotes a country in the wave 6 of World Values Survey.

Table 2: Classification of Countries by Executive Recruitment

Open	Mixed		Closed
Always Closed Executive Recruitment	Multiple Changes (Number of Upward Changes, Number of Downward Changes)		Always Open Executive Recruitment
Morocco*† Oman	Afghanistan (1,1) Argentina*† (4,4) Austria (3,2) Belgium (2,1) Bolivia (6,6) Brazil*† (3,2) Chile*† (2,2) China*† (2,1) Colombia*† (1,1) Denmark (2,1) Dominican Republic (5,5) Ecuador† (4,3) El Salvador (3,2)	Hungary* (3,2) Iran*† (2,1) Italy* (2,1) Liberia (2,2) Mexico*† (3,2) Nepal (2,1) Netherlands*† (2,1) Nicaragua (2,2) Norway* (2,1) Peru*† (8,7) Portugal (3,2) Paraguay (2,2) Romania (2,2)	Canada*† Costa Rica New Zealand† Switzerland* United States*†
Permanent Switch to Closed Executive Recruitment (Year of Switch)	Ethiopia*† (3,2) France* (5,4) Germany*† (1,1) Greece (4,4) Guatemala (5,4) Haiti (6,6) Honduras (6,6)	Russia*† (1,0) Serbia* (3,3) Spain*† (4,3) Thailand*† (6,5) Turkey*† (2,1) Uruguay* (2,1) Venezuela (1,1)	Permanent Switch to Open Executive Recruitment (Year of Switch)
			Japan*† (1952) Sweden*† (1917) United Kingdom*† (1837)

Notes: Sample is 50 countries which appear in the PolityIV data base as independent countries in 1875. The data set covers the period 1800 to 2011 and Table 3 displays when each country first entered the data. Data for Germany are for unified Germany; West Germany had strong executive constraints from 1950 onwards. A * denotes a country in wave 5 and a † denotes a country in the wave 6 of World Values Survey.

Table 3: Timing of Institutional Reforms

Country	First year in Data	First year Open	First year Strong	Country	First year in Data	First year Open	First year Strong
Afghanistan	1800	1989	-	Japan	1800	1952	1868
Argentina	1825	1825	1983	Liberia	1847	1847	1847
Austria	1800	1920	1946	Mexico	1822	1867	-
Belgium	1830	1919	1853	Morocco	1800	-	-
Bolivia	1825	1825	1982	Nepal	1800	2006	-
Brazil	1824	1894	1946	Netherlands	1815	1917	1889
Canada	1867	1867	1867	New Zealand	1857	1857	1857
Chile	1818	1818	1891	Nicaragua	1838	1838	1995
China	1800	1914	-	Norway	1814	1898	1884
Colombia	1832	1832	1867	Oman	1800	-	-
Costa Rica	1838	1838	1875	Paraguay	1811	1811	1992
Denmark	1800	1915	1915	Peru	1821	1822	1990
Dominican Republic	1844	1844	1962	Portugal	1800	1911	1836
Ecuador	1830	1901	1979	Romania	1859	1859	2004
El Salvador	1841	1903	-	Russia	1800	1992	-
Ethiopia	1855	1946	1855	Serbia	1830	1838	1838
France	1800	1800	1877	Spain	1800	1873	1871
Germany	1800	1919	1990	Sweden	1800	1917	1917
Greece	1827	1827	1864	Switzerland	1848	1848	1848
Guatemala	1839	1876	-	Thailand	1800	1935	1992
Haiti	1820	1918	-	Turkey	1800	1923	1961
Honduras	1839	1839	-	United Kingdom	1800	1837	1800
Hungary	1867	1948	1990	United States	1800	1800	1800
Iran	1800	1982	-	Uruguay	1830	1904	1985
Italy	1861	1928	1948	Venezuela	1830	1830	-

Notes: Sample is 50 countries which appear in the PolityIV data base as independent countries in 1875. The data base covers the period 1800 to 2011. Data for Germany are for unified Germany; West Germany had strong executive constraints from 1950 onwards.

Table 4: Determinants of Individual Values

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Major Oil Producer	-0.166*** (0.059)	-0.073 (0.053)			-0.089* (0.051)			
Soviet Influence			-0.270*** (0.079)	-0.148** (0.065)	-0.196** (0.084)			
German Occupation						0.087** (0.0411)	0.085** (0.032)	
Settler Mortality								-0.005** (0.002)
Individual & country controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Countries	28	62	28	62	28	14	26	12
Sample	Countries in Table 1 and in WVS 5 or 6	All Countries in WVS 5 or 6 and PIV	Countries in Table 1 and in WVS 5 or 6	All Countries in WVS e 5 or 6 and PIV	Countries in Table 1 and in WVS 5 or 6	Countries in Mixed Category in Table 1 in WVS 5 or 6	All Countries in Mixed Category in WVS 5 or 6	All Countries in Mixed Category in WVS 5 or 6
Observations	57690	117133	57690	117133	57690	30956	55834	25298
R²	0.037	0.025	0.039	0.027	0.041	0.030	0.032	0.033

Notes: This table reports results from individual-level regressions. The dependent variable is a dummy, which equals one if the individual expresses Support for Democracy (on a 10 point Scale) at 9 or 10. Major Oil Producer is a dummy variable which is equal to one if a country has a oil rents in excess of 10% of GDP according to the World Bank. German Occupation is equal to one if the country was occupied n WWII and had a prior history of strong executive constraints. Settler Mortality is obtained from Acemoglu, Johnson and Robinson (2001). Individual controls are ten dummies for income group, three for education group, gender and three age bands. Country level controls are for GDP, human capital and the cumulative history of executive constraints. Standard errors are adjusted for clustering at the country level. A "*" denotes significant at 10%, a "**" significant at 5% and "***" significant at 1%.

On Line Appendix Material

A   Proofs

Proof of Proposition 1.   First define

$$\Gamma\left(\mu, x\right) = u^I\left(x, 1\right) - V\left(\mu, x\right)$$

as the utility difference between the two institutions. Hence, $D\left(\mu, x\right) = 1$ if and only if $\Gamma\left(\mu, x\right) \geq 0$. Now observe that $\Gamma\left(1, \overline{x}\right) > 0$ using (6) and $\Gamma\left(0, \underline{x}\right) > 0$ using (2). Since $\Gamma\left(\mu, x\right)$ is continuous and increasing in $\mu \in [0, 1]$, the intermediate-value theorem implies that there exists μ^H such that

$$\Gamma\left(\mu^H, \overline{x}\right) = 0.$$

Next define μ^L from

$$\Gamma\left(\mu^L, \underline{x}\right) = 0$$

Since $\Gamma\left(\mu, x\right)$ is increasing in μ, it follows that for $\mu \in \left[\mu^L, \mu^H\right]$, there exists a critical value $\widehat{x}\left(\mu\right) \in [x, \overline{x}]$ such that

$$\Gamma\left(\mu, \widehat{x}\left(\mu\right)\right) = 0.$$

∎

Proof of Proposition 2.   This rests on the complementarity such that $\Delta_\mu\left(\mu\right) \geq 0$. To see this, note that (7) and (8) imply that:

$$\Delta_\mu\left(\mu\right) = \begin{cases} \int_{\underline{x}}^{\overline{x}} \gamma\left(x\right) dH\left(x\right) & \mu \geq \mu^H \\ \rho \int_{\widehat{x}\left(\mu\right)}^{\overline{x}} \frac{dp\left(\mu, f^*\left(\mu, x\right)\right)}{d\mu} \gamma\left(x\right) dH\left(x\right) + & \\ \left[\gamma\left(x\right) + \lambda\left(\mu, \widehat{x}\left(\mu\right)\right)\right] h\left(\widehat{x}\left(\mu\right)\right) \frac{\partial \widehat{x}\left(\mu\right)}{\partial \mu} & \mu \in \left[\mu^L, \mu^H\right] \\ \rho \int_{\underline{x}}^{\overline{x}} \gamma\left(x\right) \frac{dp\left(\mu, f^*\left(\mu, x\right)\right)}{d\mu} dH\left(x\right) & \mu \leq \mu^L. \end{cases} \qquad \text{(A.1)}$$

That is to say, we get $\Delta_\mu\left(\mu\right) \geq 0$ for all $\mu \in [0, 1]$ after observing that

$$\frac{\partial \widehat{x}\left(\mu\right)}{\partial \mu} = \frac{f^*\left(\mu, \widehat{x}\left(\mu\right)\right)}{\frac{\partial p\left(\mu, f^*\left(\mu, \widehat{x}\left(\mu\right)\right)\right)}{\partial \mu} \widehat{x}\left(\mu\right)} > 0.$$

∎

B   Strategic Socialization

We now show that the key equation (3) can be derived from a model, in which matching is assortative and socialization is purposeful. This follows the approach of Cavalli-Sforza and Feldman (1981) as adapted by Bisin and Verdier (2001). Socialization would then have two parts:

1. Direct Socialization: A parent may directly socialize a child into being a concerned citizen, depending on parental effort.

2. Oblique Socialization: If this is unsuccessful, the child may become socialized by society at large becoming a concerned citizen with probability μ_t.

We focus on the case where marriages are perfectly assortative and each pair of parents has two kids. Let $e \in \{0, 1\}$ be the effort put into socializing kids as concerned at cost C. Also, let the probability of successful socialization depend on $e + \varphi$ where φ is a stochastic socialization shock uniformly distributed on $\left[-\frac{1}{L}, \frac{1}{L}\right]$. Then, we have:

$$\text{Prob[concerned: } e] = \frac{1}{2} + Le.$$

Finally, as in our canonical model, let η be an idiosyncratic shock to parental preferences. They now choose socialization effort:

$$e^* = \arg\max \left\{ \left(\frac{1}{2} + Le\right) [\Delta(\mu) + \eta] - Ce \right\}.$$

This defines a threshold

$$\hat{\eta} = \nu - \Delta(\mu),$$

where $\nu = C/L$ such that $e^* = 1$ if and only $\eta \geq \hat{\eta}$.

For the children of concerned parents, the probability of a child being socialized as concerned is $G(\Delta(\mu_t) - \nu)$. For those who are not directly socialized, the probability of oblique socialization into being concerned is $(1 - G(\Delta(\mu_t) - \nu))\mu_t$.

Adding these expressions, the overall probability that the kids of concerned parents are concerned is:

$$G(\Delta(\mu_t) - \nu) + (1 - G(\Delta(\mu_t) - \nu))\mu_t. \tag{B.2}$$

If a child is born to passive parents, we assume she is never directly socialized into being concerned. However, she is socialized as passive with probability $(1 - G(\Delta(\mu_t) - \nu))$. The fraction of such children who are obliquely socialized as concerned is therefore:

$$G(\Delta(\mu_t) - \nu)\mu_t. \tag{B.3}$$

The overall fraction of concerned citizens in the next generation becomes

$$
\begin{aligned}
\mu_{t+1} &= \mu_t \left[G(\Delta(\mu_t) - \nu) + (1 - G(\Delta(\mu_t) - \nu))\mu_t \right] + (1 - \mu_t) \left[G(\Delta(\mu_t) - \nu)\mu_t \right] \\
&= (\mu_t)^2 + 2(1 - \mu_t)\mu_t G(\Delta(\mu_t) - \nu),
\end{aligned}
$$

which is essentially the same expression as in the canonical model. The only difference is that costly effort of being socialized as passive reduces the probability of concerned citizens in the population relative to our case in Section 4, which has $\nu = 0$. This is the special case when $C = 0$ – i.e., when the effort by parents into socializing their child is costless.

In this setting, the candidate for an interior steady state is:

$$\Delta(\hat{\mu}) = \nu,$$

but when $\Delta_\mu(\mu) \geq 0$ this is unstable and the basic thrust of the canonical-model analysis goes through unscathed.

C Comparison with Replicator Dynamics

Suppose that concerned and passive citizens are two behavioral types. Standard replicator dynamics would follow:

$$
\begin{aligned}
\mu_{t+1} - \mu_t &= \mu_t \left[\frac{[(\text{Utility Concerned:}\mu_t) - (\text{Average Utility:}\mu_t)]}{\text{Maximum Utility of Concerned Citizen}} \right] \\
&= \mu_t (1 - \mu_t) \left[\frac{[(\text{Utility Concerned:}\mu_t) - (\text{Utility Passive:}\mu_t)]}{\text{Maximum Utility of Concerned Citizen}} \right]
\end{aligned}
$$

where we have chosen to normalize by "Maximum Utility of Concerned Citizen" so that the term following $\mu_t (1 - \mu_t)$ remains in the unit interval. Let $\pi(x, \mu)$ be the probability that $D = 1$ given $\{x, \mu\}$. This expression boils down to

$$
\begin{aligned}
\mu_{t+1} - \mu_t &= \mu_t (1 - \mu_t) \frac{\int_{\underline{x}}^{\bar{x}} [\pi(x, \mu_t) \gamma(x) - (1 - \pi(x, \mu_t) \lambda(x, \mu_t))] \, dH(x)}{\int_{\underline{x}}^{\bar{x}} [2u^m(x, 1) - u^m(x, 0)] \, dH(x)} \\
&= \mu_t (1 - \mu_t) \frac{\Delta(\mu_t)}{\int_{\underline{x}}^{\bar{x}} [2u^m(x, 1) - u^m(x, 0)] \, dH(x)}.
\end{aligned}
$$

Then the tipping point for the dynamics would be $\Delta(\hat{\mu}) = 0$, which would be similar to our analysis. Moreover, as long as $\Delta_\mu(\mu) \geq 0$, the dynamics would be qualitatively the same as in the canonical model.

D Executive Constraints and Openness

In this Appendix, we discuss some of the detailed modeling of executive recruitment and constraints.

New timing in the sequential-choice model In the sequential model, each generation goes through the following 6 steps.

1. An interim leader from generation t is chosen from the ruling group at $t - 1$. The popularity shock of the incumbent group z_t is realized.

2. The interim leader chooses O_t and resources to fight F_t at step 3.

3. If $O_t = 1$, the interim leader is confirmed in power with probability z_t and a leader from each one of the other groups is selected with probability $(1 - z_t)/M$.

3´. If $O_t = 0$, (concerned) citizens get an opportunity to protest with probability ρ. If the protest is unsuccessful, then the interim leader is confirmed in power. If the process is successful, then the interim leader is removed from power and a leader is selected at random from one of the other groups.

4. The revenue shock, n_t, is realized and the leader from step 3 chooses X_t and f_t

5. If $X_t = 1$, at stage 1 of the legislative bargaining game the leader proposes an allocation S_t and the other groups vote whether to accept or reject it. If a majority rejects it, the legislative bargaining game goes on to stage 2 as described in the text.

5´. If $X_t = 0$, (concerned) citizens get an opportunity to protest with probability ρ. If the protest is unsuccessful, then $s_i = 1$ and $s_m = 0$ for $\mathcal{M}_{-i_t}$. If it is successful, then $s_i = 0$, and $s_m = \frac{1}{M}$ for $\mathcal{M}_{-i_t}$.

6. Payoffs are realized, a new generation is born and socialized by their parents. Parents die.

As before, we solve the model in period t backwards.

Proof of Corollary   As in the basic model, the choice of executive constraints hinges on n. Hence, for $\mu \in \left[\mu_X^L, \mu_X^H\right]$ define $\hat{n}(\mu)$ from:

$$\left[\frac{1}{2} - \rho p\left(\mu, f^*\left(\mu; \hat{n}(\mu)\right)\right)\right]\left[\tau w + \hat{n}(\mu)\right] = w f^*\left(\mu; \hat{n}(\mu)\right).$$

with $\hat{n}\left(\mu_X^H\right) = \bar{n}$ and $\hat{n}\left(\mu_X^L\right) = \underline{n}$.   Then

$$X(\mu, n) = \begin{cases} 1 & \text{if } \mu \geq \mu_X^H \\ 1 & \text{if } \mu \in \left[\mu_X^L, \mu_X^H\right] \text{ and } n \leq \hat{n}(\mu) \\ 0 & \text{otherwise.} \end{cases}$$

Now consider two distributions $K^1(n), K^0(n)$ such that $K^1(n) \leq K^0(n)$ for all $n \in [\underline{n}, \bar{n}]$ and let:

$$
\begin{aligned}
b^j(\mu) &= \int_{\underline{n}}^{\bar{n}}\left(X(\mu, n)\frac{M-1}{2M}[\tau w + n]\right. \\
&\quad \left. + \left[1 - X(\mu, n)\right]\left[V(\mu, n) - \frac{\rho p\left(\mu, f^*(\mu, n)\right)}{2M}[\tau w + n]\right]\right)dK^j(n)
\end{aligned}
$$

for $j \in \{0, 1\}$.   Then $b^1(\mu) \geq b^0(\mu)$ since $\left[V(\mu, n) - \frac{\rho p(\mu, f^*(\mu, n))}{2M}[\tau w + n]\right]$ and $\frac{M-1}{2M}[\tau w + n]$ are both increasing in n. The inequality is strict if $\mu \in \left[\mu_X^L, \mu_X^H\right]$ and there exists $n' \in [\underline{n}, \bar{n}]$ such that $K^1(n) < K^0(n)$.

Now for $\mu \in \left[\mu_O^L, \mu_O^H\right]$, we can define a critical threshold $\hat{z}(\mu; b)$ from the equilibrium condition:

$$\left(\left[1 - \rho p\left(\mu, F^*\left(\mu; b^j(\mu)\right)\right)\right] - \hat{z}\left(\mu; b^j(\mu)\right)\right) b^j(\mu) = w F^*\left(\mu; b^j(\mu)\right).$$

with $\hat{z}\left(\mu_O^H, b^j\left(\mu_O^H\right)\right) = \underline{z}$ and $\hat{z}\left(\mu_O^L, b^j\left(\mu_O^L\right)\right) = \bar{z}$. Consider

$$O(\mu, z) = \begin{cases} 1 & \text{if } \mu \geq \mu_O^H \\ 1 & \text{if } \mu \in \left[\mu_O^L, \mu_O^H\right] \text{ and } z \geq \hat{z}\left(\mu, b^j(\mu)\right) \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\hat{z}\left(\mu; b^1(\mu)\right) \geq \hat{z}\left(\mu; b^0(\mu)\right)$. All else equal, a first-order stochastically dominating shift in n thus reduces the range of z such that $O = 1$.

The model induces two probability distributions over institutions, which we denote by

$$o(z; \mu) = O(\mu, z) + \left[1 - \rho p\left(\mu, F^*(\mu; b(\mu))\right)\right]\left[1 - O(\mu, z)\right],$$

which is increasing in z, and

$$e(n; \mu) = X(\mu, n) + \left[1 - \rho p\left(\mu, f^*(\mu; n)\right)\right]\left[1 - X(\mu, n)\right],$$

40

which is increasing in n. Then

$$\int_{\underline{n}}^{\bar{n}} e\left(n;\mu\right) dK^1\left(n\right) \geq \int_{\underline{n}}^{\bar{n}} e\left(n;\mu\right) dK^0\left(n\right)$$

Now consider two distributions $Q^1\left(z\right), Q^0\left(z\right)$ such that $Q^1\left(z\right) \leq Q^0\left(z\right)$ for all $z \in [\underline{z}, \bar{z}]$ then

$$\int_{\underline{z}}^{\bar{z}} o\left(z;\mu\right) dQ^1\left(z\right) \geq \int_{\underline{z}}^{\bar{z}} o\left(z;\mu\right) dQ^0\left(z\right)$$

since $O\left(\mu, z\right)$ is a (weakly) increasing function of z. Note that the inequality is strict if $\mu \in \left[\mu_O^L, \mu_O^H\right]$ and there exists $z' \in [\underline{z}, \bar{z}]$ such that $Q^1\left(z\right) < Q^0\left(z\right)$.

With these preliminaries, the corollary stated in Subsection 5.3 follows. For a mixed regime with $X = 0$ and $O = 1$, we require that z and n be sufficiently high. Moreover, the likelihood of this outcome is higher with first-order stochastically dominating shifts in $K\left(n\right)$ or $Q\left(z\right)$.

Value dynamics in the sequential-choice model   We now compute the terms which drive the dynamics of democratic values. Let $\omega^m\left(O, X; z, n\right)$ be the total expected utility of the subordinated citizens from a particular configuration of institutions, and let $v(M; z, n)$ be the expected material utility from policy under full democracy. Then

$$
\begin{aligned}
\omega^m\left(0, 0; z, n\right) &= 0 \\
\omega^m\left(0, 1; z, n\right) &= \frac{1}{2}\left(\tau w + n\right) \\
\omega^m\left(1, 1; z, n\right) &= \omega^m\left(0, 1; z, n\right) + \frac{(1-z)}{2M}\left[M - 1\right]\left(\tau w + n\right) \\
&= \omega^m\left(0, 1; z, n\right) + v(q, M; z, n) \\
\omega^m\left(1, 0; z, n\right) &= \left(1 - z\right)\left(\tau w + n\right).
\end{aligned}
$$

Note that $\nu\left(1; z, n\right) = 0$, i.e., with only two groups, open executive recruitment has no additional value over and above strong executive constraints.

We can use these to compute gains γ and losses λ, where the reference point for gains is $X_t = O_t = 0$ and for losses is $X_t = O_t = 1$. As $\omega^m\left(0, 0; z, n\right) = 0$, $\gamma\left(O, X; z, n\right) = \omega^m\left(O, X; z, n\right)$.

Turning to losses, we have:

$$
\begin{aligned}
\lambda\left(0, 0; z, n\right) &= -\omega^m\left(1, 1; z, n\right) \\
\lambda\left(1, 1; z, n\right) &= 0 \\
\lambda\left(0, 1; z, n\right) &= -\nu\left(q, M; z, n\right) \\
\lambda\left(1, 0; z, n\right) &= \omega^m\left(1, 0; z, n\right) - \omega^m\left(1, 1; z, n\right).
\end{aligned}
$$

The critical terms to study the cultural dynamics are the sums of the gain and loss in each state:

$$
\begin{aligned}
\delta\left(0, 0; z, n\right) &= -\omega^m\left(1, 1; z, n\right) \\
\delta\left(1, 1; z, n\right) &= \omega^m\left(1, 1; z, n\right) \\
\delta\left(0, 1; z, n\right) &= \omega^m\left(1, 1; z, n\right) - 2\nu\left(M; z, n\right) \\
\delta\left(1, 0; z, n\right) &= 2\left(1 - z\right)\left(\tau w + n\right) - \omega^m\left(1, 1; z, n\right).
\end{aligned}
$$

Moreover, we have shown that an increase in μ raises the probability that $O_t = 1$ and $X_t = 1$. Then

$$\Delta(\mu) = \int_{\underline{n}}^{\bar{n}} \int_{\underline{z}}^{\bar{z}} \left([o(z;\mu) e(n;\mu)] \delta(1,1;z,n) + \right.$$

$$[[1-o(z;\mu)][1-e(n;\mu)]] \delta(0,0;z,n)$$

$$+ [[1-o(z;\mu)] e(n;\mu)] \delta(0,1;z,n)$$

$$\left. + [o(z;\mu)[1-e(n;\mu)]] \delta(1,0;z,n) \right) dK(n) \, dQ(z).$$

For $\Delta(\mu)$ to be everywhere increasing, a sufficient condition is that

$$\delta(1,1;z,n) > \max\{\delta(1,0;z,n), \delta(0,1;z,n)\}$$

and

$$\min\{\delta(1,0;z,n), \delta(0,1;z,n)\} > \delta(0,0;z,n)$$

for all (z,n). Necessary and sufficient for this are:

$$\omega^m(1,1;z,n) > \nu(M;z,n)$$

and

$$\omega^m(1,1;z,n) > (1-z)(\tau w + n). \tag{D.4}$$

The first of these always holds in our framework. However, the second requires that

$$\left(1 - \frac{1-z}{M}\right) \frac{1}{2M} + \frac{1-z}{2M} > \frac{1-z}{M},$$

or that $z \geq 1/(M+1)$. So the complementarity is present if $\underline{z} \geq 1/(M+1)$, which will hold if M is large enough.