Dynamics of Political Systems[*]

Lukas Buchheim

University of Munich

Robert Ulbricht

Toulouse School of Economics

May 14, 2016

Abstract

We develop a model of regime dynamics that embeds the principal transition scenarios examined by the literature. Political systems are *a priori* unrestricted and dynamics emerge through the combination and interaction of transition events over time. The model attributes a key role to beliefs held by political outsiders about the vulnerability of regimes, governing the likelihood and outcome of transitions. In equilibrium, transition likelihoods are declining in a regime's maturity, generating episodes of political stability alternating with rapid successions of revolts, counter revolts, and reforms. The stationary distribution of regimes is bimodal. The model dynamics match the data remarkably well.

Keywords: Critical junctures, democratization, invariant distribution, iron law of oligarchy, regime dynamics, revolts.
JEL Classification: D74, D78, P16.

1 Introduction

A growing literature in political economy explores the causes and circumstances of political transitions. So far, this literature has focused on explaining specific patterns of regime changes, initiated by either reforms or revolts. In contrast, the unfolding dynamics of political systems, which result from the combination and interaction of individual transition events over time, have received little attention. In this paper, we take a step towards filling this gap, placing the dynamic process of political systems at the center of analysis.

In particular, we provide a unified framework that integrates the principal transition scenarios discussed in the previous literature: (i) democratization (a move from autocratic regimes to democratic ones); (ii) reversely, the collapse of democracies; and (iii) the replacement of autocracies by other autocracies. Although simple in its nature the framework is rich enough to generate observable statistics of the transition process of political systems, allowing us not only to study regime dynamics unfolding in an environment that jointly allows for the main transition events, but also to quantitatively compare them to the data.

Following the literature, we focus on reforms and revolts as means of political transitions, but generalize earlier models to endogenize the outcome of political transitions and to ensure that there are no exogenously absorbing political systems. Specifically, our model builds on the seminal paper by Acemoglu and Robinson (2000b), which posits that the threat of revolts is an important cause for democratizing reforms. To endogenize the outcome of reforms, we allow franchise extensions to be of arbitrary scope. For revolts, we model a coordination process that determines the mass of revolting outsiders, who in turn form the newly emerging regime if a revolt succeeds. In consequence, our model defines a continuous space of political systems, ranging from single-person dictatorships to full-scale democracies. All political regimes within this space are, in principle, attainable through political transitions.[1]

Two important features of political transitions in the data are that (i) transitions tend to be clustered over time, with the flipside that both autocratic and democratic

---

[1]As is common in the politico-economic literature, we characterize political systems by the fraction of the population with access to political power. Examples for regimes where political power is concentrated in the hands of exclusive elites are, e.g., Chile (1973–90) and today's North Korea. The majority of the population is, by contrast, enfranchised in most Western democracies. Regimes between the two extremes, where parts of the population is deprived from political rights in an otherwise inclusive system, are, e.g., Hungary (1921–31) and Madagascar (1960–72).

regimes tend to stabilize in the long run, and (ii) the co-existence of both reforms and revolts as means of transition. The model sketched so far does not capture these facts well. We propose that these features can be accounted for in a parsimonious way if political outsiders know less about the regime's vulnerability to a revolt than political insiders. On the one hand, such an information asymmetry will lead insiders to sometimes take "tough stance" rather than to negotiate on moderate reforms when facing revolutionary pressure, opening the door for significant revolts along the equilibrium path. On the other hand, we will show that learning about a regime's vulnerability across periods also provides a natural explanation for both the emergence of episodes of political instability as well as the long-run stabilization of political regimes.

In addition to generating realistic patterns for the likelihoods of transitions, asymmetric information also helps generating properties for the outcomes of transitions that are consistent with the data. In particular, we show that the modal reform leads to a democratic political system, while the typical revolt establishes an autocracy. We proceed by outlining the intuition for our main findings in greater detail.

Regime dynamics Because of asymmetric information, the likelihood of transitions crucially hinges on how vulnerable the regime is perceived to be by political outsiders. To illustrate this as cleanly as possible, we first consider a baseline version of our model where we treat the prior of political outsiders regarding the regime's vulnerability as exogenous. In this environment, if it is *a priori* unlikely that the regime is vulnerable to a revolt, few outsiders find it worthwhile to support a revolt, posing a negligible threat to the regime, and reforms and revolts are ultimately unlikely. When, by contrast, a regime is perceived to be vulnerable, more outsiders are in principle willing to support a revolt, which in turn also increases the regime's incentives to reform. In equilibrium, this generates a non-trivial likelihood of either type of transition.

The link between prior beliefs and transition likelihoods suggests a crucial role for learning dynamics in shaping the timing of transitions. Specifically, if the institutional characteristics underlying the vulnerability of a regime are persistent, then outsiders may not only learn from the regime's contemporaneous actions, but also from the history of political transitions (or their absence). In our main exploration of regime dynamics, we use a version of the model where we allow for such learning dynamics.

The key insight regarding the timing of regime changes is a negative relationship between the likelihood of observing a transition and the maturity of a regime, which

accounts for the long-run stabilization of regimes seen in the data. The reason for this finding is that both reforms and revolts are more likely to occur when a regime is vulnerable, whereas the absence of a transition is a sign of internal stability. Accordingly, outsiders gradually become more and more convinced that a regime is invulnerable as it matures, reducing their willingness to revolt. Once a transition eventually occurs, however, outsiders rationally believe the new regime to be relatively more vulnerable, entailing a rise in the likelihood of further transitions.

A consequence of the negative relation between a regime's hazard rate and its maturity is that transition events tend to be clustered across time, giving rise to episodes of political stability alternating with episodes of political turbulence. Accounting for the rich transition patterns that can be observed in the data, episodes of political turbulence can be composed of rapid successions of revolts, of alternations between reforms and counter-revolts, as well as of gradual democratization episodes through a series of reforms.

One aspect recently highlighted is the existence of critical junctures for the evolution of political systems, where small differences in initial conditions can lead to diverging paths with long-lasting consequences (Acemoglu and Robinson, 2012). In the model, political turbulences pose such critical junctures. Whereas the outcome of such episodes is largely determined by small and random variations in current states, the type of political system that eventually survives an episode of political turbulence is likely to persist for a long time.

Another prediction of the model is what is sometimes labeled as the "iron law of oligarchy". Because outsiders believe mature regimes to be invulnerable, mature regimes in turn find it generally attractive to abstain from reforms regardless of their true vulnerability. Accordingly, mature regimes are bound to eventually fall by means of a revolt and to be succeeded by an autocratic regime.

The flipside of this result is that the typical path to democracy starts with a revolt triggering a critical juncture. However, because revolts are likely to be small, the event that ultimately establishes a democracy is usually a reform. This is consistent, e.g., with the observation of Karl (1990, p. 8) that no stable South American democracy has been the result of mass revolutions (see also Rustow, 1970; O'Donnell and Schmitter, 1973; Huntington, 1991).

Invariant distribution   At an aggregate level, perhaps the most salient characteristic of the process defining the dynamics of regimes is its invariant distribution. Matching the data, the model's distribution of political systems across time is bimodal with mass concentrated on autocratic and democratic political systems and with little mass on intermediate polities. The model identifies two forces underlying the bimodal shape of the invariant distribution.

First, there is a polarization of political systems during their emergence, with transitions resulting in regimes that tend to be either autocratic or democratic. In particular, observing concessions in the form of a reform, outsiders in the model conclude that the regime is weak. Accordingly, small reforms fuel coordination amongst outsiders along the intensive margin, while doing little to reduce revolutionary pressure along the extensive margin. To be effective, reforms thus need to be far-reaching, leading to the establishment of fairly democratic regimes with little opposition.

In contrast, revolts in the model result in fairly autocratic regimes. This is because revolts with widespread support that are likely to succeed would be preempted by insiders. On the other hand, revolts that are likely to fail cannot grow too large either, since only a small set of outsiders with sufficiently high gains from revolting would be willing to take the risk of supporting an ill-fated revolt.

The second force underlying the bimodal shape of the long-run distribution is a persistence of both autocracies and democracies relative to intermediate types of systems. Democracies are more likely to survive (though not indefinitely[2]) due to a lack of meaningful opposition. Autocracies, by contrast, face higher transition probabilities than democracies, in particular when they are young, but these transitions are likely to lead to a succession of an autocratic regime by another autocracy. While the identity of autocratic leaders may change more frequently, autocratic *systems* therefore tend to survive over time.

Comparison to the data   Our model breaks down the dynamics of political systems into statistics describing the short-run outcomes of transition events, the stability of political regimes over time, as well as the long-run distribution of political systems. We contrast these statistics as generated by the model with their empirical counterparts using panel data on regime dynamics for the majority of countries from 1919 onwards.

---

[2]Even though democratic regimes enfranchise the majority of the population, there is typically a small group of oppositional "hardliners" in equilibrium that will eventually succeed in overthrowing even the most democratic political system.

Considering that the model's predictions are derived from a simple framework, the match of the model to the data is remarkably good, both qualitatively and quantitatively

Related literature To study the dynamics of political system, our work builds on a number of ingredients that have previously been studied mainly in isolation. In particular, the preemptive logic of reforms is based on the seminal work by Acemoglu and Robinson (2000b), Conley and Temini (2001), as well as Boix (2003).[3] Closely related to our paper is the analysis by Acemoglu and Robinson (2001) where preemptive reforms co-exist with coups against democracies. Similarly, Acemoglu, Ticchi and Vindigni (2010), Acemoglu and Robinson (2000a), and Ellis and Fender (2011) explore settings that feature both reforms and revolts against autocracies, where the latter two are particularly close to us as they also make use of asymmetric information to achieve co-existence of reforms and revolts. However, all these papers focus on specific, exogenously imposed transition patterns and, with the exception of Acemoglu and Robinson (2001), rule out repeated transitions.

Regarding revolts, our modeling strategy builds on a long tradition of using coordination games in conflict situations (see, e.g., Granovetter, 1978, Kuran, 1989, and the discussion of Morris and Shin, 1998 by Atkeson, 2000). Depending on the context, information has been noted to be able to both spur or hinder coordination (e.g. Lohmann, 1994, Chwe, 2000, Bueno de Mesquita, 2010, Fearon, 2011). In such situations, political leaders have obvious incentives to manipulate public signals directly, as in the works of Angeletos, Hellwig and Pavan (2006) and Edmond (2013), or to carefully consider the information costs associated with their policies, as in our model. Supporting the mechanism proposed in this paper, Finkel, Gehlbach and Olsen (2015) present empirical evidence that halfhearted reforms may fuel revolts by raising the expectations of success among disenfranchised parts of the population.

Our paper also relates to a theoretical political economy literature that studies environments characterized by rich polity spaces. This includes the work on gradual enfranchisement by Justman and Gradstein (1999), Jack and Lagunoff (2006), and

---

[3]A related strand of the democratization literature argues that reforms may also be reflective of situations where autocratic decision makers are better off in a democratized political system than under the status quo (e.g., Bourguignon and Verdier, 2000; Lizzeri and Persico, 2004; and Llavador and Oxoby, 2005). On the empirical side, Przeworski (2009), Aidt and Jensen (2014), and Aidt and Franck (2015) provide evidence suggesting that preemptive reforms are indeed the driving force behind democratization. In a similar spirit, Besley, Persson and Reynal-Querol (2014) show theoretically and empirically that a higher risk to lose political power induces leaders to conduct constitutional reforms.

Gradstein (2007), as well as a number of recent papers analyzing different manifestations of the dynamic tradeoff that arises when the future distribution of power, and thus future policy, is driven by current political decisions (Roberts, 1999*a*,*b*; Barberà, Maschler and Shalev, 2001; Lagunoff, 2009; Acemoglu, Egorov and Sonin, 2008, 2012, 2016; Bai and Lagunoff, 2011; Prato, 2016). Due to their theoretical focus, these papers, however, either restrict transitions to be reforms (often modeled by means of dynamic voting games) or do not specify the exact transition mechanism in detail.

Relative to all the aforementioned literature, a key novelty of this paper is its focus on providing a rationale for the dynamics of political systems observed in aggregate data. As such, this paper adds to the literature by, first, making explicit predictions regarding the relative importance of the most frequent types of transitions—reforms and revolts—, and by, second, relating the mode of transitions to the characteristics of the resulting regimes.[4] Both these steps are necessary to obtain a quantifiable model of aggregate political dynamics.

In addition to the aforementioned predictions, this paper also provides an explanation for the decreasing hazard rates of regimes, which has been previously documented by, e.g., Bienen and van de Walle (1989, 1992) and Svolik (2008). Earlier theories have attributed this fact to a consolidation of power over time (e.g., Svolik, 2008, 2009) and to the establishment of institutions that allow for a credible sharing of power amongst autocratic elites (Gandhi and Przeworski, 2007; Magaloni, 2008; Boix and Svolik, 2013; Francois, Rainer and Trebbi, 2015). We complement these rationales by highlighting how, in the presence of asymmetric information, learning over time naturally leads to a stabilization of regimes as they mature.

Layout   The rest of the paper is organized as follows. Section 2 introduces the baseline model with exogenous priors. Section 3 characterizes the likelihood and outcomes of transitions conditional on the prior. Section 4 extends the model to the case with endogenous priors. Sections 5 and 6 contain our main results on regime dynamics and the long-run properties of political systems. Section 7 compares the theoretical predictions to the data, and Section 8 concludes. Technical details are confined to the appendix.

---

[4]In this sense, we also relate to Cervellati, Fortunato and Sunde (2012, 2014) who show that consensual transitions foster civil liberties and property rights provisions in contrast to violent transitions.

6

2 The model with exogenous priors

In this section we set up a simple, dynamic model of repeated political transitions that are driven by both reforms and revolts. Political systems are defined by the fraction of the population with access to power and can attain any value on $[0, 1]$.

2.1 Setup

We consider an infinite horizon economy with a continuum of two-period lived agents. Each generation has a mass equal to 1. At time t, fraction λ_t of the population has the power to implement political decisions, whereas the remaining agents are excluded from political power. We refer to these two groups as (political) "insiders" and "outsiders".

When born, the distribution of political power among the young is inherited from their parent generation; that is, λ_t agents are born as insiders, while $1 - \lambda_t$ agents are born as outsiders. However, agents who are born as outsiders can attempt to overthrow the current regime and thereby acquire political power. To this end, outsiders choose individually and simultaneously whether or not to participate in a revolt. Because all political change will take effect at the beginning of the next period (see below), only young outsiders have an interest in participating in a revolt. Accordingly, we denote young outsider i's choice by $\phi_{it} \in \{0, 1\}$ and use the aggregated mass of supporters, $s_t = \int \phi_{it} \, di$, to refer to the size of the resulting revolt.

Given the mass of supporters s_t, the probability that a revolt is successful is given by

$$p(\theta_t, s_t) = \theta_t h(s_t), \tag{1}$$

where $\theta_t \in \Theta$ is a random state of the world that reflects the vulnerability of the current regime or their ability to put down a revolt, and h is an increasing and twice differentiable function, $h : [0, 1] \to [0, 1]$, with $h(0) = 0$. That is, the threat of a revolt to the current regime is increasing in the mass of its supporters and in the vulnerability of the regime. When a revolt has no supporters ($s_t = 0$) or the regime is not vulnerable ($\theta_t = 0$), it fails with certainty.

The purpose of θ_t in our model is to introduce asymmetric information between insiders and outsiders that, as will become clear below, ensures that revolts are prevalent along the equilibrium path. Formally we have (for now) that the state θ_t is exogenously distributed on $\Theta = [0, 1]$, is i.i.d. from one period to the next, and is revealed to insiders

at the beginning of each period. Outsiders only know the prior distribution of θ_t. We assume that θ has a differentiable distribution function F with $F'(\theta) > 0$ for all θ in the interior of Θ.

After they learn θ_t, insiders may try to alleviate the threat of a revolt by conducting reforms. We follow Acemoglu and Robinson (2000b) by modeling these reforms as an extension of the franchise to outsiders, which is effective in preventing them from supporting a revolt.[5] Aiming to endogenize the political system λ_t, we, however, generalize this mechanism by allowing insiders to continuously extend the regime by any fraction, $x_t - \lambda_t$, of young outsiders, where $x_t \in [\lambda_t, 1]$ is the reformed political system. Because preferences of insiders will be perfectly aligned, there is no need to specify the decision making process leading to x_t in detail.

Given the (aggregated) policy choices s_t and x_t, and conditional on the outcome of a revolt, the political system evolves as follows:

$$\lambda_{t+1} = \begin{cases} s_t & \text{if the regime is overthrown, and} \\ x_t & \text{otherwise.} \end{cases} \tag{2}$$

When a revolt fails (indicated by $\eta_t = 0$), reforms take effect and the old regime stays in power. The resulting political system in $t+1$ is then given by x_t. In the complementary case, when a revolt succeeds ($\eta_t = 1$), those who have participated will form the new regime. Accordingly, after a successful revolt, the fraction of insiders at $t+1$ is equal to s_t. Note that this specification prevents non-revolting outsiders from reaping the benefits from overthrowing a regime so that there are no gains from free-riding in our model.

To complete the model description, we still have to specify how payoffs are distributed across the two groups of agents at t. As for outsiders, we assume that they receive a constant per period payoff of γ_{it} which is privately assigned to each agent at birth and is drawn from a uniform distribution on $[0, 1]$. This heterogeneity is meant to reflect differences in the propensity to revolt, possibly resulting from different degrees of economical or ideological adaption to a regime.

In contrast, insiders enjoy per period payoffs $u(\lambda_t)$, where u is twice differentiable, $u' < 0$, and $u(1)$ is normalized to unity. We think of $u(\cdot)$ as a reduced form function that captures the various benefits of having political power (e.g., from extracting a

---

[5]That it is indeed individually rational for enfranchised outsiders to not support a revolt is shown in Appendix B.

common resource stock, implementing preferred policies, etc.).[6] One important feature of u is that it is decreasing in the current regime size and, hence, extending the regime is costly for insiders (e.g., because resources have to be shared, or preferences about policies become less aligned). Another thing to note is that $u(\lambda_t) \geq \gamma_{it}$ for all λ_t and γ_{it}; that is, being part of the regime is always desirable. In the case of full democracy ($\lambda_t = 1$) all citizens are insiders and enjoy utility normalized to the one of a perfectly adapted outsider.

To simplify the analysis, we assume that members of an overthrown regime and participants in a failed revolt are worst-adapted to the new regime ($\gamma_{it} = 0$).

For the upcoming analysis it will be convenient to define the (future) utility of agents that are born at time t, which is given by:

$$V^I(\eta_t, x_t) = (1 - \eta_t)u(x_t),$$ (3)

$$V^O(\eta_t, \gamma_{it}, s_t, \phi_{it}) = \phi_{it}\eta_t u(s_t) + (1 - \phi_{it})\gamma_{it},$$ (4)

where superscript I and O denote agents that are born as (or were enfranchised) insiders and outsiders, respectively. In both cases, the terms correspond to the second period payoffs accruing from date $t + 1$ (which are a function of date-t choices), omitting the first period payoffs (which are unaffected by the policy choices of generation-t). Since agents do not face an intertemporal tradeoff, we do not need to define a discount rate here.

The timing of events within one period can be summarized as follows:

1. The state of the world θ_t is revealed to insiders.

2. Insiders may extend political power to a fraction $x_t \in [\lambda_t, 1]$ of the population.

3. Observing x_t, outsiders individually and simultaneously decide whether or not to participate in a revolt.

4. Transitions according to (1) and (2) take place, period $t + 1$ starts with the birth of a new generation, and payoffs determined by λ_{t+1} are realized.

---

[6]More specifically, u should be interpreted as a value function where all policy choices associated with having political power—except enfranchising political outsiders—are replaced by optimal policy rules. In particular, this applies to all question about the organization of the economy, resource reallocation, or (similarly) the design of political institutions used to enfranchise outsiders. Subsuming these issues into u allows us to tractably focus on the inherent dynamics of political systems spanned by political reforms and revolts. Notice, however, that all other policy choices still affect our analysis inasmuch as they determine the shape of u (see also the discussion in Footnote 13).

We characterize the set of perfect Bayesian equilibria. To increase the predictive power of our model, we impose two equilibrium refinements. First, we rule out "instable" coordination outcomes where an infinitesimal perturbation of the conjectured equilibrium support $\hat{s}_t$ would result in a first-order shift in s_t.[7] Second, we limit attention to equilibria that are consistent with the D1 criterion introduced by Cho and Kreps (1987), a standard refinement for signaling games. The D1 criterion restricts outsiders to believe that whenever they observe a reform x' that is not conducted in equilibrium, the reform has been implemented by a regime with vulnerability θ', for which a deviation to x' would be most attractive.[8]

Anticipating some equilibrium properties, we simplify our notation as follows. First, outsiders' beliefs regarding the regime's vulnerability will be uniquely determined in our setup. We therefore denote the commonly held belief by $\hat{\theta}_t$, dropping the index i. Second, there are no nondegenerate mixed strategy equilibria in our game (see the proofs to Propositions 1 and 2). Accordingly, we restrict notation to pure strategies.

This leads to the following definition of equilibrium for our economy.

Definition. Given a history $\delta = \{\lambda_0\} \cup \{\{\phi_{i\tau}\}, \theta_\tau, x_\tau, \eta_\tau\}_{\tau=0}^{t-1}$, an equilibrium in this economy consists of strategies $x_\delta : (\theta, \lambda) \mapsto x$ and $\{\phi_{i\delta} : (\hat{\theta}, x) \mapsto \phi_i\}$, and beliefs $\hat{\theta}_\delta(\lambda, x) \mapsto \hat{\theta}$, such that for all possible histories δ:

a. Reforms x_δ maximize expected insider's utility $\mathbb{E}\{V^I(\cdot)\}$ given states (θ_t, λ_t), and outsiders' beliefs $\hat{\theta}_\delta$ and strategies $\{\phi_{i\delta}\}$;

b. Each outsider's revolt choice $\phi_{i\delta}$ maximizes $\mathbb{E}\{V^O(\cdot)\}$ given insiders' reforms x_δ, other outsiders' revolt choices $\{\phi_{j\delta}\}$, and beliefs $\hat{\theta}_\delta$;

---

[7]This rules out coordination on $s_t = 0$ supported by the belief that $\hat{s}_t = 0$ (implying a zero probability of success), but where an infinitesimal small chance of success would persuade a non-marginal mass of outsiders to revolt. In a previous version of this paper (Buchheim and Ulbricht, 2014), we demonstrate that this restriction is formally equivalent to characterizing the set of trembling-hand perfect equilibria (at the expense of additional notation). An alternative (and outcome-equivalent) approach to rule out these instabilities would be to restrict attention to equilibria which are the limit to a sequence of economies with a finite number of outsiders, where each agent's decision has non-zero weight on s_t.

[8]More precisely, beliefs are attributed to the state in which a deviation to x' is attractive for the largest set of possible inferences about the regime's vulnerability. Formally, let $V^*(\theta) \equiv \mathbb{E}\{V^I(\eta, x^*(\theta, \lambda))|\theta\}$ be the insiders' expected payoff in state θ under a candidate equilibrium x^*. Then the D1 criterion restricts beliefs for off-equilibrium events x' to states θ' that maximize $D_{\theta', x'} = \{\hat{\theta} : \mathbb{E}\{V^I(\eta, x')|\theta', s = s(\hat{\theta}, x')\} \geq V^*(\theta')\}$, where $s(\hat{\theta}, x)$ is the mass of outsiders that revolt given a belief $\hat{\theta}$ under the conjectured equilibrium ($D_{\theta', x'}$ is maximal here if there is no θ'', such that $D_{\theta', x'}$ is a proper subset of $D_{\theta'', x'}$).

c. Beliefs $\hat{\theta}_\delta$ are obtained using Bayes rule given x_δ, and $\hat{\theta}_\delta$ satisfies the D1 criterion;

d. States (λ_t, η_t) are consistent with (1) and (2);

e. Coordination among outsiders is stable; i.e., perturbing perceived coordination $\hat{s}_t$ by ϵ shifts the coordination outcome s_t by at most ν where $\nu \to 0$ as $\epsilon \to 0$.

2.2 Political equilibrium

We now derive the equilibrium strategies of insiders and outsiders, pinning down the political equilibrium in the model economy. Our analysis is simplified by the overlapping generations structure of the model, which gives rise to a sequence of "generation games" between young insiders and young outsiders. Since (for now) the distribution of political power at time t captures all payoff-relevant information of the history up to t, the only link between generations is λ_t. We can therefore characterize the set of equilibria in the baseline model by characterizing the equilibria of the generation games as a function of λ_t. All other elements of the history up to time t may affect the equilibrium at t only by (hypothetically) selecting between multiple equilibria (if the equilibrium in the generation game would not be unique).

The generation game consists of two stages that determine the political system at $t + 1$. In the second stage, outsiders have to choose whether or not to support a revolt. Because the likelihood that a revolt succeeds depends on the total mass of its supporters, outsiders face a coordination problem in their decision to revolt. In the first stage, prior to this coordination problem, insiders decide on the degree to which political power is extended to outsiders. On the one hand this will decrease revolutionary pressure along the extensive margin by contracting the pool of potential insurgents. On the other hand, extending the regime may also contain information about the regime's vulnerability. As a result, reforms may also increase revolutionary pressure along the intensive margin by increasing coordination among outsiders who are not subject to reforms. Insiders' policy choices will therefore be governed by signaling considerations.

We proceed by backward induction in solving for the equilibrium of the generation game, beginning with the outsiders' coordination problem.

Stage 2: Coordination among outsiders Consider the outsiders' coordination problem at time t. Without loss of generality, define $\hat{\theta}_t \equiv \mathbb{E}\{\theta_t\}$.[9] For any given belief, $(\hat{\theta}_t, \hat{s}_t) \in \Theta \times [0,1]$, individual rationality requires all outsiders to choose a ϕ_{it} that maximizes their expected utility $\mathbb{E}\{V^O(\cdot)\}$. At time t, outsider i with opportunity cost γ_{it} will therefore participate in a revolt if and only if

$$\gamma_{it} \leq p(\hat{\theta}_t, \hat{s}_t)\, u(\hat{s}_t) \equiv \bar{\gamma}(\hat{s}_t). \tag{5}$$

In equilibrium $\bar{\gamma}(\hat{s}_t)$ is the expected benefit of participating in a revolt that is supported by a mass $\hat{s}_t$ of outsiders. Since $\bar{\gamma}(\hat{s}_t)$ is independent of γ_{it}, it follows that in any equilibrium the set of outsiders who support a revolt at t is given by the agents who are least adapted to the current regime. Suppose for the time being that $\bar{\gamma}(\hat{s}_t) \leq 1$. Then, $\bar{\gamma}(\hat{s}_t)$ defines the fraction of young outsiders that participates in a revolt, and, therefore, the size of a revolt, s_t, that would follow from $\bar{\gamma}(\hat{s}_t)$ is given by

$$s_t = (1 - x_t)\, \bar{\gamma}(\hat{s}_t). \tag{6}$$

Further note that in any equilibrium it must hold that $s_t = \hat{s}_t$. Therefore, as long as $\bar{\gamma}(\hat{s}_t) \leq 1$, the share of outsiders that support a revolt at t has to be a fixed point to (6). To guarantee that this is always the case and to further ensure that a well-behaved fixed point exists, we impose the following assumption.

Assumption A1. *For $\psi(s) \equiv h(s) \cdot u(s)$,*

    a. $\psi' \geq 0$ and $\psi'' \leq 0$;

    b. $\lim_{s \to 0} \psi'(s) = \infty$.

Intuitively, Assumption A1 states that participating in a revolt becomes more attractive as the total share of supporters grows; i.e., the participation choices of outsiders are strategic complements. This requires that the positive effect of an additional supporter on the success probability outweighs the negative effect of being in a slightly larger regime after a successful revolt. To ensure existence, we further require that the positive effect of an additional supporter is sufficiently strong when a revolt is smallest, and is decreasing as revolts grow larger.

---

[9]From our specification of p, $\mathbb{E}\{V^O\}$ is linear in θ_t so that $\mathbb{E}\{\theta_t\}$ is a sufficient statistic for the full posterior distribution of θ_t.

Using Assumption A1, the above discussion together with our stability requirement leads to the following proposition.

Proposition 1. *In any equilibrium, the mass of outsiders supporting a revolt at time t is uniquely characterized by a time-invariant function, $s : (\hat{\theta}_t, x_t) \mapsto s_t$, which satisfies $s(0, \cdot) = s(\cdot, 1) = 0$, increases in $\hat{\theta}_t$, and decreases in x_t.*

All proofs are in the appendix. Proposition 1 establishes the already discussed tradeoff of conducting reforms: On the one hand, reforms reduce support for a revolt along the extensive margin. In particular, in the limit where regimes reform to a full-scaled democracy, any threat of revolt is completely dissolved. On the other hand, if reforms signal that a regime is vulnerable, they may backfire by increasing support along the intensive margin.

Stage 1: Reforms by insiders We now turn to the insiders' decision problem. Since more vulnerable regimes have higher incentives to reform than less vulnerable ones, conducting reforms will be associated with being intrinsically weak and, therefore, indeed increases coordination along the intensive margin. For the benefits along the extensive margin to justify these costs, reforms have to be far-reaching, inducing regimes to enfranchise a large proportion of the population whenever they conduct reforms. The following proposition summarizes the resulting equilibrium schedule of reforms.

Proposition 2. *In any equilibrium, policy choices of insiders and beliefs of outsiders are uniquely[10] characterized by time-invariant functions $x : (\theta_t, \lambda_t) \mapsto x_t$, $\xi : \theta_t \mapsto \xi_t$, $\hat{\theta} : (\lambda_t, x_t) \mapsto \hat{\theta}_t$, and $\bar{\theta} : \lambda_t \mapsto \bar{\theta}_t$, such that*

$$x(\theta_t, \lambda_t) = \begin{cases} \lambda_t & \text{if } \theta_t < \bar{\theta}(\lambda_t) \\ \xi(\theta_t) & \text{if } \theta_t \geq \bar{\theta}(\lambda_t), \end{cases}$$

---

[10]Uniqueness results from the D1 equilibrium refinement. However, even without D1 reforms are always weakly increasing, starting from a strictly positive pool at $x_t = \lambda_t$ and having a discontinuity at the marginally reforming regime $\bar{\theta}_t$. Accordingly, the D1 refinement merely pins down a unique shape of ξ—i.e., x_t conditional on that there is a reform (see the proof for details)—ensuring global uniqueness of reforms.

and

$$\hat{\theta}(\lambda_t, x_t) = \begin{cases} \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}(\lambda_t)\} & \text{if } x_t = \lambda_t \\ \bar{\theta}(\lambda_t) & \text{if } \lambda_t < x_t < \xi(\bar{\theta}(\lambda_t)) \\ \xi^{-1}(x_t) & \text{if } \xi(\bar{\theta}(\lambda_t)) \leq x_t \leq \xi(1) \\ 1 & \text{if } x_t > \xi(1), \end{cases}$$

where $\xi' > 0$ with $\xi(\theta_t) > \lambda_t$ for all $\theta_t > \bar{\theta}(\lambda_t)$ and $\bar{\theta}(\lambda_t) > 0$ for all λ_t.

Proposition 2 defines insiders' policy choices for generation t as a function of (θ_t, λ_t). Because the logic behind these choices is the same for all values of λ_t, we can discuss the underlying intuition keeping λ_t fixed. To this end, Figure 1 plots reform choices (left panel) and the implied probability to be overthrown (right panel) for a given λ_t. It can be seen that whenever a regime is less vulnerable than $\bar{\theta}(\lambda_t)$, insiders prefer to not conduct any reforms (i.e., $x_t = \lambda_t$), leading to a substantial threat for regimes with θ_t close to $\bar{\theta}(\lambda_t)$. Only if $\theta_t \geq \bar{\theta}(\lambda_t)$, reforms will be conducted ($x_t = \xi(\theta_t)$), which in equilibrium effectively mitigates the threat to be overthrown, ruling out marginal reforms where $\xi(\theta_t) \to \lambda_t$.

To see why marginal reforms are not effective in reducing revolutionary pressure consider Figure 2. Here we plot equilibrium beliefs (left panel) and the corresponding mass of insurgents (right panel) as functions of x_t. If the political system is left unchanged by insiders, outsiders only know the average state $\hat{\theta}_t^{\text{pool}} \equiv \mathbb{E}\{\theta | \theta \leq \bar{\theta}(\lambda_t)\}$ of all regimes that pool on $x_t = \lambda_t$ in equilibrium. On the other hand, every extension of the regime—how small it may be—leads to a non-marginal change in outsiders' beliefs from $\hat{\theta}_t^{\text{pool}}$ to $\hat{\theta}_t \geq \bar{\theta}(\lambda_t)$ and, hence, results in a non-marginal increase in revolutionary pressure along the intensive margin. It follows that there exists some $\tilde{x}(\lambda_t)$, such that for all $x_t < \tilde{x}(\lambda_t)$ the increase of pressure along the intensive margin dominates the decrease along the extensive margin. Thus, reforms smaller than $\tilde{x}(\lambda_t)$ will backfire and *increase* the mass of insurgents (as seen in the right panel of Figure 2), explaining why effective reforms have to be non-marginal.

Furthermore, optimality of reforms requires that the benefit of reducing pressure compensates for insiders' disliking of sharing power. Because $\tilde{x}(\lambda_t) > \lambda_t$, it follows that $u(\tilde{x}(\lambda_t)) < u(\lambda_t)$. Moreover, any reform marginally increasing the regime beyond $\tilde{x}(\lambda_t)$ leads only to a marginal increase in the likelihood to stay in power. Hence, there exists a non-empty interval, given by $[\tilde{x}(\lambda_t), \xi(\bar{\theta}(\lambda_t))]$, in which reforms are effective, yet
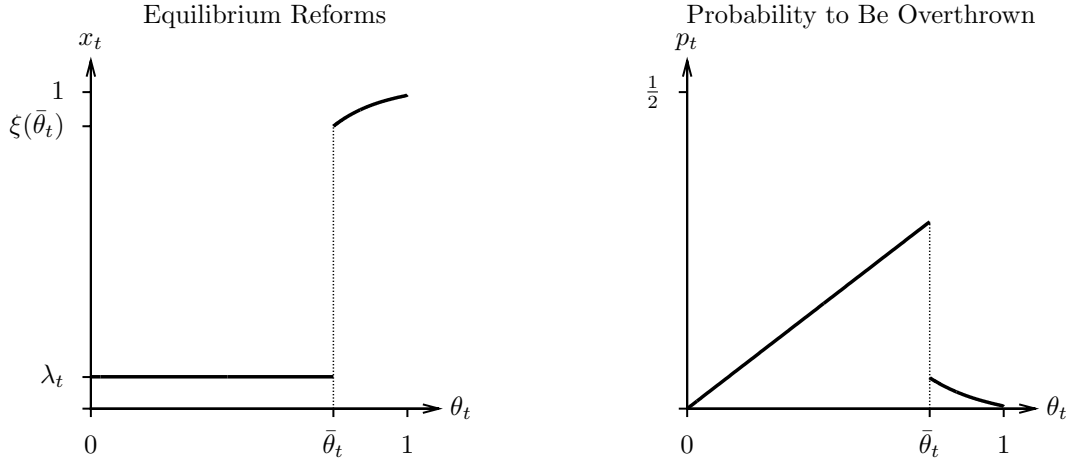
14

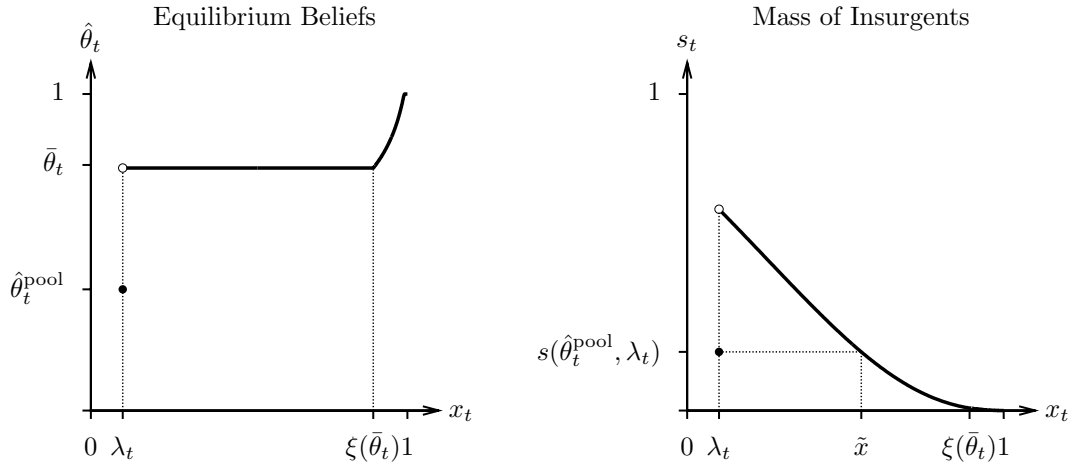**Figure 1.** Equilibrium reforms and implied probability to be overthrown.

Figure 2. Equilibrium beliefs and implied mass of insurgents.

insiders prefer to gamble for their political survival in order to hold on to the benefits of not sharing power in case they survive. This explains the substantial threat for regimes with θ_t close to $\bar{\theta}(\lambda_t)$, as seen in the right panel of Figure 1.[11]

Existence and uniqueness of equilibrium Propositions 1 and 2 uniquely pin down all policy choices in every state, which in turn determine the evolution of political systems. We conclude that there is no scope for multiple equilibria in our model. Verifying that an equilibrium exists, then permits us to reach the following conclusion.

---

[11]More precisely, gambling for survival increases the likelihood to be overthrown in two ways. First, since at the margin more vulnerable regimes join the pool at $x_t = \lambda_t$, these regimes obviously face a high threat by not conducting reforms. Second, since these regimes also shift the pooling belief towards pooling regimes being more vulnerable, the threat further increases for regimes of all vulnerabilities in the pool (reflected by an upward-rotation of the revolt probability for pooling regimes around zero).

Proposition 3. *There exists an equilibrium, in which for all histories δ, policy mappings x_δ and $\{\phi_{i\delta}\}_{i=0}^1$, as well as beliefs $\hat{\theta}_\delta$ correspond to the time-invariant mappings underlying Propositions 1 and 2. Furthermore, for any given initial political system λ_0, the equilibrium is unique.*

3 Likelihood and outcomes of transitions

We begin our analysis of regime dynamics by investigating the properties of political transitions conditional on the current regime λ_t (and conditional on the exogenous prior of outsiders). By Proposition 3, policy mappings are time-invariant, implying that (λ_t, θ_t) is a sufficient statistic for characterizing the transition dynamics of the political system from time t to $t+1$. Integrating out θ_t, political systems in the unique equilibrium follow a Markov process $Q(\lambda_t, \Lambda)$ where Q is the probability that $\lambda_{t+1} \in \Lambda$ in state λ_t. To qualitatively and quantitatively explore this process, We decompose Q into observable statistics as follows:

$$Q(\lambda_t, \Lambda) = \rho^S(\lambda_t) \times Q^S(\lambda_t, \Lambda) + \rho^R(\lambda_t) \times Q^R(\lambda_t, \Lambda)$$
$$+ \{1 - \rho^S(\lambda_t) - \rho^R(\lambda_t)\} \times \mathbb{1}_{\lambda_t \in \Lambda}. \quad (7)$$

Here, ρ^S and ρ^R are the probabilities that in state λ_t a transition occurs via revolts or reforms; Q^S and Q^R are conditional transition functions (specifying the probability that in state λ_t the system $\lambda_{t+1} \in \Lambda$ emerges from a revolt or reform); and $\mathbb{1}$ is an indicator function equal to unity whenever $\lambda_t \in \Lambda$.[12] Accordingly, the first term in (7) defines the probability that system $\lambda_{t+1} \in \Lambda$ emerges through a revolt, the second term defines the probability that $\lambda_{t+1} \in \Lambda$ emerges from a reform, and the third term indicates the event of no transition.

---

[12]Formally, let $\tilde{s}(\lambda, \theta) \equiv s(\hat{\theta}(\lambda, x(\lambda, \theta)), x(\lambda, \theta))$ and $\tilde{p}(\lambda, \theta) \equiv p(\theta, \tilde{s}(\lambda, \theta))$. Then:

$$\rho^S(\lambda) = \int_0^1 \tilde{p}(\lambda, \theta) \, \mathrm{d}F(\theta) \qquad Q^S(\lambda, \Lambda) = \rho^S(\lambda)^{-1} \int_{\{\tilde{s}(\lambda,\theta) \in \Lambda\}} \tilde{p}(\lambda, \theta) \, \mathrm{d}F(\theta)$$

$$\rho^R(\lambda) = \int_{\tilde{\theta}(\lambda)}^1 (1 - \tilde{p}(\lambda, \theta)) \, \mathrm{d}F(\theta) \qquad Q^R(\lambda, \Lambda) = \rho^R(\lambda)^{-1} \int_{\{x(\lambda,\theta) \in \Lambda \setminus \{\lambda\}\}} (1 - \tilde{p}(\lambda, \theta)) \, \mathrm{d}F(\theta).$$

Parametrization   Before proceeding, let us introduce a simple parametric example, which we will use for all simulations throughout the paper. For this, let

$$u(\lambda_t) = -\exp(\lambda_t \beta_1) + \beta_0 \quad \text{and} \quad h(s_t) = s_t^{\alpha}.$$

Here one may think of β_0 as a common resource stock or some other private benefits that decline at an exponential rate β_1 as power is shared with more insiders. The parameter β_1 hence measures the costs of enfranchising political outsiders. In practice, we expect these costs to be high for resource-rich and less developed economies.[13] To reduce the number of free parameters, further suppose that $\psi'(1) = 0$; i.e., the strategic effect of an additional outsider supporting a revolt becomes negligible when revolts are supported by the full population. Together with our assumptions on u and h, this pins down α and β_0 in terms of β_1, which is restricted to approximately satisfy $\beta_1 \in (0, 0.56)$.[14]

While the qualitative predictions of our model are similar for different values of β_1, the magnitude of β_1 will control the relative frequency of reforms to revolts. For β_1 close to its lower bound, reforms are essentially free and there will be a single, comprehensive reform at $t = 0$ that leads to an (almost) universal democracy. For β_1 close to its upper bound, reforms will be prohibitively expensive and regime dynamics will be exclusively driven by revolts. For all illustrations throughout the main paper, we set β_1 to an intermediary value of 0.385 that is chosen to roughly match the empirical distribution of autocratic regimes relative to democratic ones (see Section 6 for details). Alternative parametrization are discussed in Appendix D.

In the remainder of this section we explore the properties of ρ^S, ρ^R, Q^S and Q^R and their dependence on the exogenous prior of outsiders. While it highlights some of the forces that eventually shape dynamics in the endogenous prior model, readers mainly interested in the dynamics under endogenous priors may wish to skip to Section 4.

Outcomes of transitions   Using (7), the type of political systems that emerge from transitions are defined by the conditional transition functions Q^S and Q^R. From Proposition 2 it is clear that reforms in state λ_t will be bounded below by $\xi(\bar{\theta}(\lambda_t))$,

---

[13]In particular, we expect that in modern production economies with strong labor complementarities and high capital returns a commitment to honor property rights by enfranchising outsiders generates positive effects on aggregate income. In particular, property rights are likely to encourage enfranchised outsiders to acquire human capital, to supply high-skilled labor, or to invest their savings. We therefore expect u' to be small when elites profit from these benefits, mitigating the costs of sharing power.

[14]In particular, $\alpha = \beta_1 \exp(\beta_1)$ and $\beta_0 = \exp(\beta_1) + 1$, restricting $\beta_1 \in (0, \exp(-\beta_1)) \approx (0, 0.56)$.
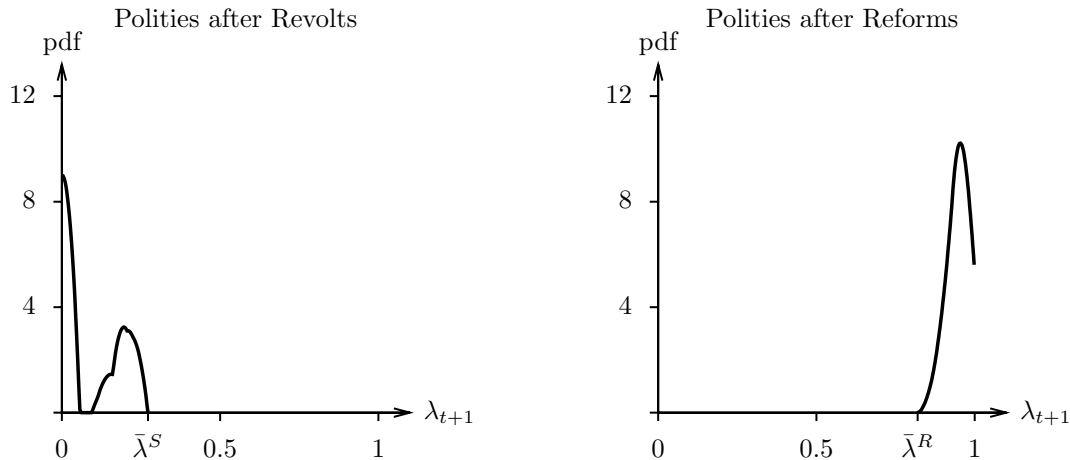
Polities after Revolts   Polities after Reforms

Figure 3. Distribution of political systems after revolts and reforms (uniform prior).

since smaller reforms would be ineffective in reducing revolutionary pressure as they would increase coordination of outsiders along the intensive margin. Accordingly, there is an interval $[0, \bar{\lambda}_\lambda^R]$ that is not attained by reforms originating in state λ; i.e., $Q^R(\lambda, [0, \bar{\lambda}_\lambda^R]) = 0$. Similarly, revolts cannot grow too large, since otherwise insiders would prefer to preempt them.[15] In particular, effectiveness of reforms implies that revolts will be bounded above by $s(\hat{\theta}_t^{\text{pool}}, \lambda_t)$, implying $Q^S(\lambda, [\bar{\lambda}_\lambda^S, 1]) = 0$ for some $\bar{\lambda}_\lambda^S$.

In a previous version of this paper (Buchheim and Ulbricht, 2014, Proposition 4), we have shown for uniform F that $\bar{\lambda}_\lambda^R > 1/2$ and $\bar{\lambda}_\lambda^S < 1/2$ for all λ. That is, regardless of the originating regime, reforms would always result in regimes where a majority of citizens holds political power, whereas revolts would always result in autocracies with a small elite ruling over a majority of political outsiders. Figure 3 illustrates the uniform case using the parametric example introduced above. Here we plot the conditional distributions of political systems that emerge from reforms and from revolts.[16]

From the left panel, it becomes apparent that approximately two types of autocracies emerge after revolts: dictatorships, corresponding to regimes that emerge after revolts against democracies, and autocracies which emerge after succeeding other non-democratic regimes. The right panel, in turn, displays the distribution of political

---

[15]More precisely, insiders would preempt large revolts if they are vulnerable. Accordingly, outsiders know the regime to be resistant in the absence of reforms, so that joining a revolt becomes risky and only outsiders with high potential gains from revolting—i.e., low realizations of γ_i—are willing to take the risk. This further reduces the chances of success, reducing the support even more, *etc.*

[16]The reported distributions weight the transition functions, $Q^S(\lambda_t, \cdot)$ and $Q^R(\lambda_t, \cdot)$, with the invariant distribution of λ_t. E.g., letting Ψ denote the invariant distribution, the distribution of political systems after reforms is given by $\int_0^1 Q^R(\lambda_t, \lambda_{t+1}) \, d\Psi(\lambda_t)$.
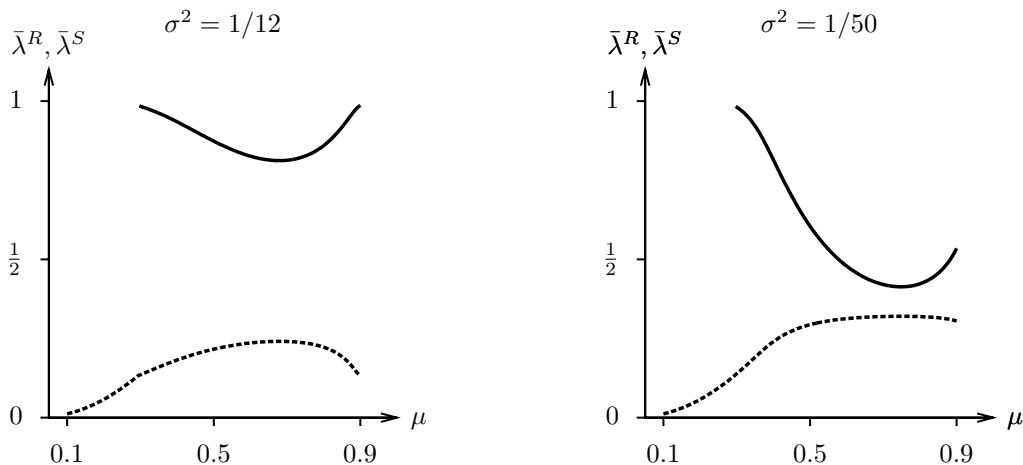
Figure 4. Smallest reforms (solid) and largest revolts (dashed) for different distributions of θ_t.

systems after reforms, which only has positive weight on fairly democratic political systems. Clearly visible, there is a set of intermediate political systems, reaching from $\bar{\lambda}^S \approx 0.22$ to $\bar{\lambda}^R \approx 0.87$, that do neither emerge from reforms, nor from revolts.

While the finding that $\bar{\lambda}^R_\lambda > 1/2$ and $\bar{\lambda}^S_\lambda < 1/2$ does not generalize to non-uniform distributions of θ, the observed polarization pattern generally pertains. Figure 4 illustrates this for the case where θ is drawn from a Beta distribution. The figure plots the smallest reforms (solid lines) and largest revolt (dashed lines) that may occur along the equilibrium path, whereas each point on the curves represents a different parametrization of the Beta distribution defining F. In particular, the left panel shows how the bounds change for different unconditional means of θ, holding the variance of θ fixed at its uniform value of $1/12$ (e.g., for $\mu = 1/2$, the plot shows $\lambda^S \approx 0.22$ and $\lambda^S \approx 0.87$ as seen in Figure 3[17]). The right panel repeats the exercise for a more precise distribution of θ where we set the variance to $1/50$.

In both panels, for low values of μ, outsiders' expect the regime to be strong, implying small revolts and the absence of reforms ($\bar{\theta} = 1$). As prior mass is shifted towards intermediate levels of θ_t, the impact of revealing θ_t by conducting reforms on beliefs becomes smaller, so that conducting small reforms becomes increasingly attractive for sufficiently high levels of θ_t. Eventually, as prior mass is shifted towards unity (and to the right side of $\bar{\theta}$), the impact of conducting reforms on beliefs becomes again larger, so that again only large reforms are effective in reducing revolutionary pressure. This explains the U-shape of $\bar{\lambda}^R$ and inverse U-shape of $\bar{\lambda}^S$. Comparing the

---

[17]The uniform distribution is a special case of the Beta distribution with shape parameters $a = b = 1$ or, equivalently, with moments $\mu = 1/2$ and $\sigma^2 = 1/12$.

left panel to the right, it can be seen that for more precise priors, both larger revolts and smaller reforms are feasible along the equilibrium path whenever the regime is *a priori* likely to be vulnerable.

An interesting implication of these results is that democracies tend to arise only by means of reforms. By contrast, even the largest revolts typically lead to at most intermediate-sized regimes that require further reforming in order to become fully democratic. In that sense the commonly made assumption in the previous literature that democracies are established by means of reforms is an endogenous outcome in our model. In Section 7 we will see that this is also largely in line with the data.

Likelihood of transitions   Along with the conditional transition functions, regime dynamics are defined by the likelihoods of reforms and revolts. In the model, revolutionary pressure becomes naturally negligible as regimes become fully inclusive ($\lim_{\lambda \to 1} \rho^S(\lambda) = 0$), which in turn reduces incentives to reform ($\lim_{\lambda \to 1} \bar{\theta}(\lambda) = 1$).[18] These forces are amplified by the intensive margin of revolutionary pressure which, conditional on beliefs, reinforces any change in the extensive margin. That is, as regimes become more inclusive and prospective support for revolts falls, also incentives to support a revolt decline, leading to correspondingly less supporters and hence even less revolutionary pressure. In sum, there is a strong force of stabilization for democratic regimes.

Yet, the likelihood of political transitions for democracies is not necessarily zero. This is because the limit case of a fully inclusive regime might never emerge along the equilibrium path. In general, whether or not $\lambda = 1$ is emerging along the equilibrium path depends on whether $\xi(1) = 1$ or $\xi(1) < 1$. Whenever $\xi(1) < 1$, there always remains a small fraction of outsiders that in principle is willing to participate in subversive attempts, implying a small but positive probability of a regime reversal. In our example we have $\xi(1) = 0.988$, so that there is indeed a small probability of observing reversals— even for the largest feasible democracy along the equilibrium path.[19] In the next sections, we will demonstrate that young democracies are particular prone to such reversals, once we allow for θ_t to persist from one period to the next.

In contrast to democracies, the likelihood of transitions is generally bounded away

---

[18]To see this, observe that $\bar{\theta}(\lambda)$ defines the largest $\bar{\theta}$ such that $\mathbb{E}\{V^I(\eta, \lambda)|\bar{\theta}, s = s(\mathbb{E}\{\theta|\theta \leq \bar{\theta}\}, \lambda)\} \geq \mathbb{E}\{V^I(\eta, \xi(\bar{\theta}))|\bar{\theta}, s = s(\bar{\theta}, \xi(\bar{\theta}))\}$. Since both sides of the inequality are continuous in $\bar{\theta}$ and λ, so is $\lambda \mapsto \bar{\theta}$, implying the result as it clearly holds that $\bar{\theta}(1) = 1$.

[19]Evaluated at the invariant distribution of polities, the implied probability of a regime reversal against democracies is 0.42 percent per period.
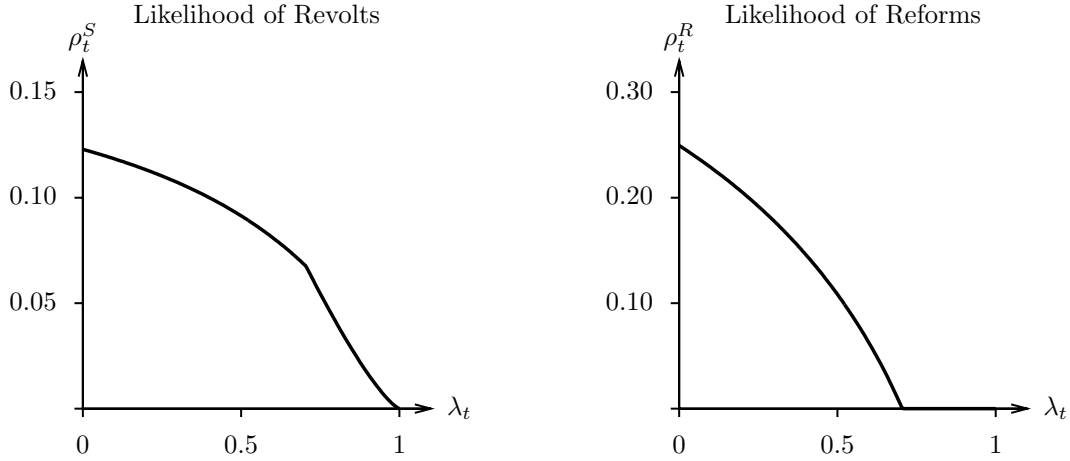
Figure 5. Likelihood of revolts and reforms for different political systems λ_t (uniform prior).

from zero for autocracies. On the one hand, a large fraction of outsiders poses significant threats, which due to the discussed incentives of small regimes to "gamble for their survival" translate into sizable equilibrium revolts. On the other hand, the flipside of large equilibrium threats is that once regimes become sufficiently vulnerable ($\theta > \bar{\theta}(\lambda)$), they face strong incentives to conduct reforms.[20]

Figure 5 illustrates these points, plotting the transition likelihoods of revolts ρ^S and reforms ρ^R as a function of the current regime type λ_t. Both mappings are decreasing, so that autocracies are significantly more likely than democracies to experience a transition of either type. Accordingly, autocracies are on average relatively short-lived due to their high transition probabilities. (This last point will be qualified in the next section, where autocracies can become stable once they become sufficiently mature.)

Nevertheless, even if individual autocracies are short-lived, there is a tendency for autocracies to persist *across* regimes. This is because after a revolt against an autocratic regime, the succeeding regime will be very similar to its predecessor as becomes evident from Figure 3. Hence, while the identity of autocratic leaders may change frequently over time, autocratic *systems* tend to be persistent across regimes.

Finally, consider the impact of distributional shifts in the prior towards more vulnerable states. Similar to how such a shift affects Q^S and Q^R, it also increases the likelihood of transitions: If a regime appears to be immune to revolts, outsiders consider it indeed unattractive to revolt; accordingly, the regime has no incentives to reform. As a regime is perceived to be more vulnerable, both the probability of revolts and

---

[20]For priors F which place a lot of mass on stable regimes, reforms will be completely off the equilibrium path; i.e., $\bar{\theta}_t(\lambda) = 1$ for all $\lambda \in [0, 1]$ (c.f., the left panel of Figure 4).
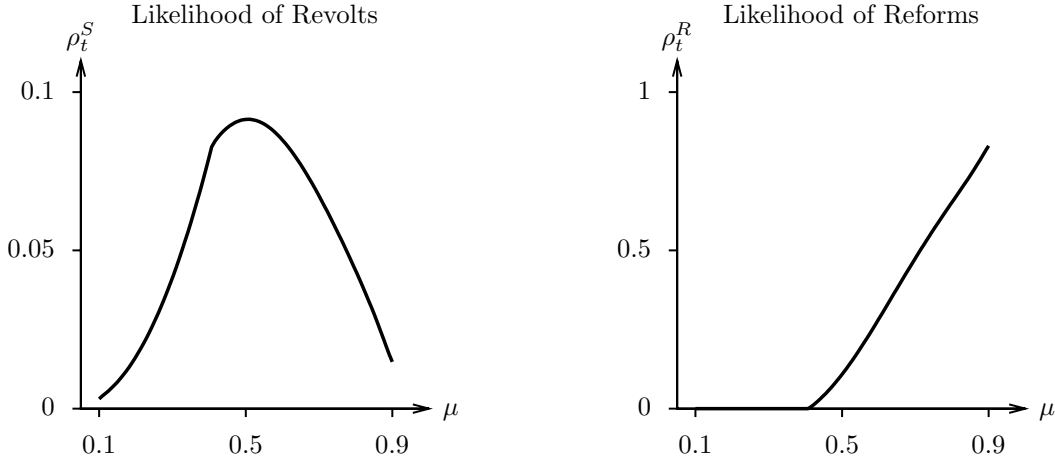
Figure 6. Likelihood of revolts and reforms for different distributions of θ_t. The likelihoods are plotted for $\lambda = 0.5$ and $\sigma^2 = 1/12$. The likelihoods for other values of λ and σ are similar in shape, whereas magnitudes decrease in σ and λ.

reforms initially increase until eventually the threat grows so large that the regime conducts inclusive reforms in almost every state and revolts disappear in equilibrium. Importantly, due to a regime's incentive to gamble for its survival, there is an interim region, where both revolts and reforms co-exists with significant probabilities. Figure 6 illustrates this relation.

4   The model with endogenous priors

The preceding analysis suggests that the prior beliefs of outsiders are an important determinant for the likelihood of transitions and their outcome. So far we have exogenously specified the prior of outsiders by imposing a distribution for θ_t that is i.i.d. across time. We now relax this assumption, allowing the institutional characteristics underlying a regime's vulnerability to persist from one period to the next. Assuming that outsiders learn the public history of transitions when they are born, this will naturally give rise to fluctuations in the prior beliefs of outsiders over time.[21]

Let $F(\theta_t|\delta_{t-1})$ be the exogenous cdf of θ_t for a given history δ_{t-1}, and let $\delta_t^p \equiv \{\lambda_0\} \cup \{x_\tau, s_\tau, \eta_\tau\}_{\tau=0}^t$ denote the publicly observable partition of the history at the end of date t. Accordingly, the conditional belief $\theta_t|\delta_{t-1}^p$ defines the prior of outsiders born at date t. One technical challenge is that priors of generation t, $\theta_t|\delta_{t-1}^p$, will generally

---

[21]Here we use the term "prior" to refer to the beginning-of-period t beliefs regarding θ_t, which is a combination of the unconditional prior about θ_t implied by F and the information inferred from observing the public history of transition up to date $t-1$.

not conjugate with the ones of generation $t - 1$, making it difficult to keep track of beliefs in a dynamic setting. We address this challenge by backward-engineering F, such that the prior of outsiders is always Beta-distributed.

In particular, let $\hat{\mu}_t$ and $\hat{\sigma}_t^2$ denote the first two moments of $\theta_t | \delta_t^p$—i.e., the belief regarding θ_t conditional on information available at the beginning of $t + 1$. Similarly, let μ_{t+1} and σ_{t+1}^2 denote the first two moments of outsiders' prior $\theta_{t+1} | \delta_t^p$ at date $t + 1$. In general, to obtain the latter, one would form $\theta_t | \delta_t^p$ using Bayes rule and then use the exogenous distribution F to project the prior regarding θ_t onto θ_{t+1}. Our assumption is that F is such that the prior of outsiders can be parametrized by a Beta distribution, so that

$$\mu_{t+1} = \pi \hat{\mu}_t + (1 - \pi)\mu_0 \tag{8}$$

$$\sigma_{t+1}^2 = \pi \hat{\sigma}_t^2 + (1 - \pi)\sigma_0^2 + \pi(1 - \pi)(\hat{\mu}_t - \mu_t)^2 \tag{9}$$

for some $\pi \in (0, 1)$ and $\mu_0, \sigma_0 > 0$ where $\sigma_0^2 < \mu_0(1 - \mu_0)$. Intuitively this states that the first two moments of outsiders' beliefs evolve *as if* the state θ_{t+1} is left unchanged with probability π, and is otherwise drawn from a fixed distribution with mean μ_0 and variance σ_0^2.[22]

Learning dynamics   Introducing learning affects the equilibrium dynamics by adding the prior moments μ_t and σ_t^2 as additional state variables to the Markov process defined in (7). Conditional on (μ_t, σ_t^2) the previous equilibrium characterization in Section 2.2 remains fully valid.

In particular, it holds that for any prior (μ_t, σ_t^2), the transition process at date t is described by the (previously time-invariant) versions of Q^S, Q^R, ρ^S and ρ^R that would arise when the previously exogenous prior is replaced by the now endogenous prior (i.e., the Beta-cdf with mean and variance (μ_t, σ_t)). The equilibrium dynamics are therefore completely determined by (7) and the law of motion for μ_t and σ_t^2 that is implicit in (8) and (9).

---

[22]Note that given our specification of outsiders' priors there is no need to pin down the precise shape of F. One possible way for F to implement the specified beliefs, would be to set $F(\cdot | \delta_t)$ equal to the cdf of a Beta distribution with moments equal to μ_{t+1} and σ_{t+1}^2. With such a choice of F, outsiders could trivially infer the current distribution of θ_{t+1} from observing δ_t^p, whereas by design the resulting first two prior moments are consistent with the moments that *would* follow from agents invoking Bayes law in the presence of the described mixture process for θ_t. An alternative interpretation would be to let θ_t indeed follow a mixture process, but use an approximation similar to Krusell and Smith (1998) by matching the first two moments of outsiders' beliefs to a Beta distribution.

Appendix C provides a detailed characterization of μ_t and σ_t^2 as a function of $(\mu_{t-1}, \sigma_{t-1}^2)$ and the events at date $t-1$. Intuitively, μ_t is small when there is no transition event and is high after a transition is observed. Specifically, reforms and revolts against reforming regimes fully reveal the state θ_t, which conditional on a reform is larger than $\bar{\theta}_{t-1}$, so that $\mu_t > \pi\bar{\theta}_{t-1} + (1-\pi)\mu_0$. Similarly, Bayesian updating implicates that the regime is likely to be vulnerable when a revolt is observed in the absence of reforms. In contrast, when neither a reform nor a revolt are observed, Bayesian updating implies that θ_t is likely to be low. In sum, for a given state $(\lambda_{t-1}, \mu_{t-1}, \sigma_{t-1}^2)$, it holds that

$$\mu_t|(\text{reform}_{t-1}) \geq \mu_t|(\text{revolt}_{t-1}) > \mu_t|(\text{no transition}_{t-1}). \tag{10}$$

Regarding σ_t^2, we have that uncertainty is smallest after reforms and revolts against reforming regimes. The ordering of σ_t^2 between no transition event and revolts in the absence of reforms depends on the precise prior distribution.

5 Regime dynamics

We are now ready to explore the emerging dynamics of political regimes. Relative to the conditional properties explored in Section 3, learning across periods now adds an implicit dependence of transition likelihoods and outcomes on the current regime's maturity. The next subsection characterizes this relation, before discussing the emerging transition patterns in Subsection 5.2.

5.1 Regime maturity and likelihood of transitions

In Section 3 we have seen that regimes that are perceived to be vulnerable (large μ_t) are more likely to face reforms or revolts than regimes that are perceived as invulnerable. Combining this with (10), young regimes tend to be less stable than more mature ones conditional on λ_t.

Figure 7 illustrates this distinction between young and mature regimes, plotting the likelihoods of reforms ρ_t^R (left panel) and revolts ρ_t^S (right panel) as a function of λ_t and conditional on whether there was a reform (solid), a revolt (dashed), or no transition event (thin dotted lines) at date $t-1$.[23] It is apparent how the likelihood of either

---

[23]The figure is plotted using the same parametrization introduced in Section 3. Throughout, the
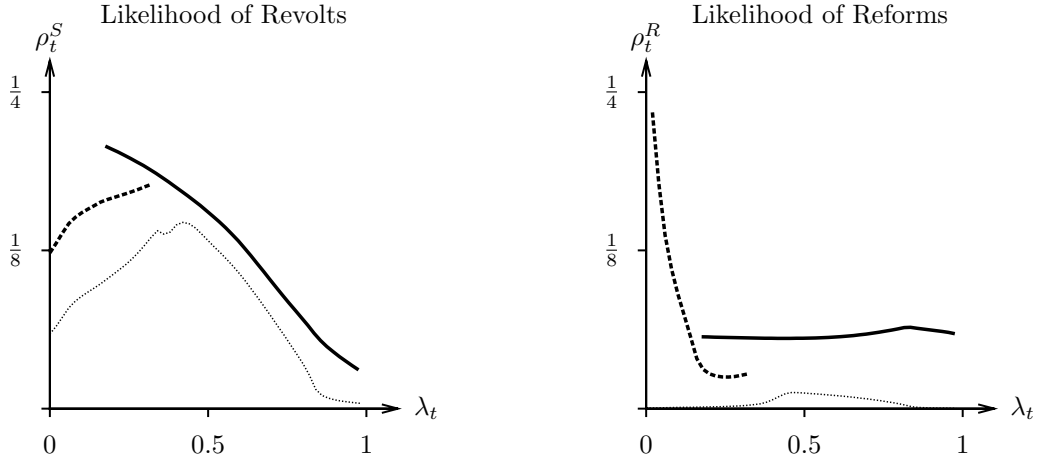
Figure 7. Likelihood of revolts (left panel) and reforms (right panel) conditional on λ_t and conditional on that there was a reform (solid line), revolt (dashed line), or no transition event (thin dotted line) at date $t-1$.
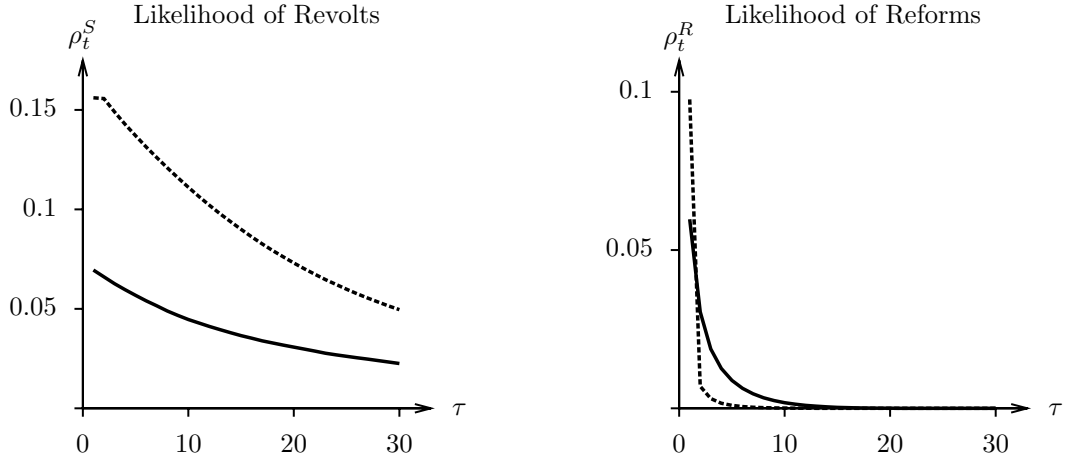

Figure 8. Likelihood of revolts (left panel) and reforms (right panel) conditional on the maturity τ of the current regime and conditional on that the regime was originally established via reform (solid line) or via revolt (dashed line).

transition type is higher immediately after a transition compared to when the regime was already in place the previous period (see below for an intuition about the shapes).

Figure 8 further illustrates this point by averaging the likelihoods across λ_t and plotting them against the maturity of a regime. It can be seen that the hazard rates for either transition type are generally decreasing in the maturity of the regime in line

learning parameters are set to $\pi = 0.99$, $\mu_0 = 2/5$ and $\sigma_0 \approx 1/4$ (or, equivalently, $a_0 = 1, b_0 = 3/2$), implying a highly persistent distribution of θ_t with a slow drift towards a moderate vulnerability of $2/5$. All distributions and conditional likelihoods in Figures 7–12 are evaluated at the invariant equilibrium distribution over regimes and priors.
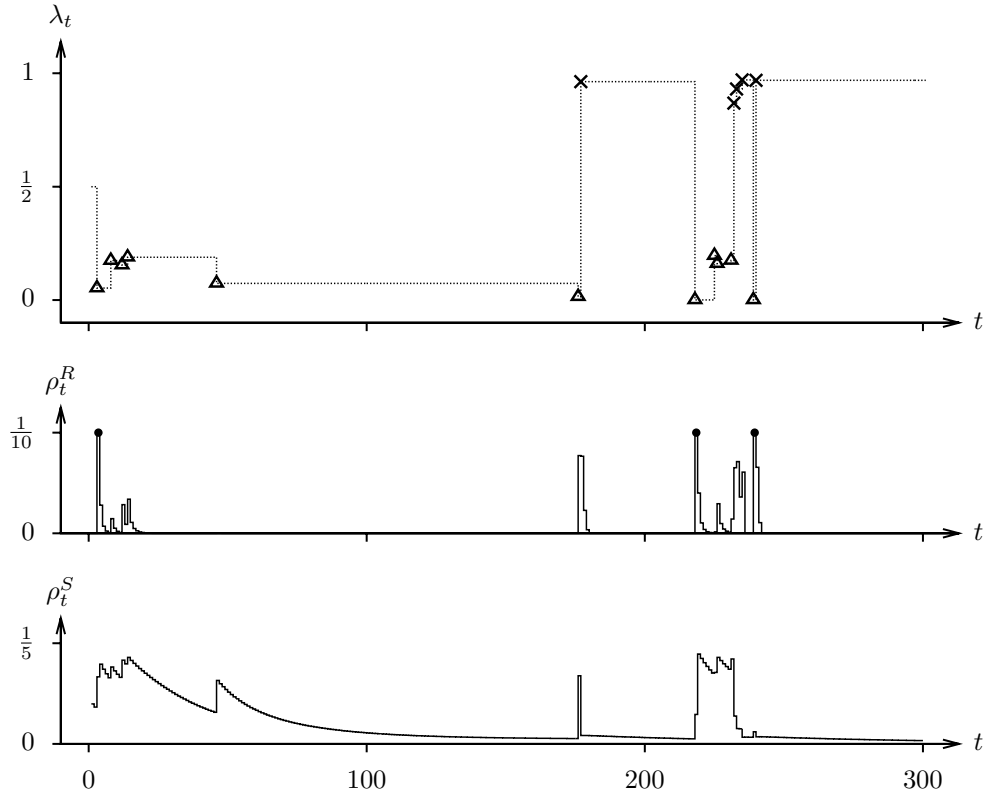
25

Figure 9. Simulated time series of the model with learning. *Notes:* Reforms are marked by "×", successful revolts are marked by "△". Middle and lower panels display the hazard rates of reforms and revolts. Black dots in the graph of ρ^R indicate truncation of 0.35, 0.80 and 0.96 at 0.1, respectively.

with empirical findings that regimes become more stable as they age (see Section 7 and, e.g., Bienen and van de Walle, 1989, 1992; Svolik, 2008). In our model, this is because outsiders' prior means μ_t converge towards μ_0 in each period without a transition event, while observing any type of transition is a signal of political vulnerability.

5.2 Patterns of regime changes

We now discuss the types of transition patterns that can emerge in equilibrium. To have an example at hand, we simulated the model to generate a random time series of 300 periods. Figure 9 shows the resulting regime dynamics. The top panel plots the political system, λ_t, at time t and indicates the dates where transitions occur via revolts (marked by Δ) and reforms (marked by ×). The middle and bottom panel further plot the corresponding hazard rates, ρ^R and ρ^S.

Critical junctures From Figure 9 it is evident how transition events tend to be clustered across time, giving rise to episodes of political stability (periods 45–173, 176–215 and 239 to end) and episodes of political turbulence that are characterized by rapid successions of revolts, counter revolts, and reforms (2-44, 174–175 and 216–238). This clustering is a direct consequence of the decline of hazard rates in a regime's maturity.

In the model, political turbulences are triggered by small probability transition events against mature regimes.[24] Once triggered, such episodes pose critical junctures in the sense that small and random variations in current conditions may cause persistent differences in future political systems (e.g., notice how the similar looking critical junctures starting in period 2 and period 216 eventually lead to a stable autocracy and democracy, respectively). On the one hand, this is because the outcome of political turbulences is largely determined by the random variables θ_t and η_t and further hinges on small variations in the current state λ_t, μ_t and σ_t^2 (e.g., whether $\bar{\theta}(\lambda_t, \mu_t, \sigma_t^2)$ is slightly above or below θ_t). On the other hand, because democracies and autocracies both stabilize once they become mature, the type of political system that eventually survives an episode of political turbulence is likely to persist for a long time.

Gradual reforms and counter revolts From the previous discussion it follows that newly established democracies first go through a phase of instability, before they eventually stabilize. Specifically, young democracies face heightened threats of counter revolts (see the solid line in the right panel of Figure 7 in comparison to the dotted line). While fully consolidated democracies tend to be stable, "transitional" democracies are therefore prone to regime reversals (period 237 in Figure 9).

The flipside of these regime reversals is that young democracies have strong incentives to conduct further reforms, in particular if the initial reform was small (see solid lines in the right panels of Figures 7 and 8). The model can thus generate patterns of gradual democratization (periods 230–233 in Figure 9) that are similar to the predictions by Jack and Lagunoff (2006), Justman and Gradstein (1999), and Gradstein (2007).[25]

---

[24]More generally, political turbulences are ultimately caused in our model by a change in outsiders' sentiment that causes them to perceive the regime to be more vulnerable. Accordingly, critical junctures could also be triggered by exogenous shifts in outsiders' beliefs, for instance caused by the deaths of political leaders which are known to increase the likelihood of transitions (Jones and Olken, 2009; Besley, Persson and Reynal-Querol, 2014). While it would be straightforward to incorporate the possibility of such belief shifts into our model, we abstract from this possibility for the sake of simplicity.

[25]Underlying the possibility for gradual enfranchisement in our model is that reforms reveal the current vulnerability of the regime, increasing coordination along the intensive margin. Accordingly,

Revolutions and democratization Based on the model, what are viable paths to democracy? We have seen in Figure 4 that the most inclusive political systems are always established via reform. Yet, in the model with learning, revolts can become quite large in some states of the world.[26] In particular after counter revolts against reforming regimes, the newly emerged autocracy is known to be vulnerable with high precision, helping outsiders to coordinate and thus leading to large revolts (c.f., the right panel of Figure 4).

However, even though there are large revolts, regimes that emerge from revolts are far away from being inclusive and, in addition, are known to be vulnerable. Whether or not a major revolt may ultimately lead to democratization therefore depends on whether the resulting regime chooses to reform before it falls to a counter revolt. As it turns out, the probability of observing a reform after a revolt is decreasing in the size of the preceding revolt (the dashed line in the right panel of Figure 7), reflecting again that more inclusive regimes are less vulnerable to revolutionary threats.

As the flipside of a downward sloping probability for reforms is an upward sloping likelihood of revolts (the dashed line in the left panel of Figure 7), it is, ultimately, unlikely that a large revolt leads to democratization. This is consistent with, e.g., the observation of Karl (1990, p. 8) that no stable South American democracy has been the result of mass revolutions (see also Rustow, 1970; O'Donnell and Schmitter, 1973; Huntington, 1991). Below, we will provide further empirical evidence that mass revolutions leading to democracies are indeed the rare exception.

Iron law of oligarchy Finally, the characterization above implies that regimes consolidate their power as they mature, meaning that μ_t converges to its baseline value μ_0 over time. Once μ_t is sufficiently low, however, regimes prefer to abstain from reforms for (almost) all instability levels and the probability of reforms drops (close) to zero (c.f., right panel of Figure 8). Accordingly, the likelihood of a "direct" route to democratization is low and (mature) autocracies are bound to be eventually succeeded

---

the likelihood of survival of future regimes becomes highly sensitive to adverse changes in θ_t, providing strong incentives to conduct further reforms in such an event. Our model thus provides a completely different rational for gradual enfranchisement compared to the literature on dynamic voting games. In these models, the current median voter needs to delegate power to another median voter to implement her preferred policy, but chooses a path of gradual enfranchisement due to the anticipation of further franchise extensions conducted by her "preferred" median voter.

[26]In our parametrization the largest revolt along the equilibrium path is supported by approximately one third of the entire population. While equally sized revolutions are also possible when θ_t is i.i.d., this would require a parametrization in which transitions occur almost exclusively via revolts.
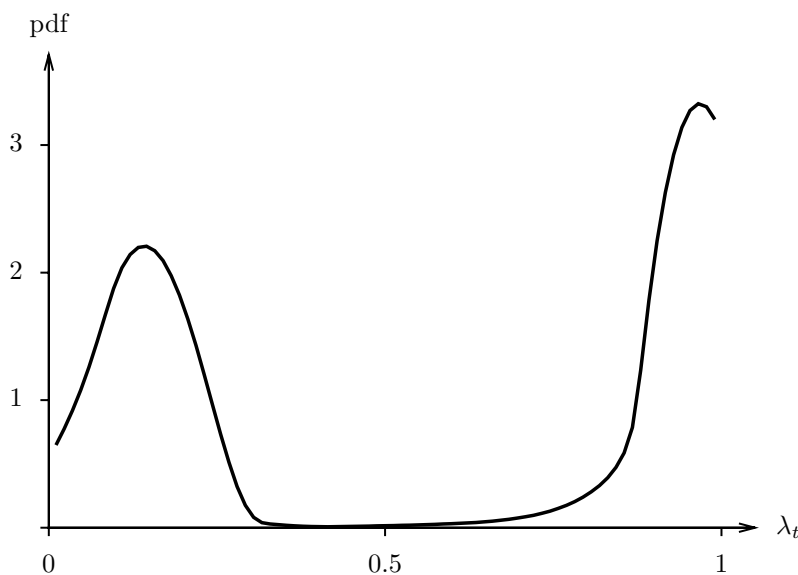
Figure 10. Invariant distribution of political systems in the model with learning.

by another autocratic regime—a pattern that is sometimes dubbed as the "iron law of oligarchy". In the context of our model, the path to democratization therefore necessarily leads through a critical juncture, whereas mature regimes are bound to fall—if ever—to a revolt (periods 174 and 216 in Figure 9).[27]

6   Long-run distribution of political systems

We now discuss the distribution of political systems that should be expected in the long-run. The key implication of both the model with and without learning is that the long-run distribution of political systems is bimodal with mass being concentrated on the extremes. Figure 10 displays the long-run distribution for the model with learning. Similar results hold for the model without learning and are discussed in Appendix D.[28] Specifically, it can be seen that our parametrization (see Section 3) yields about equal mass on both autocratic and democratic regimes. While different parametrizations shift mass between autocratic and democratic systems, the overall bimodal shape is always

---

[27]In line with our discussion in Footnote 24, critical junctures are not necessarily tied to transition events. A shift in beliefs may thus well lead to enfranchisement directly out of a state of oligarchy.

[28]While both versions of the model yield qualitatively similar results, the model without learning requires parametrizations with significant higher frequencies of revolts in order to generate nontrivial mass on the autocratic side of the political spectrum. Since learning introduces a mechanism for autocracies to stabilize, the model naturally gives rise to more mass on the autocratic spectrum, while keeping the overall frequency of revolts low.

preserved.

To see what is underlying the bimodal shape recall from Section 3 that the model gives rise to a polarization of political regimes during their emergence and a tendency for both autocratic and democratic *systems* to persist across time. For an illustration consider Figures 11 and 12 where we display the conditional distributions of newly emerging regimes and the marginal likelihood of reforms and revolts as a function of λ_t (the counterparts to Figures 3 and 5 in our baseline model). Clearly, the model with learning shows the same type of polarization forces discussed in Section 3, which lays out the grounds for the bimodal shape of the long-run distribution.[29]

The initial polarization of regimes is reinforced by the higher persistence of extreme political systems as compared to intermediate ones, which is driven by two effects. First, as in the baseline model, repeated successions of autocratic regimes via revolts introduce a persistence of autocratic *systems* that exceeds the stability of individual autocratic regimes. Second, the model with learning also gives rise to hump-shaped likelihoods of reforms and revolts as visible in Figure 12. This is because intermediate political systems are statistically most likely to inherit an intermediate vulnerability from their predecessor, since predecessors with low values of θ_t are unlikely to transform and predecessors with high values of θ_t are choosing farther-reaching reforms. Due to their intermediate stability and their limited inclusiveness, these regimes face a substantial threat of revolts and, as a consequence, have high incentives to reform.

7   A look at the data

Our model makes a number of predictions about statistics that describe the dynamics of political systems: the frequency of transitions conditional on the current political system and its age, the political systems emerging from transitions, as well as the distribution of political systems in the long-run. While these predictions are based on a simple framework that emphasizes the importance of asymmetric information and learning for political dynamics, they are rich enough to allow for a sensible quantitative comparison with aggregate data. In this section, we contrast the model's predictions with data on regime dynamics for the majority of countries from 1919 onwards. While we make no claim on causality, we find that the model is able to capture important

---

[29]Due to the nature of the time-varying priors in the model with learning and the resulting shifts in Q^R and Q^S the outcomes, in particular for reforms, are more smooth and have wider support compared to the model without learning.
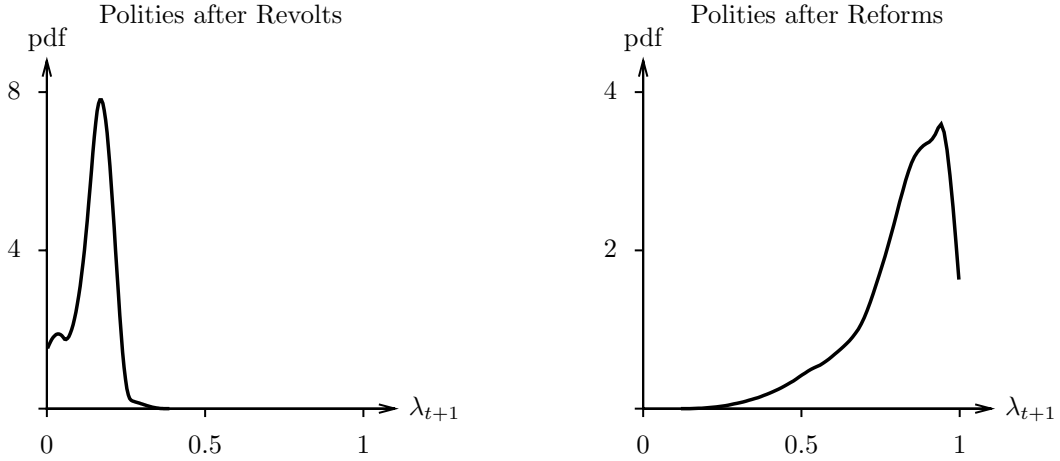
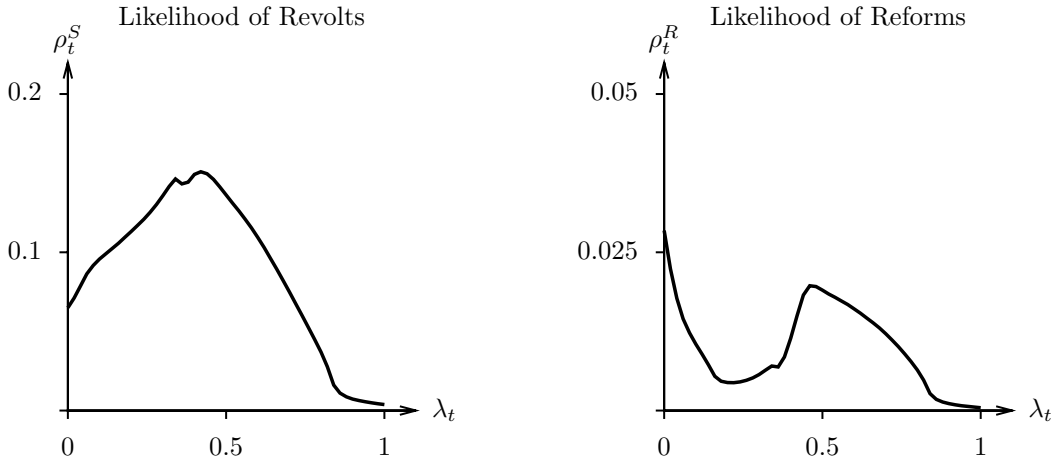**Figure 11.** Distribution of political systems after revolts and reforms in the model with learning.

Figure 12. Likelihood of revolts and reforms for different regimes λ_t in the model with learning.

features of regime dynamics in the data.

7.1 Data

To construct empirical counterparts to the model statistics that characterize the regime dynamics, we combine information on the inclusiveness of political systems (the empirical counterpart to λ_t) with information on transition events. As a measure for the inclusiveness, we use the *polity* variable, scaled to $[0, 1]$, from the Polity IV Project (Marshall and Jaggers, 2002), which ranks political regimes on a 21 point scale between autocratic and democratic.[30]

---

[30]In contrast to our model, the *de facto* distribution of political power may sometimes differ from the *de jure* scope of the franchise. For this reason, we use the polity index to measure political power as it

To classify successful revolts, we use the Archigos Dataset of Political Leaders (Goemans, Gleditsch and Chiozza, 2009). The dataset is available for the years between 1919 and 2004, which defines the overall time frame of our panel. We record a successful revolt if a leader is irregularly removed from office due to domestic popular protest, rebel groups, or military actors (defined by Archigos' *exitcodes* 2, 4 and 6), and if at the same time the leader's successor takes office in irregular manner (defined by an *entrycode* 1). Furthermore, we take a revolt to be causal for a change in the political system if a change in the political system is recorded in the Polity IV database within a two week window of the revolt.[31]

Finally, we use the dataset on the Chronology of Constitutional Events from the Comparative Constitution Project (Elkins, Ginsburg and Melton, 2010) to classify reforms. We define reforms by a constitutional change (*evnttype* equal to *new*, *reinstated*, or *amendment*) accompanied by a change in the political system (as indicated by the variable *durable* from the Polity IV Project) that is not matched to a revolt or another irregular regime change from the Achigos Dataset. To be consistent with the model's definition of reforms, we restrict attention to positive changes.

The resulting dataset is a daily panel on the country level, which covers 175 countries and records 251 revolts and 97 reforms.

7.2 Empirical properties of political systems and transitions

At the aggregate, the long-run distribution generated by our model matches the shape of the empirical distribution of political systems in our dataset (shown in Figure 13). Comparing Figure 13 with Figure 10, both distributions are bimodal, with mass concentrated mainly on autocratic and democratic political systems.

The model identifies two forces underlying the bimodal shape of the long-run distribution. First, political transitions are subject to polarization: Reforms establish predominantly democratic political systems, while revolts mainly establish autocracies. Figure 14—the empirical counterpart to Figure 11—shows that political systems emerg-

attempts to measure the *de facto* scope of the franchise—which is the relevant aspect of the model's concept of political power—using various proxies regarding the openness of political institutions to participation. Alternatives based on a *de jure* interpretation of constitutions such as data on the formal right to vote from the Comparative Constitution Project (Elkins, Ginsburg and Melton, 2010) or data on electoral turnout from the Polyarchy Index (Vanhanen, 2000) are less useful for this purpose.

[31]To prevent the data from generating a mechanical correlation between revolts and small polity scores, we use the polity score at the end of periods indicated by the Polity IV Project to be "transitionary" to code the outcome of transitions.
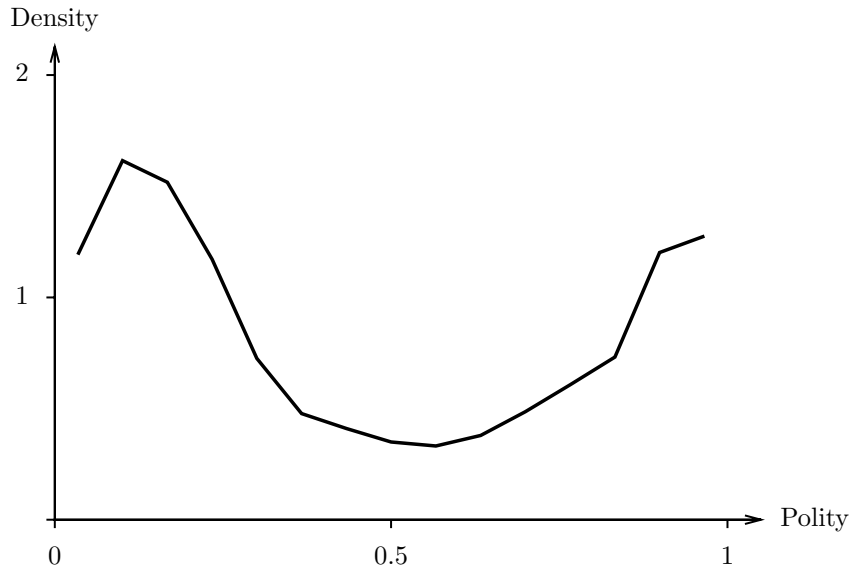
Figure 13. Distribution of political systems since World War I.

ing from revolts (left panel) and reforms (right panel) share a similar shape with their theoretical counterparts. In particular, revolts are indeed by and large autocratic, while the modal political system established via reforms is democratic (right panel).

Second, democratic and autocratic regimes are more stable than intermediate types of political systems, reflected in theoretical transition likelihoods that are hump-shaped in the inclusiveness of the political systems (Figure 12). Figure 15 shows that a similar pattern can be seen in the data.

In the presence of learning, our theory also predicts a negative correlation between transitions and the maturity of a regime. Analogously to Figures 7 and 8, Figures 16 and 17 relate the empirical frequency of transitions to the maturity of regimes. Specifically, Figure 16 displays local polynomial estimates of the empirical likelihoods of either transition for young regimes (between two and five years of age) conditional on being established via reform (solid lines) or revolt (dashed lines), as well as for mature regimes (older than five years; dotted lines).[32] It can be seen that transitions are far more frequent when there was a recent transition compared to the case when a regime is mature. Similarly, Figure 17 relates the frequency of reforms and revolts directly to the age of a regime, showing that consistent with the theoretical predictions the frequencies are (i) decreasing, with (ii) similar levels for reforms (right panel) and (iii) higher

---

[32]To avoid conclusions from very small samples, we restrict the support of young regimes to systems for which we observe the equivalent of at least 12 country-years.
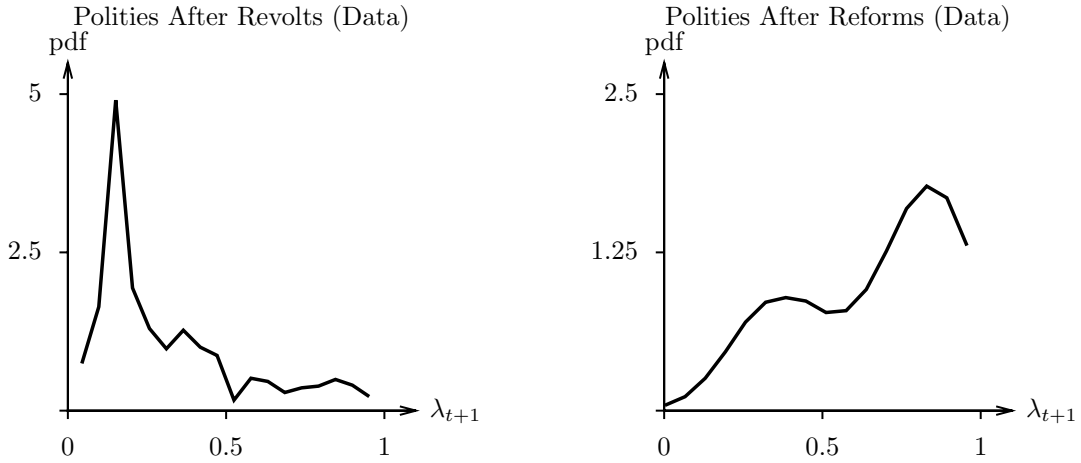
33

Polities After Revolts (Data)

pdf

Polities After Reforms (Data)

pdf

Figure 14. Empirical distribution of political systems after revolts and reforms.

Likelihood of Revolts (Data)

ρ_t^S
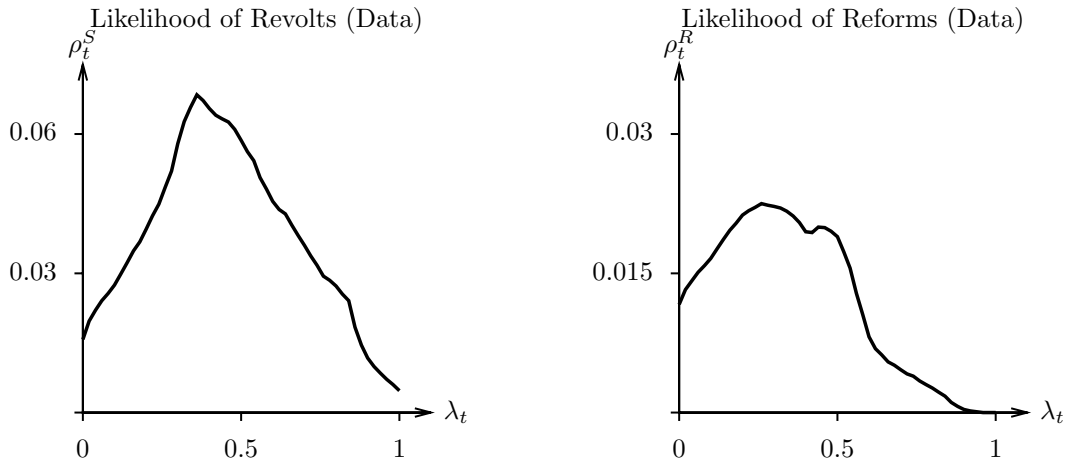
Likelihood of Reforms (Data)

ρ_t^R

Figure 15. Annual empirical likelihood of revolts and reforms.

probabilities to observe revolts after revolts than after reforms (left panel).

In sum, the model does a remarkably good job in matching central observable properties of the data, even at a quantitative level.

One objection to this conclusion, however, could be that the descriptive statistics shown so far are mainly driven by the cross-sectional variation in the data, while the model predicts these to arise through transition dynamics of individual countries over time. In the remainder of this section, we hence complement our graphical analysis by panel regressions including year and country fixed effects and show that our results do not rely on either times of global political instability or country specific effects.

Columns 1 and 2 of Table 1 substantiate the finding that revolts and reforms lead to a polarization of political systems. Specifically, Column 1 shows that revolts lead, on
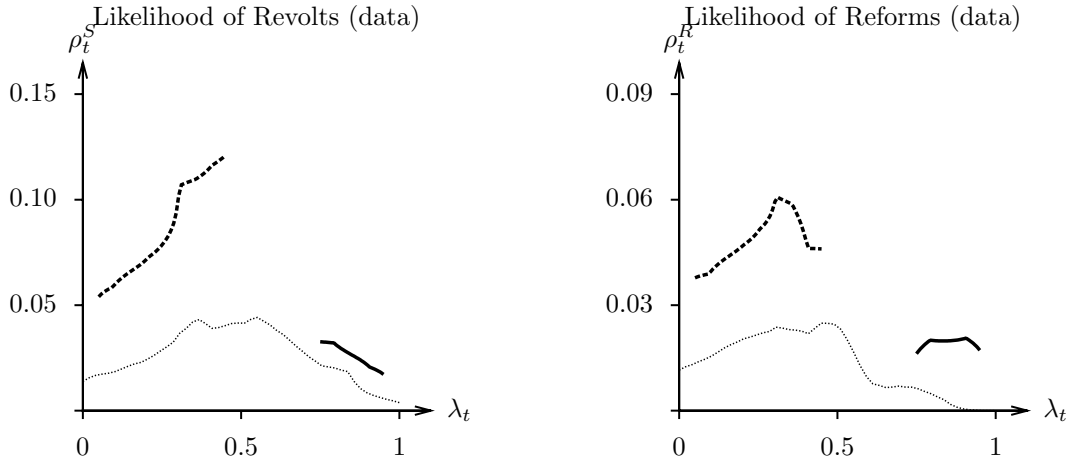
Figure 16. Annual empirical likelihood of revolts (left panel) and reforms (right panel) for political systems between two and five years of age established via reform (solid lines) or revolt (dashed lines) as well as for political systems five years or older (dotted). See the text for details.
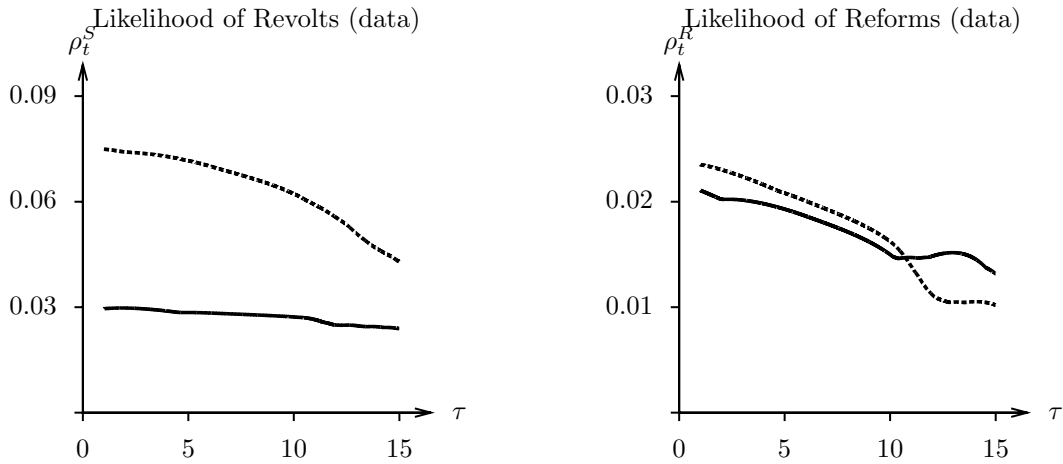

Figure 17. Annual empirical likelihood of revolts (left panel) and reforms (right panel) for political systems established via reform (solid lines) or revolt (dashed lines) as function of regime age in years.

average to a decrease in the polity index of 0.06 points, while reforms lead to an increase of 0.39 points. Column 2 further dissects these average effects by conditioning on the polity index of the originating regime. It can be seen that revolts against all regimes with an index value greater than $0.2/0.7 \approx 0.29$ have a negative effect on the future polity with large reverting effects against democratic regimes. Reforms, in contrast, do not affect the political system by much when they are conducted within already democratic societies ($0.56 - 0.59 \times \text{polity}_t = 0$ for $\text{polity}_t = 0.94$), but represent a major push towards democracy otherwise. In the absence of transition events, the current political system has naturally no impact on political change.

Table 1. Empirical results controlling for country and year fixed effects

Dependent Var.	(1) $\Delta\,\text{Polity}_t$	(2) $\Delta\,\text{Polity}_t$	(3) Revolt_t	(4) Reform_t	(5) Revolt_t	(6) Reform_t	(7) Revolt_t	(8) Reform_t
Revolt_t	−0.062***	0.202***						
	(0.016)	(0.028)						
Reform_t	0.386***	0.564***						
	(0.024)	(0.038)						
$\text{Polity}_t \times \text{Revolt}_t$		−0.699***						
		(0.082)						
$\text{Polity}_t \times \text{Reform}_t$		−0.594***						
		(0.066)						
Polity_t	−0.000***	−0.000***	0.002	−0.052***	0.196***	0.009		
	(0.000)	(0.000)	(0.017)	(0.008)	(0.059)	(0.027)		
Polity_t^2					−0.200***	−0.062**		
					(0.056)	(0.028)		
$\text{Polity}_t \times \text{Revolt-Episode}$							0.083*	−0.052***
							(0.043)	(0.014)
$\text{Polity}_t \times \text{Reform-Episode}$							−0.044	−0.071**
							(0.027)	(0.030)
Age in 25 yrs			−0.001	−0.005**	−0.002	−0.005**	−0.001	−0.003
			(0.004)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
				Control variables				
Episode-Dummy	No	No	No	No	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes.— All parameters estimated via OLS. Number of observations are $3\,289\,400$ country-days. Standard errors clustered at the country level are reported in parentheses. The dependent variable in Columns 1 and 2 is the change in the polity index between date t and $t+1$. The dependent variables in Columns 3 to 8 are dummies indicating whether a revolt or reform is observed at date t. Revolt-Episodes (Reform-Episodes) indicate regimes that are established via revolt (reform). Coefficients and standard errors in Columns 3–8 are multiplied by 365.25 to indicate annual likelihoods. Level of significance: * p<0.1, ** p<0.05, *** p<0.01.

In the remaining columns of Table 1 we investigate the proximate determinants of the likelihood of regime transitions identified by the model. First, note that, in general, transitions become less likely with the age of the regime. However, this effect is statistically significant for the probability of reforms only, owing possibly to the limited number of regimes observed for each country.

Second, Columns 3 and 5 confirm that the probability of revolts is a hump-shaped function of the polity index with its maximum at the center of the polity scale, as shown in the left panel of Figure 14. Columns 4 and 6 verify the observation from the right panel of Figure 14 that the probability of reforms has a similar shape but a somewhat stronger negative slope overall.

Finally, we examine the more subtle prediction of the model that autocratic regimes—

typically established via revolts—are more inclined to gamble for political survival if they represent a larger share of the population, while democratic regimes—established via reforms—become uniformly more stable the more inclusive they are. This prediction implies that the probability of revolt is increasing in the polity index for regimes that originate from a revolt and decreasing in the polity index for regimes originating from a reform. The probability of reform, in contrast, should be declining in the inclusiveness of the regime in revolt-episodes and be approximately constant in reform-episodes (Figure 7). The results in Column 7 and 8 of Table 1 show that, with the exception of the constant reform hazard in reform episodes, these predictions can be reconciled with within-country variation in the frequency of regime transitions. Overall, the panel regressions thus confirm the close match between even quite subtle predictions of our model and the data.

8   Concluding remarks

This paper explores regime dynamics in a simple model where transitions (or their absence) are governed by signaling and learning considerations. Although simple in its nature, the model provides a unified framework of political transitions, combining all the principal transition scenarios previously considered in isolation. In particular, the model accounts for (possibly gradual) democratization processes, regime reversals against both transitional and mature democracies, and power struggles amongst autocratic regimes.

The model predicts a number of central properties of regime dynamics. First, political dynamics are characterized by a Markov process where the likelihood of transitions decreases in a regime's age, giving rise to extended periods of political stability that alternate with politically turbulent times. The model is thus able to explain the high failure rate of transitional democracies as well as why—even though autocracies tend to be overall less stable than democracies—some autocratic leaders have been in power for a long time. Second, the model emphasizes that political transitions lead to a polarization of political regimes explaining the bimodal distribution of political systems in the data, which is reinforced by transition likelihoods being hump-shaped in the political system in place (conditional on a regime's age). Third, the model also predicts that for mature regimes, the only transitions that occur with positive probability are revolts establishing autocracies. This result gives an underpinning to what is sometimes called the "iron law of oligarchy".

An important feature of the model is the quantitative character of its predictions. Exploiting these, we also provide a first assessment of whether the specific features that shape the model dynamics are present in the data. To this end, we dissect the process of political systems into various conditional statistics, which we compare to their empirical counterparts using data on political systems and transitions. Even though the model is rather stylized, its predictions are remarkably close to the observations from the data.

The good empirical fit suggests that the model may represent a useful foundation for more quantitative studies of regime dynamics. For instance, it may be worth to explore potential microfoundations for the gains from political power u, the opportunity cost of revolting γ, or the benefits of coordination h. Relating these primitives of the model to, e.g., the presence of resource rents, the degree of economic development, or communication technologies the model could be used to examine differences in regime dynamics across geography or time.

References

Acemoglu, Daron, and James A. Robinson. 2000a. "Democratization or Repression?" *European Economic Review*, 44: 683–693.

Acemoglu, Daron, and James A. Robinson. 2000b. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics*, 115(4): 1167–1199.

Acemoglu, Daron, and James A. Robinson. 2001. "A Theory of Political Transitions." *American Economic Review*, 91(4): 938–963.

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty.* New York:Crown Publishers.

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni. 2010. "A Theory of Military Dictatorships." *American Economic Journal: Macroeconomics*, 2(1): 1–42.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2008. "Coalition Formation in Non-Democracies." *Review of Economic Studies*, 75(4): 987–1009.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2012. "Dynamics and Stability of Constitutions, Coalitions, and Clubs." *American Economic Review*, 102(4): 1446–1476.

Acemoglu, Daron, Georgy Egorov, and Konstantin Sonin. 2016. "Political Economy in a Changing World." *Journal of Political Economy*, forthcoming.

Aidt, Toke S., and Peter S. Jensen. 2014. "Workers of the World, Unite! Franchise Extensions and the Threat of Revolution in Europe, 1820–1938." *European Economic*

Review, 72: 52–75.

Aidt, Toke S., and Raphaël Franck. 2015. "Democratization Under the Threat of Revolution: Evidence From the Great Reform Act of 1832." *Econometrica*, 83(2): 505–547.

Angeletos, George-Marios, Christian Hellwig, and Alessandro Pavan. 2006. "Signaling in a Global Game: Coordination and Policy Traps." *Journal of Political Economy*, 114(3): 452–484.

Atkeson, Andrew. 2000. "[Rethinking Multiple Equilibria in Macroeconomic Modeling]: Comment." *NBER Macroeconomics Annual*, 15: 162.

Bai, Jinhui H., and Roger Lagunoff. 2011. "On the Faustian Dynamics of Policy and Political Power." *Review of Economic Studies*, 78(1): 17–48.

Barberà, Salvador, Michael Maschler, and Jonathan Shalev. 2001. "Voting for Voters: A Model of Electoral Evolution." *Games and Economic Behavior*, 37(1): 40–78.

Besley, Timothy, Torsten Persson, and Marta Reynal-Querol. 2014. "Resilient Leaders and Institutional Reform: Theory and Evidence." *Working Paper*.

Bienen, Henry, and Nicolas van de Walle. 1989. "Time and Power in Africa." *The American Political Science Review*, 83(1): 19.

Bienen, Henry, and Nicolas van de Walle. 1992. "A Proportional Hazard Model of Leadership Duration." *The Journal of Politics*, 54(3): 685–717.

Boix, Carles. 2003. *Democracy and Redistribution.* Cambridge (UK):Cambridge University Press.

Boix, Carles, and Milan W. Svolik. 2013. "The Foundations of Limited Authoritarian Government: Institutions, Commitment, and Power-Sharing in Dictatorships." *The Journal of Politics*, 75(2): 300–316.

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

Buchheim, Lukas, and Robert Ulbricht. 2014. "Emergence and Persistence of Extreme Political Systems." *TSE Working Paper*, 464.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review*, 104(03): 446–466.

Cervellati, Matteo, Piergiuseppe Fortunato, and Uwe Sunde. 2012. "Consensual and Conflictual Democratization." *The B.E. Journal of Theoretical Economics (Contributions)*, 12(1).

Cervellati, Matteo, Piergiuseppe Fortunato, and Uwe Sunde. 2014. "Violence During Democratization and the Quality of Democratic Institutions." *European Economic Review*, 66: 226–247.

Cho, In-Koo, and David M. Kreps. 1987. "Signaling Games and Stable Equilibria." *Quarterly Journal of Economics*, 102(2): 179–221.

Chwe, Michael Suk-Young. 2000. "Communication and Coordination in Social Networks." *Review of Economic Studies*, 67(1): 1–16.

Conley, John P., and Akram Temini. 2001. "Endogenous Enfranchisement when Groups' Preferences Conflict." *Journal of Political Economy*, 109(1): 79–102.

Edmond, Chris. 2013. "Information Manipulation, Coordination, and Regime Change." *Review of Economic Studies*, 80(4): 1422–1458.

Elkins, Zachary, Tom Ginsburg, and James Melton. 2010. "Chronology of Constitutional Events, Version 1.1."

Ellis, Christopher J., and John Fender. 2011. "Information Cascades and Revolutionary Regime Transitions." *Economic Journal*, 121(553): 763–792.

Fearon, James D. 2011. "Self-Enforcing Democracy." *Quarterly Journal of Economics*, 126(4): 1661–1708.

Finkel, Evgeny, Scott Gehlbach, and Tricia D. Olsen. 2015. "Does Reform Prevent Rebellion? Evidence From Russia's Emancipation of the Serfs." *Comparative Political Studies*, 48(8): 984–1019.

Francois, Patrick, Ilia Rainer, and Francesco Trebbi. 2015. "How Is Power Shared in Africa?" *Econometrica*, 83(2): 465–503.

Gandhi, Jennifer, and Adam Przeworski. 2007. "Authoritarian Institutions and the Survival of Autocrats." *Comparative Political Studies*, 40(11): 1279–1301.

Goemans, Henk E., Kristian S. Gleditsch, and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders." *Journal of Peace Research*, 46(2): 269–283.

Gradstein, Mark. 2007. "Inequality, Democracy and the Protection of Property Rights." *Economic Journal*, 117(516): 252–269.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology*, 83(6): 1420–1443.

Huntington, Samuel P. 1991. *The Third Wave: Democratization in the Late Twentieth Century.* Norman (OK):University of Oklahoma Press.

Jack, William, and Roger Lagunoff. 2006. "Dynamic Enfranchisement." *Journal of Public Economics*, 90(4): 551–572.

Jones, Benjamin F., and Benjamin A. Olken. 2009. "Hit or Miss? The Effect of Assassinations on Institutions and War." *American Economic Journal: Macroeconomics*, 1(2): 55–87.

Justman, Moshe, and Mark Gradstein. 1999. "The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19th-Century Britain." *Explorations in Economic History*, 36(2): 109–127.

Karl, Terry Lynn. 1990. "Dilemmas of Democratization in Latin America." *Comparative Politics*, 23(1): 1–21.

Krusell, Per, and Anthony A. Jr. Smith. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106(5): 867.

Kuran, Timur. 1989. "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." *Public Choice*, 61(1): 41–74.

Lagunoff, Roger. 2009. "Dynamic Stability and Reform of Political Institutions." *Games and Economic Behavior*, 67: 569–583.

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(2): 707–765.

Llavador, Humberto, and Robert J. Oxoby. 2005. "Partisan Competition, Growth, and the Franchise." *Quarterly Journal of Economics*, 120(3): 1155–1189.

Lohmann, Susanne. 1994. "The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91." *World Politics*, 47(1): 42–101.

Magaloni, Beatriz. 2008. "Credible Power-Sharing and the Longevity of Authoritarian Rule." *Comparative Political Studies*, 41(4-5): 715–741.

Mailath, George J. 1987. "Incentive Compatibility in Signaling Games with a Continuum of Types." *Econometrica*, 55(6): 1349–1365.

Marshall, Monty, and Keith Jaggers. 2002. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2002."

Morris, Stephen, and Hyun Song Shin. 1998. "Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks." *The American Economic Review*, 88(3): 587–597.

O'Donnell, Guillermo, and Philipp C. Schmitter. 1973. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies.* Baltimore (MD):John Hopkins University Press.

Prato, Carlo. 2016. "Electoral Competition and Policy Feedback Effects." *Working Paper.*

Przeworski, Adam. 2009. "Conquered or Granted? A History of Suffrage Extensions." *British Journal of Political Science*, 39(2): 291.

Roberts, Kevin. 1999*a*. "Dynamic Voting in Clubs." *Working Paper.*

Roberts, Kevin. 1999*b*. "Voting in Organizations and Endogenous Hysteresis." *Working Paper.*

Rustow, Dankwart A. 1970. "Transitions to Democracy: Toward a Dynamic Model." *Comparative Politics*, 2(3): 337–363.

Svolik, Milan W. 2008. "Authoritarian Reversals and Democratic Consolidation." *American Political Science Review*, 102(02).

Svolik, Milan W. 2009. "Power Sharing and Leadership Dynamics in Authoritarian Regimes." *American Journal of Political Science*, 53(2): 477–494.

Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy, 1810-1998." *Journal of*

A   Equilibrium characterization (proofs)

A.1   Proof of Proposition 1

We first establish that any solution to the outsiders' coordination problem is a fixed point to equation (6). From our discussion in the main body of the paper it is clear that this is the case if and only if $\bar{\gamma}(\hat{s}_t) \leq 1$ for all $\hat{s}_t$. From Assumption A1 it follows that $\bar{\gamma}$ is increasing in $\hat{s}_t$, and therefore $\bar{\gamma}(\hat{s}_t) \leq 1$ holds if $\bar{\gamma}(1) = p(\hat{\theta}_t, 1) \, u(1) \leq 1$. Since $u(1) = 1$ and $p(\cdot) \in [0, 1]$ this is indeed the case.

Let $f(\hat{s}_t) \equiv (1 - x_t) \, \bar{\gamma}(\hat{s}_t)$. Then, since $f(0) = 0$ for all $(\hat{\theta}_t, x_t) \in \Theta \times [0, 1]$, there always exists a fixed point to (6) at $\hat{s}_t = 0$. When $\hat{\theta}_t = 0$ or $x_t = 1$, then $f(\hat{s}_t) = 0$ for all $\hat{s}_t$, and therefore $\hat{s}_t = 0$ is obviously the only—and therefore stable—fixed point to (6). On the other hand, when $\hat{\theta}_t \neq 0$ and $x_t \neq 1$, then from Assumption A1 $f'(0) > 1$, so that iteratively best responding to any perceived $\hat{s}_t = \varepsilon > 0$ leads to the stable equilibrium $s_t^* > 0$ defined below.

Having ruled out $\hat{s}_t = 0$ as a solution to the coordination problem for $\hat{\theta}_t \neq 0$ and $x_t \neq 1$, we now show that there is a unique, stable $\hat{s}_t > 0$ solving (6) for $\hat{\theta}_t \neq 0$ and $x_t \neq 1$. From $\bar{\gamma} \in [0, 1]$ it follows that f is bounded by its support, $[0, 1 - x_t]$. Moreover, by Assumption A1 we have that $\lim_{\hat{s} \to 0} \psi'(\hat{s}) = \infty$, implying that $\lim_{\hat{s} \to 0} f'(\hat{s}) = \infty$. Hence, there exists a $\tilde{s} > 0$, such that $f(\tilde{s}) > \tilde{s}$. Together with continuity of ψ (and thus of f), it follows that there exists a strictly positive fixed point to (6), which by concavity of ψ (and thus of f) is unique on $(0, 1]$. Let $s_t^* = f(s_t^*)$ denote this fixed point. Clearly, it must hold that $f'(s_t^*) < 1$, and so s_t^* is stable.

The above arguments establish that s_t is uniquely determined by a (time-invariant) function $s : (\hat{\theta}_t, x_t) \to s_t$. It remains to be shown that $\partial s / \partial \hat{\theta}_t \geq 0$ and $\partial s / \partial x_t \leq 0$. Given that s_t is a fixed point to (6), we have that

$$\pi(s_t, x_t) \equiv s_t - (1 - x_t) \, \hat{\theta}_t \, \psi(s_t) = 0.$$

Implicit differentiation implies that

$$\frac{\partial s_t}{\partial x_t} = -\hat{\theta}_t \, \psi(s_t) \times \left(\frac{\partial \pi_t}{\partial s_t} \right)^{-1}$$

and

$$\frac{\partial s_t}{\partial \hat{\theta}_t} = (1 - x_t) \, \psi(s_t) \times \left(\frac{\partial \pi_t}{\partial s_t} \right)^{-1},$$

where

$$\frac{\partial \pi_t}{\partial s_t} = -(1 - x_t)\frac{\partial \bar{\gamma}}{\partial s_t} + 1.$$

Since ψ is bounded by $\psi(1) = 1$, (6) implies that $\lim_{\hat{\theta}_t \to 0} s_t^* = \lim_{x_t \to 1} s_t^* = 0$, and therefore the case where $\hat{\theta}_t = 0$ or $x_t = 1$ is a limiting case of $\hat{\theta} \neq 0$ and $x_t \neq 1$. From the implicit function theorem it then follows that s is differentiable on its whole support. Moreover, the previous arguments imply that $f(\tilde{s}) > \tilde{s}$ for all $\tilde{s} < s_t^*$ and $f(\tilde{s}) < \tilde{s}$ for all $\tilde{s} > s_t^*$, implying that $f'(s_t^*) < 1$ or, equivalently, $\partial\bar{\gamma}/\partial s_t < (1 - x_t)^{-1}$ at s_t^*. Thus $\partial \pi_t/\partial s_t > 0$ for all $(\hat{\theta}_t, x_t) \in \Theta \times [0, 1]$, which yields the desired results.

Finally, while we focus on pure strategies above, it is easy to see that the proposition generalizes to mixed strategies. By the law of large numbers, in any mixed strategy equilibrium, beliefs about s are of zero variance and, hence, the arguments above apply, implying that all outsiders, except a zero mass i with $\gamma_i = \bar{\gamma}(s_t^*)$, strictly prefer $\phi_i = 0$ or $\phi_i = 1$. We conclude that there is no scope for (nondegenerate) mixed best responses.

A.2 Proof of Proposition 2

The proof proceeds by a series of lemmas. To simplify notation, in what follows we drop λ_t as an argument of x and $\hat{\theta}$ where no confusion arises. Furthermore, we use $\tilde{V}^I(\theta_t, \hat{\theta}_t, x_t) = (1 - \theta_t h(s_t))\, u(x_t)$ to denote insider's indirect utility, as follows from $s_t = s(\hat{\theta}_t, x_t)$ given Proposition 1.

Lemma 1. *x is weakly increasing in θ_t.*

Proof. Suppose to the contrary that $x(\theta'') < x(\theta')$ for $\theta' < \theta''$. Let $x' \equiv x(\theta')$, $x'' \equiv x(\theta'')$, $u' \equiv u(x')$, $u'' \equiv u(x'')$, $h' \equiv h(s(\hat{\theta}(x'), x'))$, and $h'' \equiv h(s(\hat{\theta}(x''), x''))$. Optimality of x' then requires that $\tilde{V}^I(\theta', \hat{\theta}(x''), x'') \leq \tilde{V}^I(\theta', \hat{\theta}(x'), x')$, implying $u'h' - u''h'' \leq (u' - u'')/\theta' < (u' - u'')/\theta''$, where the last inequality follows from $\theta' < \theta''$ and $u' < u''$. Hence, $\tilde{V}^I(\theta', \hat{\theta}(x''), x'') \leq \tilde{V}^I(\theta', \hat{\theta}(x'), x')$ implies that $\tilde{V}^I(\theta'', \hat{\theta}(x''), x'') < \tilde{V}^I(\theta'', \hat{\theta}(x'), x')$, contradicting optimality of x'' for θ''. $\square$

Lemma 2. *Suppose x is discontinuous at θ', and define $x^- \equiv \lim_{\varepsilon \uparrow 0} x(\theta' + \varepsilon)$ and $x^+ \equiv \lim_{\varepsilon \downarrow 0} x(\theta' + \varepsilon)$. Then for any $x' \in (x^-, x^+)$, the only beliefs consistent with the D1 criterion are $\hat{\theta}(x') = \theta'$.*

Proof. Let $\theta'' > \theta'$, and let $x'' \equiv x(\theta'')$. Optimality of x'' then requires that $\tilde{V}^I(\theta'', \hat{\theta}(x''), x'') \geq$

$\tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+)$ and, thus for any $\tilde{\theta}$,

$$\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x''), x'') \quad \text{implies that}$$
$$\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+).$$

Moreover, arguing as in the proof of Lemma 1,

$$\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+) \quad \text{implies that}$$
$$\tilde{V}^I(\theta', \tilde{\theta}, x') > \tilde{V}^I(\theta', \hat{\theta}(x^+), x^+).$$

Hence, if $\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+) = \bar{V}^I(\theta'')$, then $\tilde{V}^I(\theta', \tilde{\theta}, x') > \tilde{V}^I(\theta', \hat{\theta}(x^+), x^+) = \bar{V}^I(\theta')$. Therefore, $D_{\theta'',x'}$ is a proper subset of $D_{\theta',x'}$ if $\theta'' > \theta'$. (For the definition of $D_{\theta,x}$, see Footnote 8.) A similar argument establishes that $D_{\theta'',x'}$ is a proper subset of $D_{\theta',x'}$ if $\theta'' < \theta'$ and, thus, the D1 criterion requires that $\hat{\theta}(x') = \theta'$ for all $x' \in (x^-, x^+)$. $\qquad\square$

Lemma 3. *There exists $\bar{\theta}(\lambda_t) > 0$, such that $x(\theta_t, \lambda_t) = \lambda_t$ for all $\theta_t < \bar{\theta}(\lambda_t)$. Moreover, $x(\theta'') > x(\theta') > \lambda_t + \mu$ for all $\theta'' > \theta' \geq \bar{\theta}(\lambda_t)$ and some $\mu > 0$.*

Proof. First, consider the existence of a connected pool at $x_t = \lambda_t$. Because for $\theta_t = 0$, $x_t = \lambda_t$ dominates all $x_t > \lambda_t$, we have that $x(0) = \lambda_t$. It follows that there exists a pool at $x_t = \lambda_t$, because otherwise $\hat{\theta}(\lambda_t) = 0$ and, therefore, $p(\cdot, s(\hat{\theta}(\lambda_t), \lambda_t)) = 0$, contradicting optimality of $x(\theta) > \lambda_t$ for all $\theta > 0$. Moreover, by Lemma 1, x is increasing, implying that any pool must be connected. This proves the first part of the claim.

Now consider $x(\theta'') > x(\theta')$ for all $\theta'' > \theta' \geq \bar{\theta}(\lambda_t)$ and suppose to the contrary that $x(\theta'') \leq x(\theta')$ for some $\theta'' > \theta'$. Since x is increasing, it follows that $x(\theta) = x^+$ for all $\theta \in [\theta', \theta'']$ and some $x^+ > \lambda_t$. W.l.o.g. assume that θ' is the lowest state in this pool. Then Bayesian updating implies that $\theta^+ \equiv \hat{\theta}(x^+) \geq \mathbb{E}\{\theta_t | \theta'' \geq \theta_t \geq \theta'\} > \theta'$ and, therefore, $\tilde{V}^I(\theta', \theta^-, x^+) > \tilde{V}^I(\theta', \theta^+, x^+)$ for all $\theta^- \leq \theta'$. Hence, because θ' prefers x^+ over $x(\theta^-)$, it must be that $x(\theta^-) \neq x^+$ for all $\theta^- \leq \theta'$ and, hence, $x(\theta^-) < x^+$ by Lemma 1. Accordingly, let $x^- \equiv \max_{\theta^- \leq \theta'} x(\theta^-)$. Then from continuity of $\tilde{V}^I$ and $\theta^+ > \theta'$ it follows that there exists an off-equilibrium reform $x' \in (x^-, x^+)$ with $\tilde{V}^I(\theta', \theta', x') > \tilde{V}^I(\theta', \theta^+, x^+)$. Hence, to prevent θ' from choosing x' it must be that $\hat{\theta}(x') > \theta'$. However, from Lemma 2 we have that $\hat{\theta}(x') = \theta'$, a contradiction.

Finally, to see why there must be a jump-discontinuity at $\bar{\theta}(\lambda_t)$ note that $\tilde{V}^I(\bar{\theta}(\lambda_t), \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}(\lambda_t)\}, \lambda_t) = \tilde{V}^I(\bar{\theta}(\lambda_t), \bar{\theta}(\lambda_t), x(\bar{\theta}(\lambda_t)))$; otherwise, there necessarily exists a θ in the neighborhood of $\bar{\theta}(\lambda_t)$ with a profitable deviation to either λ_t or $x(\bar{\theta}(\lambda_t))$. From the continuity of $\tilde{V}^I$ and the non-marginal change in beliefs from $\mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}(\lambda_t)\}$ to $\bar{\theta}(\lambda_t)$ it follows that $x(\bar{\theta}(\lambda_t)) > \lambda_t + \mu$ for all λ_t and some $\mu > 0$. $\qquad\square$

Lemma 4. *x is continuous and differentiable in θ_t on $[\bar{\theta}(\lambda_t), 1]$.*

Proof. Consider continuity first and suppose to the contrary that x has a discontinuity at $\theta' \in (\bar{\theta}(\lambda_t), 1)$. By Lemma 1, x is monotonically increasing in θ_t. Hence, because x is defined on an interval, it follows that for any discontinuity θ', $x^- \equiv \lim_{\varepsilon \uparrow 0} x(\theta')$ and $x^+ \equiv \lim_{\varepsilon \downarrow 0} x(\theta')$ exist, and that x is differentiable on $(\theta' - \varepsilon, \theta')$ and $(\theta', \theta' + \varepsilon)$ for some $\varepsilon > 0$. Moreover, from Lemmas 2 and 3 it follows that in equilibrium $\hat{\theta}(x') = \theta'$ for all $x' \in [x^-, x^+]$. Hence, $\tilde{V}^I(\theta', \theta', x^-) = \tilde{V}^I(\theta', \theta', x^+)$, since otherwise there necessarily exists a θ in the neighborhood of θ' with a profitable deviation to either x^- or x^+. Accordingly, optimality of $x(\theta')$ requires $\tilde{V}^I(\theta', \theta', x') \leq \tilde{V}^I(\theta', \theta', x^-)$ and, thus, $\tilde{V}^I(\theta', \theta', x^-)$ must be weakly decreasing in x. Therefore, $\partial \tilde{V}^I / \partial \hat{\theta}_t < 0$ and $\lim_{\varepsilon' \downarrow 0} \partial \hat{\theta}(x^- - \varepsilon') / \partial x_t > 0$ (following from Lemma 3) imply that $\lim_{\varepsilon' \downarrow 0} \partial \tilde{V}^I(\theta', \hat{\theta}(x^- - \varepsilon'), x^- - \varepsilon') / \partial x_t < 0$. Hence, a profitable deviation to $x^- - \varepsilon'$ exists for some $\varepsilon' > 0$, contradicting optimality of $x(\theta')$.

We establish differentiability by applying the proof strategy for Proposition 2 in Mailath (1987). Let $g(\theta, \hat{\theta}, x) \equiv \tilde{V}^I(\theta, \hat{\theta}, x) - \tilde{V}^I(\theta, \theta', x(\theta'))$, for a given $\theta' > \bar{\theta}(\lambda_t)$, and let $\theta'' > \theta'$. Then, optimality of $x(\theta')$ implies $g(\theta', \theta'', x(\theta'')) \leq 0$, and optimality of $x(\theta'')$ implies that $g(\theta'', \theta'', x(\theta'')) \geq 0$. Letting $a = (\alpha\theta' + (1-\alpha)\theta'', \theta'', x(\theta''))$, for some $\alpha \in [0, 1]$ this implies

$$0 \geq g(\theta', \theta'', x(\theta'')) \geq -g_\theta(\theta', \theta'', x(\theta''))(\theta'' - \theta') - \tfrac{1}{2} g_{\theta\theta}(a)(\theta'' - \theta')^2,$$

where the second inequality follows from first-order Taylor expanding $g(\theta'', \theta'', x(\theta''))$ around $(\theta', \theta'', x(\theta''))$ and rearranging the expanded terms using the latter optimality condition. Expanding further $g(\theta', \theta'', x(\theta''))$ around $(\theta', \theta', x(\theta'))$, using the mean value theorem on $g_\theta(\theta', \theta'', x(\theta''))$, and noting that $g(\theta', \theta', x(\theta')) = g_\theta(\theta', \theta', x(\theta')) = 0$, these inequalities can be written as

$$
\begin{aligned}
0 \geq\ & g_{\hat{\theta}}(\theta', \theta', x(\theta')) + \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} \times [g_x(\theta', \theta', x(\theta')) \\
& + \tfrac{1}{2} g_{xx}(b(\beta))(x(\theta'') - x(\theta')) + g_{\hat{\theta}x}(b(\beta))(\theta'' - \theta')] + \tfrac{1}{2} g_{\hat{\theta}\hat{\theta}}(b(\beta))(\theta'' - \theta') \\
& \geq -[g_{\theta\hat{\theta}}(b(\beta')) + \tfrac{1}{2} g_{\theta\theta}(a)](\theta'' - \theta') - g_{\theta x}(b(\beta'))(x(\theta'') - x(\theta')),
\end{aligned}
$$

for $b(\beta) = (\theta', \beta\theta' + (1-\beta)\theta'', \beta x(\theta') + (1-\beta)x(\theta''))$ and some $\beta, \beta' \in [0, 1]$. Because $\tilde{V}^I$ is twice differentiable, all the derivatives of g are finite. Moreover, continuity of x implies that $x(\theta'') \to x(\theta')$ as $\theta'' \to \theta'$ and, therefore, for $\theta'' \to \theta'$,

$$0 \geq g_{\hat{\theta}}(\theta', \theta', x(\theta')) + \lim_{\theta'' \to \theta'} \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} g_x(\theta', \theta', x(\theta')) \geq 0.$$

By Lemma 3, x and, hence, $\hat{\theta}$ are strictly increasing for all $\theta \geq \bar{\theta}(\lambda_t)$. Arguing similarly as we did to show continuity, optimality of x, therefore, requires that $g_x = \partial \tilde{V}^I / \partial x_t \neq 0$ and, hence,

the limit of $(x(\theta'') - x(\theta'))/(\theta'' - \theta')$ is well defined, yielding

$$\frac{\mathrm{d}x}{\mathrm{d}\theta_t} = -\frac{\partial \tilde{V}^I/\partial \hat{\theta}_t}{\partial \tilde{V}^I/\partial x_t}. \tag{11}$$

$\square$

Lemma 5. $x(\theta_t, \lambda_t) = \xi(\theta_t)$ for all $\theta_t > \bar{\theta}(\lambda_t)$, where ξ is unique and $\partial \xi/\partial \theta_t > 0$.

Proof. From Lemma 4 we have that ξ is differentiable, and by Lemma 3, $\partial \xi/\partial \theta_t > 0$. We thus only need to show that ξ is unique. By the proof to Lemma 4, $dx/d\theta_t$ is pinned down by the partial differential equation (11), which must hold for all $x_t \geq x(\bar{\theta}(\lambda_t))$. Moreover, whenever $\bar{\theta}(\lambda_t) < 1$, in equilibrium $\hat{\theta}(x(1)) = 1$ and, therefore, it obviously must hold that $x(1, \lambda_t) = \arg\max_{x_t} \tilde{V}^I(1, 1, x_t)$, providing a boundary condition for (11). Because $\tilde{V}^I$ is independent of λ_t, it follows that $x(\theta_t, \lambda_t)$ is uniquely characterized by a function, i.e., $\xi : \theta_t \mapsto x_t$, for all $\theta_t \geq \bar{\theta}(\lambda_t)$. $\square$

Lemma 6. $\bar{\theta}(\lambda_t)$ *is unique.*

Proof. Suppose to the contrary that $\bar{\theta}(\lambda_t)$ is not unique. Then there exist $\bar{\theta}'' > \bar{\theta}'$, defining two distinct equilibria for a given λ_t. By Lemma 5, there is a unique $\xi(\theta)$ characterizing reforms outside the pool for both equilibria. Optimality for type $\theta \in (\bar{\theta}', \bar{\theta}'')$ then requires $\tilde{V}^I(\theta, \theta, \xi(\theta)) \geq \tilde{V}^I(\theta, \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}'\}, \lambda_t)$ in the equilibrium defined by $\bar{\theta}'$, and $\tilde{V}^I(\theta, \theta, \xi(\theta)) \leq \tilde{V}^I(\theta, \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}''\}, \lambda_t)$ in the equilibrium defined by $\bar{\theta}''$. However, $\tilde{V}^I(\theta, \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}'\}, \lambda_t) > \tilde{V}^I(\theta, \mathbb{E}\{\theta_t | \theta_t \leq \bar{\theta}''\}, \lambda_t)$, a contradiction. $\square$

This establishes uniqueness of $x(\theta_t, \lambda_t)$, with all properties given by Lemmas 3 and 5, and the corresponding beliefs $\hat{\theta}(\lambda_t, x_t)$ following from Lemma 2 and Bayesian updating. Again, for the purpose of clarity we have established this proposition by focusing on pure strategy equilibria. In the following we outline how the proof generalizes to mixed strategy equilibria; a detailed version of these steps can be attained from the authors on request.

Replicating the proof of Lemma 1, it is trivial to show that if $\tilde{V}^I(\theta', \hat{\theta}(x'), x') = \tilde{V}^I(\theta', \hat{\theta}(x''), x'')$, then $\tilde{V}^I(\theta'', \hat{\theta}(x'), x') < \tilde{V}^I(\theta'', \hat{\theta}(x''), x'')$ for all $\theta' < \theta''$ and $x' < x''$. It follows that (i) supports, $\mathcal{X}(\theta)$, are non-overlapping, and (ii) $\min \mathcal{X}(\theta'') \geq \max \mathcal{X}(\theta')$. Moreover, noting that $\tilde{x}(\theta) \equiv \max \mathcal{X}(\theta)$ has a jump-discontinuity if and only if type θ mixes in a nondegenerate way, (ii) further implies that there can be only finitely many types that mix on the closed interval $[0, 1]$. The logic of Lemmas 2, 3, and 4 then apply, ruling out any jumps of $\tilde{x}$ on $[\bar{\theta}(\lambda_t), 1]$. This leads to the conclusion that at most a mass zero of types (i.e., $\theta_t = \bar{\theta}(\lambda_t)$) could possibly mix in any equilibrium (with no impact on $\hat{\theta}$) and, thus, there is no need to consider any nondegenerate mixed strategies.

46

A.3   Proof of Proposition 3

From the discussion in the main body of the paper it is clear that the equilibrium is uniquely pinned down by the time-invariant mappings given by Propositions 1 and 2 if it exists. Existence requires that the candidate equilibrium also is consistent with the D1 criterion. This is true by construction and can be seen from the proof of Proposition 2 where we apply Lemma 2 to restrict off-equilibrium beliefs, such that $\hat{\theta}$ is necessarily consistent with the D1 criterion.

B   Becoming an insider is optimal

Here we show formally that outsiders have no incentives to ever refuse becoming enfranchised. To show this, we need to show that

$$(1 - p(\cdot, x_t))u(x_t) \geq \max\{\hat{\theta}_t\psi(s_t), \gamma_{it}\}.$$

A lower bound on the utility as an enfranchised insider is $u(1)$, since $x_t = 1$ is in the choice set of insiders; i.e., by revealed preferences it holds that $(1 - p(\cdot, x_t))u(x_t) \geq (1 - p(\cdot, 1))u(1) = u(1)$. When the best outside option is to not support a revolt, the result trivially follows from $u(1) \geq \gamma_{it}$ for all i and t. For the case, where an outsider's best outside option is to revolt, an upper bound on the utility is given by $\psi(1) = h(1)u(1)$, since by Assumption A1 revolts are more rewarding when they have more supporters; i.e., $\hat{\theta}_t\psi(s_t) \leq \psi(s_t) \leq \psi(1)$. Noting that $h(1) \leq 1$ gives the result.

C   Learning dynamics

In this appendix we characterize the evolution of outsiders' priors $\theta_{t+1}|\delta_t^p$, which jointly with (7) define the dynamics in the model with learning.

Given our specification of F, it is sufficient to derive the first two moments for $\theta_t|\delta_t^p$. Once we have $\hat{\mu}_t$ and $\hat{\sigma}_t^2$, we can derive μ_{t+1} and σ_{t+1}^2 from (8) and (9), which then pin down the shape parameters of the prior at date $t + 1$:

$$a_{t+1} = \mu_{t+1}\left(\frac{\mu_{t+1}(1 - \mu_{t+1})}{\sigma_{t+1}^2} - 1\right)$$

$$b_{t+1} = (1 - \mu_{t+1})\left(\frac{\mu_{t+1}(1 - \mu_{t+1})}{\sigma_{t+1}^2} - 1\right)$$

To obtain $\hat{\mu}_t$ and $\hat{\sigma}_t^2$, we need to consider 3 cases. First, whenever insiders conduct reforms $x_t > \lambda_t$, the state is fully revealed so that $\hat{\mu}_t = \theta_t$ and $\hat{\sigma}_t^2 = 0$. The same is true when there is

a revolt against a separating regime. Second, whenever insider abstain from reforms ($x_t = \lambda_t$) and a successful revolt is observed ($\eta_t = 1$), we use Bayes law to compute

$$g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 1) = \frac{\theta h(s_t)}{\int_0^{\bar{\theta}_t} \theta h(s_t)g_{a_t,b_t}(\theta)\,\mathrm{d}\theta}g_{a_t,b_t}(\theta),$$

where $g_{a_t,b_t}(\cdot)$ denotes the prior pdf with shape parameters a_t, b_t, and $g_{a_t,b_t}(\cdot|\theta \leq \bar{\theta}_t, \eta_t = 1)$ is the resulting posterior pdf when conditioning on $(\theta \leq \bar{\theta}_t, \eta_t = 1)$. It follows that

$$\hat{\mu}_t = \int_0^{\bar{\theta}_t} \theta g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 1)\,\mathrm{d}\theta = \frac{M^2_{a_t,b_t}(\bar{\theta}_t)}{M^1_{a_t,b_t}(\bar{\theta}_t)}$$

and

$$\hat{\sigma}_t^2 = \int_0^{\bar{\theta}_t} (\theta - \hat{\mu}_t)^2 g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 1)\,\mathrm{d}\theta = \frac{M^3_{a_t,b_t}(\bar{\theta}_t)}{M^1_{a_t,b_t}(\bar{\theta}_t)} - \hat{\mu}_t^2.$$

Here $M^i_{a_t,b_t}(\bar{\theta}_t)$ denotes the i-th raw moment of the $\bar{\theta}_t$-truncated Beta-distribution,

$$M^i_{a_t,b_t}(\bar{\theta}_t) \equiv \mathbb{E}\{\theta^i|\theta \leq \bar{\theta}_t\} = B(\bar{\theta}_t, a_t + i, b_t)/B(\bar{\theta}_t, a_t, b_t),$$

where B is the incomplete Beta function and where a_t and b_t are the shape parameters of the prior at t. Finally, when insiders abstain from reforms and no revolt is observed, we can similarly use Bayes law to obtain

$$g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 0) = \frac{1 - \theta h(s_t)}{\int_0^{\bar{\theta}_t}[1 - \theta h(s_t)]g_{a_t,b_t}(\theta)\,\mathrm{d}\theta}g_{a_t,b_t}(\theta),$$

so that

$$\hat{\mu}_t = \int_0^{\bar{\theta}_t} \theta g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 0)\,\mathrm{d}\theta = \frac{M^1_{a_t,b_t}(\bar{\theta}_t) - h(s_t)M^2_{a_t,b_t}(\bar{\theta}_t)}{1 - h(s_t)M^1_{a_t,b_t}(\bar{\theta}_t)}$$

and

$$\hat{\sigma}_t^2 = \int_0^{\bar{\theta}_t} (\theta - \hat{\mu}_t)^2 g_{a_t,b_t}(\theta|\theta \leq \bar{\theta}_t, \eta_t = 0)\,\mathrm{d}\theta = \frac{M^2_{a_t,b_t}(\bar{\theta}_t) - h(s_t)M^3_{a_t,b_t}(\bar{\theta}_t)}{1 - h(s_t)M^1_{a_t,b_t}(\bar{\theta}_t)} - \hat{\mu}_t^2.$$

It remains to be checked that the resulting moments are consistent with a Beta distribution. In general, any combination of μ_t and σ_t is consistent with some a_t and b_t (and uniquely so),

if $\sigma_t^2 < \mu_t(1 - \mu_t)$, or equivalently

$$\pi\hat{\sigma}_t^2 + (1 - \pi)\sigma_0^2 + \pi(1 - \pi)(\hat{\mu}_t - \mu_0)^2$$
$$< [\pi\hat{\mu}_t + (1 - \pi)\mu_0] \times [\pi(1 - \hat{\mu}_t) + (1 - \pi)(1 - \mu_0)]$$
$$= \pi^2\hat{\mu}_t(1 - \hat{\mu}_t) + (1 - \pi)^2\mu_0(1 - \mu_0)$$
$$+ \pi(1 - \pi)[\hat{\mu}_t(1 - \mu_0) + (1 - \hat{\mu}_t)\mu_0]. \quad (12)$$

By assumption, $\sigma_0^2 < \mu_0(1 - \mu_0)$. Further, suppose for a moment that $\hat{\sigma}_t^2 \leq \hat{\mu}_t(1 - \hat{\mu}_t)$. Then subtracting the known inequalities and dividing by $\pi(1 - \pi)$, (12) simplifies to

$$\hat{\sigma}_t^2 + \sigma_0^2 < \hat{\mu}_t(1 - \hat{\mu}_t) + \mu_0(1 - \mu_0),$$

which is true under the maintained assumption. Hence, a sufficient condition for μ_t and σ_t to be Beta-implementable is that $\hat{\sigma}_t^2 \leq \hat{\mu}_t(1 - \hat{\mu}_t)$ or, equivalently, $\mathbb{E}\{\theta_t^2|\delta_t^p\} \leq \mathbb{E}\{\theta|\delta_t^p\}$. Given that $\theta_t \in [0, 1]$, this is trivially true, which concludes the proof. (Note how the addition of nondegenerate noise in the event of a redraw suffices to retain the strict inequality for the prior, even when the posterior has zero variance and unit mean).

To summarize, outsiders' beliefs at date t can be recursively computed, where we use the updating formulas derived above to go from (μ_t, σ_t^2)—or, equivalently, (a_t, b_t)—to $(\hat{\mu}_t, \hat{\sigma}_t^2)$, and then apply (8) and (9) to go to $(\mu_{t+1}, \sigma_{t+1}^2)$ and (a_{t+1}, b_{t+1}).

D Cost of reforms and equilibrium dynamics

In this appendix we show how variations in the cost of reforming, β_1, impact the equilibrium dynamics and the long-run distribution of political systems. For simplicity, we focus on the model with exogenous priors, but analogous conclusions apply to the case with learning.

As alluded to in the main text, variations in the cost of reforms affect equilibrium dynamics by changing the frequency of reforms relative to the frequency of revolts. To see this, consider Figure 18. Here we display a simulated time series for different values of β_1 and for 300 periods each. To make the three parametrization comparable, we keep the sequence of $\{\theta_t\}$ fixed across all three specifications, which is drawn from a uniform distribution F. Similar, we fix the random sequence of quantile ranks that determine the realizations of $\{\eta_t\}$, so that any differences in transition dynamics are purely driven by deterministic changes in Q^S, Q^R, ρ^S and ρ^R that are due to the variations in β_1.

For each time path, we plot the political system, λ_t, at time t, and indicate the dates where transitions occur via revolts (marked by Δ) and reforms (marked by $\times$). It can be seen that low costs of reforms in Setting 1 ($\beta_1 = 0.35$) result in immediate democratic reforms
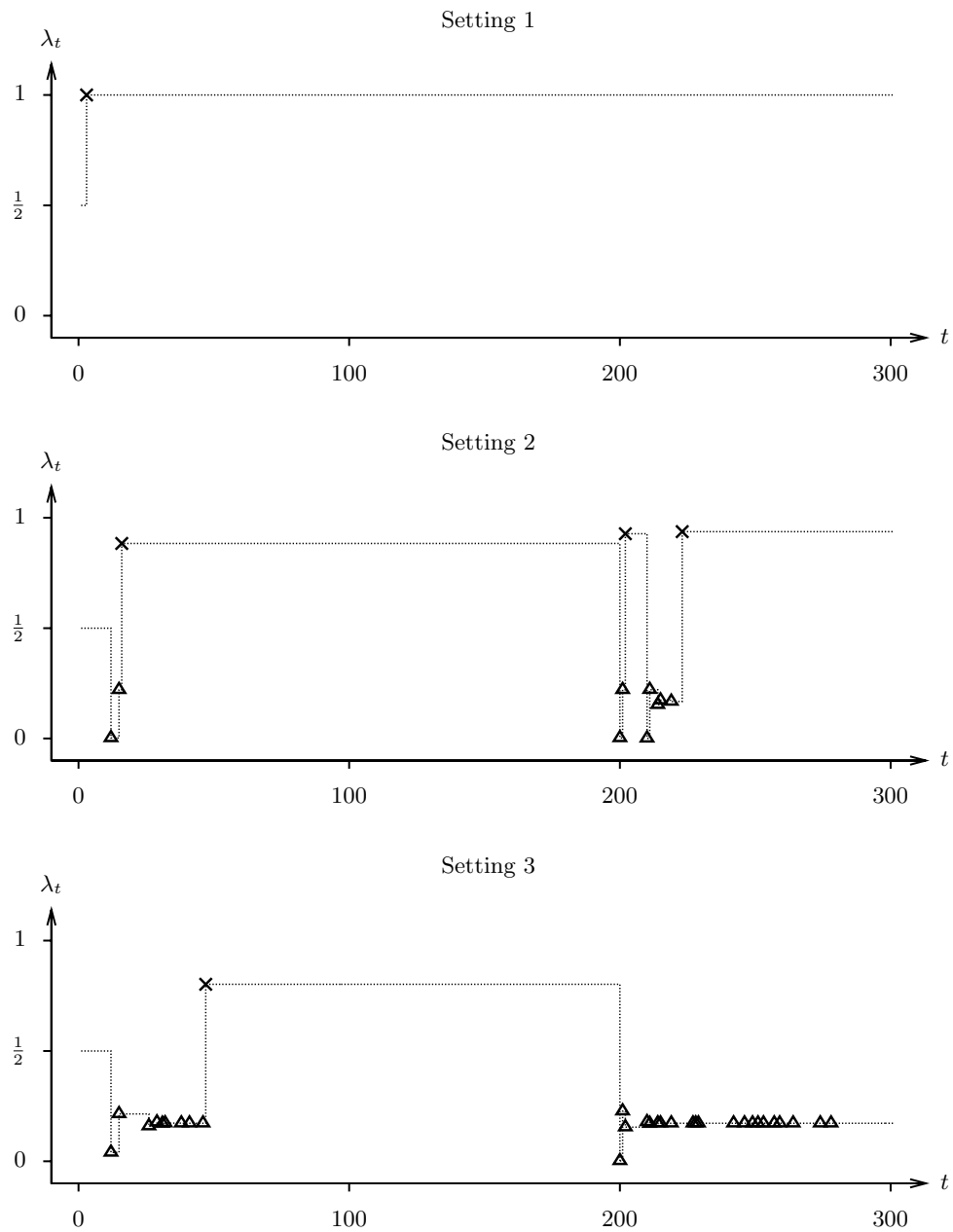
Figure 18. Simulated time series of the model with exogenous priors. *Notes:* Reforms are marked by "×", successful revolts are marked by "△". Costs of reforms (β_1) are increasing from Setting 1 to 3.

50

and the absence of successful subversive attempts. As the costs of reforms are increasing in Setting 2 ($\beta_1 = 0.40$) and Setting 3 ($\beta_1 = 0.45$), insiders initially prefer to abstain from reforms and gamble for their survival—despite facing the same sequence of θ_t. Their gambling for survival eventually leads to a successful revolt in Settings 2 and 3 in periods 12 and 15. Given the realization of θ_{16}, insiders then conduct reforms in period 16 for intermediate costs β_1, but continue to gamble for their survival in the case of high costs. The paths converge back towards each other in period 47, when insiders eventually reform in the case of high costs (the convergence is not perfect though, since given high costs of reforming the reforms will be less inclusive for larger values of β_1).[33] Around period 200, we then observe a reversal for intermediate and high values of β_1, while the more inclusive democracy in the case of small costs is sufficiently stable to survive the threat. The subsequent periods then show similar patterns, where insiders conduct reforms for intermediate values of β_1 and abstain in the high value case.

Despite these differences in the frequency of reforms and revolts, the observed patterns of stable democracies, instable autocracies, and polarization are similar across specifications. Mirroring our results in the model with learning (Section 6), the long-run distribution with exogenous priors is hence bimodal with mass concentrated on the extremes. Variations in β_1 thereby manifest themselves in shifts between the long-run mass on autocratic and democratic regimes. Figure 19 shows this, plotting the invariant distribution of political systems for various values of β_1 obtained from running a kernel density regression on simulated time series of 3.2 Million observations each.[34] For low values of β_1 (Settings 1 and 2), reforms are likely relative to revolts such that mass is mainly concentrated on democratic systems. The converse is true when the costs of conducting reforms are high (Settings 3 and 4).

---

[33]The inclusiveness of reformed regimes also differs across specifications, since reforms $\xi(\theta_t)$ depend on the precise realization of θ_t which differs across dates.

[34]To retain a constant scale across all settings, we smooth the simulated distribution using a bandwidth of 0.025. Somewhat hidden by this is that in Setting 4 all mass is collapsed into a single mass point at $\lambda = 0.12$, which in Setting 4 is absorbing. More generally, there are two scenarios under which a certain political system can be absorbing. First, if $\xi(1) = 1$, then λ arbitrary closely to 1 is reached in equilibrium, which is almost surely absorbing. However, since $\xi(1) < 1$ in all of the reported settings, we do not observe $\lambda \to 1$ along any of the equilibrium paths. Second, if there exists a $\tilde{\lambda}$, such that $\bar{\theta}(\tilde{\lambda}) = 1$ and $s(\mu, \tilde{\lambda}) = \tilde{\lambda}$, then the system $\lambda = \tilde{\lambda}$ is locally attracting and absorbing (despite frequent regime changes), as is the case in Setting 4.
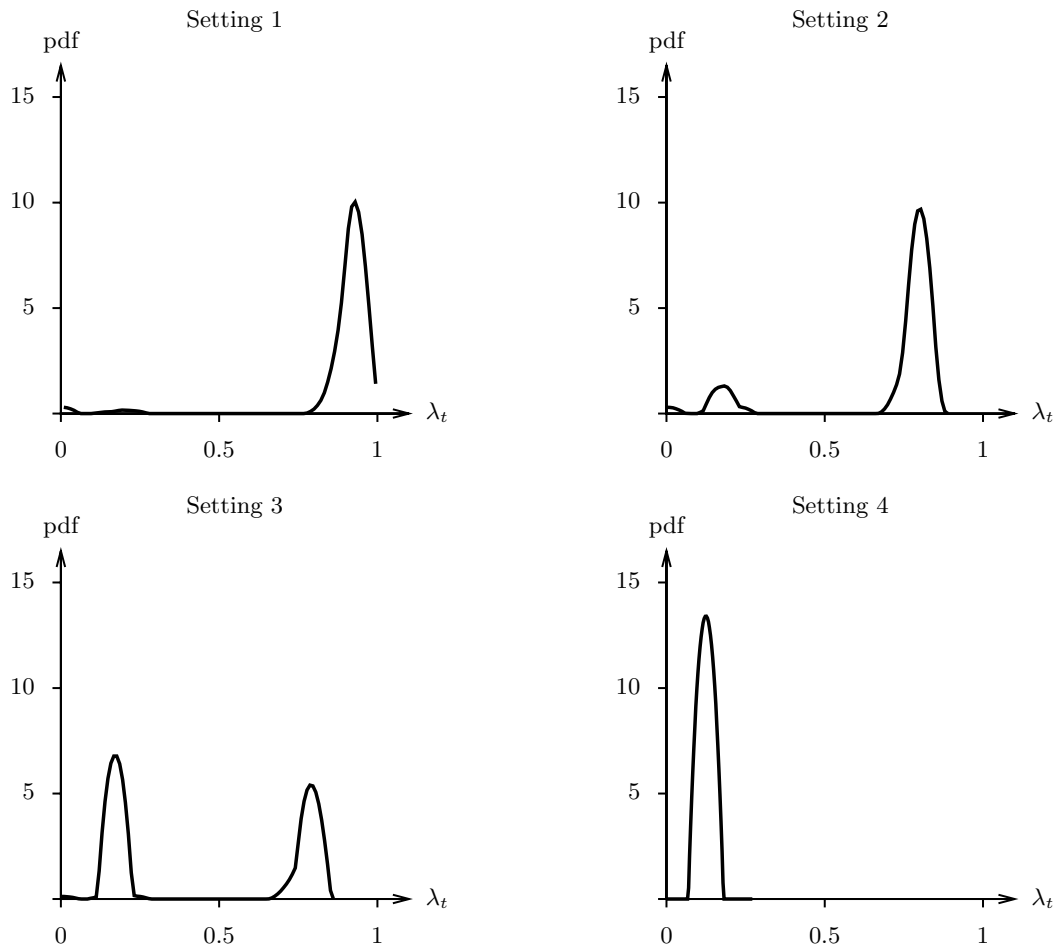
Figure 19. Invariant distribution of political systems. *Note:* Costs of reforms (β_1) are increasing from Setting 1 to 4.