WORKING PAPERS

N° TSE - 464

Revised October 2018

"Dynamics of Political Systems"

Lukas Buchheim and Robert Ulbricht

Toulouse School of Economics

Dynamics of Political Systems

Lukas Buchheim
LMU Munich

Robert Ulbricht
Toulouse School of Economics

October 17, 2018

Abstract

We develop a quantitative theory of repeated political transitions driven by revolts and reforms. In the model, the beliefs of disenfranchised citizens play a key role in determining revolutionary pressure, which in interaction with preemptive reforms determines regime dynamics. We estimate the model structurally, targeting key moments of the data. The estimated model generates a process of political transitions that looks remarkably close to the data, replicating the empirical shape of transition hazards, the frequency of revolts relative to reforms, the distribution of newly established regime types after revolts and reforms, and the unconditional distribution over regime types. Using the estimated model, we also explore circumstances of successful democratization, finding that the sentiment of political outsiders is key for creating a window of opportunity, whereas the scope of the initial democratic reform is key for the survival of young democracies.

Keywords: Democratic reforms, regime dynamics, revolts, structural estimation, transition hazards.

JEL Classification: D74, D78, P16.

1 Introduction

This paper develops a quantitative theory of political transitions based on the evolution of beliefs regarding the regime's strength. Traditionally, the literature has focused on explaining specific patterns of regime changes, focusing on isolated transition episodes. In this paper, we shift the focus to a macro perspective, aiming to account for a number of stylized facts in a unified framework.

Specifically, Section 2 of this paper documents five empirical regularities, which motivate the theoretical framework.

1. The evolution of political systems is shaped by both revolts and democratic reforms, with revolts being about three times as likely as reforms. Other modes of transition are secondary.

2. Transition hazards are declining in regime maturity. Newly established regimes are about three times as likely to be overthrown by a revolt and about six times as likely to implement a democratic reform compared to regimes older than 10 years.

3. Transition hazards are inverse "J-shaped" in the inclusiveness of political systems: Political systems at the extremes of the autocracy-democracy spectrum have smaller transition hazards than regimes near the center of the spectrum; full-scale democracies are overall most stable.

4. Revolts establish autocratic regimes; reforms establish democracies. Political systems near the center of the autocracy-democracy spectrum are unlikely to arise from either mode of transition.

5. The distribution of regime types is bi-modal, with mass concentrated towards the extremes of the autocracy-democracy spectrum.

This paper puts forward a theory of political transitions, which accounts for all five stylized facts above. In the model, the inclusiveness of a political system is defined by the fraction of the population with access to political power ("political insiders"). Transitions are governed by three main ingredients. First, reforms are rationalized by a preemptive logic as in the seminal works by Acemoglu and Robinson (2000*b*), Conley and Temini (2001), and Boix (2003). Second, revolts are the outcome of a coordination game among the disenfranchised ("political outsiders"), introducing an intensive margin to revolting, defined by the degree of equilibrium coordination among outsiders. Finally, the degree of coordination is shaped by the beliefs of outsiders regarding the regime's strength, which is privately observed by insiders at the beginning of each period.

In combination, the intensive margin of revolting and learning imply that revolt hazards are decreasing in the regime's strength *as perceived by outsiders.* This link between outsiders' beliefs and revolt hazards is at the heart of our predictions. In particular, because in equilibrium concessions are associated with being weak, the link implies that small reforms will generally backfire and increase revolutionary pressure. Accordingly, when facing moderate threats, insiders rather take "tough stance" than preempting a subversive threat, explaining the prevalence of revolts documented in the data. Similarly, because transitions are more likely to occur when a regime is weak, outsiders rationally become more and more convinced that a regime is invulnerable as it matures, explaining the decline of transition hazards in regime maturity. The logic behind the inverse J-shape of transition hazards is a combination of two factors: On the one hand, full-scale democracies are intrinsically strong due to a lack of opposition (the extensive margin of revolting). On the other hand, similar to mature regimes, the most repressive autocracies are stable due to a low degree of coordination among outsiders (the intensive margin of revolting). This is because such regimes arise precisely from revolts that ex ante were considered as futile by outsiders, making them also less prone to future unrest. Finally, the two remaining regularities are again a consequence of small reforms backfiring and that revolts cannot grow too large as they would have been preempted otherwise.

The model is rich enough to lend itself to a quantitative exploration, mainly due to two modeling choices. First, transitions take place in a continuous polity space. This stands in contrast to the previous literature, which typically considers transitions between two or three exogenously defined political systems. Second, there are no exogenously absorbing states in our model, allowing us to compare model moments (computed at the stationary distribution) with their empirical counterparts. We demonstrate the quantitative potential through a structural estimation of our model. The model matches the data remarkably well. It is not only able to account for the above-listed regularities, but also quantitatively replicates the shape of transition hazards, conditional outcome distributions, and the stationary distribution of regime types.

We also use the estimated model to study circumstances under which successful democratization is likely. In the model, the belief or "sentiment" of outsiders is instrumental for creating a window of opportunity, in which democratization is possible. Only if outsiders perceive the regime as sufficiently vulnerable, they are likely to coordinate on large revolts and regimes are inclined to implement reforms to preempt them. However, due to the presence of asymmetric information, regimes generally do not find it optimal to completely preempt a given threat of revolt. As a result, episodes in which democratization is possible are also marked by high revolt hazards, and the political system emerging from such "critical junctures" is determined by chance and random variations in the state of the world. Moreover, because newly established democracies emerge precisely when the regime is revealed to be most vulnerable, they are highly susceptible to counter-revolts by a small but highly coordinated group of outsiders. The model thus suggests that successful democratization critically hinges on the extent of the initial push for democratization. While reforms that enfranchise between 75% and 85% of the population have a cumulative failure rate of over 80 percent after 25 periods, the failure rate drops to 12 percent if reforms initially enfranchise more than 95% of the population.

The remainder of the paper is structured as follows. The next section presents a list of empirical regularities that are the target for our theoretical model. The model itself is developed in Section 3. In Section 4, we estimate the model to demonstrate its ability to quantitatively match the data. In Section 5, we provide intuition for how the different features of the model contribute to its ability to account for the empirical regularities. In Section 6, we study the model's implications for the formation and survival of democracies. In Section 7, we relate the model's mechanism as well as its predictions to the existing literature. Section 8 concludes.

2 Evidence on Political Transitions

This section presents a list of stylized facts about political dynamics, which motivates the theoretical framework developed in the next sections.

The presented regularities are based on the universe of political regimes existing between 1946 and 2010, combining information from three distinct datasets. First, we obtain the universe of *regime spells* from Geddes, Wright and Frantz (2014), who define regimes based on the identity of the ruling groups. Second, we use the Polity IV Project's polity index (Marshall, Gurr and Jaggers, 2017) to assign a *regime type* to each regime spell, which ranks political regimes on a 21-point scale between autocratic and democratic (normalized to values between 0 and 1). Finally, we treat any substantial change in the composition of regime insiders, as indicated by the turnover dates of regime spells, as *transition events*. Whenever available, we use the classification provided by Geddes, Wright and Frantz (2014) to classify transition events. Otherwise, we match transitions to leader changes collected by the Archigos Database of Political Leaders (Goemans, Gleditsch and Chiozza, 2009) and classify them according to the nature of the observed leader change. The resulting database covers 485

Table 1: Frequency of Transition Events

Transition event	Frequency	Share	Yearly hazard
Revolt	188	0.56	0.0213
Democratic reform	66	0.20	0.0075
Autocratic consolidation	19	0.06	0.0021
Foreign imposition	16	0.05	0.0018
Other/unknown	49	0.14	0.0055
Total	338	1.00	0.0382

Notes.—The table reports number of occurrences for each transition type for all regime changes between 1946 and 2010, as well as frequencies normalized by total transitions (shares) and by country-years (yearly hazards).

regime spells and 329 transitions in 155 countries. Appendix A.1 describes the construction of the dataset in detail.

Fact 1: The most frequent modes of transition are revolts and democratic reforms, with revolts being about three times as likely as reforms.

Our definition of revolts encompasses all forms of coercive takeovers by domestic actors (popular uprisings, power struggles between competing factions, and coups[1]). Democratic reforms are peaceful transitions that lead to a more democratic political system. Together, revolts and reforms constitute 75 percent of all observed transition events. This corresponds to about .021 revolts and .008 reforms per country-year. The remaining transitions occur either via autocratic consolidations (peaceful transitions towards more autocratic polities, six percent), foreign imposition (five percent), or cannot be categorized based on the available information (14 percent). See Table 1 for further details.

Fact 2: Transition hazards are declining in regime maturity.

Figure 1 plots the transition hazards for our data.[2] Newly established regimes are about three times as likely to be overthrown via revolt compared to regimes older than 10 years, and four to six times as likely to reform as regimes older than 5 years.

[1] Geddes, Wright and Frantz (2014) define regime spells as uninterrupted reign of the same group of political elites. Accordingly, coups only constitute a transition if they substantially alter the composition of the ruling group. By contrast, coups that, e.g., replace one military leader by another from the same group of military leadership do not constitute transition events.

[2] The hazards are estimated by differencing and smoothing over Nelson-Aalen estimates for the cumulative hazard rate and are adjusted for left and right censoring. All findings are robust to controlling for the current political system and region fixed effects (see Appendix A.2 for details). Similar patterns have been documented by Sanhueza (1999) and Svolik (2008, 2015). Relatedly, Bienen and van de Walle (1989, 1992) and Bueno de Mesquita et al. (2003) find a declining risk of loosing power at the level of political leaders.

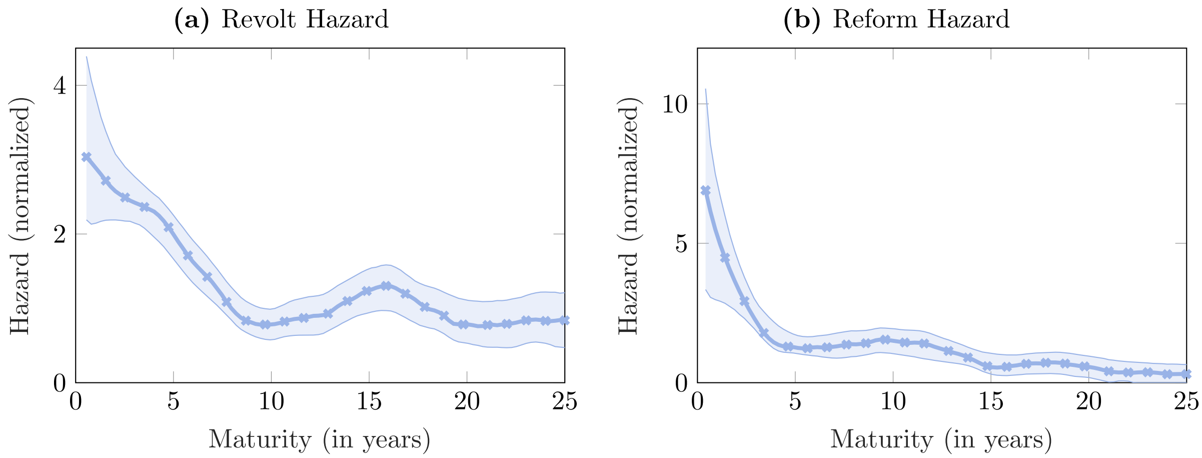

Figure 1: Empirical transition hazards for revolts (left panel) and reforms (right panel). Notes.—Hazards are normalized relative to the unconditional hazard of revolts and reforms, respectively. Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

Fact 3: Transition hazards are inverse "J-shaped" in the inclusiveness of political systems.

The regularity that regimes at the extremes of the political spectrum are most stable has been documented by a number of recent studies (e.g., Bremmer, 2006; Gates et al., 2006; Goldstone et al., 2010; Knutsen and Nygård, 2015). To evaluate the regularity in our data, we estimate a Cox model with a cubic spline in the polity dimension (see Appendix A.2 for further details). Figure 2 plots the predicted relationship between polities and hazard ratios, normalized relative to the most autocratic regimes (with polity score equal to zero). Full-scaled democracies (with polity score of one) are least vulnerable to transitions with a relative hazard of approximately 1/5. Hybrid regimes, in contrast, are on average up to four times as likely to undergo a transition compared to the most autocratic regimes.

Fact 4: Revolts establish autocratic regimes; reforms establish democratic regimes.

Figure 3 shows the conditional distribution over political systems arising from revolts and reforms. The median revolt establishes an ("autocratic") regime with a polity score of 0.2. The median reform establishes a ("democratic") regime with a polity score of 0.8. Political systems near the center of the autocracy-democracy spectrum are unlikely to arise from either mode of transition.[3]

[3] See also Gleditsch and Choung 2004; Gleditsch and Ward 2006; Celestino and Gleditsch 2013; Derpanopoulos et al. 2016. The results are also consistent with a number of qualitative studies documenting that democracies are unlikely to arise without a reform process (Rustow, 1970; O'Donnell and Schmitter, 1986; Karl, 1990; Huntington, 1991).

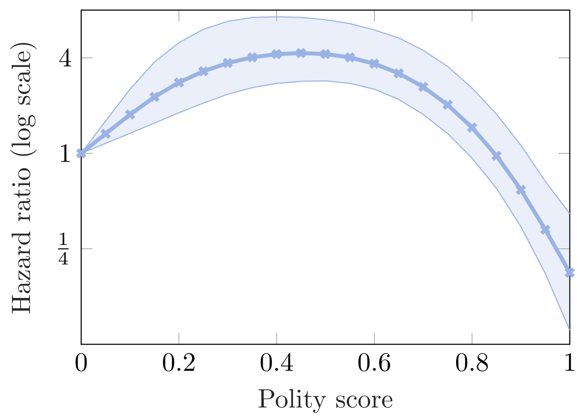

Figure 2: Estimated hazard ratios of political systems. Notes.—Hazard ratios are estimated by a Cox regression with a cubic spline in the polity dimension. All hazard rates are for the combined failure due to reform and revolt, and are normalized relative to the combined hazard of regimes with a polity score of zero. Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

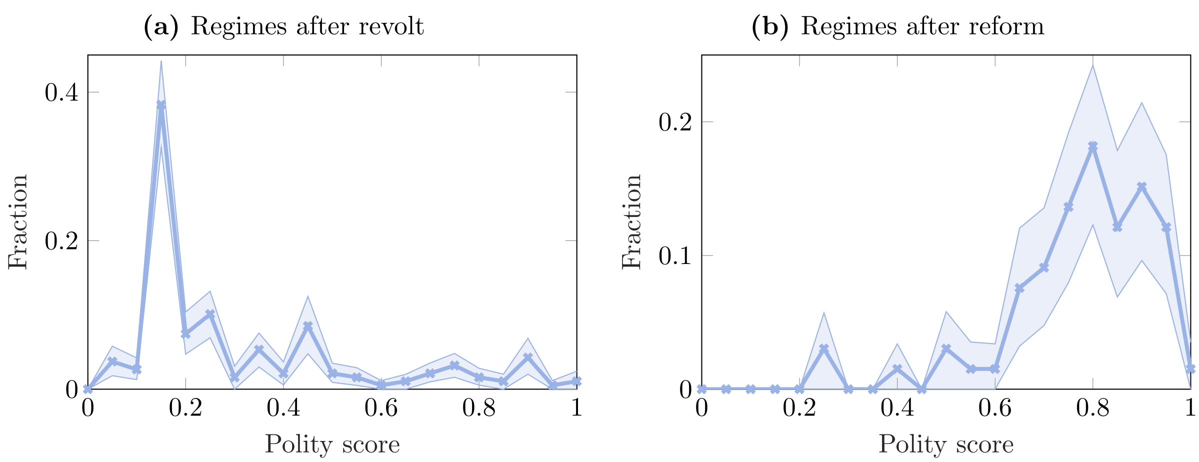

Figure 3: Empirical distribution of political systems arising from revolts (left panel) and reforms (right panel). Notes.—Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

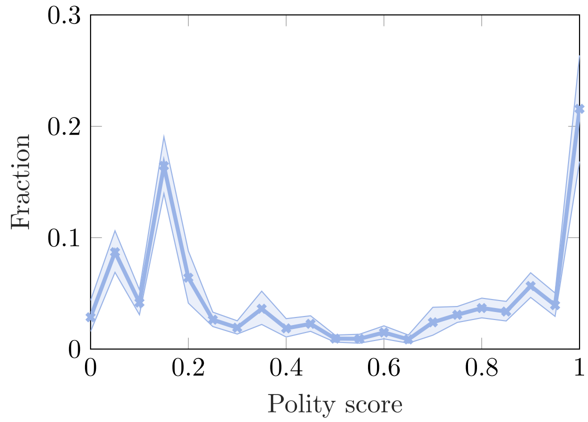

Figure 4: Empirical distribution of political systems between 1946 and 2010. Notes.—Unit of observation are country-days. Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

Fact 5: The distribution of regime types (polities) is bi-modal, with mass concentrated towards the extremes.

Finally, as illustrated in Figure 4, most mass of the empirical distribution over regime types is concentrated towards the extremes of the political spectrum: The combined mass of observations with a polity score ≤ 0.25 and a polity score ≥ 0.75 is 84 percent.

3 The Model

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

3.1 Setup

We consider an infinite horizon economy, populated by overlapping generations of two-period lived agents. Each generation consists of a continuum of agents with mass equal to 1. At time t, fraction λ_t of the population has the power to implement political decisions; 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. 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. We denote young outsider i's choice by $\phi_{it} \in \{0,1\}$ and use the aggregated mass of supporters, $s_t = \int \phi_{it}\, \mathrm{d}i$, 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 $h(0) = 0$. Intuitively, the threat of a revolt to the current regime is increasing in the mass of revolutionaries and in the regime's vulnerability. When a revolt has no supporters ($s_t = 0$) or the regime is not vulnerable ($\theta_t = 0$), the regime survives with certainty.

The state of θ_t follows a (commonly known) Markov process with c.d.f. $G(\theta_t|\theta_{t-1})$ and is assumed to have full support on $[0,1]$. At the beginning of each period, insiders learn the current realization of θ_t. By contrast, outsiders do not observe θ_t directly and instead use Bayes' law to form beliefs over its current realization based on the history of past political transitions. We use F_t to denote the belief of outsiders over θ_t at the beginning of period t.

After learning the realization of θ_t, insiders may try to alleviate the threat of a revolt by conducting reforms. We follow Acemoglu and Robinson (2000*b*) by modeling these reforms as an extension of the franchise to outsiders, which is effective in preventing them from supporting a revolt. Generalizing this mechanism to a continuous polity space, we allow insiders to continuously extend the regime by any fraction, $x_t - \lambda_t$, of young outsiders, where $x_t \in [\lambda_t, 1]$ denotes 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.

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 per period payoff of γ_{it} that 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. Outsiders' payoffs remain constant over their life if they abstain from revolting, and otherwise change conditional on the success of the revolt (detailed below).

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 common resource stock, implementing preferred policies, etc.).[4] Note that $u' < 0$ implies that extending the regime is costly for insiders (e.g., because resources have to be shared, or preferences about policies become less aligned). Also, $u(1) = 1$ implies 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 become worst-adapted to the new regime ($\gamma_{it} = 0$).[5] 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:

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

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

where superscript I and O denote agents that are born as (or are newly 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). The first period payoffs are omitted, as they are unaffected by the policy choices of generation t.

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

[4] One could microfound u as a value function where all policy choices associated with having political power—except enfranchising political outsiders—are chosen optimally. Subsuming these decisions into u allows us to tractably explore the dynamics of political systems emerging from the interplay of reforms and revolts. All other policy choices still affect our analysis inasmuch as they determine the shape of u.

[5] In our dataset, 83 percent of overthrown leaders are killed, imprisoned or sentenced to exile under the new regime. Similar punishments are common for supporters of failed insurgencies, making the assumption that the losing party is worst-adapted arguably realistic. Further note that this assumption effectively maximizes the cost of engaging in political confrontation.

1. The current state of θ_t realizes and is revealed to insiders.

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

3. Outsiders, if excluded from the reform, 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.

Two remarks The core of our model defines an interaction between revolutionary pressure and preemptive reforms in the tradition of Acemoglu and Robinson (2000*b*), Conley and Temini (2001), and Boix (2003). Implicit in the preemptive logic of reforms is the requirement that extending the franchise entails a *credible* commitment to share political power that is not easily reversible. Accordingly, our notion of inclusiveness, λ_t, is best understood as the fraction of citizens that are protected from losing political power, either because of hard-to-overturn institutional guarantees as in Acemoglu and Robinson (2000*b*), or because each insider is indispensable for the stability of the ruling coalition as in Acemoglu, Egorov and Sonin (2012). In line with this interpretation of λ_t, as well as with the low frequency of autocratic consolidations in the data, our model abstracts from the possibility of "adverse reforms".[6]

Relatedly, we assume that reforms are effective in the sense that newly enfranchised outsiders (as well as agents born as insiders) do not rebel against the regime. As formally proved in Appendix B.1, this is indeed internally consistent within our setting, as newly enfranchised outsiders (and born insiders) would never support a revolt if given the choice.

Equilibrium definition We characterize the set of perfect Bayesian equilibria subject to two equilibrium refinements. First, we rule out "instable" coordination among outsiders on $s_t = 0$, whenever an infinitesimal small chance of success would persuade a non-marginal mass of outsiders to revolt.[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

[6] In addition to the aforementioned reasons, abstracting from adverse reforms is also analytically convenient, as it allows us to treat insiders as a homogeneous group, rather than providing an explicit model of within-regime power struggles that may result in the ejection of certain subgroups.

[7] 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.

games.[8] As detailed below (see Footnote 11), the refinement improves the predictive power of our model by selecting a unique equilibrium, but is inconsequential for our main predictions.

Before defining equilibrium, it is useful to fix some notation. First, as already noted, we use F_t to denote the "prior" belief of young outsiders born at date t, which is formed using Bayes' law (if applicable) given all publicly observable information available at the *beginning* of period t. Specifically, we have

$$F_t(\vartheta) = \Pr\left[\theta \leq \vartheta | \delta_{t-1}\right] \tag{5}$$

for any publicly observable history $\delta_t \equiv \{\{\phi_{i\tau}\}, x_\tau, \lambda_\tau, \eta_\tau\}_{\tau=0}^{t}$ that is reached along the equilibrium path with strictly positive probability. As usual, off-equilibrium beliefs can be chosen freely, subject to the restrictions imposed by the D1 criterion. Similarly, we use $\hat{F}_t$ to denote the *interim* belief of outsiders, which combines F_t with the information signaled by reforms x_t:

$$\hat{F}_t(\vartheta) = \Pr\left[\theta \leq \vartheta | \delta_{t-1}, x_t\right] \tag{6}$$

for all (δ_{t-1}, x_t) reached along the equilibrium path. Here we do not index F_t and $\hat{F}_t$ by i, since they will be pinned down uniquely by the D1-refinement—even off the equilibrium path—ruling out any scope for belief heterogeneity across outsiders.

We are now ready to define the equilibrium for our model. To simplify notation, we only define pure strategies here, since only pure strategy equilibria exist in our game (see the proofs to Propositions 1 and 2).

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

a. Reforms $x_{\bar{\delta}}$ maximize insider's expected utility $V^I(p_{\bar{\delta}}, x_{\bar{\delta}})$ given outsiders' beliefs $\hat{F}_\delta$ and strategies $\{\phi_{i\delta}\}$;

b. Each outsider's revolt choice $\phi_{i\delta}$ maximizes $\mathbb{E}_{\hat{F}_\delta}\{V^O(p_{\bar{\delta}}, \gamma_{i\bar{\delta}}, s_{\bar{\delta}}, \phi_{i\delta})\}$ given insiders' reforms $x_{\bar{\delta}}$, other outsiders' revolt choices $\{\phi_{j\delta}\}_{j\neq i}$, and beliefs $\hat{F}_\delta$;[9]

[8] 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 in the sense that it is beneficial under the largest set of possible inferences $\{\hat{\theta}\}$ about the regime's vulnerability.

[9] Throughout we use subscripts to $\mathbb{E}$ to indicate the probability measure with respect to which the expectation is taken.

c. Beliefs $\hat{F}_\delta$ are obtained using (6) for all (δ_{t-1}, x_t) along the equilibrium path, and satisfy the D1 criterion otherwise;

d. The evolution of (λ_t, η_t), contained in $\bar{\delta}$, is consistent with (1) and (2);

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

3.2 Equilibrium Characterization

As a result of the overlapping generations structure of the model, the characterization of equilibrium can be separated into a sequence of "generation games" between young insiders and young outsiders. Generations are linked across periods through the evolution of the payoff-relevant state, given by $\mathcal{S}_t \equiv (\theta_t, \lambda_t, F_t)$.

The generation game at t consists of two stages. 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 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. Let $\hat{\theta}_t \equiv \mathbb{E}_{\hat{F}_t}\{\theta_t\}$ define the interim-expectation of outsiders regarding θ_t. Because $\mathbb{E}_{\hat{F}}\{V^O(\cdot)\}$ is linear in θ, $\hat{\theta}_t$ is a sufficient statistic for $\hat{F}_t$. For any belief, $(\hat{\theta}_t, \hat{s}_t) \in [0,1]^2$, individual rationality requires all outsiders to choose a ϕ_{it} that maximizes their expected utility $\mathbb{E}_{\hat{F}}\{V^O(\cdot)\}$. At time t, outsider i with opportunity cost γ_{it} will therefore participate in a revolt if and only if

$$\gamma_{it} \leq p(\hat{\theta}_t, \hat{s}_t)\, u(\hat{s}_t) \equiv \bar{\gamma}(\hat{\theta}_t, \hat{s}_t). \tag{7}$$

In equilibrium, $\bar{\gamma}(\hat{\theta}_t, \hat{s}_t)$ is the expected benefit of participating in a revolt that is supported by a mass $\hat{s}_t$ of outsiders. Since $\bar{\gamma}$ 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. For any $\bar{\gamma}$, the size of the resulting revolt is then given by

$$s_t = (1 - x_t) \min\{\bar{\gamma}(\hat{\theta}_t, \hat{s}_t), 1\}. \tag{8}$$

In equilibrium, it must hold that $s_t = \hat{s}_t$. Accordingly, the share of outsiders supporting a revolt is pinned down by the fixed point to (8). To guarantee that a well-behaved fixed point exists, we impose the following assumption.

Assumption 1. *Let $\psi(s) \equiv h(s) \cdot u(s)$. Then, $\psi' \geq 0$, $\psi'' \leq 0$ and $\lim_{s \to 0} \psi'(s) = \infty$.*

Assumption 1 imposes that 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 nonincreasing as revolts grow larger.

Equipped with Assumption 1, we obtain the following proposition.

Proposition 1. *In any equilibrium, the mass of outsiders supporting a revolt at time t is uniquely characterized by the solution to (8), given by the time-invariant mapping $s : (\hat{\theta}_t, x_t) \mapsto s_t$. The solution 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 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 portion of the population whenever they conduct reforms. The next proposition describes the equilibrium schedule of reforms.

Proposition 2. *Define insiders' expected utility as* $\tilde{V}^I(\theta, \hat{\theta}, x) \equiv V^I(\theta h(s(\hat{\theta}, x)), x)$*, and let* ξ *be the differential equation solving*[10]

$$\xi'(\theta) = -\tilde{V}_2^I(\theta, \theta, \xi)/\tilde{V}_3^I(\theta, \theta, \xi) > 0$$

with boundary condition $\xi(1) = \arg\max_{\xi \in [0,1]} \tilde{V}^I(1, 1, \xi)$*. Then, in any equilibrium, policy choices of insiders are uniquely defined by the time-invariant function,* $x : (\theta_t, \lambda_t, F_t) \mapsto x_t$*,*

$$x(\theta, \lambda, F) = \begin{cases} \lambda & \text{if } \theta \leq \bar{\theta}(\lambda, F) \\ \xi(\theta) & \text{if } \theta > \bar{\theta}(\lambda, F) \end{cases}$$

with $\xi(\theta) > \lambda$ *for all* $\theta > \bar{\theta}(\lambda, F)$*. The threshold type,* $\bar{\theta} : (\lambda_t, F_t) \mapsto \bar{\theta}_t$*, is implicitly defined by (whenever a solution exists)*

$$\tilde{V}^I(\bar{\theta}, \bar{\theta}, \xi(\bar{\theta})) = \tilde{V}^I(\bar{\theta}, \hat{\theta}(\lambda, \lambda, F), \lambda), \tag{9}$$

and is otherwise given by $\bar{\theta} = 1$*. Outsiders' interim beliefs are defined by*[11]*,* $\hat{\theta} : (\lambda_t, x_t, F_t) \mapsto \hat{\theta}_t$*, with*

$$\hat{\theta}(\lambda, x, F) = \begin{cases} \mathbb{E}_F\{\theta | \theta \leq \bar{\theta}(\lambda, F)\} & \text{if } x = \lambda \\ \xi^{-1}(x) & \text{if } \xi(\bar{\theta}(\lambda)) \leq x \leq \xi(1). \end{cases}$$

Proposition 2 describes equilibrium reforms as a function of $(\theta_t, \lambda_t, F_t)$. Because the logic behind these choices is the same for all values of λ_t and F_t, we can discuss the underlying intuition keeping (λ_t, F_t) fixed. To this end, Figure 5 plots reform choices (left panel) and the implied probability to be overthrown (right panel), fixing $\lambda_t = 0.1$ and $F_t(\theta) = \theta$. Extended versions of the figure with alternative values for λ_t and F_t can be found in Appendix E.

Whenever $\theta_t \leq \bar{\theta}(\lambda_t, F_t)$, insiders do not reform ($x_t = \lambda_t$), implying a substantial threat for regimes with θ_t close to $\bar{\theta}_t$. To see the logic behind this, first consider Figure 6. Here we plot equilibrium beliefs (left panel) and the corresponding mass of insurgents (right panel) as functions of x_t. If there are no reforms, outsiders only know the average vulnerability, $\hat{\theta}_t^{\text{pool}} \equiv \mathbb{E}_{F_t}\{\theta_t | \theta_t \leq \bar{\theta}_t\}$, of all regimes pooling on $x_t = \lambda_t$. By contrast, every extension x of the regime—how small it may be—leads to a non-marginal change in outsiders' beliefs from

[10] Throughout, we use f_i to denote the derivative with respect to the i-th argument of some function f.

[11] Off the equilibrium path, beliefs are uniquely pinned down by the D1 criterion as $\hat{\theta}_t = \bar{\theta}(\lambda_t)$ for $x \in (\lambda_t, \xi(\bar{\theta}_t))$ and $\hat{\theta}_t = 1$ for $x > \xi(1)$, contributing to the overall uniqueness of the reform schedule. However, even without D1, reforms are always increasing, starting from a strictly positive pool at $x_t = \lambda_t$ and have a discontinuity at the marginally reforming regime $\bar{\theta}_t$. Accordingly, the D1 refinement merely pins down the precise shape of ξ, but is not substantial for generating any of the main features of the reform schedule.

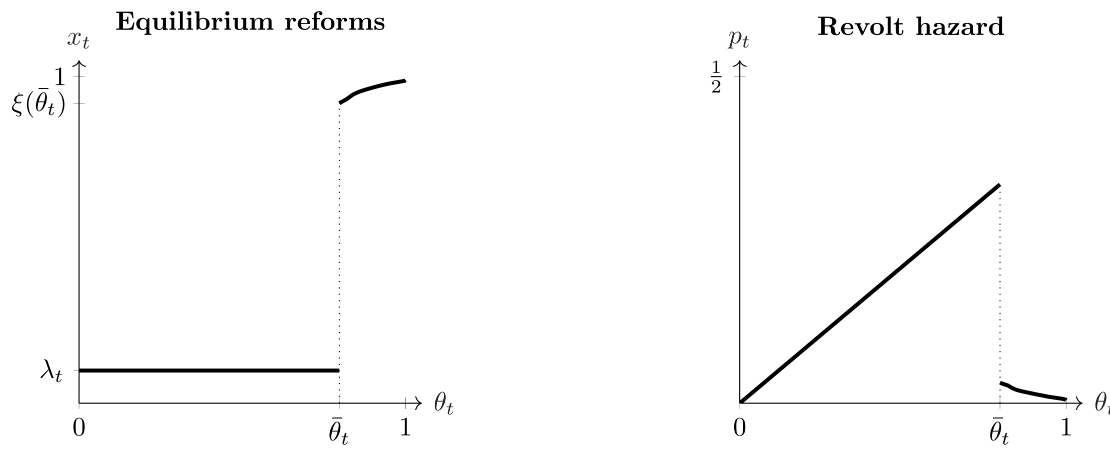

Figure 5: Equilibrium reforms and implied probability to be overthrown.

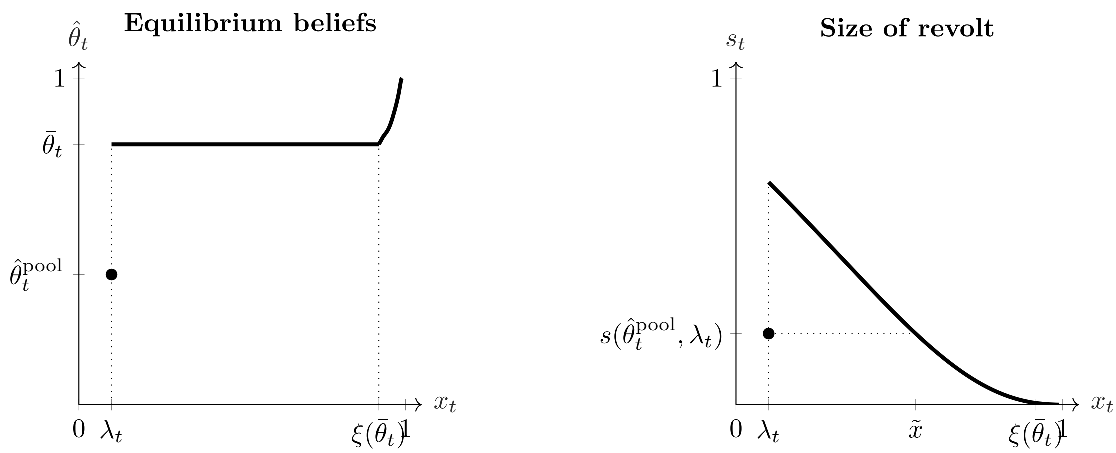

Figure 6: Equilibrium beliefs and implied mass of insurgents.

$\hat{\theta}_t^{\text{pool}}$ to $\xi^{-1}(x) \geq \bar{\theta}_t$ and, therefore, causes a non-marginal increase in revolutionary pressure along the intensive margin. It follows that small reforms will always backfire and *increase* the mass of insurgents as the increase in coordination dominates any marginal reduction in the group of potential insurgents along the extensive margin.

Furthermore, optimality of reforms requires that the benefit of reducing pressure compensates for insiders' disliking of sharing power. Hence there exists a nonempty interval, depicted by $[\tilde{x}, \xi(\bar{\theta}_t)]$ in the right panel of Figure 6, in which reforms are effective, yet insiders prefer to "gamble for their survival" in order to hold on to the benefits of not sharing power. This causes a substantial threat for regimes with θ_t close to $\bar{\theta}_t$, which can reconcile a frequent occurrence of revolts with the co-occurrence of preemptive reforms.[12]

[12]More precisely, gambling for survival increases the revolt hazard in two ways. Firstly, since at the margin more vulnerable regimes join the pool at $x_t = \lambda_t$, these regimes obviously face a high threat by not reforming. Secondly, since these regimes also deteriorate the average pooling belief towards being more vulnerable, there

Learning dynamics Propositions 1 and 2 fully characterize actions at t conditional on $\mathcal{S}_t$. To complete the characterization of equilibrium, we have to describe how $\mathcal{S}_t$ evolves over time. The evolution of θ_t and λ_t is described by the processes G and (2), leaving us to characterize the law of motion for F_t.

Let $\tilde{F}_t$ define the "posterior" belief of outsiders living at date t, formed using Bayes' law, given all publicly available information at the *end* of period t,

$$\tilde{F}_t(\vartheta) = \Pr[\theta \leq \vartheta | \delta_t].$$

Intuitively, $\tilde{F}_t$ combines the prior F_t with the information signaled by x_t (yielding the interim-belief $\hat{F}_t$) and the information contained in whether or not the regime survives, η_t.

Once we have compute $\tilde{F}_t$, we can use it to obtain the prior of the next generation, F_{t+1}, by simply "forecasting" θ_{t+1} using the law of motion for θ_t:

$$F_{t+1}(\vartheta) = \int_0^{\vartheta} \int_0^1 G'(\theta'|\theta)\, \mathrm{d}\tilde{F}_t(\theta)\, \mathrm{d}\theta' \tag{10}$$

We complete our equilibrium characterization with an explicit characterization of $\tilde{F}_t$.

Proposition 3. *Let $M_t^i(\vartheta) \equiv \mathbb{E}_{F_t}\{\theta^i | \theta \leq \vartheta\}$ define the i-th (raw) moment of $F_t(\theta|\theta \leq \vartheta)$. Then, along the equilibrium path, outsiders' posterior is given by:*

(i) if there is a reform attempt ($x_t > \lambda_t$),

$$\tilde{F}_t(\vartheta) = \begin{cases} 0 & \text{if } \vartheta < \theta_t \\ 1 & \text{else,} \end{cases}$$

(ii) if there is a revolt and no reform attempt ($x_t = \lambda_t$ and $\eta_t = 1$),

$$\tilde{F}_t(\vartheta) = \begin{cases} \frac{F_t(\vartheta)}{F_t(\bar{\theta}_t)} \frac{M_t^1(\vartheta)}{M_t^1(\bar{\theta}_t)} & \text{if } \vartheta < \bar{\theta}_t \\ 1 & \text{else,} \end{cases}$$

(iii) if there is no transition ($x_t = \lambda_t$ and $\eta_t = 0$),

$$\tilde{F}_t(\vartheta) = \begin{cases} \frac{F_t(\vartheta)}{F_t(\bar{\theta}_t)} \cdot \frac{1-h(s_t)M_t^1(\vartheta)}{1-h(s_t)M_t^1(\bar{\theta}_t)} & \text{if } \vartheta < \bar{\theta}_t \\ 1 & \text{else.} \end{cases}$$

is a further infra-marginal increase in the threat that affects all regimes without reforms.

For later reference, it is useful to also compute the first two moments of $\tilde{F}_t$. Specifically, using Proposition 3, the posterior mean and variance are given by

$$\tilde{\mu}_t = \begin{cases} \theta_t & \text{if } x_t > \lambda_t \\ \frac{M_t^2(\bar{\theta}_t)}{M_t^1(\theta_t)} & \text{if } x_t = \lambda_t \text{ and } \eta_t = 1 \\ \frac{M_t^1(\bar{\theta}_t) - h(s_t) M_t^2(\bar{\theta}_t)}{1 - h(s_t) M_t^1(\bar{\theta}_t)} & \text{if } x_t = \lambda_t \text{ and } \eta_t = 0 \end{cases} \tag{11}$$

and

$$\tilde{\sigma}_t^2 = \begin{cases} 0 & \text{if } x_t > \lambda_t \\ \frac{M_t^3(\bar{\theta}_t)}{M_t^1(\bar{\theta}_t)} - \tilde{\mu}_t^2 & \text{if } x_t = \lambda_t \text{ and } \eta_t = 1 \\ \frac{M_t^2(\bar{\theta}_t) - h(s_t) M_t^3(\bar{\theta}_t)}{1 - h(s_t) M_t^1(\bar{\theta}_t)} - \tilde{\mu}_t^2 & \text{if } x_t = \lambda_t \text{ and } \eta_t = 0. \end{cases} \tag{12}$$

Existence and uniqueness of equilibrium Propositions 1–2 uniquely pin down insiders' and outsiders' actions conditional on $\mathcal{S}_t$, whereas Proposition 3 (in conjunction with G and (2)) pins down a unique law of motion for $\mathcal{S}_t$. 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.

Proposition 4. *There exists an equilibrium, in which for all histories $\bar{\delta}$, policy mappings $x_{\bar{\delta}}$ and $\{\phi_{i\delta}\}_{i=0}^1$, as well as beliefs F_δ correspond to the time-invariant mappings underlying Propositions 1–3. Furthermore, for any given initial state $\mathcal{S}_0$, the equilibrium is unique.*

4 Model Implications for Political Dynamics

To explore the empirical performance of the model, we fit it to a few key moments of the data. We first study the implications for the frequency of transitions, hazard rates, transition outcomes, and the stationary distribution of political systems. Overall, the model fits the patterns documented in Section 2 quite well, even those that are not targeted by the estimation. Then, in the next section, we provide intuition for our results and illustrate how the different features of the model contribute to matching the data.

4.1 Parametrization

We choose the following parametrization of the model. The utility of insiders and the likelihood of a successful revolt are given by $u(\lambda) = 1 + \alpha_u(1 - \lambda)$ and $h(s) = s^{\alpha_h}$. Here, α_u is the marginal disutility of extending the regime, whereas α_h defines the elasticity of p_t with respect to an additional revolutionary. The restrictions we imposed on u and h require $\alpha_u, \alpha_h \in (0,1)$ and $\alpha_u \leq \alpha_h$. Based on some initial exploration, we found that the latter constraint is typically binding when trying to implement a stationary distribution with non-trivial mass on autocracies.[13] Accordingly, we fix $\alpha_u = \alpha_h \equiv \alpha$ to reduce the computational complexity of the estimation.

Next, we set G so that θ_t follows a truncated AR(1) process,

$$\theta_t = \min(\max(\rho\theta_{t-1} + \epsilon_t, 0), 1),$$

with persistence rate $\rho \in [0, 1)$ and innovations ϵ_t that are i.i.d. normal with mean μ_ϵ and variance σ_ϵ^2. Observe that for σ_ϵ sufficiently small, the mean and variance of F_{t+1} are approximately given by:

$$\mu_{t+1} = \rho\tilde{\mu}_t + \mu_\epsilon \tag{13}$$

$$\sigma_{t+1}^2 = \rho^2\tilde{\sigma}_t^2 + \sigma_\epsilon^2. \tag{14}$$

One challenge in simulating the model over long periods of time is that F_t typically does not stay within a given parametric family of distributions, making it difficult to keep track of beliefs over time. To address this issue, we approximate F_{t+1}, derived in (10), by a Beta distribution with mean μ_{t+1} and variance σ_{t+1}^2 matching the corresponding moments of F_{t+1} as given by (13) and (14). We explore the accurateness of the approximation in Appendix D, finding it to be extremely precise. Since the Beta distribution is fully parametrized by its first two moments, this approach allows us to efficiently keep track of outsiders beliefs using just μ_t and σ_t^2.[14]

With our approximation for F_t, the state space reduces to $\mathcal{S}_t = (\theta_t, \lambda_t, \mu_t, \sigma_t^2)$. Throughout our exploration, we will take the stand that θ_t is inherently unobserved to the statistician (as it is in the data), meaning that we will only look at moments where θ_t is marginalized out. In addition to ensuring consistency with the empirical moments, this view turns out to be also convenient, as it allows us to eliminate θ_t from $\mathcal{S}_t$ when characterizing the stationary distribution, requiring us to only keep track of $(\lambda_t, \mu_t, \sigma_t^2)$.[15]

[13]For small α_u, autocracies are less profitable and regimes tend to reform frequently, resulting in a large mass of democracies relative to autocracies.

[14]The approach is similar to the ubiquitous practice of solving models with complex state-spaces by approximating the perceived dependence on distributions by their first two moments as in Krusell and Smith (1998).

[15]In particular, exploiting that the information set of the statistician is aligned with the one of outsiders, we use a hidden state forward algorithm where instead of keeping track of θ_t, we use F_t to keep track of *distributions* over θ_t that are consistent with a particular (publicly observed) history $\delta = \{\lambda_\tau, \mu_\tau, \sigma_\tau^2\}_{\tau=0}^t$. Specifically, at

We approximate the continuous state space with a finite grid. Specifically, we approximate λ using a grid of (almost[16]) linearly spaced points $\lambda_1, \lambda_2, \ldots, \lambda_{N_\lambda}$ on $[0,1]$, where we set $N_\lambda = 21$ to match the discretization in the data. Similarly, we specify grids of linearly spaced points $\mu_1, \mu_2, \ldots, \mu_{N_\mu}$ on $[0,1]$ and log-linearly spaced points $\sigma_1, \sigma_2, \ldots, \mu_{N_\sigma}$ on $[0, 1/2]$, with $N_\mu = N_\sigma = 20$, to define the belief process.[17]

The parameterized model is described by four parameters, $\omega \equiv (\alpha, \rho, \mu_\epsilon, \sigma_\epsilon)$. We choose ω to match, as closely as possible, a list of empirical moments $\hat{M}$ (described below). Let $M(\omega)$ denote the mapping from ω to the corresponding model moments. A detailed description of the algorithm implementing M is given in Appendix C. Our estimator for ω is given by

$$\hat{\omega} = \arg\min_{\omega} \left(\hat{M} - M(\omega)\right)' \hat{V}^{-1} \left(\hat{M} - M(\omega)\right), \tag{15}$$

where $\hat{V}$ is a diagonal matrix with the bootstrapped variances of $\hat{M}$ along the diagonal. The estimated parameter values are $\hat{\alpha} = .569$, $\hat{\mu}_\epsilon/(1-\hat{\rho}) = .736$, $\hat{\sigma}_\epsilon^2/(1-\hat{\rho}^2) = .030$, and $\hat{\rho} = .9997$. These values imply that elites in an autocratic system enjoy roughly 60 percent higher value than citizen in a full-scale democracy. The process for θ is highly persistent, with unconditional mean of .736 and unconditional variance of .030.

4.2 Predicted Dynamics

The empirical and simulated moments, targeted by our estimation, are presented in Table 2. All model moments are computed at the stationary distribution. Specifically, we target nine moments, chosen to reflect the regularities presented in Section 2: (i) the co-occurrence of revolts and reforms, summarized by the ratio of revolts to reforms; (ii) the negative relation between transition hazards and regime maturity, summarized by the revolt and reform hazard for new regimes relative to the respective average hazards; (iii) the inverse J-shape of transition hazards in political inclusiveness, summarized by the hazard ratio at the peak of the inverse J-curve and at the most inclusive system ($\lambda = \lambda_{N_\lambda}$), both normalized relative to the least

each point of time, our algorithm computes the transition function $\Pr[(\lambda_{t+1}, \mu_{t+1}, \sigma_{t+1}^2)|(\lambda_t, \mu_t, \sigma_t^2)]$ by first solving the generation game conditionally on $(\theta_t, \lambda_t, \mu_t, \sigma_t)$ and then integrating over θ_t using F_t as probability measure. The resulting marginal distribution over δ—which is sufficient to compute all moments of interest to us—is identical to the one resulting from solving the model on its full state space, since Bayesian consistency requires that for any δ, the unconditional distribution over $\mathcal{S}$, denoted by $\mathcal{P}$, satisfies $\mathcal{P}(\theta|\delta) = F(\theta|\delta)$.

[16] Specifically, we chose *thresholds* $\{.025, .075, \ldots, .975\}$, defining the edges between two adjacent grid points $\{\lambda_i, \lambda_{i+1}\}$, such that the mid-points of each λ-bin, $\{.05, .1 \ldots, .95\}$, match the desired discretization of λ in the interior of the grid. At the boundaries, we obtain $\lambda_1 = .0125$ and $\lambda_{N_\lambda} = .9875$ as the mid-points of the two most extreme λ-bins.

[17] Observe that the standard deviation of the Beta distribution is bounded above by $1/2$. We chose a log-linearly spaced grid for σ_t as the distribution over σ_t is strongly right-skewed.

Table 2: Data Moments and Model Simulated Moments

Fitted Moments	Data	Model
Revolt-reform ratio	2.85	2.82
	(0.39)	
Revolt-hazard for new regimes/avg. hazard	3.04	2.04
	(0.86)	
Reform-hazard for new regimes/avg. hazard	6.90	8.95
	(2.72)	
Total transition hazard by λ: peak/autocracy	4.29	3.21
	(1.92)	
Total transition hazard by λ: democracy/autocracy	0.18	0.04
	(0.16)	
Median revolt	0.20	0.20
	(0.02)	
Median reform	0.80	0.80
	(0.02)	
Unconditional mass on $\lambda \leq 0.25$	0.42	0.47
	(0.03)	
Unconditional mass on $\lambda \geq 0.75$	0.42	0.47
	(0.04)	

Notes.—Bootstrap standard errors for the data, clustered at the country-level, are in parenthesis. All model moments are computed at the stationary distribution. The empirical moments are based on the data presented in Section 2.

inclusive system ($\lambda = \lambda_1$); (iv) the polarization of new regimes, summarized by the median revolt and reform; and (v) the concentration of mass towards the extremes of the stationary distribution $\mathcal{P}(\lambda)$, summarized by $\mathcal{P}(\lambda \leq 0.25)$ and $\mathcal{P}(\lambda \geq 0.75)$.

Overall, the model fits the targeted moments quite well, with most model moments being within one standard deviation of their empirical counterpart. The exception is the stationary distribution, where the model over-predicts the concentration towards the extremes by 1.5 standard deviations, and the revolt hazard for new regimes, which the model under-predicts relative to the corresponding average rate. Despite some discrepancies, the estimated model is clearly able to replicate the documented regularities.

For further evaluation of the model fit, we next study how well the model matches the precise shape of the transition hazards, conditional outcome distributions, and stationary distribution depicted in Figures 1–4. Beyond targeting the statistics in Table 2, none of these shapes are targeted by our estimation.

Figure 7 shows the corresponding relations for the estimated model. For convenience, the graphs also include the empirical relations from Section 2. Overall the model fits the data

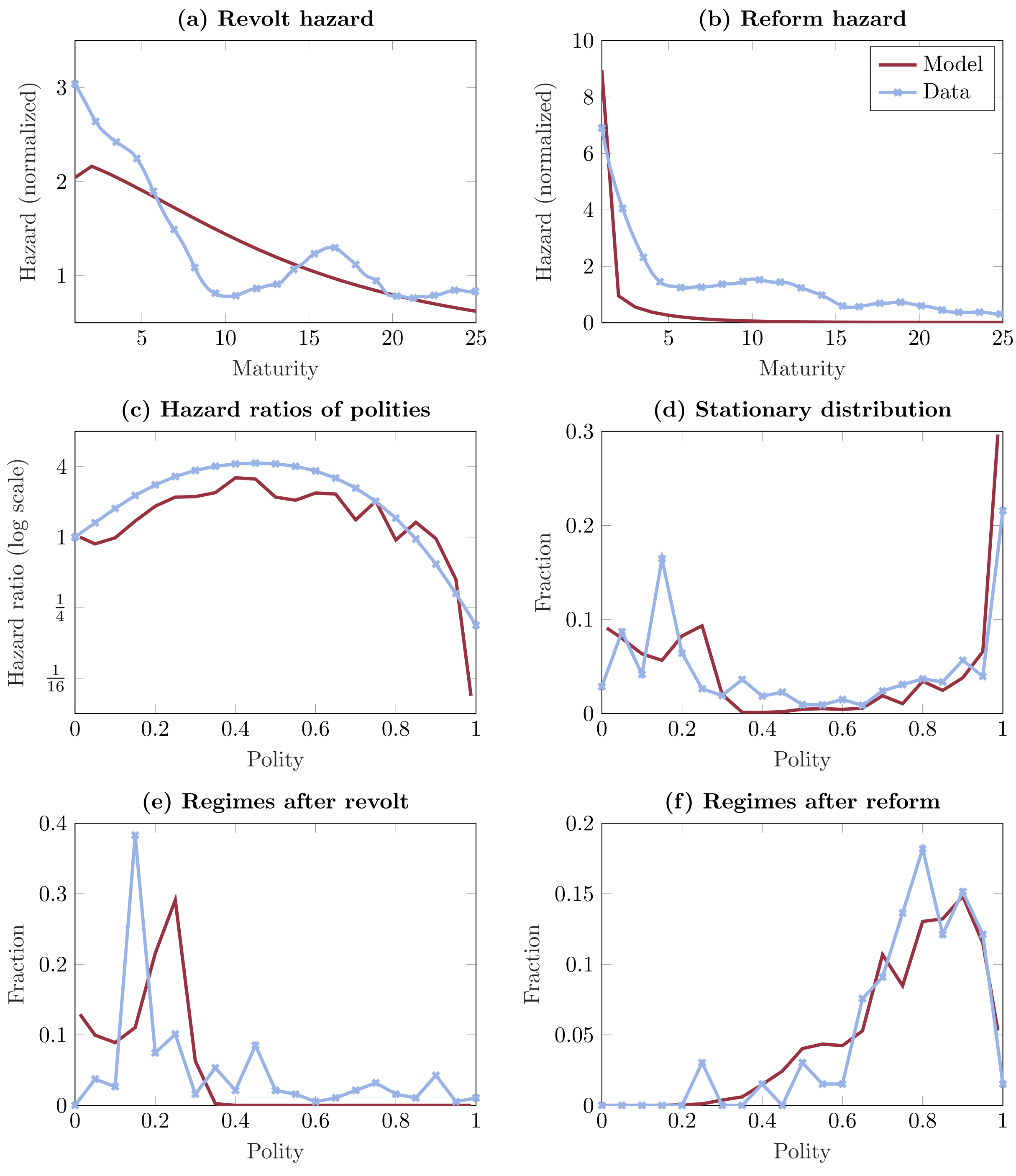

Figure 7: Comparison between model (solid red lines) and data (crossed blue lines). None of the depicted relations are directly targeted by the estimation.

very well. We do not fully capture the shape of the relation between revolt hazard and regime maturity and we slightly underpredict the reform hazard for very mature regimes, but we capture the average rates at which these hazards decline in maturity—steeply for reforms and relatively slowly for revolts. Similarly, the model captures the inverse J-shape of transition hazards in inclusiveness, although it slightly underpredicts the hazard for the most inclusive regimes. Finally, the fit of the conditional outcome distributions and stationary distribution is almost perfect.

5 Understanding the Key Features of the Model

In this section, we provide intuition for how the different features of the model contribute to explaining the empirical facts.

Co-occurrence of revolts and reforms In the estimated model, revolts are almost three times as likely as reforms. Why are there so many coercive transitions if regimes could preempt any revolt by extending political power to outsiders?

There are two reasons. First, reforms are costly so that regimes are willing to tolerate some risk of failure in order to hold on to power. If sharing power would bear no costs ($\alpha_u \to 0$), then clearly any regime would immediately transform to a perfectly inclusive democracy and there were no incentives to ever revolt. Second, as detailed in Proposition 2, asymmetric information reduces the effectiveness of reforms and detains regimes from conducting small reforms altogether. If instead the realization of θ_t would be observed by outsiders, reforms have no signaling value and simply solve

$$x^{\text{sym}}(\lambda, \theta) = \arg \max_{x \in [\lambda, 1]} \tilde{V}^I(\theta, \theta, x).$$

Generally, x^{sym} lies strictly above the equilibrium schedule characterized in Proposition 2. I.e., not only are small reforms precluded by asymmetric information, but more generally they are biased downwards. Asymmetric information therefore reduces the likelihood of reforms and tends to increase the likelihood of revolts (see Figure 20 in the appendix for an illustration).

To gauge the quantitative importance of these factors, we re-solve the model for different values of α_u and for the case with symmetric information. All other parameters remain fixed at their estimated values. Table 3 reports the resulting revolt-reform ratios and the mass on "autocracies" (with $\lambda \leq .25$) relative to "democracies" (with $\lambda \geq .75$) at the stationary distribution. Clearly, both of the aforementioned factors are crucial for revolts to be the

Table 3: Frequency of Revolts for Alternative Parameters and Without Asymmetric Information

Cost of sharing power (α_u)	Asymmetric information: Revolt-reform ratio	Asymmetric information: Autocracy-democracy ratio	Symmetric information: Revolt-reform ratio	Symmetric information: Autocracy-democracy ratio
0.15	0.20	0.03	0.01	0.03
0.30	0.95	0.30	0.18	0.20
0.45	1.86	0.39	1.07	0.52
0.57	2.82	0.99	1.48	0.46

Notes.—The model is solved for different values of α_u. All other parameters are fixed at their estimated values. The autocracy-democracy ratio defines the mass of regimes with $\lambda \leq .25$ relative to the mass of regimes with $\lambda \geq .75$ at the stationary distribution.

dominant driver of transitions. If information were symmetric, then for any value of α_u,[18] the revolt-reform ratio drops below 1.5 (compared to 2.85 in the data), and $\mathcal{P}(\lambda \geq .75)$ exceeds $\mathcal{P}(\lambda \leq .25)$ by a factor of at least two (compared to equal shares in the data).

Transition hazards and maturity Consider next the declining shape of transition hazards in regime maturity. The driving force behind this is a perceived "stabilization", reflected in a decline in outsiders' prior mean, μ_t, as a regime becomes more mature. Specifically, from equation (11), it follows that for any $\mathcal{S}_t$,[19]

$$\mu_{t+1}|(\text{reform}_t) \geq \mu_{t+1}|(\text{revolt}_t) > \mu_{t+1}|(\text{no transition}_t). \tag{16}$$

After reforms (and revolts against reforming regimes), outsiders fully learn θ_t, which conditionally on a reform is larger than $\bar{\theta}_t$. Similarly, Bayesian updating implies that θ_t is likely to be high 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. As a regime ages, it is therefore perceived to be less and less vulnerable. Accordingly, joining a revolt becomes less and less attractive, reducing both the number of outsiders supporting a revolt and the incentives of insiders to reform.[20]

[18]Recall that α_u is bounded above by α_h, limiting the permissible range of α_u to $[0, \alpha_h]$. While increasing both α_u and α_h jointly may in principle increase the relative frequency of revolts, it does not do so in our case as an higher elasticity α_h *increases* the effectiveness of reforms, thwarting the effect of increasing their costs.

[19]To see this, recall that M^2 defines the second *raw* moment, which is bounded by $(M^1)^2 < M^2 \leq M^1$ (the lower bound is strict as $\text{Var}_{\tilde{F}}[\theta_t] = M^2 - (M^1)^2 > 0$ for $x_t = \lambda_t$). Evaluating (11) at the upper and lower bound for M^2, respectively, and combining with (13) yields the two inequalities stated in (16).

[20]If θ_t is unobserved to the statistician (as it is both in our computations and in the data), the belief effect is further strengthened by statistical selection, which, similarly to outsiders' beliefs, places more probability mass on stable realizations of θ_t for older regimes. See the second step in the proof of Proposition 5 for a

Observe that (16) holds under any F_t and does not hinge on the shape of G (or on our Beta-approximation to (10)). To strengthen this point, consider the limit where $\Pr(\theta_t = \theta_{t-1}) \to 1$ (θ is fully persistent) and F_t is computed exactly (without approximation).

Proposition 5. *Let G such that $\Pr(\theta_t = \theta_{t-1}) \to 1$. Then the revolt and reform hazards are decreasing in the maturity of a regime. Specifically, for any $\mathcal{S}_0$, if $x_s = \lambda_0$ and $\eta_s = 0$ for all $s < t$, then*

$$\Pr_t(\eta_t = 1) < \Pr_{t-1}(\eta_{t-1} = 1) \quad \textit{and} \quad \Pr_t(x_t > \lambda_t) = 0.$$

Proposition 5 proves for perfectly persistent θ that the revolt and reform hazards are declining in regime maturity. The decline in the reform hazard is especially stark, as it drops to zero for all but newly emerged regimes. While this extreme decline in the reform hazard is an artifact of $\Pr(\theta_t = \theta_{t-1}) \to 1$, it is reminiscent of the steep decline seen in the estimated model and the data.

Transition hazards and inclusiveness The inverse J-shape of transition hazards in λ is the result of two opposing forces. On the one hand, as just explained, transition hazards are increasing in the prior mean μ_t. On the other hand, transition hazards are declining in λ_t. The logic is similar to the one driving the dependence on μ_t: As the regime becomes more inclusive, revolts are more likely to fail, which makes it even less attractive for remaining outsiders to support a revolt and further reduces incentives for insiders to reform.

These two forces are opposing, because μ_t is positively linked to λ_t through statistical selection: As further detailed below, large regimes emerge from reforms, implying that they are perceived to be weak ($\tilde{\mu}_{t-1} = \theta_{t-1} \geq \bar{\theta}_{t-1}$), whereas small regimes typically emerge from revolts against pooling regimes, implying that they are perceived to be relatively strong ($\tilde{\mu}_{t-1} \leq \bar{\theta}_{t-1}$).[21] Moreover, because s is increasing in the perceived likelihood of success (Proposition 1), it is precisely revolts that ex-ante were considered as futile that give rise to the smallest regimes. Conversely, because x is increasing in θ (Proposition 2), the largest regimes will be associated with being weakest upon their emergence.

Figure 8 illustrates these two forces. The right panel shows the statistical relation between μ_t and λ_t. The left panel plots the marginal transition probability with respect to μ_t and λ_t. The black line traces out the contour, $(\mu, \lambda) = (\mathbb{E}[\mu|\lambda], \lambda)$, of the relation shown in the right

precise decomposition of the overall decline in hazard rates into the effects of outsiders' beliefs and statistical selection.

[21] Revolts are more likely to succeed against pooling regimes, because reforms must be effective in reducing the threat of revolt to be observed in equilibrium.

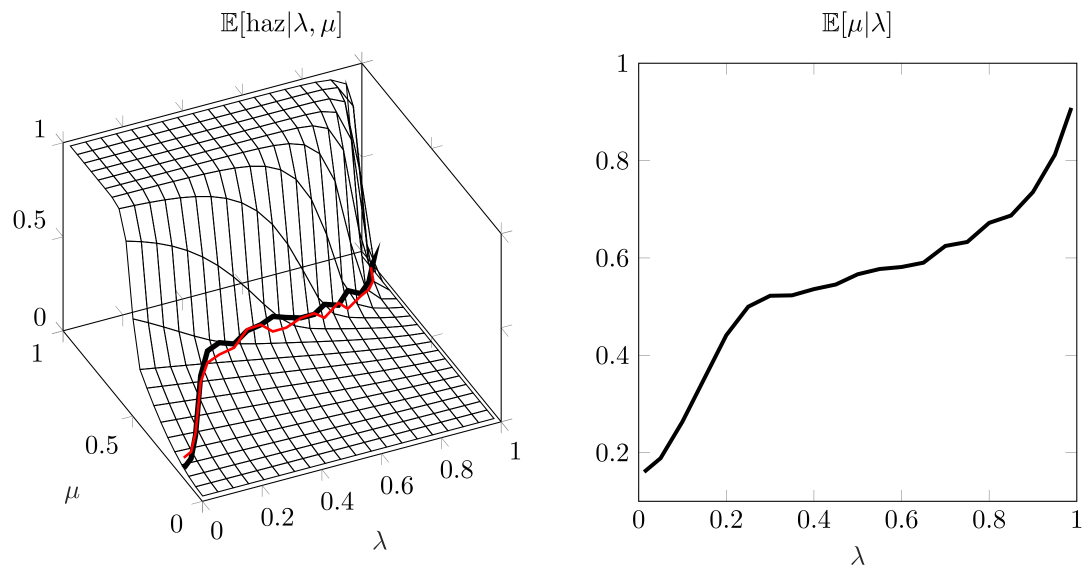

Figure 8: Intuition for inverse J-curve. Left panel: marginal transition hazard $\mathbb{E}[\text{haz}|\mu, \lambda]$; solid black line traces contour of $(\mu, \lambda) = (\mathbb{E}[\mu|\lambda], \lambda)$; red line depicts $\mathbb{E}[\text{haz}|\lambda]$. Right panel: $\mathbb{E}[\mu|\lambda]$. All expectations are computed at the stationary distribution.

panel, which closely approximates the exact J-curve $\mathbb{E}[\text{haz}|\lambda]$ (depicted in red).[22] Relative to regimes at the center of the autocracy-democracy spectrum, full-scale democracies ($\lambda \to 1$) are stable due to a lack of opposition (dominating their perceived weakness). Extreme autocracies ($\lambda \to 0$), on the other hand, are similarly uncontested due to their perceived stability implying a low degree of coordination among outsiders.

Polarization of new regimes The logic behind the polarization of new regimes is straightforward. By Proposition 2, reforms are bounded below by $\xi(\bar{\theta}_t)$, since smaller reforms would be ineffective in reducing revolutionary pressure. Conversely, revolts cannot grow too large, since otherwise insiders would prefer to preempt them if they are vulnerable. In turn, outsiders can infer the regime to be strong if it does not preempt a large revolt, making it unattractive to join such a revolt in the first place. These considerations imply state-dependent bounds $\bar{\lambda}^{\text{ref}}(\lambda_t, F_t)$ and $\bar{\lambda}^{\text{rev}}(\lambda_t, F_t)$ such that for all $\theta_t \in [0, 1]$,

$$s_t \leq \bar{\lambda}^{\text{rev}}(\lambda_t, F_t) \quad \text{and} \quad x_t \geq \bar{\lambda}^{\text{ref}}(\lambda_t, F_t).$$

[22] $\mathbb{E}[\text{haz}(\mathbb{E}[\mu|\lambda], \lambda)]$ is only approximate for two reasons. First, since the transition hazard is nonlinear in μ_t, there is an approximation error associated with evaluating the hazard at the average μ for each λ as depicted in the right panel (as opposed to computing the average hazard over the conditional distribution $\mu|\lambda$). Second, for our illustration, we have abstracted from the impact of σ^2 on the transition hazard, by marginalizing the hazard with respect to μ and λ, yielding another approximation error due to nonlinearity in σ. Comparing the approximation with the exact J-curve (in red), it can be seen that the difference is small, so that the main force behind the J-curve is indeed the statistical link between μ and λ shown in the right panel.

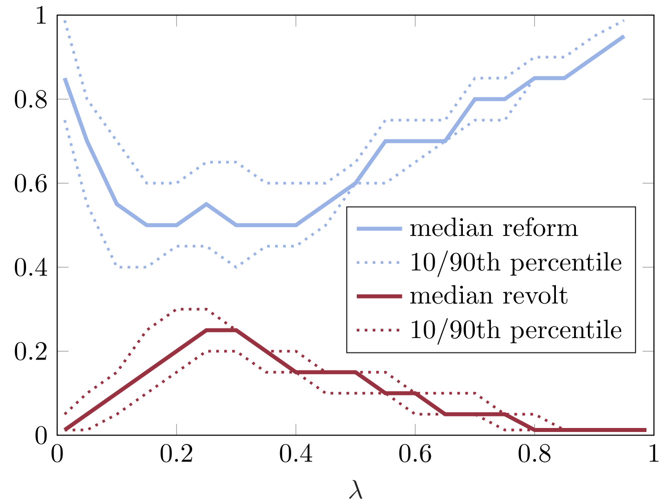

Figure 9: Median reform and revolt conditional on originating regime.

While it is difficult to characterize these bounds fully analytically, it is possible to derive somewhat more conservative bounds, as stated in the following proposition.

Proposition 6. *$\bar{\lambda}_t^{\mathrm{ref}} > 1 - (1-\lambda)M_t^1(\bar{\theta}_t)/\bar{\theta}_t$ and $\bar{\lambda}_t^{\mathrm{rev}} < (1-\lambda)M_t^1(\bar{\theta}_t)$.*

For instance, if outsiders have a uniform prior ($F_t(\theta) = \theta$), then $M_t^1(\bar{\theta}_t) = \bar{\theta}_t/2$, implying $\bar{\lambda}_t^{\mathrm{ref}} > 1 - (1-\lambda_t)/2$ and $\bar{\lambda}_t^{\mathrm{rev}} < (1-\lambda_t)/2$. For a more general illustration, consider Figure 9. Here we plot the median revolt and reform, computed conditionally on λ_t, along with the 10th and 90th percentiles. The figure reveals that the polarization is strongest for transitions originating in regimes towards the extremes of the autocracy-democracy spectrum. The underlying logic is the flipside of the inverse J-curve: as extremely autocratic and democratic regimes face small equilibrium threats, only few outsiders revolt, and consequentially only regimes with large realizations of θ_t reform. This implies low levels of s_t and large values of $x_t = \xi(\theta_t)$.

Stationary distribution Finally, the bi-modal shape of the stationary distribution over polities is a simple corollary to the polarization of new regimes, depicted in panels (e) and (f) of Figure 7, and the inverse J-shape of the transition hazard in λ, as shown in panel (c).

6 Model Implications for Democratization

In this section, we study the implications of the model for the formation and survival of democracies.

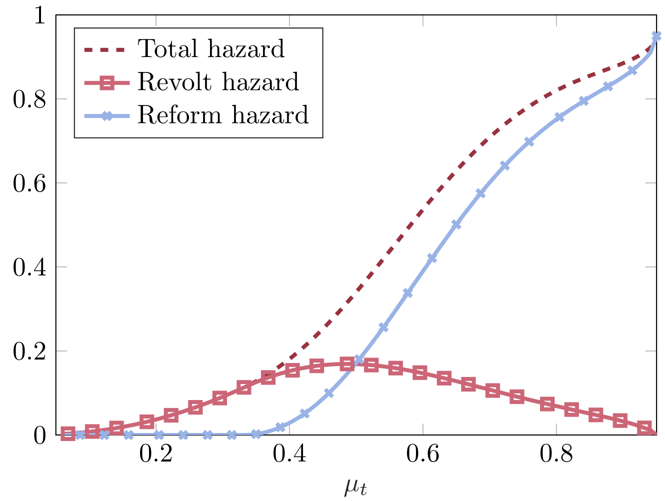

Figure 10: The role of outsiders' beliefs for transitions.

Outsiders' sentiments and critical junctures In the model, the belief or "sentiment" of outsiders is instrumental for creating a window of opportunity, in which democratization is possible. Figure 10 illustrates the role of beliefs by plotting the predicted transition hazards as a function of μ_t (for fixed values of λ and σ).[23] If outsiders perceive the regime as sufficiently stable (μ_t is small), revolts constitute little threat and insiders abstain from reforms, independently from the current realization of θ_t (i.e., $\bar{\theta}(\lambda_t, F_t) = 1$). If, by contrast, outsiders perceive the regime as vulnerable, insiders anticipate them to coordinate on potentially large revolts and are inclined to implement democratic reforms to preempt them. Because reforms are effective in reducing revolutionary pressure, the revolt hazard is hump-shaped in μ_t, even though the total transition hazard is increasing.

Interestingly, there is a region of intermediate values of μ_t, in which both transition hazards are high. This is because insiders generally do not find it optimal to fully preempt revolts (see the discussion in the previous section). Periods with intermediate values of μ_t thus constitute "critical junctures", during which small and random variations in current conditions determine whether a regime ultimately implements democratic reforms, is replaced by an autocracy, or remains unchanged (see Acemoglu et al., 2008, 2009 for empirical evidence in support of such critical junctures).[24] At the same time, because democracies and autocracies both stabilize once they mature, any system that eventually emerges at the end of a critical juncture is

[23]Here, λ and σ are fixed at .1 and .0475, respectively, but the relationship is largely insensitive in their precise values as long as λ is not too large. For large values of λ, all three hazards are significantly reduced in their magnitude, while maintaining the same qualitative shape.

[24]Observe how marginal variations in λ_t and F_t can have large and persistent effects on λ_{t+1} due to the discontinuity of $\bar{\theta}(\lambda_t, F_t)$ around θ_t. Conditionally on λ_t and F_t, outcomes are determined by the random realizations of θ_t and η_t.

likely to persist for a long time.

Illustration Figure 11 shows the dynamic responses to a counterfactual change in outsiders' beliefs, illustrating the arrival of a critical juncture and subsequent political stabilization. The time series is initialized at a "fully matured" autocracy, with $\lambda_0 = .1$ and beliefs given by their corresponding steady state values in the absence of any transition. As explained by the low value of μ_0, the initial reform and revolt hazards are close to zero. The time path shows the response to a counterfactual change in outsiders' beliefs at $t = 2$, resetting F_2 to a uniform prior.[25]

As seen in the bottom three panels, the belief shock at $t = 2$ leads to an immediate rise in the reform hazard to roughly 20 percent and the revolt hazard to roughly 15 percent. Absent any transition, μ_t drops in the sequel, causing a sharp decline in the reform hazard and a moderate decline in the revolt hazard seen at $t = 3$ and $t = 4$. If the regime is overthrown, as we assume it is at the end of $t = 4$, we see another increase in μ_t and the transition hazards. Observe that this serial correlation of transition hazards implies that critical junctures often consist of multiple transition events.

In our illustration, we consider two alternative time paths. The solid red path shows how the autocracy stabilizes in the absence of further transitions, eventually leading to a reform hazard of zero and a revolt hazard that converges to less than 1 percent after roughly 100 periods. The dashed blue path shows how the time path diverges if instead insiders implement a democratic reform at $t = 6$, which occurs at a rate of 1.5 percent. Here, the inclusion of a large fraction of the population into the regime leads to an immediate drop in the revolt hazard to 2.7 percent, despite the strong increase in μ_t. Absent further transitions, the reform hazard subsequently drops to zero and the revolt hazard eventually drops close to zero, albeit at a much smaller rate than before.[26]

[25] In the model, large belief changes are induced by (small-probability) transition events. While this means that the arrival of critical junctures is inextricably tied to regime changes, it is primarily the beliefs that are important for the subsequent dynamics. The experiment conducted here is designed to illustrate the pure impact of beliefs on transition probabilities by inducing the change in outsiders' sentiments exogenously. While absent in the estimated model, it would be straightforward to incorporate noisy signals into our framework that would rationalize such belief shifts, capturing, for instance, sentiment shifts triggered by the deaths of political leaders or by events in neighboring countries as in Buera, Monge-Naranjo and Primiceri (2011).

[26] The rate of stabilization is low due to the large value estimated for ρ, which governs the "usefulness" of past information for forming beliefs regarding θ_t. As θ_t is fully revealed through reforms, a high value of ρ implies that new information has little effect on outsiders' beliefs in the aftermath of a reform. Only over time, as the underlying state of θ_t changes (unobserved by outsiders) according to its law of motion, the precision of outsiders' beliefs eventually falls and beliefs are adjusted at a higher rate.

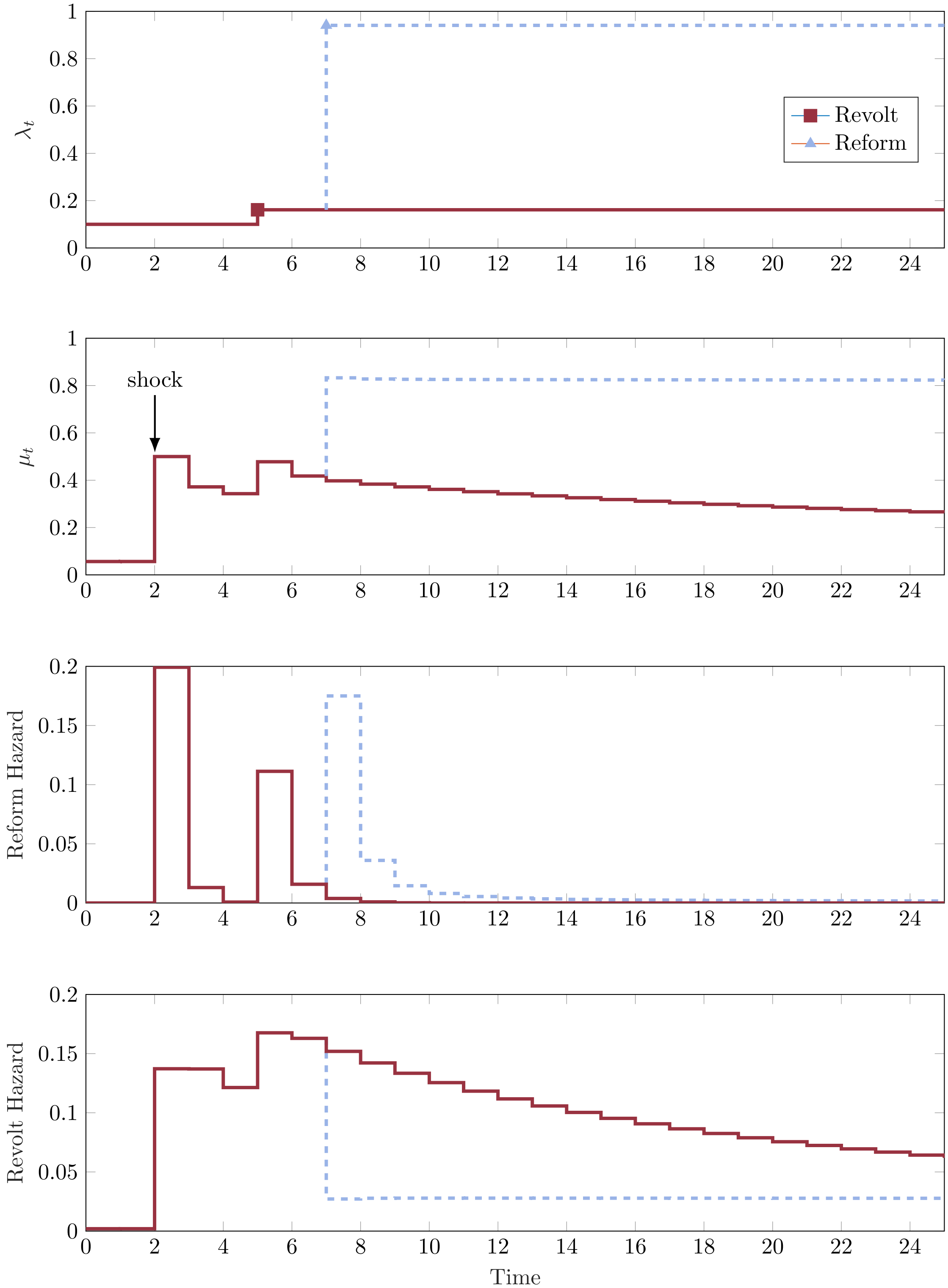

Figure 11: Critical junctures and political stabilization. Solid red lines show the dynamic response to a counterfactual change in outsiders' beliefs at $t = 2$ and a subsequent revolt at $t = 4$. Dashed blue lines show an alternative time path with an additional reform at $t = 6$. See the main text for further details.

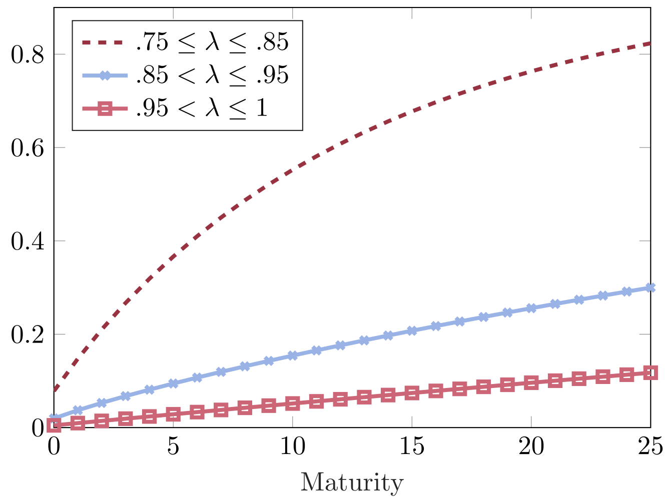

Figure 12: Cumulative reversal rate of young democracies.

Likelihood of successful democratization From the illustration in Figure 11 it is evident that newly emerging democracies face a non-trivial probability of a regime reversal. This is because outsiders excluded from the franchise extension learn the regime's vulnerability, leading to small but highly coordinated coup d'états. To study the relevance of reversals more broadly, we have simulated a time series of 10 Million observations from the estimated model. Using this time path, we compute the reversal rate of young democracies (all new regimes with $\lambda \geq .75$) as a function of their maturity and the inclusiveness of the democracy *at the time of its formation*.[27] The results are presented in Figure 12. Critical for the success of democratization is that the establishing reforms are comprehensive. Whereas the probability that a democracy with an initial polity of $\lambda \leq .85$ is overthrown in its first 25 periods is over 80 percent, the same probability drops to roughly 30 percent for democracies with an initial polity between .85 and .95, and drops to 12 percent for democracies that are initially larger than .95.

Are democracies absorbing? A related question is whether democracies are always bound to fail (albeit with a small probability), or if there is the possibility of an absorbing regime. In the model, transition hazards are strictly positive for any regime with $\lambda < 1$. Only a perfectly inclusive democracy with $\lambda = 1$ is absorbing. In the estimated version of the model, this is ruled out by our discrete approximation to λ, as $\lambda_{N_\lambda} = .9875$ (see Footnote 16 for details). But would an absorbing democracy eventually arise if we solved the model in a continuous state space? The answer depends on the value of $\xi(1)$, which determines the

[27] Here we do not count consecutive reforms as regime failures, so that the inclusiveness may change over the life time of a democracy.

largest democracy that is formed along the equilibrium path. Given the estimated value for α, we have $\xi(1) = .978$, so that fully inclusive democracies indeed do not emerge in equilibrium, even if we solve the model on a continuous polity space.

More generally, under which circumstances does an absorbing democracy emerge in equilibrium? From the boundary condition for ξ, stated in Proposition 2, $\xi(1) = 1$ if $\lim_{x\to 1} \tilde{V}_3^I(1, 1, x) \geq 0$. Intuitively, this requires $h(s(\cdot, x))$ to be sufficiently steep around $x = 1$ to compensate for the cost of reforms, $u'(1)$. With the parametrization for u and h used in the estimation, the condition reduces to a simple threshold in the elasticity of p with respect to s.

Proposition 7. *Let $u(x) = 1+\alpha_u(1-x)$ and $h(s) = s^{\alpha_h}$. Then, $\xi(1) = 1$ so that an absorbing democracy with $\lambda \to 1$ emerges along the equilibrium path (a.s.) if and only if $\alpha_h \leq .5$.*

If the success rate of revolts is relatively *inelastic* in the number of supporters ($\alpha_h \leq .5$), outsiders' coordination will not adjust strongly in response to reforms. To effectively reduce revolutionary pressure, insiders therefore mainly rely on the extensive margin of reforms, leading to (almost) absorbing democracies along the equilibrium path.[28] By contrast, if $\alpha_h > .5$, small groups of outsiders have a comparably low intensity of coordination and excluding them does not pose severe threats. In this case, democracies are always bounded away from $\lambda = 1$, so that reversals are observed with strictly positive probability against any regime.

7 Discussion

This section relates the mechanism and predictions of our model to the existing literature on political transitions and regime stability. We begin by highlighting the key components of our model and how they relate to the literature. We then discuss alternative mechanism and examine their ability to account for the empirical regularities motivating this paper.

[28]For a simple illustration, suppose there were no intensive margin of coordination; i.e., $s_t = (1 - x_t)\bar{\gamma}$ for a constant $\bar{\gamma}$. Then $\lim_{x\to 1} \tilde{V}_3^I(1, 1, x) = -\alpha_u + \bar{\gamma}\alpha_h \lim_{x\to 1} s(1, x)^{\alpha_h - 1}$, where the first term reflects the cost of reforms and the second term the reduction in the revolt threat. Clearly, as $x \to 1$ and $s \to 0$, the second term goes to ∞. Hence, absent any intensive margin, regimes with $\theta = 1$ always prefer to establish a fully inclusive regime. By contrast, if there is a (sufficiently elastic) intensive margin, then as x increases, outsiders internalize the impact on p and coordinate less intensively. As a result, insiders face lower threats for large (but not fully inclusive) regimes, reducing the marginal benefit, $-\partial h(s(\cdot, x))/\partial x$, of implementing fully inclusive reforms.

7.1 Mechanism

Transitions in our model are shaped by three main features, which in combination explain the empirical patterns documented in Section 2.

Coordinating revolts First, revolts are the outcome of a coordination game, where the degree of equilibrium coordination is crucially shaped by the beliefs of outsiders. This is in line with a long tradition in political science that views revolts as the result of individual participation choices (e.g., Tullock, 1971; Granovetter, 1978; Kuran, 1989; Casper and Tyson, 2014), and links them to the arrival of new information (e.g., Lohmann, 1994; Chwe, 2000; Bueno de Mesquita, 2010; Fearon, 2011; Shadmehr and Bernhardt, 2011; Edmond, 2013).[29]

In the context of our model, the coordination game adds an intensive margin to revolting, which together with asymmetric information prevents marginal reforms and helps explaining the prevalence of revolts (fact 1). At the same time, as emphasized in Section 5, it is precisely the drop in coordination along the intensive margin that is responsible for the stabilization of mature and small regimes, helping us to explain facts 2 and 3. Finally, absent the intensive margin, revolts would mechanically increase for small λ_t, leading to counterfactually large revolts that are at odds with fact 4.

Preemptive reforms Second, we model reforms as means to *credibly* preempt a looming revolt, as has been standard in the literature since the seminal works of Acemoglu and Robinson (2000*b*), Conley and Temini (2001), and Boix (2003).[30] Vis-à-vis these papers, a distinguishing feature of reforms in our model is that they are of endogenous scope, which is at the core of our quantitative exploration.

Asymmetric information Finally, we introduce asymmetric information, which is instrumental for the prevalence of revolts (fact 1), as well as the belief dynamics explaining the stabilization of mature regimes (fact 2). While the dynamic implications are new, it is relatively standard in the literature to introduce asymmetric information—and, hence, the need for regimes to consider the signaling value of their politics—to generate costly conflict along the equilibrium path (e.g., Acemoglu and Robinson, 2000*a*; Boix, 2003; Hirshleifer,

[29]In recent work, Aidt, Leon and Satchell (2017) empirically show that the spread of social unrest is fundamentally caused by information flows. They also provide an extensive review of the empirical literature regarding the role of coordination and information for the outbreak of social conflicts.

[30]Przeworski (2009), Aidt and Jensen (2014), and Aidt and Franck (2015) provide empirical evidence suggesting that preemptive reforms are indeed an important driving force of democratization.

Boldrin and Levine, 2009; Ellis and Fender, 2011).[31]

7.2 Predictions

This paper is first to develop a unified framework that can account for the transition patterns documented in Section 2. In the following, we explain how each of the model's predictions, in isolation, relates to the existing literature and discuss alternative mechanism.

Co-occurrence of revolts and reforms There are three main reasons present in the existing literature that explain co-existence between revolts and reforms along the equilibrium path. First, as commented on above, revolts can be rationalized by asymmetric information, which reduces the effectiveness of preemptive reforms (e.g., Acemoglu and Robinson, 2000*a*; Ellis and Fender, 2011; Boix, 2003). Second, if benefits from holding on to power are sufficiently large relative to the likelihood that a revolt succeeds, regimes may also find it optimal to "gamble for their survival" as discussed in Section 5 (see also Besley and Persson, 2018). A third reason, not present in our model, is limited commitment, which limits the compensation that can be credibly offered to outsiders so that conflict may arise whenever the constraint becomes binding (Acemoglu and Robinson, 2001; Acemoglu, Ticchi and Vindigni, 2010; Chassang and Padró i Miquel, 2009). Also, a particularly stark variant of limited commitment arises when elites have no means to appease potential rebels at all as in Ticchi, Verdier and Vindigni (2013).

A key difference between those works and this paper is that, with three exceptions, the above papers do not allow for repeated transitions. The exceptions are Acemoglu and Robinson (2001), Ticchi, Verdier and Vindigni (2013) and Besley and Persson (2018), who allow for random oscillations between autocratic and democratic regimes. However, as transitions are exogenously restricted to occur between two regimes of fixed size, these papers mechanically fix the revolt-reform ratio at unity[32] in addition to fully predetermining transition outcomes, preventing a quantitative analysis along the lines of this paper.

Transition hazards declining in maturity In our model, transition hazards decline in regime maturity due to a reduced coordination among outsiders caused by learning.[33] The

[31]Supporting the idea that politics may have a signaling value, 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.

[32]Besley and Persson (2018) consider two different modes of democratization. However, the relative frequency of democratization to regime reversals is similarly fixed at unity.

[33]Our argument is related to Gallego and Pitchik (2004), who previously pointed out that autocratic leaders with long tenure are likely to have low costs of averting coups. As noted earlier, such unobserved heterogeneity

literature has identified three alternative explanations, which in reduced form can be mapped into our framework as exogenous variations in θ, γ and u.

Specifically, the first strand of the literature has argued that young regimes are intrinsically more vulnerable compared to more mature ones (amounting to a drop in θ over time), because emerging democracies first need to establish institutions to disempower military leaders (Acemoglu, Ticchi and Vindigni, 2010) whereas emerging autocracies first need to establish institutions to effectively distribute economic rents to supporters (Svolik, 2009). Relatedly, a second strand of the literature has argued that societies become increasingly supportive of the current regime as political values adjust to political realities (Ticchi, Verdier and Vindigni, 2013; Besley and Persson, 2018), which in our framework could be interpreted as a shift in the distribution over political adjustment $\{\gamma_{i,t}\}$. Finally, Przeworski and Limongi (1997) offer empirical evidence that economic growth in the aftermath of democratization leads to political stabilization (see, however, Acemoglu et al., 2009). In the context of our framework, one possible interpretation would be that democratization may create an institutional environment supporting growth (e.g., Acemoglu and Robinson, 2008; Acemoglu, Ticchi and Vindigni, 2011), which over time increases the flow rents under democracy $u(\lambda \to 1)$ relative to other regime types, reducing the chance of regime reversals.[34]

One distinctive feature of our mechanism compared to these alternatives is that, with the exception of Besley and Persson (2018), the above explanations are specific to either autocratic or democratic consolidations. The belief-driven explanation in this paper, by contrast, applies to all regime types, explaining the equally universal decline in hazards seen in the data (see Appendix A.2 for details).

Inverse J-shape of transition hazards The observation that regime instability is inverse J-shaped in inclusiveness goes back to Bremmer (2006). In his monograph, Bremmer explains the "J-curve" with the ability of elites to control the information flow across society, which immanently varies across different regime types. While information is highly restricted in autocratic societies, inhibiting the coordination of revolts, it is precisely the free flow of information that enables democratic institutions to peacefully resolve any looming conflict. By contrast, intermediate regimes lack the institutions to preempt conflict whereas they are also ill-equipped to contain the spread of subversive ideas.[35] An alternative account is given

reinforces the belief dynamics emphasized in this paper through statistical selection.

[34] At the level of individual leaders, Ales, Maziero and Yared (2014) explain a decline in exit rates through (self-enforcing) contracts, where re-elected leaders are those who get rewarded for compliant behavior, increasing their flow utility and hence their propensity of future compliance.

[35] Relatedly, Gates et al. (2006) point out that it may not be the better access to information but the better access to societal resources that facilitate political change in intermediate regimes. They note that, compared

by Bueno de Mesquita et al. (2003), who refer to the distribution of wealth among elites across regime types to explain the J-curve. Specifically, Bueno de Mesquita et al. argue that democratic elites are more wealthy than elites in other societies due to a more efficient provision of public goods. Autocratic elites, by contrast, are similarly (albeit less) wealthy due to an efficient distribution of rents. Core supporters of intermediate regimes, by contrast, are lass wealthy, because they lack both efficiency in the provision of public goods and in the distribution of rents. Accordingly, intermediate regimes are less supported, which is interpreted as instability.

Both of these arguments essentially imply that a regime's polity λ and its internal weakness θ are inextricably linked. In contrast, our model explains the inverse "J-curve" through an endogenous correlation between λ and θ. Through the lens of our model, the wealth of autocratic elites and their tight grip on information flows that explain the stability of autocratic systems in Bueno de Mesquita et al. and Bremmer may hence be the manifestation—instead of the source—of their internal strength.

Polarization of new regimes A key novelty in this paper is the adoption of a continuous polity space, which enables us to make predictions about outcome distributions. The closest existing works are by Acemoglu and Robinson (2000*a*), Bourguignon and Verdier (2000), Lizzeri and Persico (2004), and Llavador and Oxoby (2005), which endogenize the scope of franchise extension but completely abstract from any regime dynamics that are at the center of our contribution.

Our approach to endogenize the outcomes of political transitions also relates to a growing literature on dynamic voting games, which studies stable equilibrium coalitions in rich state spaces (e.g., Lagunoff, 2009; Acemoglu, Egorov and Sonin, 2008, 2012). In particular related are Justman and Gradstein (1999); Gradstein (2007); Acemoglu, Egorov and Sonin (2015), who characterize transition paths leading to an (eventually) absorbing stable coalition.[36] However, given their focus on the *composition* of coalitions, these papers typically do not pin down a unique mode of transition between regimes, nor do they explicitly allow for violent conflict along the equilibrium path.

Bimodal distribution To the best of our knowledge, this paper is first to make predictions about the stationary distribution over regime types. The only other papers that we are aware

to heavily autocratic systems, the expansion of political participation gives "the opposition a better base from which to demand further decentralization" (p. 895).

[36]See also, Jack and Lagunoff (2006) and Bai and Lagunoff (2011) for dynamic voting games that do not necessarily lead to a stable coalition.

of that can account for the bimodal distribution seen in the data, albeit mechanically, are Acemoglu and Robinson (2001) and Ticchi, Verdier and Vindigni (2013), in which regimes oscillate randomly between autocracy and democracy.

8 Final Remarks

We have developed a quantitative theory of repeated political transitions based on the evolution of beliefs regarding the regime's strength. The model is distinguished from the existing literature by its ability to generate various patterns of regime change in a unified framework, including (possibly gradual) democratization processes, regime reversals against both emerging and mature democracies, and power struggles amongst autocratic regimes. We demonstrated the quantitative potential of the framework in a structural estimation targeting several key moments from the data. The estimation results suggest that a simple model based on the interplay between revolutionary pressure and preemptive reforms can generate a process of political transitions that looks remarkably close to the data. Crucial for the close fit is the presence of an intensive margin to revolting, which links revolutionary pressure to the beliefs of outsiders regarding the regime's strength.

References

Acemoglu, Daron, and James A. Robinson. 2000*a*. "Democratization or Repression?" *European Economic Review*, 44(4-6): 683–693.

Acemoglu, Daron, and James A. Robinson. 2000*b*. "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. 2008. "Persistence of Power, Elites, and Institutions." *The American Economic Review*, 98(1): 267–293.

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

Acemoglu, Daron, Davide Ticchi, and Andrea Vindigni. 2011. "Emergence and Persistence of Inefficient States." *Journal of the European Economic Association*, 9(2): 177–208.

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. 2015. "Political Economy in a Changing World." *Journal of Political Economy*, 123(5): 1038–1086.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. "Income and Democracy." *The American Economic Review*, 98(3): 808–842.

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2009. "Reevaluating the Modernization Hypothesis." *Journal of Monetary Economics*, 56(8): 1043–1058.

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.

Aidt, Toke S., Gabriel Leon, and Max Satchell. 2017. "The Social Dynamics of Riots: Evidence from the Captain Swing Riots, 1830-31." *Working Paper.*

Ales, Laurence, Pricila Maziero, and Pierre Yared. 2014. "A Theory of Political and Economic Cycles." *Journal of Economic Theory*, 153: 224–251.

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

Besley, Timothy, and Torsten Persson. 2018. "Democratic Values and Institutions." *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.

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

Bremmer, Ian Arthur. 2006. *The J Curve: A New Way to Understand Why Nations Rise and Fall.* New York:Simon & Schuster.

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

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. *The Logic of Political Survival.* Cambridge, Mass.:MIT Press.

Bueno de Mesquita, Ethan. 2010. "Regime Change and Revolutionary Entrepreneurs." *American Political Science Review*, 104(3): 446–466.

Buera, Francisco J., Alexander Monge-Naranjo, and Girorgio E. Primiceri. 2011. "Learning the Wealth of Nations." *Econometrica*, 79(1): 1–45.

Casper, Brett Allen, and Scott A. Tyson. 2014. "Popular Protest and Elite Coordination in a Coup d'état." *The Journal of Politics*, 76(2): 548–564.

Celestino, Mauricio Rivera, and Kristian Skrede Gleditsch. 2013. "Fresh Carnations or All Thorn, No Rose? Nonviolent Campaigns and Transitions in Autocracies." *Journal of Peace Research*, 50(3): 385–400.

Chassang, Sylvain, and Gerard Padró i Miquel. 2009. "Economic Shocks and Civil War." *Quarterly Journal of Political Science*, 4(3): 211–228.

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.

Derpanopoulos, George, Erica Frantz, Barbara Geddes, and Joseph Wright. 2016. "Are Coups Good for Democracy?" *Research & Politics*, 3(1): 205316801663083.

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

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.

Gallego, Maria, and Carolyn Pitchik. 2004. "An Economic Theory of Leadership Turnover." *Journal of Public Economics*, 88(12): 2361–2382.

Gates, Scott, Håvard Hegre, Mark P. Jones, and Håvard Strand. 2006. "Institutional Inconsistency and Political Instability: Polity Duration, 1800–2000." *American Journal of Political Science*, 50(4): 893–908.

Geddes, Barbara, Joseph Wright, and Erica Frantz. 2014. "Autocratic Breakdown and Regime Transitions: A New Data Set." *Perspectives on Politics*, 12(2): 313–331.

Gleditsch, Kristian Skrede, and Jinhee Lee Choung. 2004. "Autocratic Transitions and Democratization." *mimeo.*

Gleditsch, Kristian Skrede, and Michael D. Ward. 2006. "Diffusion and the International Context of Democratization." *International Organization*, 60(4): 141.

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.

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science*, 54(1): 190–208.

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.

Harrell, Frank E. 2001. *Regression Modeling Strategies.* New York, NY:Springer New York.

Hirshleifer, Jack, Michele Boldrin, and David K. Levine. 2009. "The Slippery Slope of Concession." *Economic Inquiry*, 47(2): 197–205.

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.

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.

Knutsen, Carl Henrik, and Håvard Mokleiv Nygård. 2015. "Institutional Characteristics and Regime Survival: Why Are Semi-Democracies Less Durable Than Autocracies and Democracies?" *American Journal of Political Science*, 59(3): 656–670.

Krusell, Per, and Anthony A. Smith, Jr. 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(2): 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.

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

Marshall, Monty G., Ted Robert Gurr, and Keith Jaggers. 2017. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016."

O'Donnell, Guillermo, and Philipp C. Schmitter. 1986. *Transitions from Authoritarian Rule: Tentative Conclusions about Uncertain Democracies.* Baltimore (MD):John Hopkins University Press.

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

Przeworski, Adam, and Fernando Limongi. 1997. "Modernization: Theories and Facts." *World Politics*, 49(2): 155–183.

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

Sanhueza, Ricardo. 1999. "The Hazard Rate of Political Regimes." *Public Choice*, 98(3/4): 337–367.

Shadmehr, Mehdi, and Dan Bernhardt. 2011. "Collective Action with Uncertain Payoffs: Coordination, Public Signals, and Punishment Dilemmas." *American Political Science Review*, 105(4): 829–851.

Svolik, Milan W. 2008. "Authoritarian Reversals and Democratic Consolidation." *American Political Science Review*, 102(2).

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

Svolik, Milan W. 2015. "Which Democracies Will Last? Coups, Incumbent Takeovers, and the Dynamic of Democratic Consolidation." *British Journal of Political Science*, 45(4): 715–738.

Ticchi, Davide, Thierry Verdier, and Andrea Vindigni. 2013. "Democracy, Dictatorship and the Cultural Transmission of Political Values." *IZA Discussion Paper*, 7441.

Tullock, Gordon. 1971. "The Paradox of Revolution." *Public Choice*, 11(1): 89–99.

A Data and Additional Empirical Results

A.1 Data

Our data combines information from three sources. It covers all *regime spells,* as collected by Geddes, Wright and Frantz (2014, GWF), that existed on January 1st of each year between 1946 and 2010 in countries with more than one million inhabitants. GWF define regimes by "the rules that identify the group from which leaders can come and [that] determine who influences leadership choice and policy" (p. 314).[37] Since GWF lists non-autocratic regimes only at a yearly frequency, we impute the begin (end) dates for non-autocratic regimes from (i) the end (begin) dates of the previous (next) regime and (ii) the begin date of the nearest Polity IV case within the same year. If (i) and (ii) yield no match, we encode the begin dates as July 1st and the end dates as June 30th.

We measure the inclusiveness of regimes using the polity score, normalized between 0 and 1, from the Polity IV Project (Marshall, Gurr and Jaggers, 2017), which ranks political regimes on a 21 point scale between autocratic and democratic. Specifically, we merge all polity spells listed in the "Polity IV Polity-Cases" dataset to our sample of GWF regime spells, harmonizing start dates on the basis of GWF spells whenever the start date of a polity case is within half a year (183 days) of a GWF start date. Otherwise, we keep track of changing polity scores by subdividing GWF spells into subspells.[38]

Third, we classify the GWF regime transitions primarily based on information provided by GWF (variable "howend"). If the information in GWF is unavailable, we match the GWF transitions to the nearest leader exit, taken from the Archigos database of political leaders by Goemans, Gleditsch and Chiozza, 2009, within half a year, and use the variables on the types of exit and entry to label the regime transitions. All violent regime transitions that are accompanied by popular protest, civil war, or coups are classified as *revolts.* Peaceful transitions where political insiders either actively change rules or newly allow for competitive elections or where there is no irregular leader change, are labeled *democratic reforms* when accompanied by an increase in the polity score and *autocratic consolidations* when accompanied by a decreasing polity score. All transitions influenced by foreign governments are called *foreign imposition.* All remaining transitions are collected in the residual

[37]Note that by focusing on the ruling group, the definition allows for leadership succession within regimes (if the identity of the ruling group remains unchanged) as well as regime changes without leadership replacement (if the leader stays in power despite a change in the ruling group, e.g., via reforms). Similarly, the definition allows for transitions that lead to a succession of regimes with similar scores of political inclusiveness.

[38]For some polity spells, Polity IV assigns "standardized authority scores" that do not fall into the autocracy-democracy range. The score of -66 encodes foreign "interruption", which we encode as missing. The polity scores of -77 ("interregnum") and -88 ("transition") identify transitional episodes. We interpret GWF regime transitions that occur during a transitional polity episode as the event defining the polity transition. Accordingly, transitional episodes just before a GWF transition are encoded with the last non-transitional polity score within the old GWF regime, while instances of transitional episodes just after a GWF transition are encoded via the first instance of a non-transitional polity score within the new GWF regime. (If the new GWF regime does not include a non-transitional polity score, we use the date of the next GWF transition to assign a date to the polity transition.) Finally, transitional episodes within a given GWF regime spell are encoded using the subsequent polity score.

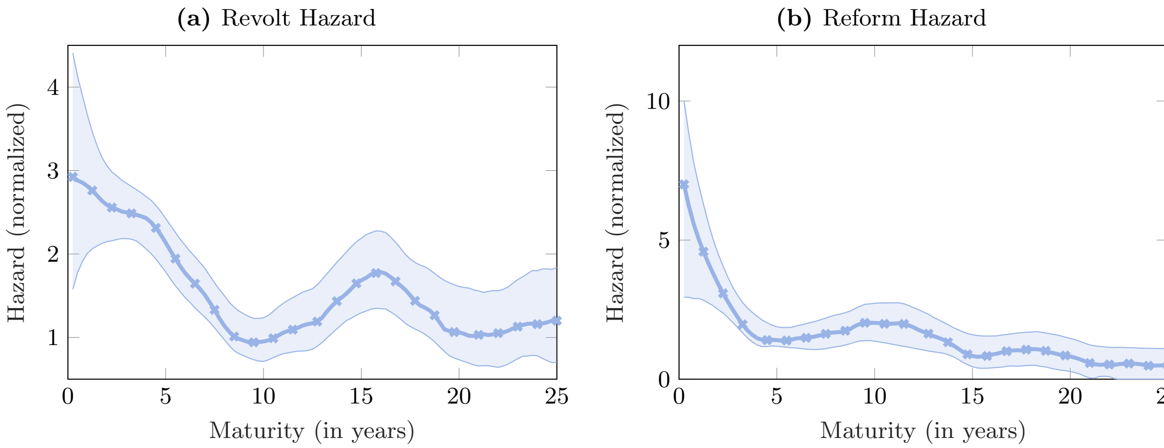

Figure 13: Empirical transition hazards for revolts (left panel) and reforms (right panel). Notes.—Hazards are the baseline hazards estimated via (17), normalized relative to the unconditional hazard of revolts and reforms, respectively. Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

category *other*.

The resulting database covers 494 regime spells in 155 countries covering a total of 8843.87 country-years.

A.2 Estimation of Transition Hazards

Transition hazards and regime maturity The hazards, reported in Figure 1, are estimated by differencing and smoothing over Nelson-Aalen estimates for the cumulative hazard rate, correcting for left and right censoring. Here we explore the robustness of the findings controlling for polity and region fixed effects.[39] Specifically, we use a Cox proportional hazard model, with hazard rate

$$p^s(\tau_{i,t}|\lambda_{i,t}, r_i) = h(\tau_{i,t}) \exp(f(\lambda_{i,t}) + r_i) \quad \text{for } s \in \{\text{reform, revolt}\}, \tag{17}$$

where $h(\tau_{i,t})$ is the baseline hazard, identified non-parametrically as a function of maturity $\tau_{i,t}$, f is a cubic spline in polity $\lambda_{i,t}$, and r_i are the region fixed effects.[40] Figure 13 plots the baseline hazard rates h for revolts and reforms, respectively. The results are similar to the ones in Figure 1, albeit with slightly larger confidence intervals (the loss in precision is expected given the small number of transition events and the large number of explanatory variables included in the current specification).

[39] Region definitions are based on the United Nations geoscheme, which we use to define 10 distinct regions in total (Eastern Europe, Eastern and Central Asia, Middle America, Northern Africa and Arabic Peninsula, South America, South-Eastern Asia, Western and Central Africa, Western Europe, Western Offshoots). Note that disentangling the geographic controls further is likely to cause incidental parameters problems (biasing our estimators given their nonlinearity), as already the region fixed effects turn out to be only weakly identified in our specifications.

[40] Following the recommendations of Harrell (2001), the cubic spline has five knots located at the 5th, 27.5th, 50th, 72.5th and 95th percentile of the distribution of polity (corresponding to the values 0.05, 0.15, 0.45, 0.90, and 1, respectively).

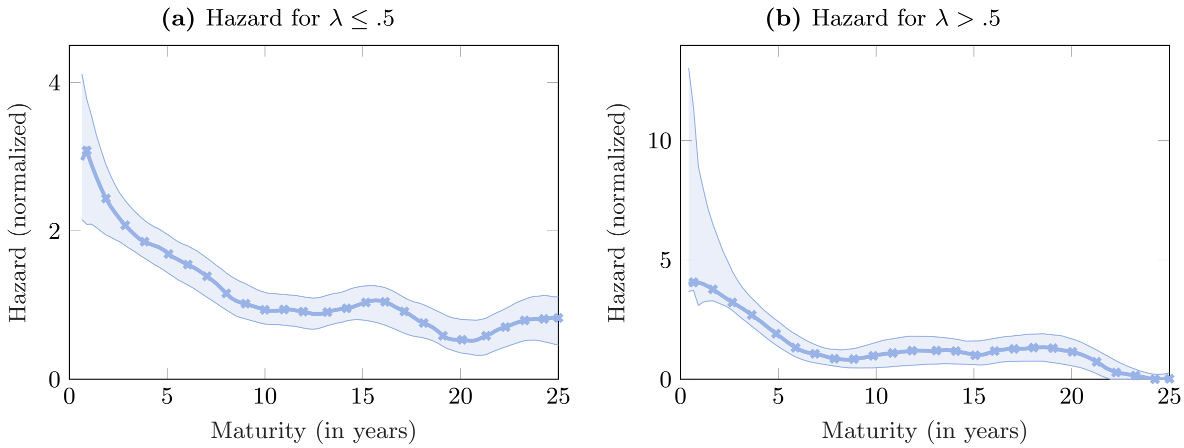

Figure 14: Empirical transition hazards for regimes on autocratic side (left panel) and democratic side (right panel) of polity spectrum. Notes.—Hazards are normalized relative to the unconditional hazard for autocracies and democracies, respectively. Shaded bands correspond to 80 percent bootstrap confidence intervals, clustered at the country level.

Transition hazards by regime type Specification (17) already controls for the political system, but continues to impose a baseline hazard h that is independent of λ. To explore inasmuch the documented stabilization patterns equally apply to autocracies and democracies, we also compute the hazard rate separately for regimes with $\lambda \leq .5$ and $\lambda > .5$. The results for the combined revolt and reform hazards are shown in Figure 14 (all results continue to hold if we separate the hazard by both origin and mode of transition). It is evident that the stabilization equally applies to regimes on the autocratic and democratic side of the spectrum.

Hazard ratios of political systems To estimate the relation between political inclusiveness and transition hazards, reported in Figure 2, we re-estimate (17) for the combined failures due to reforms and revolts. The estimated relationship is given by the cubic spline f.

B Mathematical Appendix

B.1 Effectiveness of Reforms

Here we show formally that outsiders have no incentives to ever refuse becoming enfranchised. The argument also implies that agents born as insiders never choose to rebel if given the choice. 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 $u(1)$, because $x_t = 1$ is in the choice set of insiders; i.e., $(1 - p(\cdot, x_t))u(x_t) \geq (1 - p(\cdot, 1))u(1) = u(1)$. 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 $\psi(1) = h(1)u(1)$, because by Assumption 1 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.

B.2 Proof of Proposition 1

Using (7), we can rewrite (8) as

$$\pi(s, x, \hat{\theta}) \equiv s - \min\{f(s, (1-x)\hat{\theta}), 1-x\} = 0 \tag{18}$$

for

$$f(s, y) \equiv y\psi(s).$$

Recall that $u(1) = 1$ and $h(\cdot) \in [0, 1]$. By Assumption 1, f is therefore increasing and (weakly) concave in both arguments. Accordingly, $f(s, y) \leq y$, allowing us to drop the min-operator from (18).

As $f(0, y) = 0$ for all $y \in [0, 1]$, there always exists a solution to (18) at $s = 0$. We distinguish two cases. First, let $y = 0$ (i.e,. when $\hat{\theta} = 0$ or $x = 1$). Then $f(s, y) = 0$ for all s, so that $s = 0$ is the unique stable solution to (18).

Second, let $y > 0$. By Assumption 1, $f_1(0, y) > 1$, so that $s = 0$ is unstable.[41] We now show the existence of a unique stable fixed point $s > 0$. Specifically, $f_1(0, y) > 1$ implies that $f(\tilde{s}, y) > \tilde{s}$ for $\tilde{s} \searrow 0$ and any $y > 0$. On the other hand, as noted above, $f(\tilde{s}, y) \leq y \leq \tilde{s}$ for $\tilde{s} \nearrow 1$. Continuity of ψ (and thus of f), therefore imply the existence of a fixed point $s^* > 0$. Monotonicity and concavity of f further imply that s^* is unique on $(0, 1]$. Clearly, it must hold $f_1(s^*, y) < 1$, and so s^* is stable.

The above arguments establish that s_t is uniquely determined by a (time-invariant) function $s : (\hat{\theta}_t, x_t) \mapsto s_t$. It remains to be shown that $\partial s/\partial\hat{\theta}_t \geq 0$ and $\partial s/\partial x_t \leq 0$. Implicit differentiation on (18) implies that

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

and

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

where

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

Since ψ is bounded by $\psi(1) = 1$, (8) 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 $\hat{\theta}_t \neq 0$ and $x_t \neq 1$. From the implicit function theorem it

[41] I.e., iteratively best responding to any perceived $\hat{s} = \varepsilon > 0$ leads to a distinct equilibrium $s^* > 0$ described in the following.

then follows that s is differentiable on its whole support. As noted above, $f_1(s_t, y) < 1$, implying $\bar{\gamma}_2(\hat{\theta}_t, s_t) < (1 - x_t)^{-1}$ at $s_t = s^*$. 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 developed the proof for 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.

B.3 Proof of Proposition 2

The proof proceeds by a series of lemmas. To simplify notation, in what follows we drop (λ_t, F_t) as arguments of x, $\hat{\theta}$ and $\bar{\theta}$ where no confusion arises. Furthermore, we use $\tilde{V}^I(\theta, \hat{\theta}, x) \equiv V^I(\theta h(s(\hat{\theta}, x)), x) = (1 - \theta h(s(\hat{\theta}, x)))\, u(x)$ to denote insider's expected indirect utility, where s is as given by 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''. □

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 $\tilde{\theta}$,

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

Moreover, arguing as in the proof of Lemma 1,

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

Hence, if $\tilde{V}^I(\theta'', \tilde{\theta}, x') \geq \tilde{V}^I(\theta'', \hat{\theta}(x^+), x^+) = V^*(\theta'')$, then $\tilde{V}^I(\theta', \tilde{\theta}, x') > \tilde{V}^I(\theta', \hat{\theta}(x^+), x^+) = V^*(\theta')$. Therefore, $D_{\theta'', x'}$ is a proper subset of $D_{\theta', x'}$ if $\theta'' > \theta'$, ruling out $\hat{\theta}(x') > \theta'$ by the D1 criterion.[42]

[42] The D1 criterion requires that 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

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 $x' \in (x^-, x^+)$. $\square$

Lemma 3. *There exists $\bar{\theta}_t > 0$, such that $x_t = \lambda_t$ for all $\theta_t \leq \bar{\theta}_t$. Moreover, $x(\theta'') > x(\theta') > \lambda_t + \mu$ for all $\theta'' > \theta' > \bar{\theta}_t$ 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 $\theta'' > \theta' > \bar{\theta}_t$ 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 $\theta^+ \equiv \hat{\theta}(x^+) \geq \mathbb{E}_{F_t}\{\theta_t | \theta'' \geq \theta_t \geq \theta'\} > \theta'$ 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 $\hat{\theta}(\lambda_t, x', F_t) > \theta'$. However, from Lemma 2 we have that $\hat{\theta}(x') = \theta'$, a contradiction.

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

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

Proof. Consider continuity first and suppose to the contrary that x has a discontinuity at $\theta' \in (\bar{\theta}_t, 1)$. 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 $x(\theta')$.

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 maximizes $D_{\theta',x'} = \{\hat{\theta} : \mathbb{E}\{V^I(\eta, x') | \theta', s = s(\hat{\theta}, x')\} \geq V^*(\theta')\}$ in the sense that there is no θ'' such that $D_{\theta',x'}$ is a proper subset of $D_{\theta'',x'}$.

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 $\theta' > \bar{\theta}_t$, 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

$$0 \geq g(\theta', \theta'', x(\theta'')) \geq -g_1(\theta', \theta'', x(\theta''))(\theta'' - \theta') - \tfrac{1}{2} g_{11}(a)(\theta'' - \theta')^2,$$

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 $g_1(\theta', \theta'', x(\theta''))$, and noting that $g(\theta', \theta', x(\theta')) = g_1(\theta', \theta', x(\theta')) = 0$, these inequalities can be written as

$$\begin{aligned} 0 &\geq g_2(\theta', \theta', x(\theta')) + \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} \times [g_3(\theta', \theta', x(\theta')) \\ &\quad + \tfrac{1}{2} g_{33}(b(\beta))(x(\theta'') - x(\theta')) + g_{23}(b(\beta))(\theta'' - \theta')] + \tfrac{1}{2} g_{22}(b(\beta))(\theta'' - \theta') \\ &\geq -[g_{12}(b(\beta')) + \tfrac{1}{2} g_{11}(a)](\theta'' - \theta') - g_{13}(b(\beta'))(x(\theta'') - x(\theta')), \end{aligned}$$

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 $\theta'' \to \theta'$,

$$0 \geq g_2(\theta', \theta', x(\theta')) + \lim_{\theta'' \to \theta'} \frac{x(\theta'') - x(\theta')}{\theta'' - \theta'} \, g_3(\theta', \theta', x(\theta')) \geq 0.$$

By Lemma 3, x and, hence, $\hat{\theta}$ are strictly increasing for all $\theta \geq \bar{\theta}(\lambda_t, F_t)$. Arguing similarly as we did to show continuity, optimality of x, therefore, requires that $g_3 = \tilde{V}_3^I(\theta, \theta, x) \neq 0$ and, hence, the limit of $(x(\theta'') - x(\theta'))/(\theta'' - \theta')$ is well defined, yielding

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = -\frac{\tilde{V}_2^I(\theta, \theta, x)}{\tilde{V}_3^I(\theta, \theta, x)}. \tag{19}$$

□

Lemma 5. *$x(\theta_t) = \xi(\theta_t)$ for all $\theta_t > \bar{\theta}_t$, 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, $\mathrm{d}x/\mathrm{d}\theta_t$ is pinned down by the partial differential equation (19), which must hold for all $\theta_t > \bar{\theta}_t$. Moreover, whenever $\bar{\theta}_t < 1$, in equilibrium $\hat{\theta}(x(1)) = 1$ and, therefore, it obviously must hold that $x(1) = \arg\max_{x_t} \tilde{V}^I(1, 1, x_t)$, providing a boundary condition for (19). Because $\tilde{V}^I$ is independent of (λ_t, F_t), it follows that $x(\theta_t)$ is uniquely characterized by a function, i.e., $\xi : \theta_t \mapsto x_t$, for all $\theta_t > \bar{\theta}_t$. □

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

Proof. Suppose to the contrary that $\bar{\theta}(\lambda_t, F_t)$ 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 $\tilde{V}^I(\theta, \theta, \xi(\theta)) \geq \tilde{V}^I(\theta, \mathbb{E}_{F_t}\{\theta_t | \theta_t \leq \bar{\theta}'\}, \lambda_t)$ in the equilibrium defined by $\bar{\theta}'$, and $\tilde{V}^I(\theta, \theta, \xi(\theta)) \leq \tilde{V}^I(\theta, \mathbb{E}_{F_t}\{\theta_t | \theta_t \leq \bar{\theta}''\}, \lambda_t)$ in the equilibrium defined by $\bar{\theta}''$. However, $\tilde{V}^I(\theta, \mathbb{E}_{F_t}\{\theta_t | \theta_t \leq \bar{\theta}'\}, \lambda_t) > \tilde{V}^I(\theta, \mathbb{E}_{F_t}\{\theta_t | \theta_t \leq \bar{\theta}''\}, \lambda_t)$, a contradiction. □

This establishes uniqueness of $x(\theta_t, \lambda_t, F_t)$, with all properties given by Lemmas 3 and 5, and the corresponding beliefs $\hat{\theta}(\lambda_t, x_t, F_t)$ 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 non-degenerate 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 $[\bar{\theta}(\lambda_t, F_t), 1]$. This leads to the conclusion that at most a mass zero of types (i.e., $\theta_t = \bar{\theta}(\lambda_t, F_t)$) could possibly mix in any equilibrium (with no impact on $\hat{\theta}$) and, thus, there is no need to consider any non-degenerate mixed strategies.

B.4 Proof of Proposition 3

Case (i) follows trivial, as here the state is revealed through insiders' reforms. Cases (ii) and (iii) are a straightforward application of Bayes' law. In particular, for any $\vartheta \leq \bar{\theta}_t$, we get

$$\tilde{F}_t(\vartheta | \eta_t = 1, x_t = \lambda_t) = \frac{\int_0^{\vartheta} p(\theta, s) \mathrm{d}F_t(\theta)}{\int_0^{\bar{\theta}_t} p(\theta, s) \mathrm{d}F_t(\theta)} = \frac{F_t(\vartheta)}{F_t(\bar{\theta}_t)} \frac{M_t^1(\vartheta)}{M_t^1(\bar{\theta}_t)}$$

and

$$\tilde{F}_t(\vartheta | \eta_t = 0, x_t = \lambda_t) = \frac{\int_0^{\vartheta} (1 - p(\theta, s_t)) \mathrm{d}F_t(\theta)}{\int_0^{\bar{\theta}_t} (1 - p(\theta, s_t)) \mathrm{d}F_t(\theta)} = \frac{F_t(\vartheta)}{F_t(\bar{\theta}_t)} \cdot \frac{1 - h(s_t) M_t^1(\vartheta)}{1 - h(s_t) M_t^1(\bar{\theta}_t)}.$$

Note that by letting $\int \mathrm{d}F$ denote the Lebesgue integral, the derivation applies for arbitrary, not necessarily continuous, probability measures F_t.

B.5 Derivation of Equations (11) and (12)

In case (i), $\hat{F}_t | (x_t > \lambda_t)$ is a single mass point on θ_t, so trivially $\tilde{\mu}_t = \theta_t$ and $\tilde{\sigma}_t^2 = 0$.

Consider, case (ii) next. From Proposition 3, we have that for all $\theta \le \bar{\theta}_t$,

$$\mathrm{d}\tilde{F}_t(\theta) = \frac{1}{F_t(\bar{\theta}_t)M_t^1(\bar{\theta}_t)}\mathrm{d}\big(F_t(\theta)M_t^1(\theta)\big)$$

where, using the definition of M_t^1,

$$\mathrm{d}\big(F_t(\theta)M_t^1(\theta)\big) = \mathrm{d}\int_0^\theta \vartheta \mathrm{d}F_t(\vartheta) = \theta \mathrm{d}F_t(\theta).$$

Computing the i-the raw moment of $\tilde{F}_t$, we have

$$\begin{aligned}\mathbb{E}_{\tilde{F}_t}\{\theta^i|\theta \le \bar{\theta}_t\} &= \frac{1}{F_t(\bar{\theta}_t)M_t^1(\bar{\theta}_t)}\int_0^{\bar{\theta}_t} \theta^{i+1}\mathrm{d}F_t(\theta) \\ &= \frac{M_t^{i+1}(\bar{\theta})}{M_t^1(\bar{\theta})}\end{aligned}$$

and, accordingly,

$$\begin{aligned}\tilde{\mu}_t|(\eta_t = 1, x_t = \lambda_t) &= \frac{M_t^2(\bar{\theta})}{M_t^1(\bar{\theta})} \\ \tilde{\sigma}_t^2|(\eta_t = 1, x_t = \lambda_t) &= \frac{M_t^3(\bar{\theta})}{M_t^1(\bar{\theta})} - \tilde{\mu}_t^2.\end{aligned}$$

Case (iii) is analyzed analogously. For all $\theta \le \bar{\theta}_t$, the probability measure is given by

$$\mathrm{d}\tilde{F}_t(\theta) = \frac{1}{F_t(\bar{\theta}_t)\left(1 - h(s_t)M_t^1(\bar{\theta}_t)\right)}\mathrm{d}\Big(F_t(\theta)\left(1 - h(s_t)M_t^1(\theta)\right)\Big)$$

where

$$\mathrm{d}\Big(F_t(\theta)\left(1 - h(s_t)M_t^1(\theta)\right)\Big) = \mathrm{d}F_t(\theta) - h(s_t)\mathrm{d}\int_0^\theta \vartheta \mathrm{d}F_t(\vartheta) = (1 - \theta h(s_t))\mathrm{d}F_t(\theta).$$

The i-th raw moment is thus given by

$$\begin{aligned}\mathbb{E}_{\tilde{F}_t}\{\theta^i|\theta \le \bar{\theta}_t\} &= \frac{1}{F_t(\bar{\theta}_t)\left(1 - h(s_t)M_t^1(\bar{\theta}_t)\right)}\int_0^{\bar{\theta}_t} (\theta^i - \theta^{i+1}h(s_t))\mathrm{d}F_t(\theta) \\ &= \frac{M_t^i(\bar{\theta}_t) - h(s_t)M_t^{i+1}(\bar{\theta}_t)}{1 - h(s_t)M_t^1(\bar{\theta}_t)}\end{aligned}$$

and, hence,

$$\tilde{\mu}_t|(\eta_t = 0, x_t = \lambda_t) = \frac{M_t^1(\bar{\theta}_t) - h(s_t)M_t^2(\bar{\theta}_t)}{1 - h(s_t)M_t^1(\bar{\theta}_t)}$$
$$\tilde{\sigma}_t^2|(\eta_t = 0, x_t = \lambda_t) = \frac{M_t^2(\bar{\theta}_t) - h(s_t)M_t^3(\bar{\theta}_t)}{1 - h(s_t)M_t^1(\bar{\theta}_t)} - \tilde{\mu}_t^2.$$

B.6 Proof of Proposition 4

The proposition is a straightforward corollary to Propositions 1–3: From Proposition 1 and 2, there exists a unique mapping from $\mathcal{S}_t$ to $\{\{\phi_{it}\}_{i\in[0,1]}, s_t, x_t\}$, which further implies a unique (stochastic) mapping from $\mathcal{S}_t$ to η_t. Proposition 3, in turn, implies that there exists a unique mapping from $(\mathcal{S}_t, x_t, \eta_t)$ to $\mathcal{S}_{t+1}$. As $\mathcal{S}_t$ is purely-backward looking, we conclude that for any $\mathcal{S}_0$ there exists a unique stochastic equilibrium path.

B.7 Proof of Proposition 5

Given $\Pr(\theta_t = \theta_{t-1}) \to 1$, we have that $\theta_0 = \theta_1 = \cdots \equiv \theta$ almost surely. We prove two versions of the proposition. Our preferred version amounts to the case where θ is fixed across time, but is unobserved by the statistician. To estimate transition hazards, the statistician treats θ as hidden state and refines his estimate for θ based on the realizations of (x_t, s_t, η_t). Accordingly, in our preferred version of the proposition, the statistical probability measure at date t coincides with outsiders' prior F_t. However, the proposition also holds conditional on θ_0. To show this, we first prove the result for a fixed θ_0, and then derive the more general result where θ_0 is treated as hidden state as a corollary.

Step 1 (fixed θ_0). Fix some $(\theta_0, \lambda_0, F_0)$, and let $\bar{\theta}_0$ define the pooling threshold as in Proposition 2. We tacitly assume $\theta_0 < \bar{\theta}_0$, so that there is indeed no reform at $t = 0$.

Consider any $t > 0$ and suppose there was no transition until until $t-1$. From $\Pr(\theta_t = \theta_{t-1}) \to 1$, $F_t(\bar{\theta}_{t-1}) = \tilde{F}_{t-1}(\bar{\theta}_{t-1})$. Proposition 3 then implies

$$F_t(\bar{\theta}_{t-1}) = \tilde{F}_{t-1}(\bar{\theta}_{t-1}) = 1, \tag{20}$$

so that

$$M_t^1(\bar{\theta}_{t-1}) = \mathbb{E}_{F_t}\{\theta|\theta \le \bar{\theta}_{t-1}\} = \mathbb{E}_{F_t}\{\theta\}. \tag{21}$$

Moreover, from (11),

$$\mathbb{E}_{F_t}\{\theta\} = \frac{M_{t-1}^2(\bar{\theta}_{t-1})}{M_{t-1}^1(\bar{\theta}_{t-1})}. \tag{22}$$

Combining (21) and (22) and noting that $\mathrm{Var}_{F_t}[\theta|\theta \leq \vartheta] = M_t^2(\vartheta) - [M_t^1(\vartheta)]^2 > 0$ implies $M_t^2(\vartheta) > [M_t^1(\vartheta)]^2$, we have

$$M_t^1(\bar{\theta}_{t-1}) = \frac{M_{t-1}^1(\bar{\theta}_{t-1}) - h(s_{t-1})M_{t-1}^2(\bar{\theta}_{t-1})}{1 - h(s_{t-1})M_{t-1}^1(\bar{\theta}_{t-1})} < M_{t-1}^1(\bar{\theta}_{t-1}). \tag{23}$$

Further noting that from (20), $M_t^1(\bar{\theta}_{t-1}) = M_t^1(\bar{\theta}_t)$ for all $\bar{\theta}_t \geq \bar{\theta}_{t-1}$, we conclude that

$$\mathrm{haz}_t^{\mathrm{rev}} = \theta_0 h(s(M_t^1(\bar{\theta}_t), \lambda_0)) < \theta_0 h(s(M_{t-1}^1(\bar{\theta}_{t-1}), \lambda_0)) = \mathrm{haz}_{t-1}^{\mathrm{rev}} \tag{24}$$

if $\bar{\theta}_t \geq \bar{\theta}_{t-1}$.

To complete the proof of the first step, we need to show that $\bar{\theta}_t \geq \bar{\theta}_{t-1}$, implying that (24) indeed holds, and further implying that

$$\mathrm{haz}_t^{\mathrm{ref}} \leq \Pr(\theta_0 \geq \bar{\theta}_t) = 0.$$

To see that this is true, note that from (24),

$$\tilde{V}^I(\bar{\theta}_{t-1}, \hat{\theta}(\lambda_0, \lambda_0, F_t), \lambda) > \tilde{V}^I(\bar{\theta}_{t-1}, \hat{\theta}(\lambda_0, \lambda_0, F_{t-1}), \lambda),$$

implying that the right-hand side of condition (9) is increased from t to $t-1$ at $\bar{\theta} = \bar{\theta}_{t-1}$. As the left-hand side of (9) is constant in t, it thus must hold that $\bar{\theta}_t > \bar{\theta}_{t-1}$.

Step 2 (θ_0 is a hidden state) Now suppose that the statistician does not know the realization of θ_0. Instead he or she filters through the realized history of the economy, summarized by $(x_\tau, s_\tau, \eta_\tau)_{\tau<t}$, to compute a probability measure for θ_0 and the corresponding transition hazards. As the realized history coincides with outsiders' information set, the statistical probability measure is simply given by F_t. Specifically, transition hazards at date t are given by

$$\begin{aligned}
\mathrm{haz}_t^{\mathrm{ref}} &= \int_{\bar{\theta}_t}^1 (1 - \theta h(s(\theta, \xi(\theta))))\,\mathrm{d}F_t(\theta) \\
\mathrm{haz}_t^{\mathrm{rev}} &= \int_0^{\bar{\theta}_t} \theta h(s(M_t^1(\bar{\theta}_t), \lambda_0))\,\mathrm{d}F_t(\theta) + \int_{\bar{\theta}_t}^1 \theta h(s(\theta, \xi(\theta)))\,\mathrm{d}F_t(\theta).
\end{aligned}$$

Step 1 immediately implies that for all $t > 0$, in the absence of any prior transition,

$$\mathrm{haz}_t^{\mathrm{ref}} = 0$$

and

$$\begin{aligned}\mathrm{haz}_t^{\mathrm{rev}} &= \int_0^{\bar{\theta}_t} \theta h(s(M_t^1(\bar{\theta}_t), \lambda_0))\,\mathrm{d}F_t(\theta) \\ &= h(s(M_t^1(\bar{\theta}_t), \lambda_0)) \cdot M_t^1(\bar{\theta}_t).\end{aligned}$$

To conclude the proof, note that from above, $M_t^1(\bar{\theta}_t) < M_{t-1}^1(\bar{\theta}_{t-1})$, so that $\mathrm{haz}_t^{\mathrm{rev}}$ is again strictly decreasing in t as both factors are decreasing in $M_t^1(\bar{\theta}_t)$. Intuitively, the first factor captures the decline in the revolt-hazard due to *outsiders* perceiving the regime as more stable (leading to a fall in s_t over time). The second term captures the uncertainty by the statistician, which similarly to outsiders, now also infers that the regime is more stable over time, further reducing the revolt-hazard over time compared to the fixed θ_0-case above.

B.8 Proof of Proposition 6

Bounding reforms Insiders' optimality condition implies that $\xi(\bar{\theta}_t)$ is effective in reducing revolutionary pressure; i.e.,

$$s(\bar{\theta}_t, \xi(\bar{\theta}_t)) < s(M_t^1(\bar{\theta}_t), \lambda_t). \tag{25}$$

From the proof of Proposition 1, we can write $s(\hat{\theta}, x) = \bar{s}(\omega)$ with $\omega = (1-x)\hat{\theta}$ and $\bar{s}' > 0$. Accordingly, (25) implies

$$(1-\xi(\bar{\theta}_t))\bar{\theta}_t < (1-\lambda_t)M_t^1(\bar{\theta}_t),$$

and so

$$x_t \geq \xi(\bar{\theta}_t) > 1-(1-\lambda_t)\frac{M_t^1(\bar{\theta}_t)}{\bar{\theta}_t} \equiv \bar{\lambda}_t^{\mathrm{ref}}.$$

Bounding revolts From (7) and (8), s_t solves the fixed-point equation

$$s_t = \omega_t \psi(s_t) \tag{26}$$

with $\omega_t = (1-x_t)\hat{\theta}_t$. Let $\omega' > \omega$ and, correspondingly, let $s' > s$ as in (26). Then

$$s = \omega\psi\left(s\right) < \omega\psi(s') = \frac{\omega}{\omega'}s'. \tag{27}$$

Evaluating (27) for $\omega = (1-\lambda_t)M_t^1(\bar{\theta}_t)$ and $\omega' = (1-\lambda_t)\bar{\theta}_t > \omega$ yields

$$s(M_t^1(\bar{\theta}_t), \lambda_t) < \frac{M_t^1(\bar{\theta}_t)}{\bar{\theta}_t}s(\bar{\theta}_t, \lambda_t). \tag{28}$$

Similarly, evaluating (27) for $\omega = (1-\lambda_t)\bar{\theta}_t$ and $\omega' = 1$, we have

$$s(\bar{\theta}_t, \lambda_t) < (1-\lambda_t)\bar{\theta}_t s(1, 0) \leq (1-\lambda_t)\bar{\theta}_t. \tag{29}$$

Combining (28) and (29), and recalling that optimization by insiders requires that s_t is weakly below $s(M_t^1(\bar{\theta}_t), \lambda_t)$, yields

$$s_t \leq s(M_t^1(\bar{\theta}_t), \lambda_t) < (1 - \lambda_t) M_t^1(\bar{\theta}_t) \equiv \bar{\lambda}_t^{\text{rev}}.$$

B.9 Proof of Proposition 7

Differentiating $\tilde{V}^I$ with respect to its third argument, we obtain

$$\lim_{x \to 1} \tilde{V}_3^I(1, 1, x) = -\alpha_u - \lim_{x \to 1} \alpha_h s(1, x)^{\alpha_h - 1} s_2(1, x),$$

or, substituting for s_2 as computed in the proof to Proposition 1 and observing that $x \to 1$ implies $s(1, x) \to 0$,

$$\lim_{x \to 1} \tilde{V}_3^I(1, 1, x) = -\alpha_u + \lim_{x \to 1} \alpha_h s(1, x)^{2\alpha_h - 1} \frac{1 + \alpha_u}{1 - (1 + \alpha_u)\alpha_h \frac{1-x}{s(1,x)^{1-\alpha_h}}}. \tag{30}$$

Using L'Hospital's Rule, we get after some algebra

$$\lim_{x \to 1} \frac{1 - x}{s(1, x)^{1-\alpha_h}} = \frac{1}{(1 - \alpha_h)(1 + \alpha_u)} - \frac{\alpha_h}{1 - \alpha_h} \lim_{x \to 1} \frac{1 - x}{s(1, x)^{1-\alpha_h}},$$

which has a unique fixed point at

$$\lim_{x \to 1} \frac{1 - x}{s(1, x)^{1-\alpha_h}} = \frac{1}{1 + \alpha_u}.$$

Substituting back into (30), we have that $\xi(1) = 1$ if and only if

$$\lim_{x \to 1} s(1, x)^{2\alpha_h - 1} \geq \frac{\alpha_u}{1 + \alpha_u} \frac{1 - \alpha_h}{\alpha_h}.$$

Note that the right side of the inequality is *strictly* between zero and unity, as $0 < \alpha_u \leq \alpha_h$ given the properties imposed on u and h. The left side of the inequality goes to zero for all $\alpha_h > .5$, goes to ∞ for all $\alpha_h < .5$, and is constant at unity for $\alpha_h = .5$. We conclude that $\xi(1) = 1$ if and only if $\alpha_h \leq .5$, implying that a regime with $\lambda \to 1$ emerges (almost surely) under the same conditions (as G has full support on $[0, 1]$).

C Numerical Implementation

This section describes the algorithm used to solve and estimate the model.

Solution to the model We first describe how to solve the model for a given parametrization ω. The solution is simplified by the block-recursivity of the overlapping generations structure, which let's us break down the algorithm into three successive steps.

Step 1 (coordination problem). We solve the functional fixed-point (8) for $s : (\hat{\theta}, x) \mapsto s$ using a spline collocation. Noting that $(1-x)\bar{\gamma}(\hat{\theta}, s) = yh(s)u(s)$ with $y = (1-x)\hat{\theta}$, we can reduce s to a univariate function $\bar{s} : y \mapsto s$. We parametrize $\bar{s}$ using a septic spline with 34 interior break points, with parameters chosen to solve (8) on a fine grid on $[0, 1]$. The procedure gives a very accurate approximation to s (evaluating (8) on an equally-spaced grid with 1000 points on $[0, 1]$, yields a maximal error below $5 \cdot 10^{-7}$).

Step 2 (signaling problem). The solution to the signaling problem characterized by Proposition 2 breaks down in two substeps. (i) Given s, we can solve for ξ using a standard solver for ordinary differential equations. (ii) Given s and ξ, $\bar{\theta}$ can be solved using a standard bisection method on $[0, 1]$.

Step 3 (stationary distribution). We approximate the stationary distribution on a $(N_\lambda \times N_\mu \times N_\sigma)$-point grid for (λ, μ, σ) with $N_\lambda = 21$, $N_\mu = 20$, $N_\sigma = 20$ (see the main body of the paper for more details). The laws of motion are given by (2), (13) and (14), with

$$M_t^i(\vartheta) = B(\vartheta, a_t + i, b_t)/B(\vartheta, a, b)$$

where B is the incomplete Beta function with shape parameters

$$a_t = \mu_t \left(\frac{\mu_t(1-\mu_t)}{\sigma_t^2} - 1 \right) \qquad b_t = (1-\mu_t) \left(\frac{\mu_t(1-\mu_t)}{\sigma_t^2} - 1 \right). \tag{31}$$

To compute the transition matrix $Q(\lambda_{t+1}, \mu_{t+1}, \sigma_{t+1} | \lambda_t, \mu_t, \sigma_t)$, we first solve the generation game conditional on $(\theta_t, \lambda_t, \mu_t, \sigma_t)$ and integrate out θ_t using F_t as probability measure (see Footnote 15 for details). For each $(\lambda_{t+1}, \mu_{t+1}, \sigma_{t+1})$, we then discretize the resulting transition probabilities to the eight adjacent grid-points, $\{\lambda_i, \lambda_{i+i}\} \times \{\mu_i, \mu_{i+1}\} \times \{\sigma_i, \sigma_{i+1}\}$, assigning probabilities proportionately to their inverse Euclidean distance to the respective corners of the cube. Once we have Q, we first verify that there exist a single recurrent class, consisting of 3322 states at our estimate (the remainder 5078 states are not reached along the equilibrium path). Finally, we iterate on Q until convergence, yielding the unique stationary distribution.

On a Thinkpad X230 with a i5-3230M, the whole process takes about 5 seconds to complete.

Estimation We use a combination of global and local minimization tools to solve (15). Specifically, we first use a particle swarm algorithm with 20 chains of 16 particles each to conduct a preliminary global search. The particles are initialized using scrambled Sobol quasi-random numbers, and evolve completely independent across the 20 chains. After running the particle swarm algorithm for up to 200 iterations, we then run 20 local optimizer, initialized at the 20 minima attained across the 16 particles by each of the 20 chains. Our estimator is the minimum across the 20 chains.

The process converged to the exact same estimate for the top 9 out of 20 chains. On two Xeon E5-2630 v4 processors (with 20 physical cores), the whole estimation took about 4.5 hours to complete.

D Accurateness of Belief Approximation

For the estimation of the model, we track outsiders' beliefs over time by approximating the one-step ahead projection from the posterior $\tilde{F}_t$ (which we compute exactly as in Proposition 3) to the prior F_{t+1} using a Beta distribution with moments matching (13) and (14). In this section, we explore the accurateness of this approximation. Overall, we find that the approximation of F_{t+1} is remarkably exact, tracing the true prior almost perfectly.

D.1 Beliefs after reforms

After an attempted or successful reform ($x_t > \lambda_t$, $\eta_t \in \{0, 1\}$), the current state of the regime is fully revealed. Accordingly, the exact prior at $t + 1$ is truncated normal with mean $\rho\theta_t + \mu_\epsilon$ and variance σ_ϵ^2. For any interior $\vartheta \in (0, 1)$, the pdf is given by

$$f_{t+1}(\vartheta) = \phi_{\rho\theta_t+\mu_\epsilon,\sigma_\epsilon^2}(\vartheta),$$

where ϕ_{μ,σ^2} denotes the density of a (μ, σ^2)-normal distribution. At the boundaries, $\vartheta \in \{0, 1\}$, F_{t+1} has mass points corresponding to the tails of f_{t+1}.

By contrast, the Beta approximation is given by

$$f_{t+1}^{\text{approx}}(\vartheta) = \beta_{\rho\theta_t+\mu_\epsilon,\sigma_\epsilon^2}(\vartheta),$$

where β_{μ,σ^2} denotes the density of a Beta distribution with mean μ and variance σ^2 (implemented by shape parameters as in (31)).

Panel (a) of Figure 15 compares f_{t+1} and f_{t+1}^{approx} for three different values of θ_t. Specifically, the values of θ_t are set to the 10th, 50th and 90th percentile of the distribution over θ_t conditional on there being a reform at t. In all three cases, the approximation traces the exact shape of f_{t+1} almost perfectly, despite being marginally skewed for the 90th percentile of θ_t. Moreover, because f_{t+1}^{approx} integrates to unity on $(0, 1)$, the close fit in the interior also implies that f_{t+1} integrates to approximately unity on $(0, 1)$, so that that the residual mass distributed as mass points on $\{0, 1\}$ is negligible.

D.2 Beliefs after revolts against pooling regimes

Consider next the case where the regime does not attempt any reform and is overthrown by a revolt. From Proposition 3, the posterior density $\tilde{f}_t = \tilde{F}'$ is given by

$$\tilde{f}_t(\vartheta) = \frac{1}{F_t(\bar{\theta}_t)} \frac{\vartheta F_t'(\vartheta)}{M_t^1(\bar{\theta}_t)}.$$

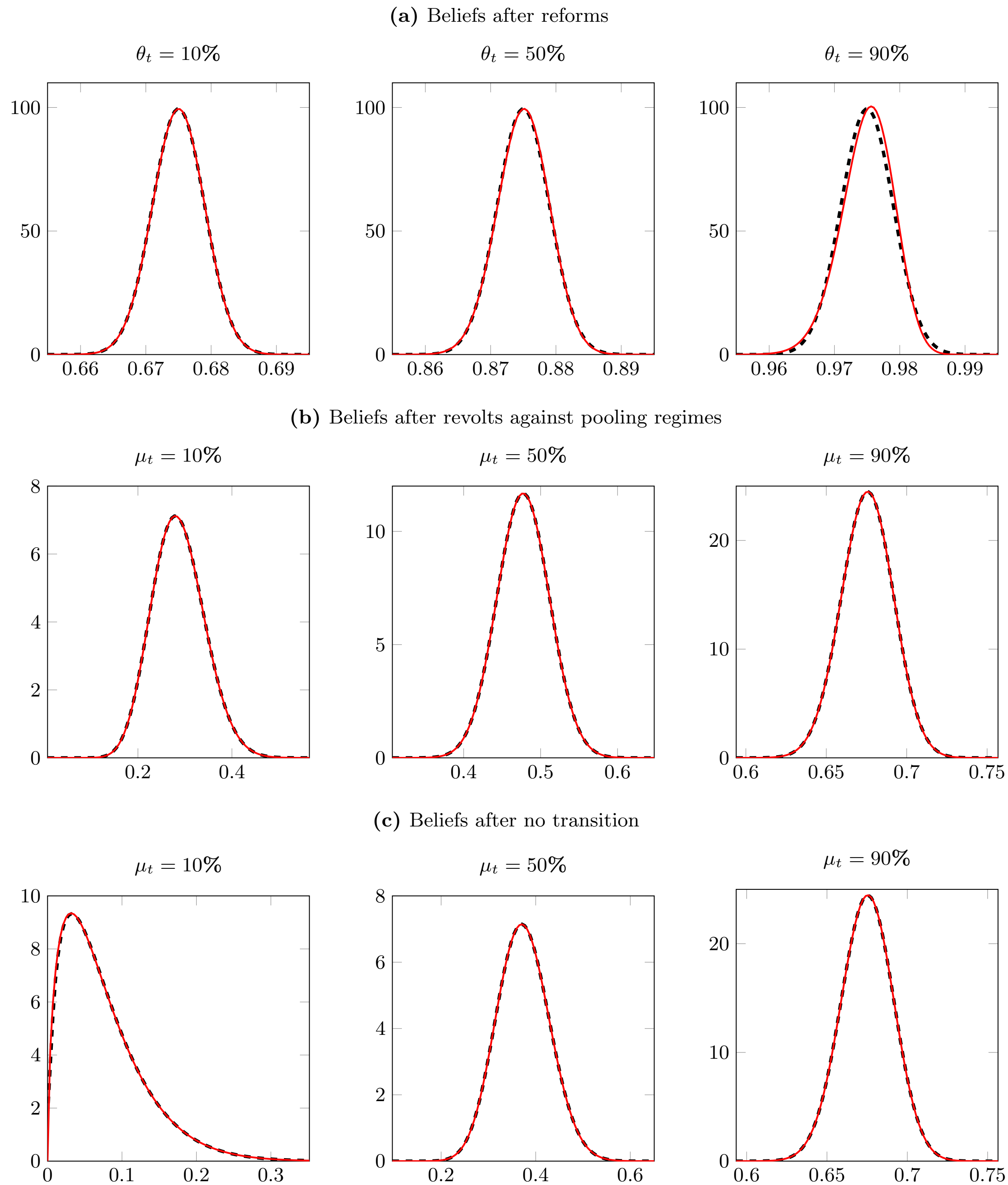

Figure 15: Accurateness of belief approximation. Black dotted lines are exact prior beliefs computed as in (10). Solid red lines approximate the one-step ahead projection using Beta-distributions with their first two moments matching (13) and (14). Top panel compares f_{t+1} with f_{t+1}^{approx} for $x_t > \lambda_t$, $\eta_t \in \{0, 1\}$ and θ_t set to the 10th, 50th and 90th percentile of $\mathcal{P}_t(\theta_t | x_t > \lambda_t)$. Middle and bottom panels compare f_{t+1} with f_{t+1}^{approx} for $x_t = \lambda_t$, $\eta_t = 1$ (middle panel) and $\eta_t = 1$ (bottom panel), μ_t set to the 10th, 50th and 90th percentile of $\mathcal{P}_t(\mu_t | x_t = \lambda_t, \eta_t)$, and λ_t and σ_t are set to the conditional (on μ_t) medians.

Substituting $\tilde{f}_t$ into (10), we obtain f_{t+1}, which for any interior $\vartheta \in (0,1)$ is given by

$$f_{t+1}(\vartheta) = \frac{1}{F_t(\bar{\theta}_t)} \frac{1}{M_t^1(\bar{\theta}_t)} \int_{-\infty}^{\infty} \phi_{\mu_\epsilon,\sigma_\epsilon^2}(\vartheta - \rho\theta)\theta f_t(\theta)\,\mathrm{d}\theta$$

where f_t is the prior density at t. The Beta approximation is given by

$$f_{t+1}^{\text{approx}}(\vartheta) = \beta_{\rho\tilde{\mu}_t+\mu_\epsilon,\rho^2\tilde{\sigma}_t^2+\sigma_\epsilon^2}(\vartheta),$$

with $\tilde{\mu}_t$ and $\tilde{\sigma}_t^2$ as in (11) and (12).

Panel (b) of Figure 15 compares f_{t+1} with its approximation f_{t+1}^{approx} for three different states $\mathcal{S}_t$. Specifically, we set μ_t to its 10th, 50th and 90th percentile conditional there being no reform and a successful revolt at t ($x_t = \lambda_t$, $\eta_t = 1$). The value of σ_t^2 is fixed at the associated median (conditional on the corresponding value for μ_t). Again, the approximation closely tracks the exact density f_{t+1} in the interior, and there is no significant mass on $\{0,1\}$.

D.3 Beliefs after no transition

Finally, consider the case of no transition. From Proposition 3, the posterior density is given by

$$\tilde{f}_t(\vartheta) = \frac{F_t'(\vartheta)}{F_t(\bar{\theta}_t)} \cdot \frac{1 - h(s_t)\vartheta}{1 - h(s_t)M_t^1(\bar{\theta}_t)},$$

yielding

$$f_{t+1}(\vartheta) = \frac{1}{F_t(\bar{\theta}_t)} \cdot \frac{1}{1 - h(s_t)M_t^1(\bar{\theta}_t)} \int_{-\infty}^{\infty} \phi_{\mu_\epsilon,\sigma_\epsilon^2}(\vartheta - \rho\theta) f_t(\theta)(1 - h(s_t)\theta)\,\mathrm{d}\theta$$

for any interior $\vartheta \in (0,1)$. The corresponding Beta approximation is given by

$$f_{t+1}^{\text{approx}}(\vartheta) = \beta_{\rho\tilde{\mu}_t+\mu_\epsilon,\rho^2\tilde{\sigma}_t^2+\sigma_\epsilon^2}(\vartheta),$$

with $\tilde{\mu}_t$ and $\tilde{\sigma}_t^2$ as in (11) and (12).

Panel (c) of Figure 15 compares f_{t+1} with its approximation f_{t+1}^{approx} for three different states $\mathcal{S}_t$. Specifically, we set μ_t to its 10th, 50th and 90th percentile conditional there being no transition at t ($x_t = \lambda_t$, $\eta_t = 0$). The values of σ_t^2 and λ_t (needed to compute $h(s_t)$) is fixed at their associated median (conditional on the corresponding value for μ_t). Again, the approximation closely tracks the exact density f_{t+1} in the interior, and there is no significant mass on $\{0,1\}$.

E Comparative Statics of the Generation Game

Comparative statics in λ Here we explore how an increase in the regime size λ affects the policy mappings depicted in Figures 5 and 6. The primary implication of an increase in λ is a reduction in potential supporters of a revolt along the extensive margin. Accordingly, absent reforms, the regime is more stable (seen in the right panel of Figure 16), which manifests itself in a reduced inclination to implement reforms ($\bar{\theta}$ is higher, see left panel of Figure 16). A second order implication then is that for increased values for $\bar{\theta}$, the pooling belief $\hat{\theta}^{\text{pool}}$ increases as well (seen in the left panel of Figure 17), which in turn increases the off-equilibrium support for revolts *conditional* on $x \in (\lambda, \xi(\bar{\theta}))$ as seen in the right panel of Figure 17.

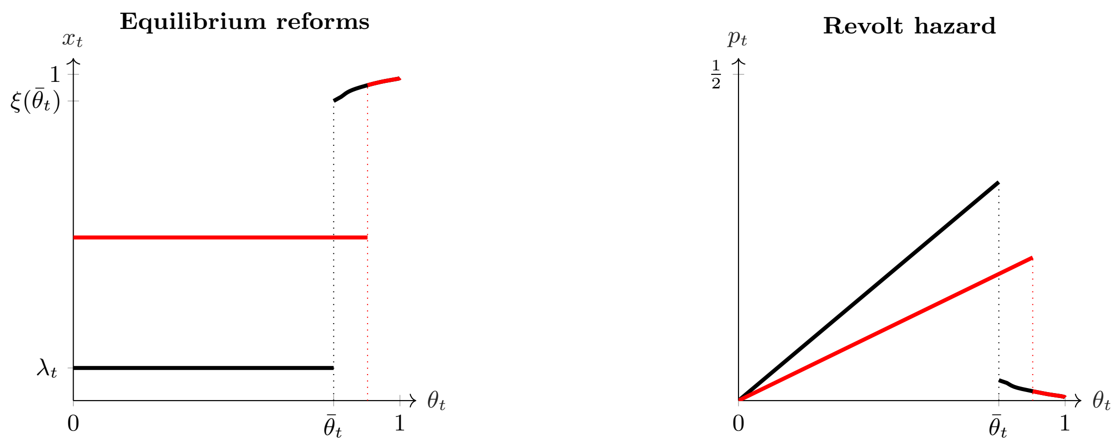

Figure 16: Effect of λ on equilibrium reforms and implied probability to be overthrown. Black lines show mappings for $\lambda = .1$, red lines show mappings for $\lambda = .5$.

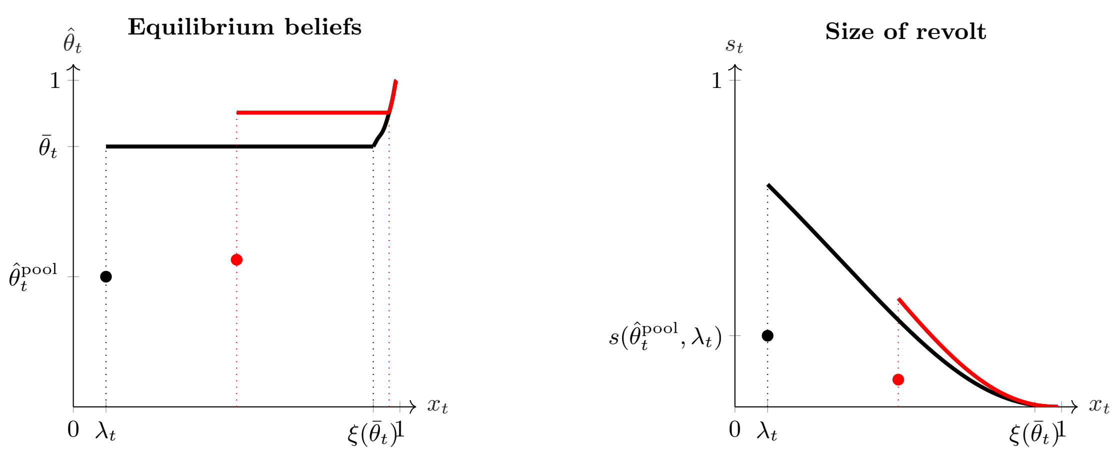

Figure 17: Effect of λ on equilibrium beliefs and implied mass of insurgents. Black lines show mappings for $\lambda = .1$, red lines show mappings for $\lambda = .5$.

Comparative statics in F To demonstrate the effect of outsiders' beliefs on F on the policy mappings, suppose F is parametrized by a Beta distribution with moments (μ, σ^2). Note that the case where F is uniform is a special of the Beta distribution where $\mu = .5$ and $\sigma^2 = 1/12$. We compare this benchmark case, depicted in the main text with the case where $\mu = .35$ and σ^2 remains fixed at the uniform value of $1/12$. The results are shown in Figures 18 and 19. It can be seen that the decline in outsiders' prior expectation (seen in the left panel of Figure 19) leads again to a drop in revolt hazard (right panels of Figures 18 and 19), which makes insiders less inclined to reform ($\bar{\theta}$ is higher, see left panel of Figure 18).

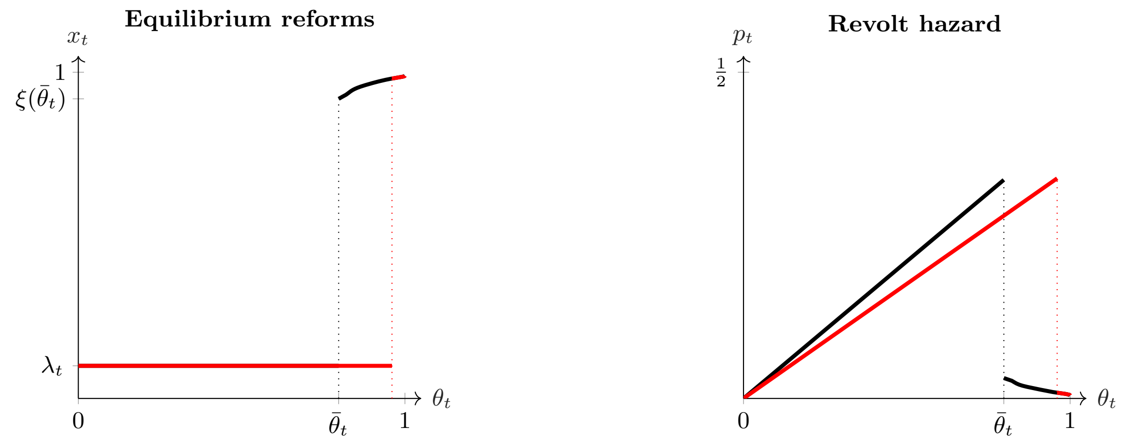

Figure 18: Effect of μ on equilibrium reforms and implied probability to be overthrown. Black lines show mappings for $\mu = .5$, red lines show mappings for $\mu = .35$.

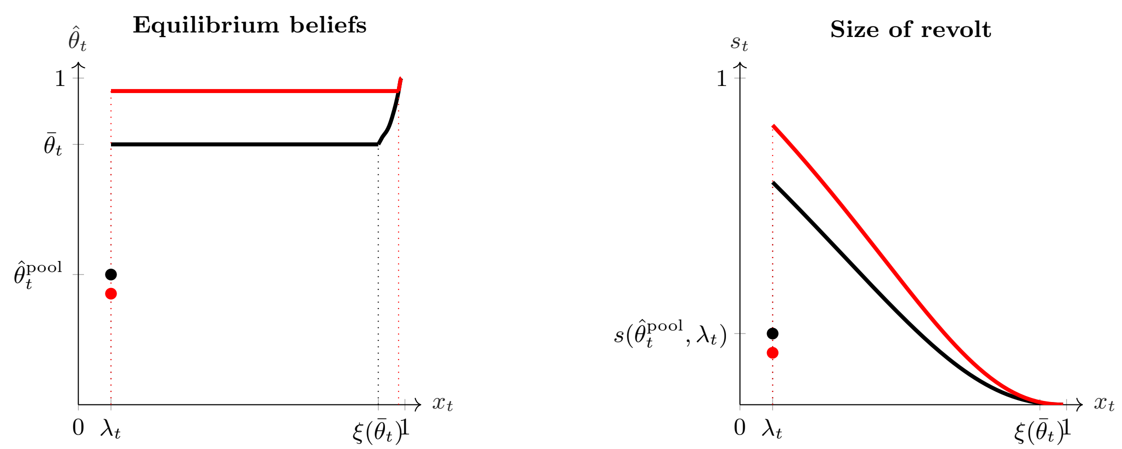

Figure 19: Effect of μ on equilibrium beliefs and implied mass of insurgents. Black lines show mappings for $\mu = .5$, red lines show mappings for $\mu = .35$.

Comparison with symmetric information case Finally, we compare the equilibrium reform mapping with the case where outsiders fully observe θ_t. Full information implies strictly more reforms by insiders compared to the asymmetric information case (see left panel of Figure 20). This is because asymmetric information essentially imposes an extra cost on reforms associated with revealing that the regime is of a higher type θ_t. On the one hand, this manifests itself in a large pool of regimes not conducting any reform, even though reforms are optimal under full information. On the other hand, since any marginal increase in reforms also implies a marginal change in outsiders' beliefs $\mathrm{d}\hat{\theta}/\mathrm{d}x = \xi^{-1}(x)$, the reform schedule itself (conditionally on conducting reforms) is biased downwards under asymmetric information. As a consequence, revolts tend to be less likely under symmetric information, even though the revolt hazard may point-wise exceed the one under asymmetric information for certain values of θ.[43] Integrating over realizations of θ (using the uniform prior as probability measure), yields an average revolt hazard of under symmetric information of 5.64 percent as opposed to 13.67 percent under asymmetric information.

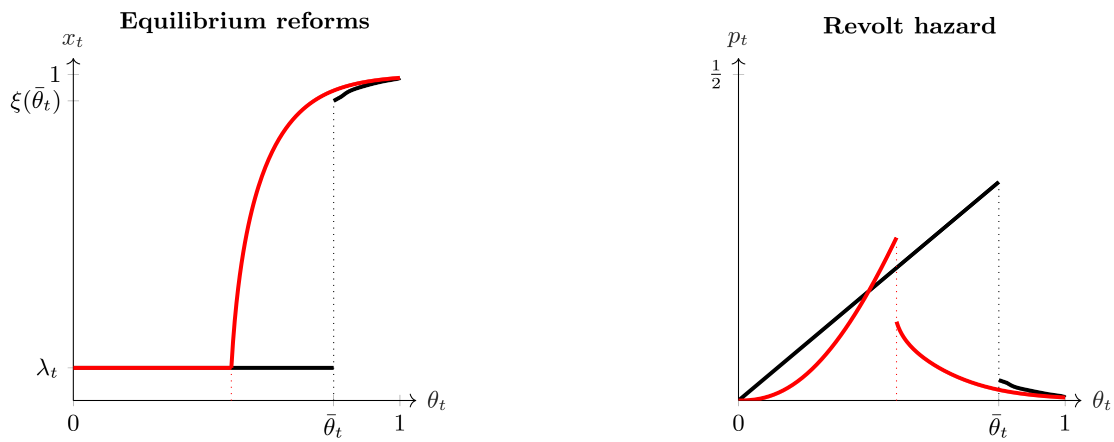

Figure 20: Equilibrium reforms and implied probability to be overthrown under symmetric information. Black lines show equilibrium mappings with asymmetric information, red lines show mappings under full information.

[43] Specifically, the hazard exceeds the one under asymmetric information for $\theta \in (\mathbb{E}_F\{\vartheta|\vartheta \leq \bar{\theta}^{\text{asym}}\}, \bar{\theta}^{\text{sym}}]$.