Review of Economic Studies (2019) **0**, 1–31 doi:10.1093/restud/rdz057

Advance access publication 8 November 2019

A Quantitative Theory of Political Transitions

LUKAS BUCHHEIM
LMU Munich
and
ROBERT ULBRICHT
Boston College

First version received May 2016*; Editorial decision August* 2019*; Accepted November* 2019 (*Eds.*)

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 determine regime dynamics. We study the quantitative implications of the model by fitting it to data on the universe of political regimes existing between 1946 and 2010. 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.

Key words: Democratic reforms, Quantitative political economy, Regime dynamics, Revolts, Transition hazards.

JEL Codes: D74, D78, P16.

1. INTRODUCTION

This article 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.[1] In this article, 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 article 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.

1. Exceptions include Acemoglu and Robinson (2001) and Besley and Persson (2018). The relation with these papers is discussed below.

The editor in charge of this paper was Michele Tertilt.

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 centre of the spectrum; full-scale democracies are overall most stable.
4. Revolts establish autocratic regimes; reforms establish democracies. Political systems near the centre 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 article 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 enfranchised fraction of the population ("political insiders"). Transitions are governed by three main ingredients. First, reforms are rationalized by a preemptive logic as in Acemoglu and Robinson (2000b). 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.

The intensive margin of revolts in combination with learning implies 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 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 stable 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 when revolts are perceived as futile, making them less prone to future unrest as well. 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 modelling 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 by fitting the model to data on the universe of political regimes existing between 1946 and 2010. 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 susceptible to counter-revolts by small but highly coordinated groups 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% after 25 periods, the failure rate drops to 12% if reforms initially enfranchise more than 95% of the population.

To the best of our knowledge, this article is first to develop a quantitative theory of political transitions. Our theory builds on several ingredients present in the existing literature. Specifically, the idea that reforms are means to *credibly* preempt a looming revolt has been standard in the literature since Acemoglu and Robinson (2000b), Conley and Temini (2001), and Boix (2003). Similarly, the view that revolts are the outcome of a coordination game has a long tradition in political science (*e.g.* Tullock, 1971; Granovetter, 1978; Kuran, 1989; Lohmann, 1994; Casper and Tyson, 2014). Finally, asymmetric information has been used to rationalize conflict along the equilibrium path in, *e.g.*, Acemoglu and Robinson (2000a), Boix (2003), Hirshleifer *et al.* (2009), Bueno de Mesquita (2010), and Ellis and Fender (2011).

Relative to existing works, a major advance of this article is the development of a unified framework of repeated political transitions driven by both reforms and revolts. Closely related to our framework, Acemoglu and Robinson (2001) and Besley and Persson (2018) study models of repeated reforms and revolts. However, because transitions are exogenously restricted to alternate between two regimes of fixed size, these papers mechanically fix the revolt–reform ratio at unity in addition to fully predetermining transition outcomes, preventing a quantitative analysis along the lines of this article.[2]

Another distinctive feature of our model is the continuity of the polity space, which is central to our predictions about the distribution of regime types. In that regard, our article relates to the literature on voluntary suffrage extensions such as Bourguignon and Verdier (2000), Lizzeri and Persico (2004), and Llavador and Oxoby (2005). These papers also endogenize the scope of reforms, but abstract from revolts and regime dynamics, which are both key to our approach.[3]

The ability of our model to account for the data is also complementary to a number of empirical studies presenting direct evidence for the ingredients at the heart of our theory. Przeworski (2009), Aidt and Jensen (2014), and Aidt and Franck (2015) present evidence supporting the preemptive logic of reforms. Aidt *et al.* (2017), Enikolopov *et al.* (2018), Manacorda and Tesei (2018), and Gonzalez (2019) present evidence supporting coordination as an essential element of revolts. Finally, Finkel *et al.* (2015) document that halfhearted reforms may fuel revolts consistent with the idea that outsiders learn about regime strength.

2. Besley and Persson (2018) consider two different modes of democratization. However, the relative frequency of democratization to regime reversals is similarly fixed at unity. Ticchi *et al.* (2013) consider an environment based on only coercive takeovers, in which regimes also alternate between autocracy and democracy.

3. Our approach to endogenize the outcomes of political transitions also relates to a growing literature on dynamic voting games, which studies equilibrium coalitions in rich state spaces (Justman and Gradstein, 1999; Jack and Lagunoff, 2006a; Gradstein, 2007; Acemoglu *et al.*, 2008, 2012, 2015; Lagunoff, 2009; Bai and Lagunoff, 2011). However, given their focus on the *composition* of coalitions, these papers typically do not pin down a specific mode of transition between regimes, nor do they allow for forcible attempts to obtain power by means of revolt.

The remainder of the article 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 and study its ability to account for the data. In Section 5, we describe the working of the model in detail and provide intuition for the forces explaining the data. In Section 6, we study the model's implications for the formation and survival of democracies. In Section 7, we revisit the regularities motivating the model and discuss inasmuch they are consistent with alternative theories. Section 8 concludes.

2. EVIDENCE ON POLITICAL TRANSITIONS

This section presents a list of stylized facts about regime 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 datasets. First, we obtain the universe of *regime spells* from Geddes *et al.* (2014), who define regimes based on the identity of the ruling group. Second, we use the Polity IV Project's polity index (Marshall *et al.*, 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 *et al.* (2014) to classify transition events. Otherwise, we match transitions to leader changes collected by the Archigos Database of Political Leaders (Goemans *et al.*, 2009) and classify them according to the nature of the observed leader change. The resulting database covers 485 regime spells and 329 transitions in 155 countries. Supplementary 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[4]). Democratic reforms are peaceful transitions that lead to a more democratic political system. Together, revolts and reforms constitute 75% of all observed transition events. This corresponds to about 0.021 revolts and 0.008 reforms per country-year. The remaining transitions occur either via autocratic consolidations (peaceful transitions towards more autocratic polities, 6%), foreign imposition (5%), or cannot be categorized based on the available information (14%). See Table 1 for further details.

Fact 2. *Transition hazards are declining in regime maturity.*

Figure 1 plots the transition hazards for our data.[5] 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.

4. Geddes *et al.* (2014) define regime spells as uninterrupted reign of the same group of political elites. By their definition, coups constitute a transition only if they substantially alter the composition of the ruling group. In this regard, power struggles that, *e.g.*, replace one military leader by another from the same group of military leadership do not constitute a transition.

5. 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 Supplementary Appendix A.2 for details). Similar patterns have been documented by

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).

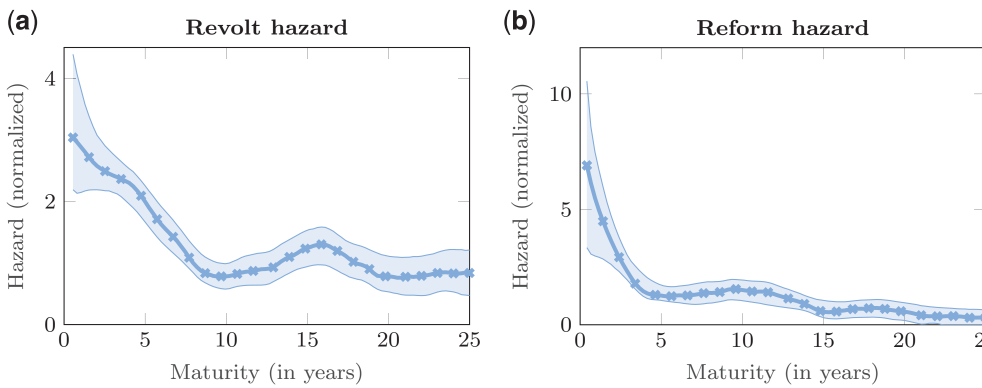

FIGURE 1

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% 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 investigate the pattern in our data, we estimate a Cox model with a cubic spline in the polity dimension (see Supplementary 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

Sanhueza (1999) and Svolik (2008, 2015). Likewise, 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.

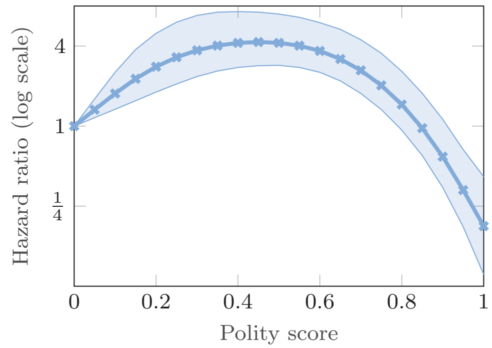

FIGURE 2

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% bootstrap confidence intervals, clustered at the country level.

near the centre of the autocracy–democracy spectrum are unlikely to arise from either mode of transition.[6]

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 80%.

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 is represented by the franchise; 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

6. 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).

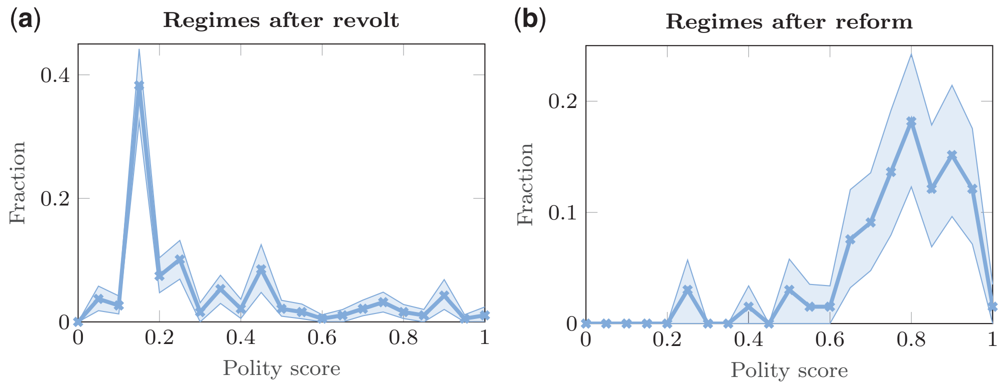

FIGURE 3

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

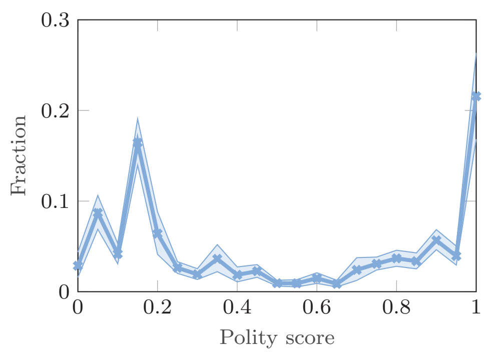

FIGURE 4

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

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

$$p(\theta_t, s_t) = \theta_t h(s_t), \tag{3.1}$$

where $\theta_t \in [0,1]$ is a random state of the world that reflects the vulnerability of the regime or its ability to withstand a revolt, and h is an increasing and twice differentiable function, $h:[0,1] \rightarrow [0,1]$, 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. In 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 (2000b) by modelling 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:

$$\lambda_{t+1} = \begin{cases} s_t & \text{if the regime is overthrown, and} \\ x_t & \text{otherwise.} \end{cases} \tag{3.2}$$

When a revolt fails, 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, 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.).[7] 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 ($\gamma_{it} = 0$).[8] 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:

$$V^I(\eta_t, x_t) = (1-\eta_t)u(x_t), \tag{3.3}$$

$$V^O(\eta_t, \gamma_{it}, s_t, \phi_{it}) = \phi_{it}\eta_t u(s_t) + (1-\phi_{it})\gamma_{it}. \tag{3.4}$$

7. 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.

8. In our dataset, 83% 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.

Here $\eta \in \{0,1\}$ is an indicator evaluating to unity if the regime is overthrown, and superscripts 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:

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 (3.1) and (3.2) take place, period $t+1$ starts with the birth of a new generation, and payoffs are realized.

3.1.1. Two remarks. The core of our model defines an interaction between revolutionary pressure and preemptive reforms in the tradition of Acemoglu and Robinson (2000b), 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 (2000b), or because each insider is indispensable for the stability of the ruling coalition as in Acemoglu *et al.* (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."[9]

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 Supplementary 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.

3.1.2. 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.[10] Second, we limit attention to equilibria that are consistent with the D1 criterion introduced by Cho and Kreps (1987), a standard refinement for signalling games.[11] As detailed below (see footnote 14), 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'

9. 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.

10. In a previous version of this article (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.

11. 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.

law (if applicable) given all publicly observable information available at the *beginning* of period t. Specifically, we have

$$F_t(\vartheta) = \Pr\left[\theta \leq \vartheta \,|\, \delta_{t-1}\right] \tag{3.5}$$

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 signalled by reforms x_t:

$$\hat{F}_t(\vartheta) = \Pr\left[\theta \leq \vartheta \,|\, \delta_{t-1}, x_t\right] \tag{3.6}$$

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 $\hat{F}_\delta$;*[12]
- *c. Beliefs $\hat{F}_\delta$ are obtained using (3.6) 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 (3.1) and (3.2);*
- *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

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

outsiders who are not subject to reforms. Insiders' policy choices will therefore be governed by signalling considerations.

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

3.2.1. Stage 2: Coordination among outsiders. 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 ϕ_{it} so as to maximize their expected utility $\mathbb{E}_{\hat{F}}\{V^O(\cdot)\}$. Specifically, an outsider with opportunity cost γ_{it} will participate in a revolt if and only if

$$\gamma_{it} \leq p(\hat{\theta}_t, \hat{s}_t)u(\hat{s}_t) \equiv \bar{\gamma}(\hat{\theta}_t, \hat{s}_t). \tag{3.7}$$

Here, $\bar{\gamma}(\hat{\theta}_t, \hat{s}_t)$ is the expected benefit of supporting a revolt of size $\hat{s}_t$. 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

$$s_t = (1 - x_t)\min\{\bar{\gamma}(\hat{\theta}_t, \hat{s}_t), 1\}. \tag{3.8}$$

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 (3.8). 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 non-increasing 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 (3.8), given by a 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 Supplementary 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.

3.2.2. 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[13]

$$\xi'(\theta) = -\tilde{V}^I_2(\theta,\theta,\xi)/\tilde{V}^I_3(\theta,\theta,\xi) > 0$$

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, $x : (\theta_t,\lambda_t,F_t) \mapsto x_t$,

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

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)

$$\tilde{V}^I(\bar{\theta},\bar{\theta},\xi(\bar{\theta})) = \tilde{V}^I(\bar{\theta},\hat{\theta}(\lambda,\lambda,F),\lambda), \tag{3.9}$$

and is otherwise given by $\bar{\theta} = 1$. Outsiders' interim beliefs are defined by[14], $\hat{\theta} : (\lambda_t,x_t,F_t) \mapsto \hat{\theta}_t$, with

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

Proposition 2 describes equilibrium reforms as a function of (θ_t,λ_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 Supplementary Appendix E.

Whenever $\theta_t \le \bar{\theta}(\lambda_t,F_t)$, 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, $\hat{\theta}^{\text{pool}}_t \equiv \mathbb{E}_{F_t}\{\theta_t|\theta_t \le \bar{\theta}_t\}$, of all regimes pooling on $x_t = \lambda_t$. In contrast, every extension $x - \lambda_t > 0$ of the regime leads to a non-marginal change in outsiders' beliefs from $\hat{\theta}^{\text{pool}}_t$ to $\xi^{-1}(x) \ge \bar{\theta}_t$ and, therefore, causes a non-marginal increase in revolutionary pressure along the intensive margin. It follows that small reforms will 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 non-empty 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.[15]

13. Throughout, we use f_i to denote the derivative with respect to the ith argument of some function f.

14. 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.

15. 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

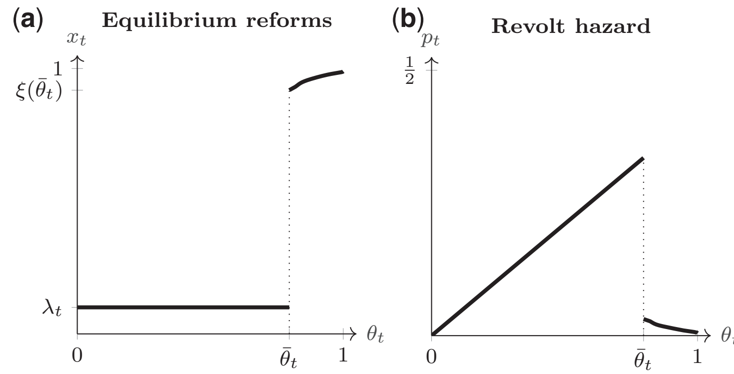

FIGURE 5
Equilibrium reforms and implied probability to be overthrown.

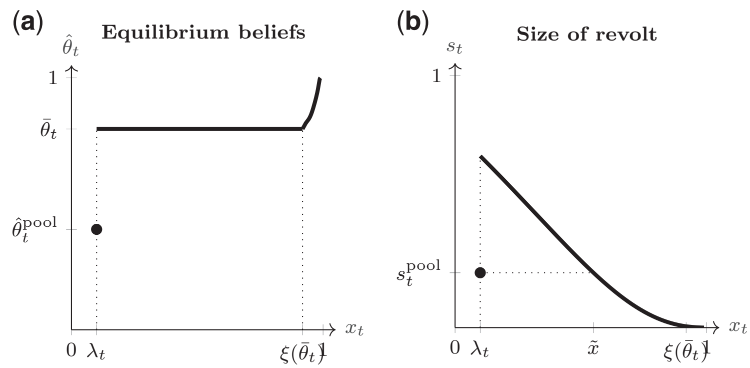

FIGURE 6
Equilibrium beliefs and implied mass of insurgents.

3.2.3. Learning dynamics. Propositions 1 and 2 fully characterize actions at t conditional on the state $\mathcal{S}_t$. 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 (3.2), 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 signalled by x_t (yielding the interim-belief $\hat{F}_t$) and the information contained in whether or not the regime is overthrown, 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:

$$F_{t+1}(\vartheta) = \int_0^{\vartheta} \int_0^1 G'(\theta'|\theta)\,\mathrm{d}\tilde{F}_t(\theta)\,\mathrm{d}\theta'. \tag{3.10}$$

these regimes also deteriorate the average pooling belief towards being more vulnerable, there is a further infra-marginal increase in the threat that affects all regimes without reforms.

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 ith (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 $\eta_t = 0$),

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

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

$$\tilde{\mu}_t = \begin{cases} \theta_t & \text{if } x_t > \lambda_t \\ \frac{M_t^2(\bar{\theta}_t)}{M_t^1(\bar{\theta}_t)} & \text{if } x_t = \lambda_t \text{ and } \eta_t = 1 \\ \frac{M_t^1(\bar{\theta}_t) - h(s_t) M_t^2(\bar{\theta}_t)}{1 - h(s_t) M_t^1(\bar{\theta}_t)} & \text{if } x_t = \lambda_t \text{ and } \eta_t = 0 \end{cases} \tag{3.11}$$

and

$$\tilde{\sigma}_t^2 = \begin{cases} 0 & \text{if } x_t > \lambda_t \\ \frac{M_t^3(\bar{\theta}_t)}{M_t^1(\bar{\theta}_t)} - \tilde{\mu}_t^2 & \text{if } x_t = \lambda_t \text{ and } \eta_t = 1 \\ \frac{M_t^2(\bar{\theta}_t) - h(s_t) M_t^3(\bar{\theta}_t)}{1 - h(s_t) M_t^1(\bar{\theta}_t)} - \tilde{\mu}_t^2 & \text{if } x_t = \lambda_t \text{ and } \eta_t = 0. \end{cases} \tag{3.12}$$

3.2.4. Existence and uniqueness of equilibrium. Propositions 1–2 uniquely pin down insiders' and outsiders' actions conditional on the state $\mathcal{S}_t$, whereas Proposition 3 (in conjunction with G and (3.2)) 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 $\{\phi_{i\delta}\}_{i \in [0,1]}$, 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. QUANTITATIVE IMPLICATIONS FOR REGIME 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 calibration. 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.[16] 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:

$$\mu_{t+1} = \rho\tilde{\mu}_t + \mu_\epsilon \tag{4.13}$$

$$\sigma_{t+1}^2 = \rho^2\tilde{\sigma}_t^2 + \sigma_\epsilon^2. \tag{4.14}$$

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 (3.10), 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 (4.13) and (4.14). We explore the accurateness of the approximation in Supplementary 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.

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 unobserved by 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 $(\lambda_t, \mu_t, \sigma_t^2)$.[17]

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

17. 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*

We approximate the continuous state space with a finite grid. Specifically, we approximate λ using a grid of (almost[18]) 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.[19]

4.2. *Calibration*

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, nine empirical moments $\hat{M}$ listed in Table 2 (further 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 Supplementary Appendix C. Our estimator for ω is given by

$$\hat{\omega}=\underset{\omega}{\arg\min}\left(\hat{M}-M(\omega)\right)'\hat{V}^{-1}\left(\hat{M}-M(\omega)\right), \tag{4.15}$$

where $\hat{V}$ is a diagonal matrix with the bootstrapped variances of $\hat{M}$ along the diagonal. The estimated parameter values are $\hat{\alpha}=0.569$, $\hat{\mu}_\epsilon/(1-\hat{\rho})=0.736$, $\hat{\sigma}_\epsilon^2/(1-\hat{\rho}^2)=0.030$, and $\hat{\rho}=0.9997$.[20] These values imply that elites in an autocratic system enjoy roughly 60% higher value than citizen in a full-scale democracy. The process for θ is highly persistent, with unconditional mean and variance of roughly 0.736 and 0.030.

4.2.1. Targeted moments.

The empirical and simulated moments, targeted in our calibration, are presented in Table 2. All model moments are computed at the stationary distribution. The target moments are 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 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 polity-distribution, $\mathcal{P}(\lambda)$, summarized by $\mathcal{P}(\lambda\leq0.25)$ and $\mathcal{P}(\lambda\geq0.75)$.

Overall, the model fits the targeted moments quite well, with most model moments being within one standard deviation of their empirical counterpart. The exceptions are the distribution

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 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 $\mathcal{P}(\theta|\delta)=F(\theta|\delta)$.

18. Specifically, we chose *thresholds* $\{0.025,0.075,\ldots,0.975\}$, defining the edges between two adjacent grid points $\{\lambda_i,\lambda_{i+1}\}$, such that the mid-points of each λ-bin, $\{0.05,0.1\ldots,0.95\}$, match the desired discretization of λ in the interior of the grid. At the boundaries, we obtain $\lambda_1=0.0125$ and $\lambda_{N_\lambda}=0.9875$ as the mid-points of the two most extreme λ-bins.

19. 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.

20. The corresponding values for $\hat{\mu}_\epsilon$ and $\hat{\sigma}_\epsilon$ are 0.197×10^{-3} and 0.0040.

TABLE 2
Data moments and model simulated moments

Fitted moments	Data	Model
Revolt–reform ratio	2.85 (0.41)	2.82
Revolt-hazard for new regimes/avg. hazard	3.04 (0.86)	2.04
Reform-hazard for new regimes/avg. hazard	6.90 (2.68)	8.95
Total transition hazard by λ: peak/autocracy	4.29 (1.91)	3.21
Total transition hazard by λ: democracy/autocracy	0.18 (0.17)	0.04
Median revolt	0.20 (0.03)	0.20
Median reform	0.80 (0.02)	0.80
Unconditional mass on $\lambda \leq 0.25$	0.42 (0.03)	0.47
Unconditional mass on $\lambda \geq 0.75$	0.39 (0.04)	0.47

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

over λ, where the model overpredicts the concentration towards the extremes by 2 standard deviations, and the revolt hazard for new regimes, which the model underpredicts relative to the corresponding average rate. Despite some discrepancies, the estimated model is clearly able to replicate the documented regularities.

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

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 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 empirical distribution over λ 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.

5.1. *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 the franchise?

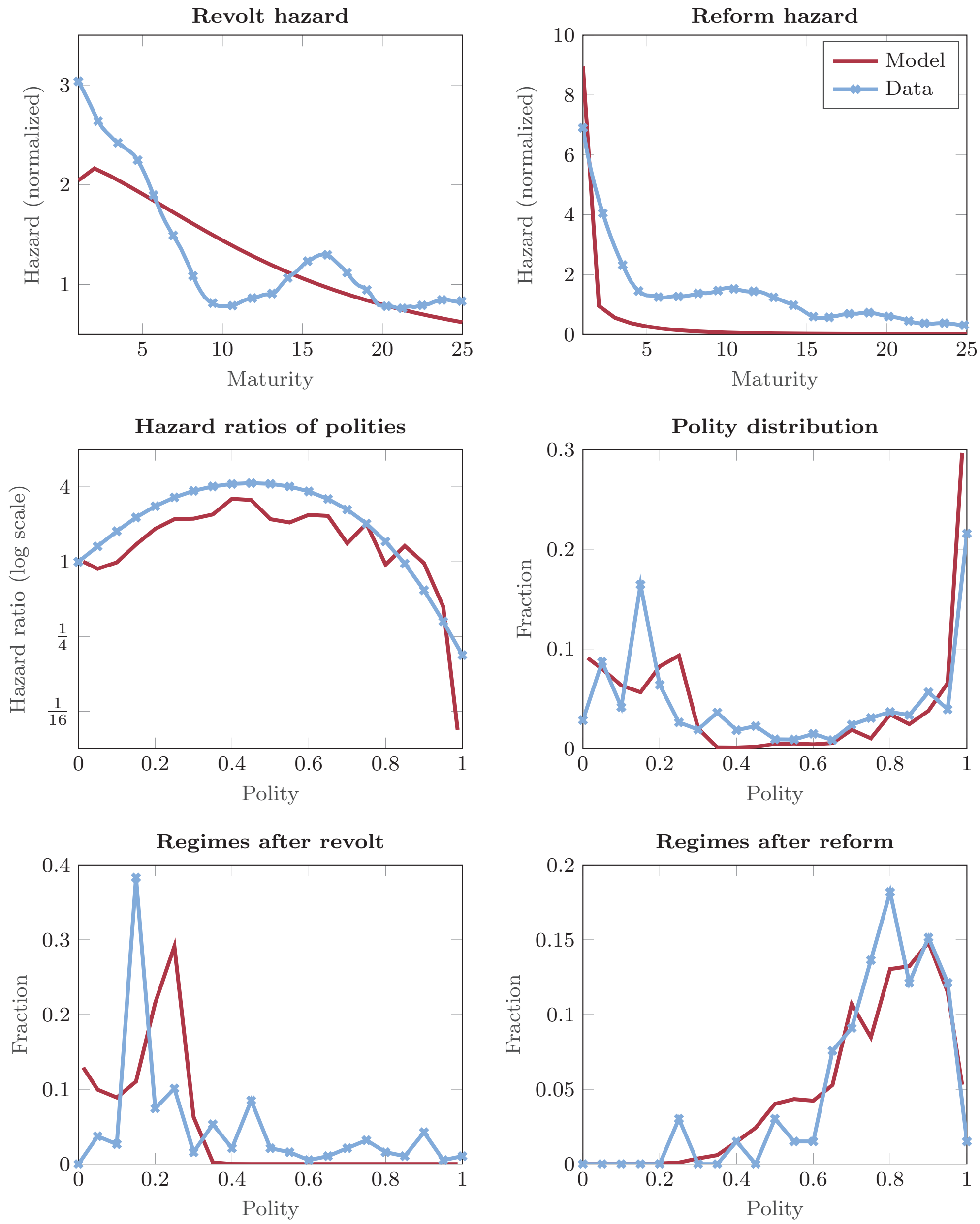

FIGURE 7

Comparison between model (solid red lines) and data (crossed blue lines). None of the depicted relations are directly targeted in the calibration.

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 cost ($\alpha_u \to 0$), 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, which further tilts the regime towards holding on to power

TABLE 3
Frequency of revolts for alternative parameters and without asymmetric information

	Asymmetric information		Symmetric information	
Cost of sharing power (α_u)	Revolt–reform ratio	Autocracy–democracy ratio	Revolt–reform ratio	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 $\lambda \leq 0.25$ relative to the mass of regimes with $\lambda \geq 0.75$ at the stationary distribution.

instead of reforming. If instead the realization of θ_t would be observed by outsiders, reforms have no signalling 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. That is, not only are small reforms precluded by asymmetric information, but more generally they are biased downwards. As such, asymmetric information reduces the likelihood of reforms and tends to increase the likelihood of revolts (see Figure E.5 in the Supplementary Appendix for an illustration).

To gauge the quantitative importance of each of these two 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 $\lambda \leq 0.25$) relative to "democracies" (with $\lambda \geq 0.75$) at the stationary distribution. For the baseline calibration ($\alpha_u = 0.57$), both of the aforementioned factors contribute roughly equally to explaining the data: if information were symmetric, then the revolt–reform ratio drops below 1.5 (compared to 2.85 in the data), and $\mathcal{P}(\lambda \geq 0.75)$ exceeds $\mathcal{P}(\lambda \leq 0.25)$ by a factor of about two (compared to roughly equal shares in the data).

5.2. *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 (3.11), it follows that for any $\mathcal{S}_t$,[21]

$$\mu_{t+1}|(\text{reform}_t) \geq \mu_{t+1}|(\text{revolt}_t) > \mu_{t+1}|(\text{no transition}_t). \tag{5.16}$$

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

21. 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 (3.11) at the upper and lower bound for M^2, respectively, and combining with (4.13) yields the two inequalities stated in (5.16).

to reform. Moreover, 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.

Observe that (5.16) holds under any F_t and does not hinge on the shape of G (or on our beta-approximation to (3.10)). 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

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

Proposition 5 proves for perfectly persistent θ that the revolt and reform hazards are declining in regime maturity (regardless of whether or not θ_t is observed by the statistician). 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 artefact of $\Pr(\theta_t = \theta_{t-1}) \to 1$, it is reminiscent of the steep decline seen in the estimated model and the data.

5.3. *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 ($\tilde{\mu}_{t-1} \leq \bar{\theta}_{t-1}$).[22] 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 bold black line traces out the contour, $(\mu, \lambda) = (\mathbb{E}[\mu|\lambda], \lambda)$, of the relation shown in the right panel, which closely approximates the exact J-curve $\mathbb{E}[\text{haz}|\lambda]$ (depicted by the thin red line).[23] Relative to regimes at the centre 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

22. 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.

23. $\mathbb{E}[\text{haz}(\mathbb{E}[\mu|\lambda], \lambda)]$ is only approximate for two reasons. First, since the transition hazard is non-linear 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 (thin red line), 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.

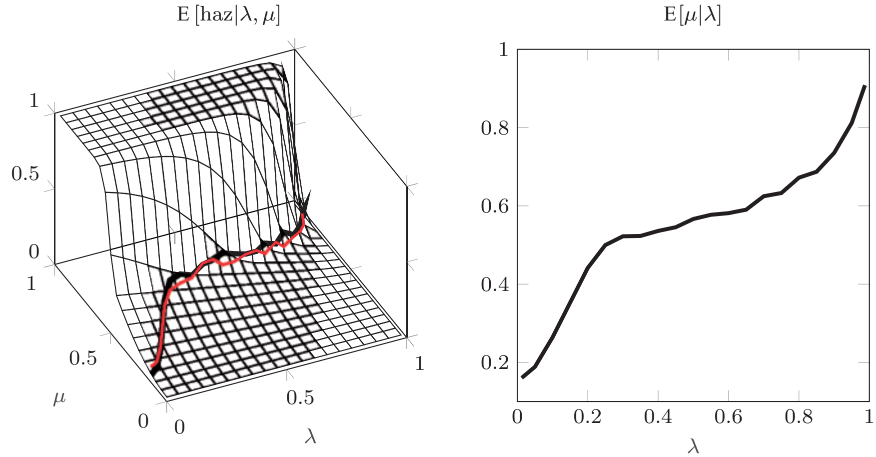

FIGURE 8

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

$(\lambda \to 0)$, on the other hand, are similarly uncontested due to their perceived strength implying a low degree of coordination among outsiders.

5.4. *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.[24] 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 $\theta_t \in [0,1]$,

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

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

For instance, if outsiders have a uniform prior $(F_t(\theta)=\theta)$, then $M_t^1(\bar{\theta}_t)=\bar{\theta}_t/2$, implying $\bar{\lambda}_t^{\text{ref}} > 1-(1-\lambda_t)/2$ and $\bar{\lambda}_t^{\text{rev}} < (1-\lambda_t)/2$. 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,

24. Note that the result is not a mechanical consequence of reforms being bounded below by λ_t, which does not rule out that reforming autocracies become marginally larger autocracies. As evident from Figure 9 below, the lower bound is far from binding for autocracies, which conduct median reforms much larger than λ_t.

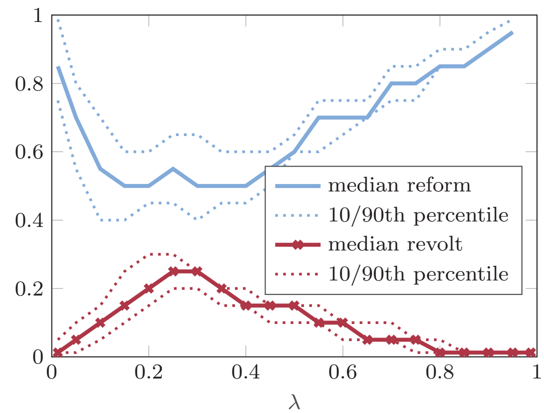

FIGURE 9
Median reform and revolt conditional on originating regime.

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 $x_t = \xi(\theta_t)$.

5.5. *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 Figure 7(e) and (f), and the inverse J-shape of the transition hazard in λ, as shown in Figure 7(c).

6. MODEL IMPLICATIONS FOR DEMOCRATIZATION

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

6.1. *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).[25] If outsiders perceive the regime as sufficiently strong (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

25. Here, λ and σ are fixed at 0.1 and 0.0475, 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.

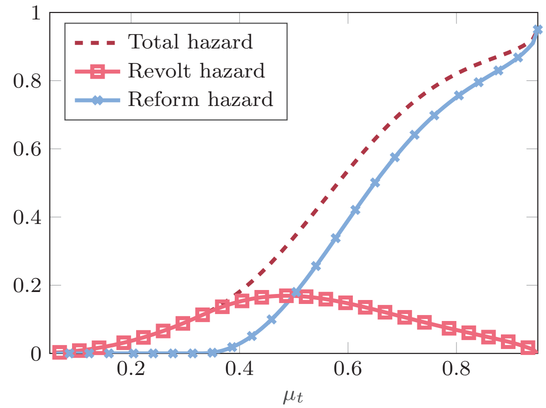

FIGURE 10
The role of outsiders' beliefs for transitions.

"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).[26] 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 likely to persist for a long time.

6.2. *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 $\lambda_0 = 0.1$ and beliefs given by their corresponding steady state values in the absence of any transition. As explained by a low implied 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.[27]

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% and the revolt hazard to roughly 15%. 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

26. 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.

27. 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 neighbouring countries as in Buera *et al.* (2011).

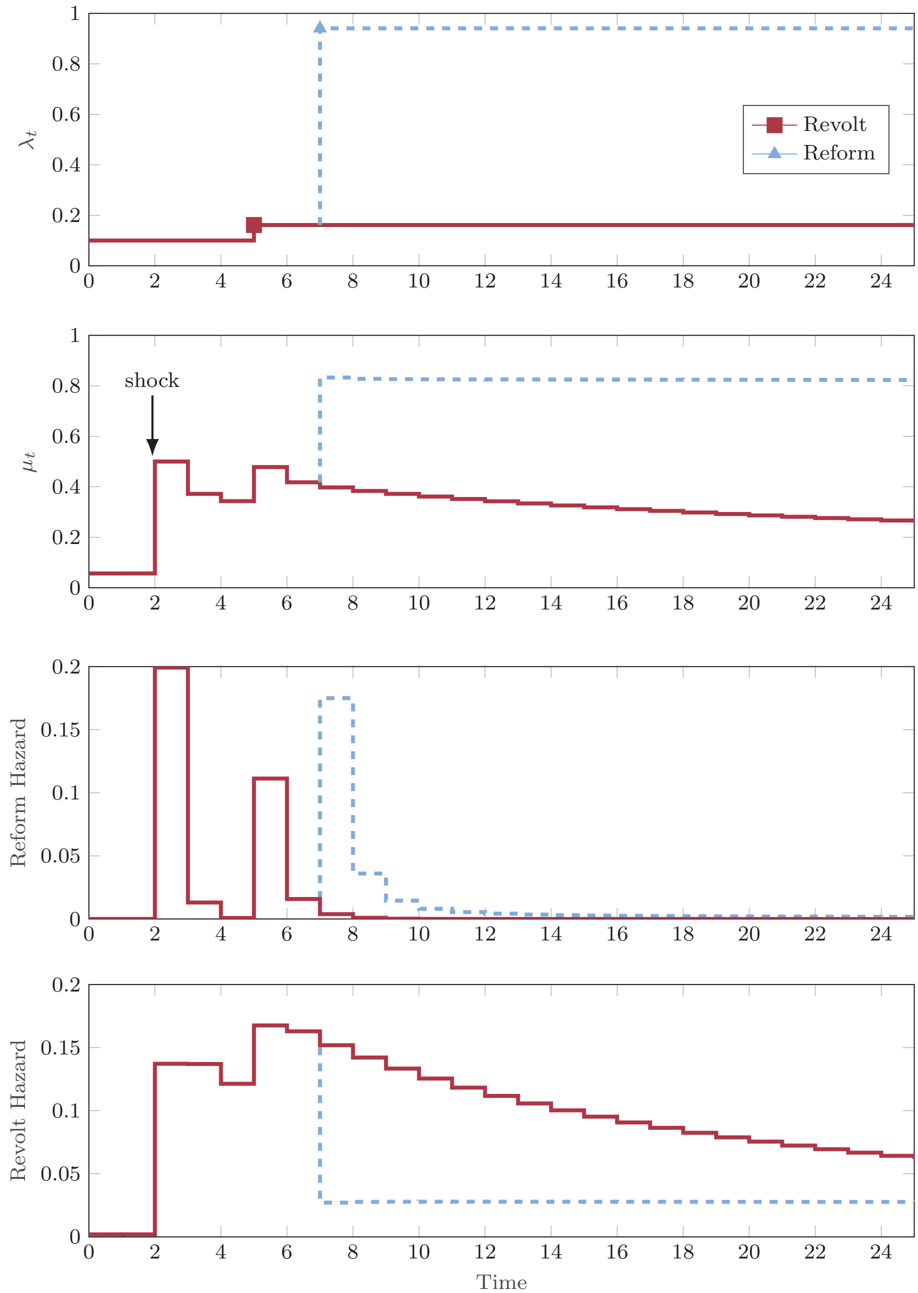

FIGURE 11

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.

of zero and a revolt hazard that converges to less than 1% 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%. 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%, 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.[28]

6.3. *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 $\lambda \geq 0.75$) as a function of their maturity and the inclusiveness of the democracy *at the time of its formation*.[29] 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 $\lambda \leq 0.85$ is overthrown in its first 25 periods is over 80%, the same probability drops to roughly 30% for democracies with an initial polity between 0.85 and 0.95, and drops to 12% for democracies that are initially larger than 0.95.

6.4. *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 $\lambda_{N_\lambda} = 0.9875$ (see footnote 18 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 largest democracy that is formed along the equilibrium path. Given the estimated value for α, we have $\xi(1) = 0.978$, 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 calibration, 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 $\alpha_h \leq 0.5$.

28. 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.

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

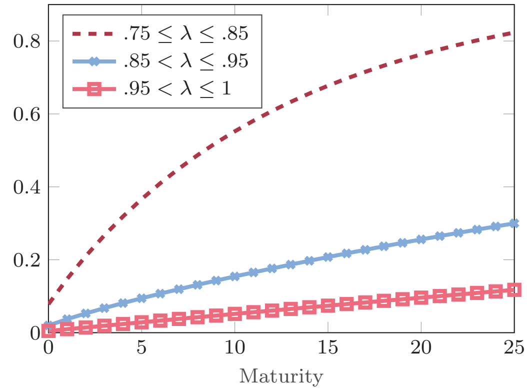

FIGURE 12
Cumulative reversal rate of young democracies.

If the success rate of revolts is relatively *inelastic* in the number of supporters ($\alpha_h \leq 0.5$), 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.[30] By contrast, if $\alpha_h > 0.5$, 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 reviews alternative mechanisms for the empirical regularities documented in Section 2 that have been suggested by the literature. Existing works generally differ from the present article in that they are not quantitative and make joint predictions for at most two of the five facts.

7.1. *Co-occurrence of revolts and reforms*

The present model already features two forces favouring the co-existence between revolts and (preemptive) reforms, both of which are also present in the existing literature.[31] Specifically, asymmetric information has been used to rationalize revolts despite preemptive reforms in, *e.g.*, Acemoglu and Robinson (2000a), Ellis and Fender (2011), and Boix (2003). Second, "gambling for survival," which emerges when the benefits from holding on to power are

30. 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. In 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.

31. Here we focus our discussion on revolts relating to *preemptive* reforms, because it is precisely the possibility of preemptive reforms that makes the prevalence of revolts puzzling in the first place. We are unaware of any work relating revolts to other (non-preemptive) types of reform.

sufficiently large relative to the likelihood that a revolt succeeds, has been previously present in Besley and Persson (2018). Finally, a third force, 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 *et al.*, 2010; Chassang and Padró i Miquel, 2009).

7.2. *Transition hazards declining in maturity*

In our model, transition hazards decline in regime maturity due to a reduced coordination among outsiders caused by learning.[32] The 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, a 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 *et al.*, 2010) whereas emerging autocracies first need to establish institutions to effectively distribute economic rents to supporters (Svolik, 2009). 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 *et al.*, 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 *et al.*, 2011), which over time increases the flow rents under democracy $u(\lambda \rightarrow 1)$ relative to other regime types, reducing the chance of regime reversals.[33]

Note that with the exception of Besley and Persson (2018), these explanations are specific to either autocratic or democratic consolidations. Our belief-driven explanation, by contrast, applies to all regime types, explaining the equally universal decline in hazards present in the data (see Supplementary Appendix A.2 for empirical hazard curves by regime type).

7.3. *Inverse J-shape of transition hazards*

The notion that regime stability is J-shaped in inclusiveness has been coined by 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. In contrast, intermediate regimes lack the institutions to preempt conflict whereas they are also ill-equipped to contain the spread of subversive ideas.[34] An alternative

32. The mechanism 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 in Section 5, a similar selection effect arises in the present framework if the econometrician does not observe θ, reinforcing the decline in hazards through learning.

33. At the level of individual leaders, Ales *et al.* (2014) explain a decline in exit rates through (self-enforcing) contracts, where re-elected leaders are those who get rewarded for compliant behaviour, increasing their flow utility and hence their propensity of future compliance.

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

account is given 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 less 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, while in our model this correlation arises endogenously. 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.

7.4. *Polarization of new regimes*

Although we are aware of no other paper that endogenously predicts polarization towards both ends of the autocracy–democracy spectrum, there are a few papers that also predict that reforms are far-reaching in equilibrium. Closely related to our approach, Acemoglu and Robinson (2000a) predict that reforms tend to be fully inclusive in the presence of asymmetric information. In addition, both Justman and Gradstein (1999) and Jack and Lagunoff (2006b) predict incremental expansions of the franchise towards full democracy, but leave the scope of each reform unspecified. In particular, Justman and Gradstein argue that technological progress lowers the costs of higher redistribution associated with becoming more inclusive relative to a (reduced-form) benefit, leading to successive extensions of the suffrage in parallel to the industrial revolution. Jack and Lagunoff provide a more abstract trade-off for monotone franchise extensions where reforms are a means to implement future policies that are preferred but cannot be credibly enacted by current leaders. As the identity of the leader shifts over time, this leads to gradually larger and larger franchises.

7.5. *Bimodal distribution*

To the best of our knowledge, this article is first to make predictions about the stationary distribution over regime types. The only other papers that we are aware of that can account for the bimodal distribution seen in the data, albeit mechanically, are Acemoglu and Robinson (2001) and Ticchi *et al.* (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 to match key facts from the data. Our 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 addition of an intensive margin of revolts, which links revolutionary pressure to the beliefs of outsiders regarding the regime's strength: precisely this link reduces the effectiveness of reforms, explaining the prevalence of revolts (fact 1). Similarly, the stabilization of mature and autocratic regimes (facts 2 and 3) are both a direct consequence of reduced coordination along the intensive margin. Finally, absent an intensive margin, revolts would also mechanically increase for small λ_t, leading to counterfactually large revolts that are at odds with fact 4.

Acknowledgments. This article supersedes an earlier working paper titled "Dynamics of Political Systems." We would like to thank the managing editor, Michèle Tertilt, four anonymous referees, Daron Acemoglu, Toke Aidt, Matthias Doepke, Christian Hellwig, Johannes Maier, Uwe Sunde, and various seminar and conference participants for valuable comments and suggestions.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ACEMOGLU, D., EGOROV, G. and SONIN, K. (2008), "Coalition Formation in Non-democracies", *Review of Economic Studies*, **75**, 987–1009.
ACEMOGLU, D., EGOROV, G. and SONIN, K. (2012), "Dynamics and Stability of Constitutions, Coalitions, and Clubs", *American Economic Review* **102**, 1446–1476.
ACEMOGLU, D., EGOROV, G. and SONIN, K. (2015), "Political Economy in a Changing World", *Journal of Political Economy* **123**, 1038–1086.
ACEMOGLU, D., JOHNSON, S., ROBINSON, J. A. *et al.* (2008), "Income and Democracy", *American Economic Review* **98**, 808–842.
ACEMOGLU, D., JOHNSON, S., ROBINSON, J. A. *et al.* (2009), "Reevaluating the Modernization Hypothesis," *Journal of Monetary Economics* **56**, 1043–1058.
ACEMOGLU, D. and ROBINSON, J. A. (2000a), "Democratization or Repression?", *European Economic Review* **44**, 683–693.
ACEMOGLU, D. and ROBINSON, J. A. (2000b), "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective", *Quarterly Journal of Economics* **115**, 1167–1199.
ACEMOGLU, D. and ROBINSON, J. A. (2001). "A Theory of Political Transitions", *American Economic Review* **91**, 938–963.
ACEMOGLU, D. and ROBINSON, J. A. (2008), "Persistence of Power, Elites, and Institutions", *American Economic Review* **98**, 267–293.
ACEMOGLU, D., TICCHI, D. and VINDIGNI, A. (2010), "A Theory of Military Dictatorships", *American Economic Journal: Macroeconomics* **2**, 1–42.
ACEMOGLU, D., TICCHI, D. and VINDIGNI, A. (2011), "Emergence and Persistence of Inefficient States", *Journal of the European Economic Association* **9**, 177–208.
AIDT, T. S. and FRANCK, R. (2015), "Democratization under the Threat of Revolution: Evidence from the Great Reform Act of 1832", *Econometrica* **83**, 505–547.
AIDT, T. S. and JENSEN, P. S. (2014), "Workers of the World, Unite! Franchise Extensions and the Threat of Revolution in Europe, 1820–1938", *European Economic Review* **72**, 52–75.
AIDT, T. S., LEON, G. and SATCHELL, M. (2017), "The Social Dynamics of Riots: Evidence from the Captain Swing Riots, 1830–31" (Working Paper).
ALES, L., MAZIERO, P. and YARED, P. (2014), "A Theory of Political and Economic Cycles", *Journal of Economic Theory* **153**, 224–251.
BAI, J. H. and LAGUNOFF, R. (2011), "On the Faustian Dynamics of Policy and Political Power", *Review of Economic Studies* **78**, 17–48.
BESLEY, T. and PERSSON, T. (2018), "Democratic Values and Institutions" (Working Paper).
BIENEN, H. and VAN DE WALLE, N. (1989), "Time and Power in Africa", *American Political Science Review 83*, 19.
BIENEN, H. and VAN DE WALLE, N. (1992), "A Proportional Hazard Model of Leadership Duration", *The Journal of Politics* **54**, 685–717.
BOIX, C. (2003), *Democracy and Redistribution* (Cambridge: Cambridge University Press).
BOURGUIGNON, F. and VERDIER, T. (2000), "Oligarchy, Democracy, Inequality and Growth", *Journal of Development Economics* **62**, 285–313.
BREMMER, I. A. (2006), *The J Curve: A New Way to Understand Why Nations Rise and Fall* (New York: Simon & Schuster).
BUCHHEIM, L. and ULBRICHT, R. (2014), "Emergence and Persistence of Extreme Political Systems", (Mimeo).
BUENO DE MESQUITA, B., SMITH, A., SIVERSON, R. M. *et al.* (2003), *The Logic of Political Survival* (Cambridge, MA: MIT Press).
BUENO DE MESQUITA, E. (2010), "Regime Change and Revolutionary Entrepreneurs", *American Political Science Review* **104**, 446–466.
BUERA, F. J., MONGE-NARANJO, A. and PRIMICERI, G. E. (2011), "Learning the Wealth of Nations", *Econometrica* **79**, 1–45.

CASPER, B. A. and TYSON, S. A. (2014), "Popular Protest and Elite Coordination in a Coup d'état", *The Journal of Politics* **76**, 548–564.

CELESTINO, M. R. and GLEDITSCH, K. S. (2013), "Fresh Carnations or All Thorn, No Rose? Nonviolent Campaigns and Transitions in Autocracies", *Journal of Peace Research* **50**, 385–400.

CHASSANG, S. and PADRÓ I MIQUEL, G. (2009), "Economic Shocks and Civil War", *Quarterly Journal of Political Science* **4**, 211–228.

CHO, I.-K. and KREPS, D. M. (1987), "Signaling Games and Stable Equilibria", *Quarterly Journal of Economics* **102**, 179–221.

CONLEY, J. P. and TEMINI, A. (2001), "Endogenous Enfranchisement When Groups' Preferences Conflict", *Journal of Political Economy* **109**, 79–102.

DERPANOPOULOS, G., FRANTZ, E., GEDDES, B. *et al.* (2016), "Are Coups Good for Democracy?", *Research & Politics* **3**, 1–7.

ELLIS, C. J. and FENDER, J. (2011), "Information Cascades and Revolutionary Regime Transitions", *Economic Journal* **121**, 763–792.

ENIKOLOPOV, R., MAKARIN, A. and PETROVA, M. (2018), "Social Media and Protest Participation: Evidence from Russia" (Working Paper).

FINKEL, E., GEHLBACH, S. and OLSEN, T. D. (2015), "Does Reform Prevent Rebellion? Evidence from Russia's Emancipation of the Serfs", *Comparative Political Studies* **48**, 984–1019.

GALLEGO, M. and PITCHIK, C. (2004), "An Economic Theory of Leadership Turnover", *Journal of Public Economics* **88**, 2361–2382.

GATES, S., HEGRE, H., JONES, M. P. *et al.* (2006), "Institutional Inconsistency and Political Instability: Polity Duration", 1800–2000. *American Journal of Political Science* **50**, 893–908.

GEDDES, B., WRIGHT, J. and FRANTZ, E. (2014), "Autocratic Breakdown and Regime Transitions: A New Data Set", *Perspectives on Politics* **12**, 313–331.

GLEDITSCH, K. S. and CHOUNG, J. L. (2004), "Autocratic Transitions and Democratization" (Working Paper).

GLEDITSCH, K. S. and WARD, M. D. (2006), "Diffusion and the International Context of Democratization", *International Organization* **60**, 141.

GOEMANS, H. E., GLEDITSCH, K. S. and CHIOZZA, G. (2009), "Introducing Archigos: A Dataset of Political Leaders", *Journal of Peace Research* **46**, 269–283.

GOLDSTONE, J. A., BATES, R. H., EPSTEIN, D. L. *et al.* (2010), "A Global Model for Forecasting Political Instability", *American Journal of Political Science* **54**, 190–208.

GONZALEZ, F. (2019), "Collective Action in Networks: Evidence from the Chilean Student Movement" (Working Paper).

GRADSTEIN, M. (2007), "Inequality, Democracy and the Protection of Property Rights", *Economic Journal* **117**, 252–269.

GRANOVETTER, M. (1978), "Threshold Models of Collective Behavior", *American Journal of Sociology* **83**, 1420–1443.

HIRSHLEIFER, J., BOLDRIN, M. and LEVINE, D. K. (2009), "The Slippery Slope of Concession", *Economic Inquiry* **47**, 197–205.

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

JACK, W. and LAGUNOFF, R. (2006a), "Dynamic Enfranchisement", *Journal of Public Economics* **90**, 551–572.

JACK, W. and LAGUNOFF, R. (2006b), "Social Conflict and Gradual Political Succession: An Illustrative Model", *Scandinavian Journal of Economics* **108**, 703–725.

JUSTMAN, M. and GRADSTEIN, M. (1999), "The Industrial Revolution, Political Transition, and the Subsequent Decline in Inequality in 19^{th}-Century Britain", *Explorations in Economic History* **36**, 109–127.

KARL, T. L. (1990), "Dilemmas of Democratization in Latin America", *Comparative Politics* **23**, 1–21.

KNUTSEN, C. H. and NYGÅRD, H. M. (2015), "Institutional Characteristics and Regime Survival: Why Are Semi-democracies Less Durable than Autocracies and Democracies?", *American Journal of Political Science* **59**, 656–670.

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

LAGUNOFF, R. (2009), "Dynamic Stability and Reform of Political Institutions", *Games and Economic Behavior* **67**, 569–583.

LIZZERI, A. and PERSICO, N. (2004), "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, with an Application to Britain's 'Age of Reform"', *Quarterly Journal of Economics* **119**, 707–765.

LLAVADOR, H. and OXOBY, R. J. (2005), "Partisan Competition, Growth, and the Franchise", *Quarterly Journal of Economics* **120**, 1155–1189.

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

MANACORDA, M. and TESEI, A. (2018), "Liberation Technology: Mobile Phones and Political Mobilization in Africa" (Working Paper).

MARSHALL, M. G., GURR, T. R. and JAGGERS, K. (2017), Polity IV Project: Political Regime Characteristics and Transitions, 1800-2016. Center for Systemic Peace, http://www.systemicpeace.org/inscrdata.html.

O'DONNELL, G. and SCHMITTER, P. C. (1986), *Transitions From Authoritarian Rule: Tentative Conclusions about Uncertain Democracies* (Baltimore, MD: John Hopkins University Press).

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

PRZEWORSKI, A. and LIMONGI, F. (1997), "Modernization: Theories and Facts", *World Politics* **49**, 155–183.

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

SANHUEZA, R. (1999), "The Hazard Rate of Political Regimes", *Public Choice* **98**, 337–367.

SVOLIK, M. W. (2008), "Authoritarian Reversals and Democratic Consolidation", *American Political Science Review* **102**, 153–168

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

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

TICCHI, D., VERDIER, T. and VINDIGNI, A. (2013), "Democracy, Dictatorship and the Cultural Transmission of Political Values" (IZA Discussion Paper 7441).

TULLOCK, G. (1971), "The Paradox of Revolution", *Public Choice* **11**, 89–99.