

Learning about Growth and Democracy*

Scott F Abramson[†]

Sergio Montero[‡]

University of Rochester

University of Rochester

April 2, 2018

Abstract

We develop and estimate a model of learning that explains the observed correlation between income and democracy as well as the clustering of democratization events. In our model, countries' own and neighbors' past experiences shape elites' beliefs about the effects of democracy on economic growth and their likelihood of retaining power. These beliefs influence the choice to transition into or out of democracy. We show that learning from past experiences is crucial to explaining observed transitions since the mid-twentieth century. Moreover, our model predicts reversals to authoritarianism if the world experienced a growth shock the size of the Great Depression.

JEL codes: D72, D83, O43

Keywords: democracy, learning, growth, diffusion

*We thank Tasos Kalandrakis, Matt Shum, Erik Snowberg, and Didac Queralt for very helpful comments, and Anna Walsdorff for excellent research assistance.

[†]Department of Political Science, University of Rochester, Rochester, NY 14627. Email: sabramso@ur.rochester.edu.

[‡]Department of Political Science and Department of Economics, University of Rochester, Rochester, NY 14627. Email: smontero@rochester.edu.

1 Introduction

Empirical scholarship on the causes of democracy has sought to understand two patterns: the strong correlative relationship between levels of material well-being and democracy (Lipset, 1959; Barro, 1996; Przeworski et al., 2000; Boix and Stokes, 2003; Acemoglu et al., 2008, 2009), and the spatial and temporal clustering of democratization events (Goldstone, 1991; Huntington, 1993; Brinks and Coppedge, 2006; Gleditsch and Ward, 2006; Ahlquist and Wibbels, 2012; Houle et al., 2016). Existing studies have treated these as distinct objects of inquiry, separately assessing the impact of domestic and international factors on the propensity to democratize. In this paper, we develop and estimate a model of elite belief formation that combines both domestic and systemic features in order to jointly explain the correlation between economic development and democracy, and the clustering of democratic transitions.

We model the choice by incumbent elites to promote or subvert democracy.¹ This choice impacts their likelihood of retaining power both directly and indirectly through its effect on economic growth. Incumbents are uncertain about the relationship between democracy and growth, and they rely upon worldwide economic history to update their beliefs. We allow beliefs to be spatially correlated so that incumbents may learn more from the experiences of more proximate (or similar) countries.² Incumbents pursue democracy or autocracy, in accordance with their beliefs, seeking to maximize the probability they remain in power. For a panel of 151 countries, we use data from 1875-1950 to calibrate initial conditions and data from 1951-2000 to estimate our model.

To assess the ability of our learning model to explain observed patterns of economic growth and democracy adoption, we conduct a series of goodness-of-fit and out-of-sample (2001-2010) prediction exercises that pit our model against a range of reduced-form panel regressions typical of the approach taken in the existing empirical literature on democratization (e.g., Acemoglu et al., 2008; Boix, 2011). We show that learning from past experiences is crucial to explaining observed transitions to and from democracy, delivering an improvement in predictive success of over 100% relative to the best-fitting specification that does not account for learning.

The success of our learning model is rooted, first, in our estimates of the political implications of economic growth. In line with a sizable empirical literature in political economy, we find that democracies tend to reward incumbents for growth by keeping them in power.³ In

¹We conceive elites broadly as the group or faction in charge of the executive. In democracies, e.g., this is typically the party of the president. See Section 2 for a more detailed discussion.

²Geographic distance is highly correlated with various measures of similarity between countries. Our results are robust to accounting directly for other types of similarity.

³See, e.g., Nordhaus (1975); Hibbs (1977); Alesina et al. (1997); Drazen (2000); Brender and Drazen

contrast, we find that growth tends to be destabilizing in autocracies. That is, our estimates are consistent with the view that rapid economic modernization in autocracies produces actors—a middle class, for example—able to challenge the group in power.⁴ Together, these results contribute to a substantial body of work on modernization and democracy. Dating to at least the mid-twentieth century, social scientists have debated whether increases in per capita income have a causal effect on the probability that a state democratizes.⁵ Our findings suggest that the observed correlation between income and democracy is largely driven by elites who seek to benefit politically from the economic consequences of institutional choice. In particular, our results indicate that the worldwide prospect for democracy depends considerably upon its ability to engender growth. Democratic incumbents will subvert democracy when they come to believe that it does not produce sufficient economic growth to win a fair election. Conversely, autocratic incumbents will only transition to democracy when expected rates of growth make them more likely to retain power via election than under continued authoritarian rule.

Importantly, we distinguish in this paper transitions of power, where only the identity of the incumbent changes, from transitions into or out of democracy, where the form of government changes. Sudden changes in economic conditions, for instance, may lead to generalized political instability regardless of the system of government in place. However, whenever an incumbent is replaced—be it through election, coup, or revolution—the new group in power again faces the choice to support or subvert democracy. Our focus is on this choice, not transitions of power between factions per se.

The second feature of our learning model that underpins its empirical success is its ability to capture the wave-like nature of democratic transitions. We find that beliefs about the relationship between democracy and economic growth are highly correlated both temporally and spatially, which provides a structural interpretation for the observed clustering of transitions. Our estimates indicate that there is an approximately 5,000 kilometer radius within which learning occurs. Outside this distance, virtually no additional information is gleaned. This advances the literature on the diffusion of democracy, which has struggled to disentangle competing mechanisms.⁶

(2008).

⁴E.g., Olson (1963); Huntington (1968); Hirschman and Rothschild (1973).

⁵The modernization hypothesis—that higher incomes per capita cause countries to democratize—dates to, at least, Lipset (1959). Versions of this argument are echoed in many seminal texts on democracy: e.g., Dahl (1973); Rueschemeyer et al. (1992); Huntington (1993). For evidence in favor, see Londregan and Poole (1996); Barro (1996); Boix and Stokes (2003); Boix (2011). Against, see Przeworski and Limongi (1997); Przeworski et al. (2000); Acemoglu et al. (2008, 2009).

⁶Proposed mechanisms include diffusion through international organizations (Pevehouse, 2002, 2005), direct emulation of neighbors (Gleditsch and Ward, 2006), diffusion through trade and economic exchange (Mansfield et al., 2000; Li and Reuveny, 2003), cultural linkages (Levitsky and Way, 2005; Wejnert, 2005),

Our approach allows us to conduct counterfactual experiments of three types, each of which highlights the importance of learning. First, it enables us to understand how systemic shocks to prosperity impact the worldwide prospect for democracy. Specifically, our model predicts considerable reversals to authoritarianism if the world were hit with a shock to growth the size of the Great Depression. Second, it allows us to ask retrospective questions about historical democratization events, such as whether Greece, Portugal, and Spain would have democratized when they did had western European democracies suffered a recession in the early 1970s. Third, our model allows us to prospectively explore conditions that would currently lead countries to transition to or from democracy. Overall, our results suggests that, ultimately, democracy is a fragile system of government, one which depends significantly upon its own economic success.

Of course, we are not the first to propose diffusion through learning as a cause of democratic transitions.⁷ Indeed, in light of the extended period of stagnant growth in Western democracies following the Great Recession, political scientist Larry Diamond argued that

“...the pace of democratic break-downs would surely accelerate and possibly gather into a potent and undeniable reverse wave, driven not only by the spread of economic crisis but also by the much deeper symbolic loss of democratic prestige in an era—were it to come—in which the rich, established, capitalist democracies proved powerless to turn back the tide of economic misfortune.”⁸

If direct and consistent measures of beliefs were available over a sufficiently long period and a wide enough set of countries, it would be conceivable to directly estimate the impact of changing beliefs on democratization.⁹ Given the current lack of systematic or reliable data, our paper represents the first attempt at estimating this impact.

We are also not the first to examine the role of learning in policymaking more broadly.¹⁰ Our paper is most closely related to Buera et al.’s (2011) structural analysis of the impact of learning on the adoption of market-oriented (versus state-interventionist) policies. We build on their framework to model the worldwide evolution and diffusion of beliefs about the economic consequences of competing policies. Yet, while they consider the problem of a welfare-maximizing social planner, we take a political economy perspective and study institutional design as the outcome of self-interested choices made by power-seeking elites.

and military coercion (Kadera et al., 2003; Gunitsky, 2014; Boix, 2011). On the inability of this literature to empirically falsify any particular mechanism of diffusion, see Torfason and Ingram (2010).

⁷E.g., Dahl (1998); Chang et al. (2007); Basora (2008); Kurlantzick and Link (2009); Jackson (2010).

⁸Diamond (2011) p. 26.

⁹For work that attempts to gauge beliefs about democracy and their consequences for democratization in a subset of countries, see Almond and Verba (1963); Norris (1999); Chen and Lu (2011).

¹⁰E.g., Besley and Case (1994); Foster and Rosenzweig (1995); Sargent (1999); Primiticri (2006); Alesina and Fuchs-Schündeln (2007); Landier et al. (2008); Conley and Udry (2010); García-Jimeno (2016).

The remainder of this paper proceeds as follows. In Section 2, we describe our model, empirical strategy, and the data used in our analysis. Next, in Section 3, we summarize our main estimation results and evaluate the explanatory power of our model. We present our counterfactual experiments in Section 4. Finally, in Section 5, we discuss the implications of our results and conclude.

2 A Learning Model of Democratization

2.1 Setup: Elites, Beliefs, and Learning

We consider the decision problem of the decisive group or political actor in power in country i at time t .¹¹ This agent faces a choice between autocracy, $D_{i,t} = 0$, and democracy, $D_{i,t} = 1$. The incumbent's objective is to retain power in period $t + 1$. Let $Y_{i,t}$ denote country i 's per capita GDP in period t , and let $y_{i,t} \equiv \log(Y_{i,t}) - \log(Y_{i,t-1})$ denote its growth rate. At the beginning of period t , to maximize her expected probability of remaining in power, the incumbent solves

$$\max_{D_{i,t} \in \{0,1\}} E_{i,t-1} \left[\frac{\exp(\alpha_i + \theta^{D=D_{i,t}} y_{i,t} - K_{i,t} D_{i,t})}{1 + \exp(\alpha_i + \theta^{D=D_{i,t}} y_{i,t} - K_{i,t} D_{i,t})} \right], \quad (1)$$

where the integrand represents the probability of retaining power in period $t + 1$ and the expectation is taken conditional on the information available in country i at the conclusion of period $t - 1$. The integrand is increasing in the index $\alpha_i + \theta^{D=D_{i,t}} y_{i,t} - K_{i,t} D_{i,t}$, where coefficient α_i establishes a baseline for country i , coefficients $\theta^{D=0}$ and $\theta^{D=1}$ respectively measure the (de)stabilizing effect of GDP growth on elite turnover under autocracy and democracy, and $K_{i,t}$ captures the political cost of democracy to the incumbent, i.e., its direct effect on her likelihood of retaining power.

The incumbent chooses $D_{i,t}$ at the start of period t , forming a subjective forecast of its effect on GDP growth, $y_{i,t}$, to solve (1).¹² Incumbents believe that the relationship between GDP growth and democracy takes the form

$$y_{i,t} = (1 - D_{i,t})\beta_i^{D=0} + D_{i,t}\beta_i^{D=1} + \epsilon_{i,t}, \quad (2)$$

where $\beta_i^{D=0}$ and $\beta_i^{D=1}$ denote country i 's long-run GDP growth rates under autocracy and democracy, respectively, and $\epsilon_{i,t}$ is an exogenous shock to growth that is uncorrelated over

¹¹Both in democracies and autocracies, we treat this as the political party or faction (when parties do not exist) in control of the executive. For a complete coding, see Table ?? in Appendix A1.

¹²In what follows, we build on the learning framework of Buera et al. (2011).

time but potentially correlated across countries. Specifically, the vector $\epsilon_t \equiv [\epsilon_{1,t}, \dots, \epsilon_{n,t}]'$ of GDP growth shocks across the n countries of the world is independently and identically distributed (i.i.d.) over time according to a mean zero Normal distribution with covariance matrix Σ , i.e.,

$$\epsilon_t \stackrel{\text{i.i.d.}}{\sim} N(0, \Sigma).$$

Incumbents do not know $\beta_i \equiv [\beta_i^{D=0}, \beta_i^{D=1}]'$, but they have perfect knowledge of all other features of the model, including $K_{i,t}$, at the time of their choice.

The timing of events is as follows. At the end of period $t - 1$, incumbents collect data on worldwide GDP growth rates and systems of government, and they update their beliefs about long-run economic growth under autocracy and democracy accordingly. At the beginning of period t , incumbents observe $K_{i,t}$ and decide what system of government, $D_{i,t}$, to adopt that period. Growth shocks and corresponding growth rates, conditional on incumbents' choices, are then realized.

Learning. In period $t = 0$, incumbents start out with a Normal prior over the vector of unknown long-run GDP growth rates $\beta \equiv [\beta_1^{D=0}, \dots, \beta_n^{D=0}, \beta_1^{D=1}, \dots, \beta_n^{D=1}]'$,

$$\beta \sim N(\bar{\beta}_0, P_0^{-1}), \tag{3}$$

where $\bar{\beta}_0$ and P_0 denote, respectively, the prior mean and precision matrix. We assume that incumbents' initial beliefs assign no correlation and the same degree of uncertainty to growth under autocracy and democracy:

$$P_0^{-1} = I_2 \otimes (V \cdot R \cdot V),$$

where $V = \text{diag}([v_1\sigma_1, \dots, v_n\sigma_n])$ is a diagonal matrix whose i th diagonal entry measures prior uncertainty (standard deviation) about country i 's long-run growth rate under autocracy/democracy, and R is the cross-country prior correlation matrix. Prior uncertainty is parameterized by $\{v_i\}_{i=1}^n$, normalized by the standard deviation of growth shocks in each country $\{\sigma_i\}_{i=1}^n$ (the square roots of the diagonal elements of Σ).

Our assumptions yield simple, recursive Bayesian-updating formulas for beliefs in each period: letting $D_t \equiv [D_{1,t}, \dots, D_{n,t}]'$ and $y_t \equiv [y_{1,t}, \dots, y_{n,t}]'$,

$$\begin{aligned} P_t &= P_{t-1} + \mathbf{D}'_t \Sigma^{-1} \mathbf{D}_t, \\ \bar{\beta}_t &= \bar{\beta}_{t-1} + P_t^{-1} \mathbf{D}'_t \Sigma^{-1} (y_t - \mathbf{D}_t \bar{\beta}_{t-1}), \end{aligned}$$

where $\mathbf{D}_t \equiv [\text{diag}(1 - D_t), \text{diag}(D_t)]$ is an $n \times (2n)$ matrix such that the i th element of the

vector $\mathbf{D}_t \bar{\beta}_{t-1}$ equals $(1 - D_{i,t}) \bar{\beta}_{i,t-1}^{D=0} + D_{i,t} \bar{\beta}_{i,t-1}^{D=1}$. The impact of new data on the posterior mean of β is determined by $P_t^{-1} \mathbf{D}'_t \Sigma^{-1}$, which depends on three key factors. First, higher initial uncertainty in beliefs (higher $\{v_i\}_{i=1}^n$) raises the relative precision of new information, increasing its impact. Second, higher correlation in growth shocks across countries (off-diagonal elements of Σ) reduces the informational content of observed growth rates and slows down learning. Lastly, higher cross-country correlation in initial beliefs (off-diagonal elements of R) increases belief responsiveness to data from other countries.

We allow incumbents to potentially learn more from neighboring (or more similar) countries than from distant countries. Letting $Z_{i,j}$ denote a vector that may include various measures of distance (geographic or otherwise) between countries i and j , we write

$$R_{i,j} = \exp(-Z'_{i,j} \gamma),$$

where γ is constrained to be nonnegative to ensure correlations between 0 and 1.¹³

Incumbents' optimal choice. While incumbents observe the political cost of democracy, $K_{i,t}$, prior to choosing $D_{i,t}$, this cost is unobservable to the researcher. We assume that $K_{i,t}$ has the following structure:

$$K_{i,t} = f_i + X'_{i,t} \xi + \kappa_{i,t}. \quad (4)$$

Coefficient f_i establishes a country-specific baseline, and the control vector $X_{i,t}$ may include various observable economic and political characteristics of country i (e.g., lagged per capita GDP or incumbents' time in power). Every period, country i also experiences an exogenous idiosyncratic shock, $\kappa_{i,t}$, to the political cost of democracy, where

$$\kappa_{i,t} \sim N(0, \zeta_i^2).$$

The volatility of shocks to the political cost of democracy, ζ_i , is allowed to be country-specific, but $\kappa_{i,t}$ is assumed to be independently distributed over time and across countries.

As discussed, when choosing $D_{i,t}$ incumbents have perfect knowledge of $K_{i,t}$ and all features of the model except for the effect of their choice on GDP growth, $y_{i,t}$. Together, (1), (2), and (4) imply that the optimal choice for country i 's incumbent at time t is

$$D_{i,t} = \mathbf{1} \left\{ E_{i,t-1} \left[\frac{\exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t})) - f_i - X'_{i,t} \xi - \kappa_{i,t}}{1 + \exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t})) - f_i - X'_{i,t} \xi - \kappa_{i,t}} \right] \right. \\ \left. > E_{i,t-1} \left[\frac{\exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))}{1 + \exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))} \right] \right\}, \quad (5)$$

¹³This formulation also guarantees the positive definiteness of R (Matérn, 1960).

where the expectations are taken only with respect to β_i and $\epsilon_{i,t}$ in accordance with the incumbent's beliefs at the conclusion of period $t - 1$.

Opposition groups and strategic experimentation. To conclude the description of our model, we briefly discuss how we account for non-elite learning and strategic interactions between incumbents and potential challengers.

The common prior assumption for incumbents in our model extends to all potential stakeholders in each country. Opposition groups (elite or non-elite) observe the same worldwide history of economic growth and democracy, and they would be faced with solving (1), in the event they came to power, using information identical to that available to the incumbent. As a result, in this shared learning environment, the identity of the incumbent only matters through its potential effect on the political cost of democracy.

For tractability, we abstract from explicitly modeling the intricacies of within-country elite turnover. Nevertheless, objective (1) can be viewed as describing an equilibrium probability of staying in power resulting from a richer interaction between the incumbent and potential challengers, both elite and non-elite. Importantly, as discussed in the Introduction, we distinguish transitions of power, where only the identity of the incumbent changes, from transitions into or out of democracy. Whenever an incumbent is overthrown by a rival elite faction or via a revolution from below, the new group in power faces the choice to support or subvert democracy.¹⁴ Understanding how this choice by self-interested elites—newly in power or entrenched—is shaped by the evolution of beliefs about the economic effects of democracy is the focus of this paper.

Finally, objective (1) precludes incumbents from adopting a form of government with negative expected consequences solely for the purpose of learning from the experience. In our setting, the prospect of losing power and thus not reaping the benefits of such experimentation limits its appeal. Relatedly, incumbents in our model are myopic and focused only on their immediate survival. While a fully dynamic version of our model with shared learning by forward-looking agents would introduce strategic experimentation incentives that would render the model virtually intractable,¹⁵ we believe our model approximates well optimal behavior by incumbents with longer time horizons.

¹⁴Historical examples abound of revolutions from below, inspired by ostensibly democratic goals, that failed to deliver on the promise of liberal democracy. For instance, the three main cases in Skocpol's (1979) seminal work on revolution—the French, Chinese, and Russian revolutions—all began as mass revolutionary movements with outwardly democratic motives and each, nonetheless, resulted in dictatorship.

¹⁵See, e.g., Bolton and Harris (1999); Bramoulle et al. (2014); Mossel et al. (2015).

2.2 Empirical Strategy

Like incumbents in our model, we adopt a Bayesian inference approach to recover the unknown structural parameters of our model, listed in Table 1.¹⁶ With a slight abuse of notation, let the vector $\boldsymbol{\theta}$ collect all the parameters in Table 1, and let $I_{i,t}$ be an indicator of whether the incumbent in country i retained power ($I_{i,t} = 1$) or not ($I_{i,t} = 0$) at the conclusion of period t . Denote by $W^T \equiv \{I_t, y_t, D_t, X_t\}_{t=1}^T$ the set of all data available up to period T , where $I_t \equiv [I_{1,t} \dots, I_{n,t}]'$ and $X_t \equiv [X_{1,t} \dots, X_{n,t}]'$. Our goal is to estimate $\boldsymbol{\theta}$ by computing the mode of the posterior distribution of the model parameters,

$$p(\boldsymbol{\theta}|W^T) \propto \mathcal{L}(W^T|\boldsymbol{\theta})\pi(\boldsymbol{\theta}),$$

given the likelihood of the data, \mathcal{L} , and our prior, π . We describe \mathcal{L} and π in turn.

Table 1: Model Parameters

$\{\alpha_i\}_{i=1}^n$:	baseline incumbent stability
$\theta^{D=0}$:	effect of GDP growth on elite turnover under autocracy
$\theta^{D=1}$:	effect of GDP growth on elite turnover under democracy
$\{\bar{\beta}_{i,0}^{D=0}\}_{i=1}^n$:	prior mean of long-run GDP growth rate under autocracy
$\{\bar{\beta}_{i,0}^{D=1}\}_{i=1}^n$:	prior mean of long-run GDP growth rate under democracy
$\{v_i\}_{i=1}^n$:	prior uncertainty about economic effects of autocracy/democracy
γ :	coefficients of cross-country correlation of prior beliefs
$\{f_i\}_{i=1}^n$:	baseline political cost of democracy
ξ :	coefficients of economic/political controls for political cost of democracy
$\{\varsigma_i\}_{i=1}^n$:	volatility of political cost of democracy

Likelihood of the data. With a slight abuse of notation, using \mathcal{L} to denote arbitrary densities of the data, the likelihood function can be written as

$$\mathcal{L}(W^T|\boldsymbol{\theta}) = \prod_{t=1}^T \mathcal{L}(W_t|W^{t-1}, \boldsymbol{\theta}),$$

where $W_t \equiv \{I_t, y_t, D_t, X_t\}$ collects the data generated in period t . Notice that, while the structure of the model described thus far specifies incumbents' beliefs about how the data are generated as well as their optimal choices given those beliefs, we have refrained from specifying the “true” data-generating process (DGP). In Section 4, to perform counterfactual

¹⁶Following Buera et al. (2011), to reduce the dimensionality of the model we set Σ equal to its estimated value from the “true” data-generating process (see Section 4).

experiments, we discuss and specify the true DGP. Next, we only make one key assumption about the true DGP that simplifies inference about the model parameters.

We assume that observed outcomes are only affected by actual choices and not by the beliefs that led to those choices. That is, transitions of power (I_t), GDP growth (y_t), and other economic and political characteristics of countries (X_t) are shaped by realized institutions (D_t), but they are not directly affected by beliefs about the potential effects of transitioning into or out of democracy. Formally, this assumption allows us to write

$$\begin{aligned}\mathcal{L}(W_t|W^{t-1}, \boldsymbol{\theta}) &= \mathcal{L}(I_t, y_t, D_t, X_t|W^{t-1}, \boldsymbol{\theta}) \\ &= \mathcal{L}(I_t|y_t, D_t, X_t, W^{t-1})\mathcal{L}(y_t|D_t, X_t, W^{t-1}) \dots \\ &\quad \mathcal{L}(D_t|X_t, W^{t-1}, \boldsymbol{\theta})\mathcal{L}(X_t|W^{t-1}),\end{aligned}$$

which implies that $\mathcal{L}(W^T|\boldsymbol{\theta}) \propto \prod_{t=1}^T \mathcal{L}(D_t|X_t, W^{t-1}, \boldsymbol{\theta})$.

To compute $\mathcal{L}(D_t|X_t, W^{t-1}, \boldsymbol{\theta})$, notice from (5) that, given $(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$, there is a threshold value of $\kappa_{i,t}$, the realized shock in period t to the political cost of democracy in country i , such that $D_{i,t} = 1$ if and only if $\kappa_{i,t}$ falls below the threshold. This threshold value, denoted $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$, is defined implicitly by

$$\begin{aligned}E_{i,t-1} \left[\frac{\exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t}) - f_i - X'_{i,t}\xi - \bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta}))}{1 + \exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t}) - f_i - X'_{i,t}\xi - \bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta}))} \right] \\ = E_{i,t-1} \left[\frac{\exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))}{1 + \exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))} \right].\end{aligned}\tag{6}$$

Since $\kappa_{i,t}$ is distributed independently across countries, the likelihood can be written as

$$\mathcal{L}(W^T|\boldsymbol{\theta}) \propto \prod_{t=1}^T \prod_{i=1}^n \mathcal{L}(D_{i,t}|X_{i,t}, W^{t-1}, \boldsymbol{\theta}),$$

where

$$\mathcal{L}(D_{i,t}|X_{i,t}, W^{t-1}, \boldsymbol{\theta}) = \Phi \left(\frac{\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})}{\varsigma_i} \right)^{D_{i,t}} \left[1 - \Phi \left(\frac{\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})}{\varsigma_i} \right) \right]^{1-D_{i,t}}$$

and Φ denotes the standard Normal cumulative distribution function (CDF).

Prior. Given the size of our model, we adopt an informative prior, π , to prevent overfitting. To do so in a principled manner, we calibrate our prior in the way agents in our model would, allowing the observed past to inform initial beliefs. We use data from 1875-1950 (excluding

the two world wars), a period that immediately precedes our main sample, to set the prior mean and precision of the model parameters so as to match analogous empirical moments. For example, we ensure that our prior over incumbents’ initial beliefs about the relationship between democracy and GDP growth is consistent with average annual growth rates among autocracies and democracies in the pre-sample period. Similarly, we use pre-sample history of elite turnover to inform our prior over the parameters describing the likelihood of retaining power. For a complete description of our prior and how it is calibrated, see Appendix A2.

Estimation and inference. Calculating $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ from (6) to evaluate the likelihood of the data is computationally expensive.¹⁷ To avoid this burden, we follow the Mathematical Programming with Equilibrium Constraints (MPEC) approach of Su and Judd (2012) to compute our maximum-a-posteriori estimator of $\boldsymbol{\theta}$. The idea behind this approach is simple: instead of calculating $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ at every trial value of $\boldsymbol{\theta}$, treat $\bar{\kappa}_{i,t}$ as an auxiliary parameter and impose (6), the optimality (or equilibrium) condition of the model, as a feasibility constraint on the log-posterior maximization program. Accordingly, we estimate $\boldsymbol{\theta}$ by solving $\max_{\boldsymbol{\theta}, \bar{\kappa}} \log(p(\boldsymbol{\theta}, \bar{\kappa}|W^T))$ subject to $\bar{\kappa}_{i,t}$ satisfies (6) for all i and t .¹⁸

As shown by Su and Judd (2012), MPEC and the standard approach of directly maximizing $\log(p(\boldsymbol{\theta}|W^T))$ yield identical estimates of $\boldsymbol{\theta}$. Computationally, MPEC’s advantage arises from the fact that modern optimization algorithms do not enforce constraints until the final iteration of the search process. Thus, the computationally expensive condition (6) is satisfied exactly once rather than at every trial value of $\boldsymbol{\theta}$. A potential disadvantage is that, by introducing $\bar{\kappa}$ as additional parameters, MPEC increases the size of the optimization problem. However, this concern is mitigated by the sparsity that results from each auxiliary parameter $\bar{\kappa}_{i,t}$ entering a single constraint.

Data. We describe our data sources in detail in Appendix A1.

3 Estimation Results: The Importance of Learning

Before summarizing our structural parameter estimates, we subject our model to a series of goodness-of-fit and out-of-sample tests that assess its ability to explain observed patterns of democracy adoption. We consider four alternative specifications of our model that differ in the number of covariates used to characterize the political cost of democracy. In our baseline

¹⁷We provide an iterative algorithm for computing $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ in Appendix A3.

¹⁸Standard errors are computed using an estimate of the inverse information matrix.

specification with no covariates, the political cost of democracy, $K_{i,t}$, consists of simply a country-specific baseline, f_i , plus an idiosyncratic shock, $\kappa_{i,t}$. We then consider specifications where we successively control for (lagged) log-GDP per capita, the incumbent’s time in power, and (lagged) trade volume as percentage of GDP. Across specifications, we use geographic distance between capitals, $Z_{i,j}$, to capture cross-country correlation in initial beliefs.¹⁹

To quantify the importance of learning for our model’s ability to fit the data, we also estimate a “no-learning” version of our model. For each specification, we constrain beliefs about long-run growth rates under autocracy and democracy to be constant over time, thus shutting down the learning mechanism. These no-learning specifications are otherwise identical to their learning counterparts.

We conduct our model performance tests as follows. With each estimated model, we compute one-year-ahead forecasts of the choice between autocracy and democracy for each country-year. That is, conditional on the state of the world at the end of year $t - 1$ as recorded in our data, we use each model to predict $D_{i,t}$ worldwide. We produce forecasts for the in-sample period used to estimate each model (1951-2000) and for 10 additional out-of-sample years (2001-2010).

In Figure 1, we plot the actual (gray) and predicted percentage of world democracies. In the top panel, predictions are generated using our baseline specification with no covariates. We plot predictions with (blue) and without (red) learning for both the in-sample (solid) and out-of-sample (dashed) periods. In the lower panel, we present the same set of estimates using our model with two covariates (lagged log-GDP per capita and incumbents’ time in power). Note the vast improvement in predictive success, both in and out of sample, when we account for learning. Unsurprisingly, the no-learning specification of our model with no covariates performs worst as it produces a constant prediction for each country.²⁰ However, while the inclusion of covariates does markedly improve the accuracy of the no-learning model, these gains pale in comparison to the role of learning. Indeed, our baseline learning model with no covariates vastly outperforms any specification that does not account for learning. As many of the covariates we condition upon are themselves outcomes of the selection process we model (e.g., per capita GDP or elite turnover), their inclusion should not yield much improvement in predictive success. Our results confirm this intuition.

In Figure 2, we present results disaggregated by four regions of the world: the Americas, Europe, Africa, and Asia-Oceania. We again compare the true proportion of democracies

¹⁹Geographic distance is highly correlated with other measures of cultural, economic, or political similarity between countries (Buera et al., 2011). Our results are virtually unchanged (available upon request) if we model the correlation in initial beliefs as a function of geographic and genetic distance between countries, the latter as measured by Spolaore and Wacziarg (2009).

²⁰The observed temporal variation is an artifact of the changing population of countries in our data.

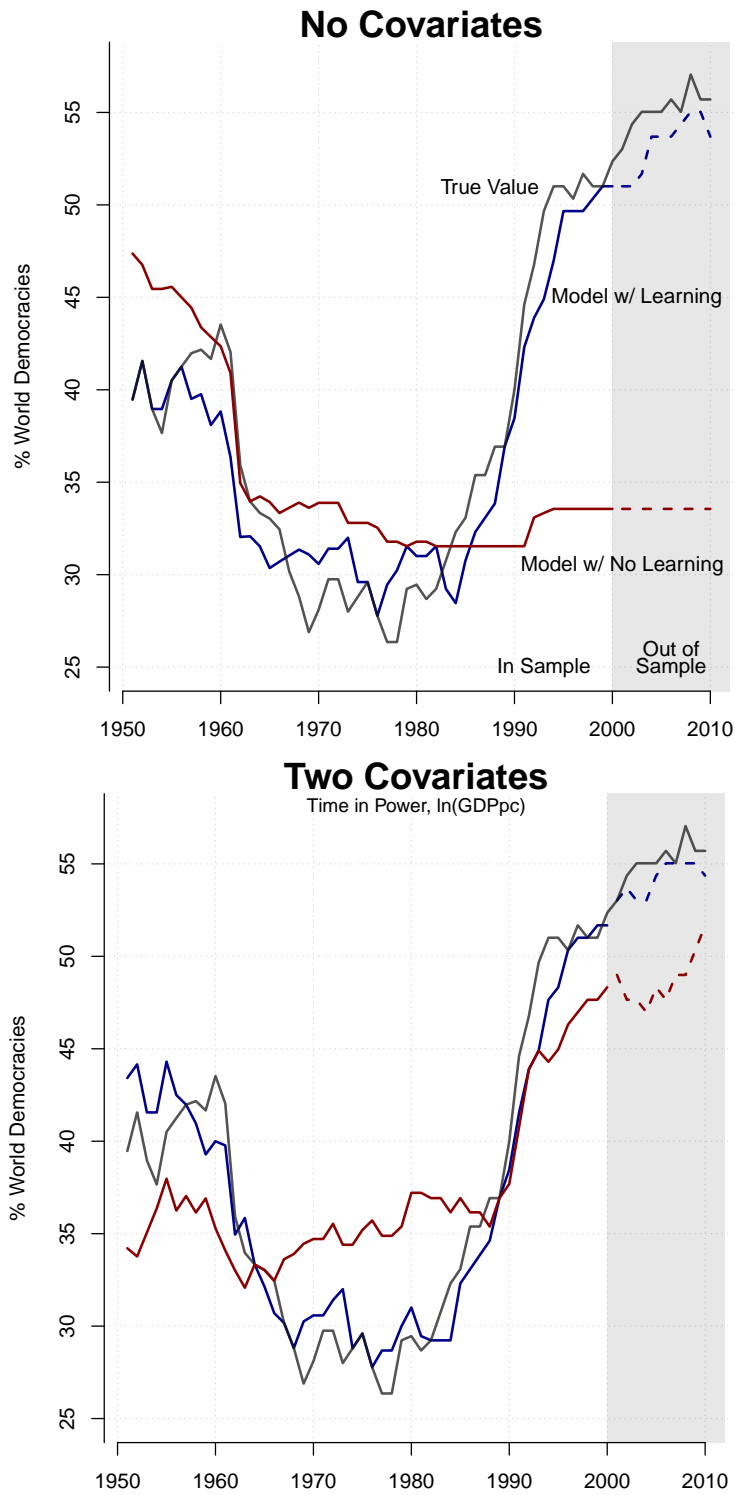


Figure 1: Observed versus Predicted Worldwide Prevalence of Democracy

Notes. This figure compares the true proportion of world democracies (gray) to in-sample (solid blue) and out-of-sample (dashed blue) estimates generated by our model. Additionally, we shut down learning in our model and produce both in-sample (solid red) and out-of sample (dashed red) predictions. In the top panel, estimates are generated using our baseline specification with no covariates. In the bottom panel, we control for lagged log-GDP per capita and incumbents' time in power.

(gray) in each region to our learning model’s predictions, both in (solid blue) and out of sample (dashed blue). The top panel presents results from our baseline model with no covariates; the bottom panel, from our model with two covariates. Both specifications perform well at this, or indeed any, level of geographic aggregation. And, while not included in Figure 2 to avoid clutter, our learning model still significantly outperforms any alternative that ignores the role of learning.

Table 2 provides a numerical summary of our goodness-of-fit tests. Each set of columns corresponds to a different model specification, with (odd columns) and without (even columns) learning. The first row gives the percentage of country-year observations each model correctly predicts. Unsurprisingly, all models perform remarkably well on this dimension. The reason is that, as transitions into or out of democracy are considerably rare (130 total in-sample events), country fixed effects go a long way in fitting the data. Indeed, our no-learning model with no covariates (second column), which produces a constant prediction for each country, has a success rate of almost 90%.

Table 2: Model Fit

	No Covariates		One Covariate		Two Covariates		Three Covariates	
	Learning	No Learning	Learning	No Learning	Learning	No Learning	Learning	No Learning
Choices (% correct)	95.2	88.9	95.7	90.5	95.9	92.0	94.8	91.9
Transitions (% correct)								
± 0 years	9.3	0.0	10.1	2.3	9.3	4.7	5.5	5.5
± 2 years	41.1	0.0	44.2	7.8	48.8	20.2	38.6	18.1
Log-likelihood	-581.4	-1,390.8	-572.5	-1,176.8	-553.6	-1,061.8	-523.6	-995.1
Observations	5,925	5,925	5,925	5,925	5,925	5,925	5,845	5,845

Notes. This table reports various goodness-of-fit measures. Each set of columns corresponds to a different specification of our model, with (odd columns) and without (even columns) learning. Models in the first two columns use only country fixed effects to characterize the political cost of democracy. Models in the third and fourth columns control for lagged log-GDP per capita. Models in the fifth and sixth columns additionally control for incumbents’ time in power. Models in the last two columns also control for trade volume as a percentage of lagged GDP. For each model, we report the percentage of correctly predicted in-sample system of government choices (first row). We similarly report the percentage of correctly predicted transitions to or from democracy within a 0-year window (second row) and a 5-year window (third row) of the event.

A much harder test—one that is considerably more revealing of the underlying causes of democracy—is whether a model can correctly predict transitions to and from democracy. In Table 2, we present two scenarios, assessing each model’s accuracy in predicting transitions within ± 0 years (second row) and ± 2 years (third row) of the event. Here, the importance of learning is striking. Models that do not account for learning perform quite poorly, matching

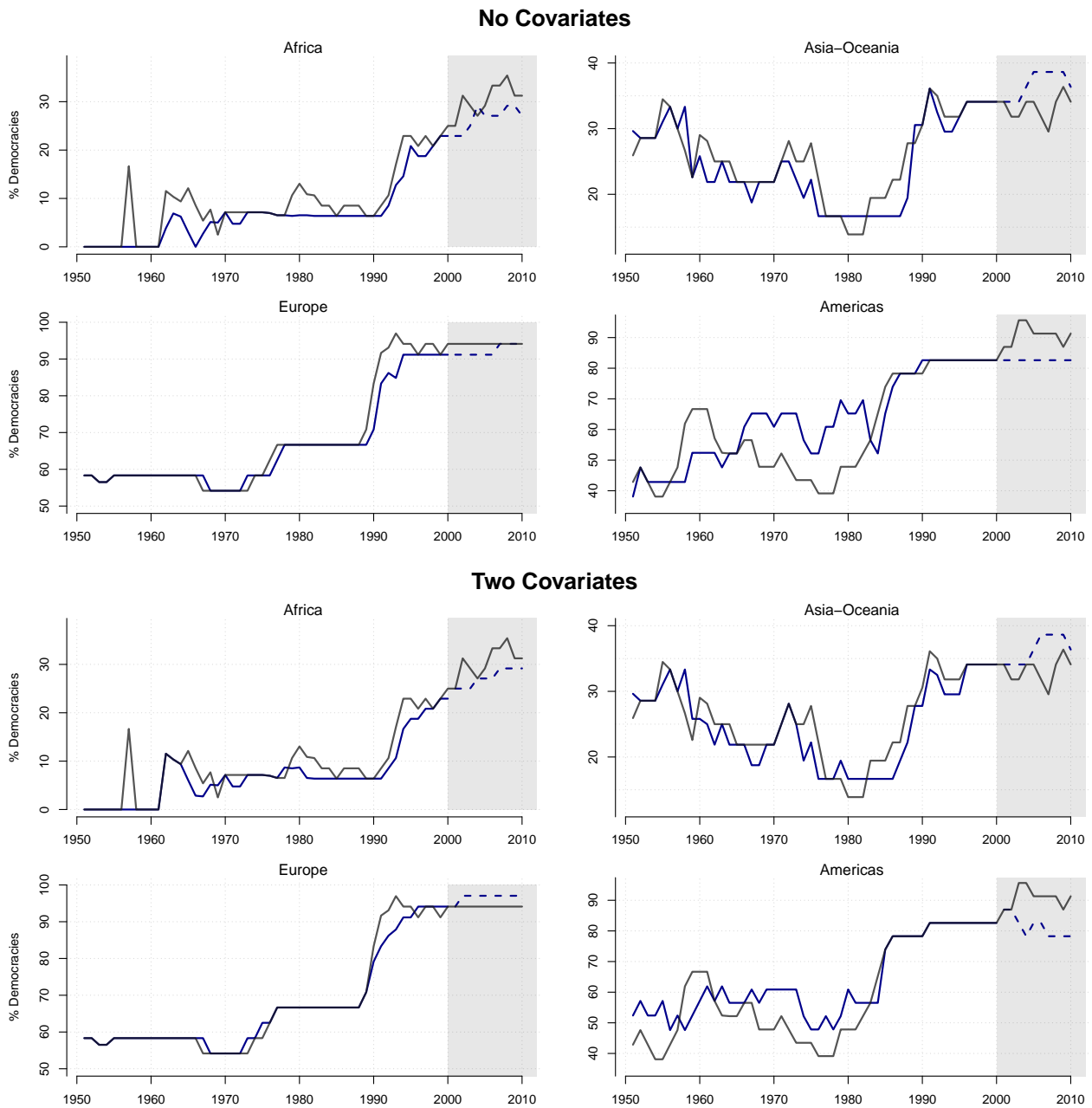


Figure 2: Observed versus Predicted Prevalence of Democracy by World Region

Notes. By world region, this figure compares the true proportion of democracies (gray) to estimates generated by our learning model for both the in-sample (solid blue) and out-of-sample (dashed blue) periods. The top panel presents results using our baseline specification with no covariates. In the lower panel, we control for lagged log-GDP per capita and incumbents' time in power.

no more than 20% of observed transitions, even within a 5-year window. And, while including additional covariates does increase accuracy, the marginal improvement is negligible. In contrast, turning on the learning mechanism in our model raises predictive success by over 100% in virtually all scenarios and all specifications. In fact, our baseline learning model with no covariates outperforms the best fitting no-learning specification by a similar rate.

To further benchmark our model, we present in Table 3 results from a series of panel regressions typical of the approach taken in the existing empirical literature on democratization. Using linear probability models, we regress our democracy measure against a full set of country fixed effects, a one-period lag of log-GDP per capita, and various controls. We exploit both annual data and, as in Acemoglu et al. (2008, 2009) and Boix (2011), five-year panels, which allow for the inclusion of covariates not available annually. As before, with each specification we produce predictions for every country-period.

Table 3: Goodness of Fit of Reduced-Form Models

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
Choices (% correct)	89.3	88.6	89.5	88.8	89.3	88.8	77.1	91.3	97.8	91.4	90.4	90.8	97.8	91.6
Transitions (% correct)														
±0 years	0.8		1.5		0.8				6.9		2.3		6.2	
±2 years	3.1	2.9	3.8	2.9	3.8	1.9	1.2	3.8	96.9	19.4	7.7	14.1	96.9	17.2
log(GDPpc) _{t-1}	0.115*** (0.010)	0.147*** (0.026)	0.116*** (0.010)	0.142*** (0.026)	0.112*** (0.011)	0.142*** (0.026)	0.033 (0.053)	0.032 (0.078)	0.024*** (0.006)	0.102*** (0.024)	0.060*** (0.014)	0.040 (0.037)	0.011 (0.008)	0.022 (0.034)
<i>Controls:</i>														
Country Effects	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Time in Power			X	X	X	X	X	X	X	X	X	X	X	X
Trade/GDP					X	X	X	X	X	X	X	X	X	X
Education							X	X						
Labor Share								X						
Democracy _{t-1}									X	X			X	X
Time Effects											X	X	X	X
Observations	5,925	1,076	5,925	1,076	5,866	1,076	649	390	5,866	1,076	5,866	1,076	5,866	1,076
Panel Length	1	5	1	5	1	5	5	5	1	5	1	5	1	5

*p<0.1, **p<0.05, ***p<0.01

Notes. This table gives linear probability estimates of the impact of (lagged) log-GDP per capita on democracy. All models include country fixed effects. We successively add in controls for the incumbent's time in power, trade as a percentage of GDP, average years of schooling from Barro (1999), and labor share of value added from Rodrik (1999). In columns XI-XIV, we include year fixed effects. In columns IX, X, XIII, and XIV, we include a one-period lag of democracy. We present estimates with annual and five-year panels. Row 1 gives the percentage of country-period observations correctly predicted by each model. For models with annual data, we also report the success rate at predicting transitions within ±0 (row 2) and ±2 (row 3) years of the event. For the five-year panels, we produce success rates at predicting transitions just within the five-year window (row 3).

Again, as with our model, it is trivial to correctly predict close to 90% of country-period observations using country fixed effects. The important departure arises when we compare

predictions of transitions derived from these reduced-form regressions to those generated by our learning model. Once more, we evaluate these predictions in exact (± 0) and five-year windows (± 2).

In terms of exact predictions, no reduced-form specification surpasses the predictive success of our baseline no-covariates learning model. With only two exceptions that include as a control a one-year lag of democracy, in the exact window our baseline model beats every reduced-form specification by well over 100%. In the five-year window, our model performs similarly well, again, save annual panel specifications with a lagged dependent variable, which correctly predict over 90% of transitions. However, aggregated over five years, the inclusion of the one-year lagged outcome leads any transition not picked up exactly by the model to be mechanically transformed into a correct prediction in subsequent years, thus yielding an extremely high predictive success rate in the five-year window and yet a low success rate within a single year. To further see this, note that, when we introduce a one-period lag in the five-year panel specifications, they successfully predict less than 20% of transitions, a rate which our model beats by more than 100%. More importantly, our model provides a structural interpretation for the observed persistence of systems of government that underpins the predictive success of these autoregressive specifications.

In sum, our goodness-of-fit tests quantify and highlight the crucial role of learning in explaining worldwide democratization events and reversals to authoritarianism, overshadowing the usefulness of other explanatory variables typically employed in the literature. As discussed, this is not surprising given that many of these controls are themselves outcomes of the learning process we model. In light of these results, we hereafter focus our attention on the baseline no-covariates specification of our learning model, using it to conduct our counterfactual experiments in Section 4.

3.1 Structural Parameter Estimates

To understand how our model is fitting the data, we summarize our main parameter estimates and discuss their implications. We begin with our estimates of the (de)stabilizing effect of GDP growth on elite turnover under autocracy and democracy, $\theta^{D=0}$ and $\theta^{D=1}$, respectively. This question, itself, has been a separate subject of close academic inquiry for decades.²¹ We find that the impact of growth on the likelihood that the incumbent group retains power indeed differs between autocracies and democracies.²² The quantitative implications of our estimates are summarized in Figure 3, which plots the estimated probability (averaged across countries) that the incumbent remains in power at different rates of GDP growth. In blue,

²¹For a summary of the early literature on the topic, see Przeworski et al. (2000) ch. 1.

²²Specifically, we estimate $\theta^{D=0} = -4.2213$ (p -value: 0.0000) and $\theta^{D=1} = 8.8279$ (p -value: 0.0000).

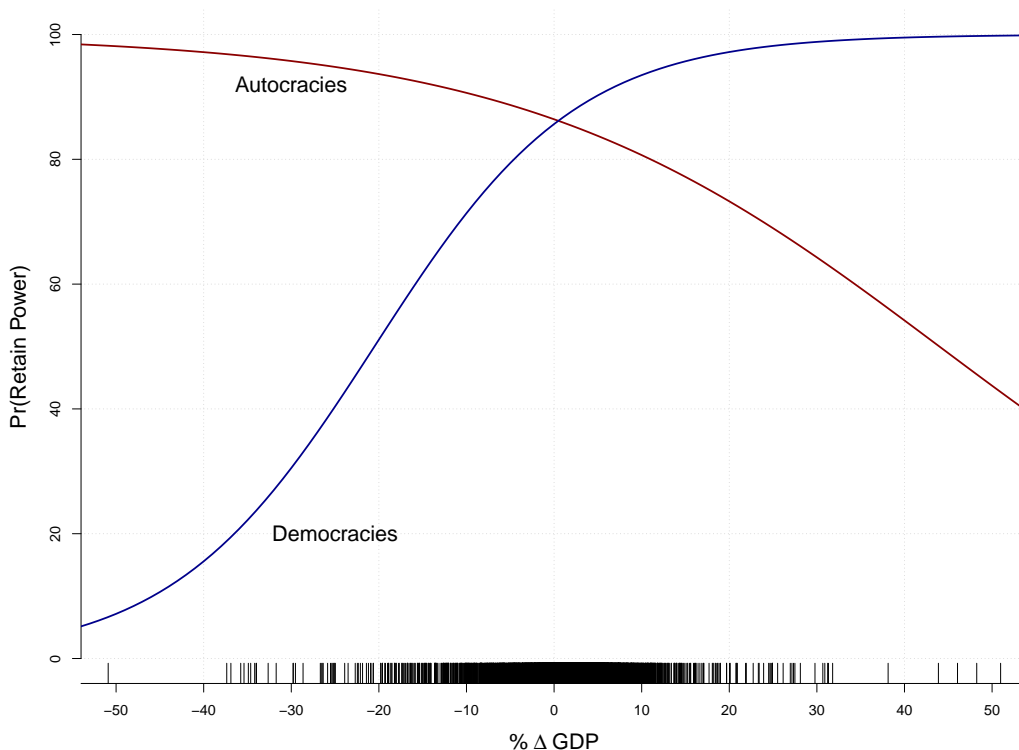


Figure 3: (De)Stabilizing Effect of GDP Growth on Elite Turnover

Notes. This figure plots the average (across countries) estimated probability of retaining power at different rates of GDP growth under autocracy (red) and democracy (blue). Observed growth rates between 1951-2000 are shown over the horizontal axis.

we plot our estimates under democracy and, in red, our estimates under autocracy.

In line with a substantial empirical literature in political economy, we find that, in democracies, economic growth is stabilizing for elites.²³ In other words, the party in government is more likely to win reelection when growth is high. In contrast, we find that growth is destabilizing in autocracies. This result comports with the view that rapid economic modernization in non-democracies creates inequalities of expectation or inequalities of outcomes that, in turn, engender attempts to subvert the political system.²⁴ That is, autocrats are less likely to remain in power when economic growth produces actors—a middle class, for example—able to place demands upon and challenge the authority of the group in power.

Coupling these results with learning helps explain the observed cross-sectional correlation between per capita GDP and democracy. Democracy becomes more appealing to incumbents

²³See, e.g., Nordhaus (1975); Hibbs (1977); Alesina et al. (1997); Drazen (2000); Brender and Drazen (2008).

²⁴See, e.g., Olson (1963); Huntington (1968); Hirschman and Rothschild (1973). The empirical evidence is mixed. For a prominent example that finds stabilizing effects of growth in non-democracies, see Londregan and Poole (1990).

as they come to believe that it is conducive to high rates of GDP growth. Conversely, autocracy entails an incentive to suppress economic growth.

Of course, there are notable exceptions. For example, over the past two decades China has experienced high rates of growth and nevertheless remained autocratic. Similarly, India for the first four decades of its independence experienced low rates of growth and yet remained democratic. Underlying these prominent cases is the country-specific political cost of democracy. We recover estimates of the structural parameters describing the baseline political cost of democracy for each country, f_i , and plot them in Figure 4. Notably, the Chinese Communist party faces, all else equal, the lowest probability of remaining in power under democracy. In contrast, Congress at India's independence had the fourth highest ex-ante probability of retaining power under democracy.

To understand the dynamics of democracy adoption, Figure 5 presents our estimates of the evolution of beliefs about the economic effects of democracy. The top panel plots the evolution of the worldwide distribution of the mean of beliefs about $\beta_i^{D=1} - \beta_i^{D=0}$, the difference in long-run GDP growth rates under democracy versus autocracy. The bottom panel shows the evolution of worldwide uncertainty (standard deviation) about these beliefs. The initial state of beliefs in 1951 simply reflects our prior calibration exercise. In the pre-sample period (1875-1950), democracies grew, on average, about 0.4% faster than autocracies. The median of mean initial beliefs in the top panel of the figure is consistent with this statistic.

While beliefs remain relatively flat for the first three decades of the in-sample period, in the 1980s and, even more dramatically, in the 1990s there is a sharp expansion of beliefs in favor of democracy's superior potential to foster economic growth. This change in worldwide beliefs, together with our estimates of the (de)stabilizing effects of GDP growth on elite turnover, helps explain our model's ability to correctly predict the striking rise in the percentage of world democracies observed in the same period (Figure 1). Both changes in the uncertainty and mean of beliefs play a role in explaining the increased prevalence of democracy in this period. By the start of the 1980s, the worldwide drop in belief uncertainty is sufficient to induce the first group of countries in this wave to democratize. Consequently, as democracies become less rare worldwide, countries sharply revise up their estimates of democracy's superior economic potential, leading to further democratization.

To conclude this section, Figure 6 presents our estimates of the spatial decay of learning, i.e., the extent to which the cross-country correlation of beliefs depends on the geographic distance between capitals. Consistent with the observed spatial clustering of democratization events noted in the literature, we find that learning is highly clustered geographically. As shown in Figure 6, at 5,000km (the approximate distance between the U.S. and Ecuador), the prior correlation of beliefs is only 0.12. At 10,000km (the distance between the U.S.

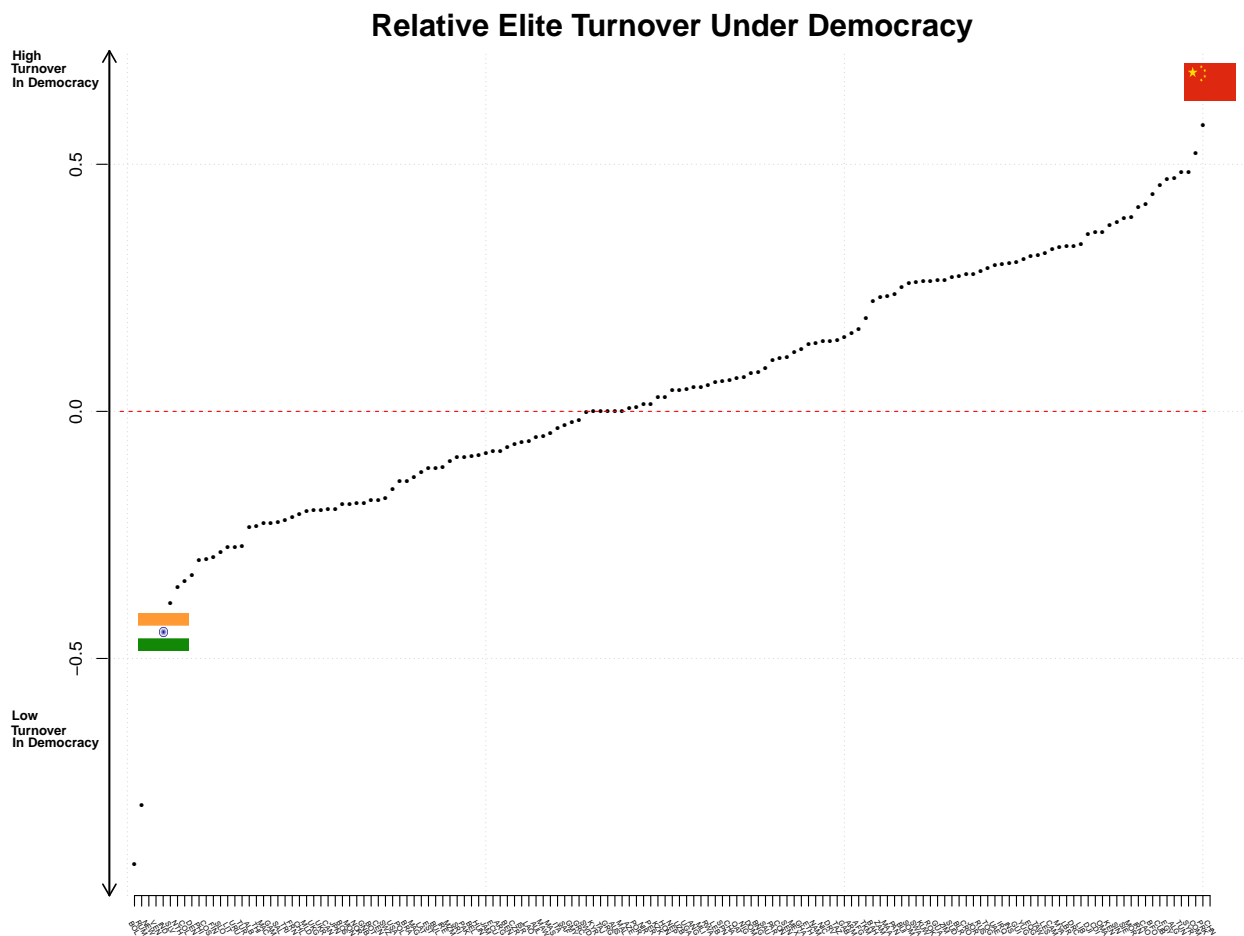


Figure 4: Political Cost of Democracy

Notes. This figure plots our estimates of the baseline political cost of democracy, f_i , across countries. Positive (negative) values imply higher (lower) elite turnover, all else equal, under democracy than autocracy.

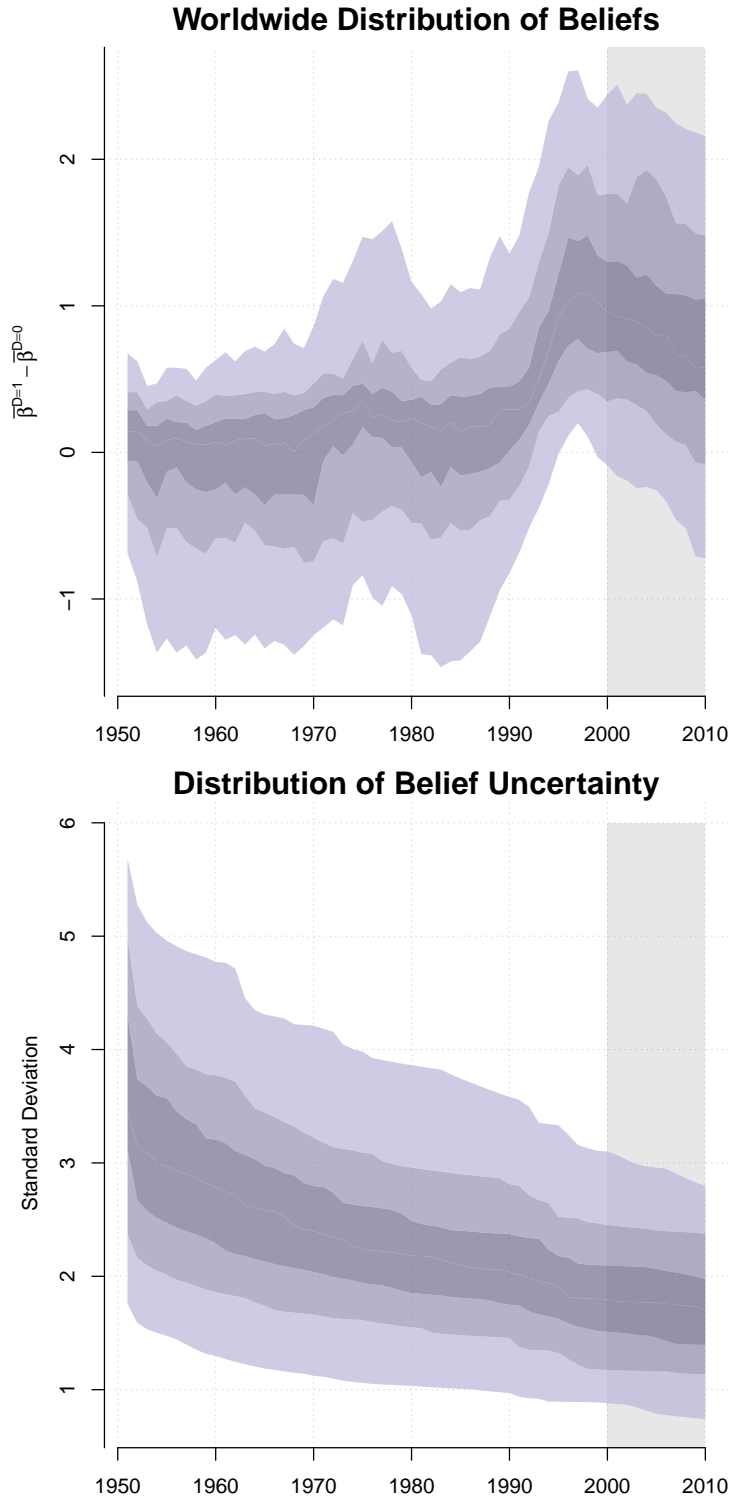


Figure 5: Worldwide Beliefs about Economic Effects of Democracy

Notes. The top panel of this figure plots estimates of the 20th-80th worldwide percentiles of the mean of beliefs about the percentage-point difference in long-run GDP growth rates under democracy versus autocracy. The lower panel shows estimates of the 20th-80th worldwide percentiles of uncertainty (standard deviation) in these beliefs.

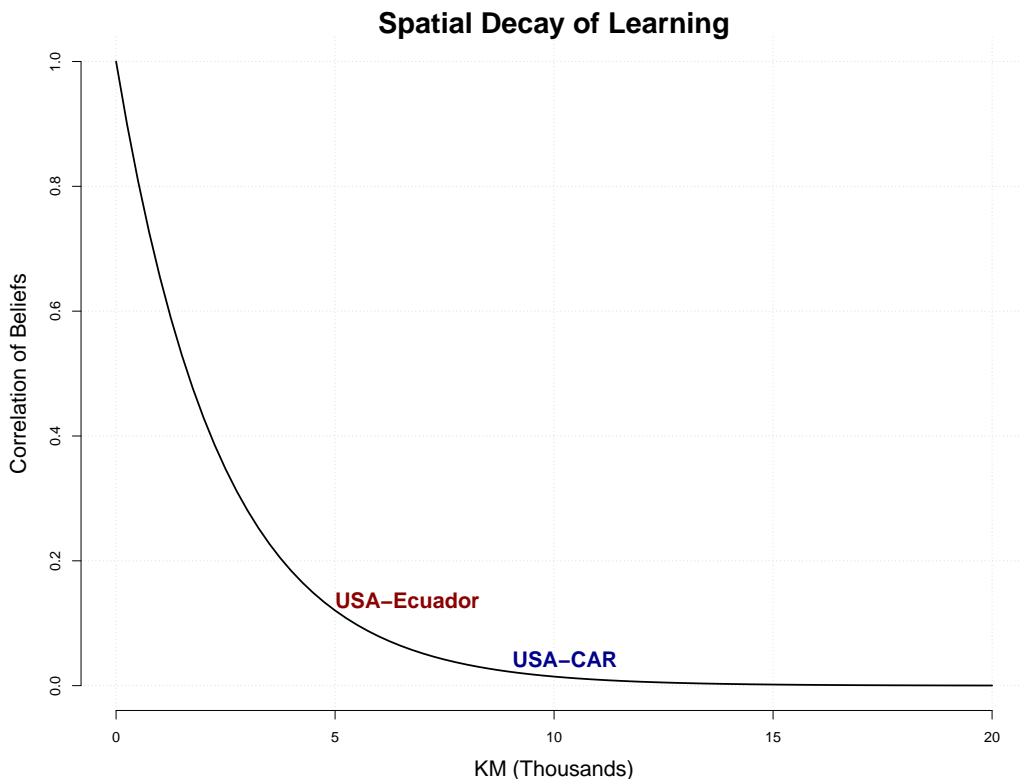


Figure 6: Cross-Country Correlation of Prior Beliefs

Notes. This figure shows the spatial decay of learning, plotting our estimate of the cross-country prior correlation of beliefs about the economic effects of democracy as a function of geographic distance between capitals (in thousands of kilometers). For reference, Ecuador and the Central African Republic are approximately five and ten thousand kilometers, respectively, from the U.S.

and the Central African Republic), the potential for learning between countries is virtually absent. This estimated feature of our model reveals that elites learn from the experiences of relatively proximate countries, and it underpins the ability of our model to explain the spatial clustering of transitions to and from democracy.

4 Counterfactuals

Our model allows us to explore counterfactual experiments of three types. First, it enables us to understand how systemic shocks to prosperity impact the worldwide prospect for democracy. Second, it allows us to ask retrospective questions about historical democratization events. Third, our model allows us to prospectively explore conditions that would lead current elites to transition to or from democracy. We present results of each type in turn.

To conduct these experiments, it is necessary to specify and estimate the “true” data-generating process. A considerable advantage of using our baseline no-covariates model to

generate these counterfactuals is that only an estimate of the true relationship between GDP growth and democracy is required. Following Buera et al. (2011), we assume that this relationship is described by a hierarchical linear model similar to (2), which we estimate using all available data between 1875-2000 (excluding the two world wars).²⁵ This specification is appealing for its flexibility and because it is consistent with our model of learning in that elites with beliefs (2) and prior (3) would eventually learn the truth over time.

4.1 A Second Great Depression

In the year before the market crash of 1929, 51% of the world’s independent states were democracies. A year later this proportion dropped to just over 43%. By 1935, the fraction of countries that remained democratic decreased by another 5%, reaching a low of 36% by 1938. In Europe, democratic backsliding was even starker. At its theretofore high in 1920, twenty-six out of twenty-eight European states were democratic, but by 1938 thirteen of these countries had transitioned away from democracy.

A substantial body of both historical and quantitative research has linked the global decline of democracy in the inter-war period directly to the economic downturn of the Great Depression.²⁶ Likewise, both in recent academic and popular discourse, the global economic recession of 2008 has been put forward as a contributing factor in the observed wave of recent democratic breakdowns.²⁷ Next, we provide evidence that these systemic economic crises indeed engender reversals to autocracy and, moreover, highlight how this is driven by changes in beliefs about the economic effects of democracy.

To that end, we simulate two sorts of crises. First, we generate a “short-deep” counterfactual crisis where, for our last in-sample year (2000), we simulate a 5.9% average worldwide contraction of per capita GDP, comparable to the worst year of the Great Depression (1931). In our second crisis, we construct a “long-shallow” counterfactual condition where we perturb growth by a smaller amount—1.7% annually (the average contraction between 1929-1933)—but extend this contraction over a five-year period (2000-2004). We present results where we concentrate these counterfactual conditions in autocratic and democratic countries, respec-

²⁵Specifically, we assume that

$$y_{i,t} = (1 - D_{i,t})b_i^{D=0} + D_{i,t}b_i^{D=1} + e_{i,t},$$

$$e_t \sim N(0, S \cdot Q \cdot S), \quad b \sim N(\bar{b}, I_2 \otimes S \cdot W \cdot S),$$

where S is a diagonal matrix, $Q_{i,j} = \exp(-Z_{i,j}\zeta_Q)$, and $W_{i,j} = \lambda \exp(-Z_{i,j}\zeta_W)$.

²⁶See, for example, Frey and Weck (1983); Payne (1996); De Bromhead et al. (2013); King et al. (2008); Berg-Schlosser and Mitchell (2000, 2002).

²⁷For scholarly work on the subject, see Bartels (2013); Bartels and Bermeo (2013); Armingeon and Guthmann (2014). For journalistic accounts, see, e.g., Friedman (5-7-2008); Douthat (6-15-2009); Krugman (12-11-2011).

tively. Under the “autocratic bias” condition, recessions are twice as deep in autocracies as in democracies, while keeping the worldwide average contraction consistent with our intervention. Conversely, under the “democratic bias” condition, recessions are twice as deep in democracies.

Model estimates of the worldwide percentage of democracies are shown in the top half of Figure 7. In the left-hand panel, we present estimates from the autocratic bias condition and, in the right-hand panel, estimates from the democratic bias condition. In both plots, the short-deep counterfactual is shown in green, and the long-shallow counterfactual is shown in purple. Note that both the short-deep and long-shallow counterfactual crises negatively impact the worldwide prevalence of democracy. While the large single-period decline in growth has a larger impact in the first year, thereafter the proportion of democracies begins to recover. On the other hand, the smaller but longer crisis has a larger overall impact, with the proportion of democracies continuing to decline through the duration of the economic contraction and recovering more slowly.

Both patterns are consistent with our actors learning about the economic effects of democracy. In both counterfactuals, the initial shock forces agents to revise their beliefs. However, in the long-shallow counterfactual, as would be expected from a continued process of learning, our agents update their beliefs following each additional negative perturbation of worldwide growth and continue to select out of democracy accordingly. In contrast, in the short-deep scenario, we observe a single large drop in the percentage of democracies. Since after the initial shock to growth there is no “new” information revealed to our agents, there is little additional updating and the worldwide percentage of democracies starts to recover.

To better see this, we present estimates of the effect size of each crisis in the lower half of Figure 7. Magnitudes are calculated by subtracting our model’s counterfactual predictions from the observed true values. Again, the reduction in world democracies following the short-deep crisis is immediate, but there is a slow recovery. In the long-shallow crisis the decline in the worldwide prevalence of democracy worsens over time. Importantly, comparing effect magnitudes across the autocratic and democratic bias conditions, it is clear that the reduction in world democracies is larger when the economic contraction is concentrated among democracies. This is consistent with the evolution of beliefs prior to our intervention, as discussed in Section 3. When democracies perform poorly, counter to the prevailing consensus, elites sharply revise their beliefs and select out of democracy.

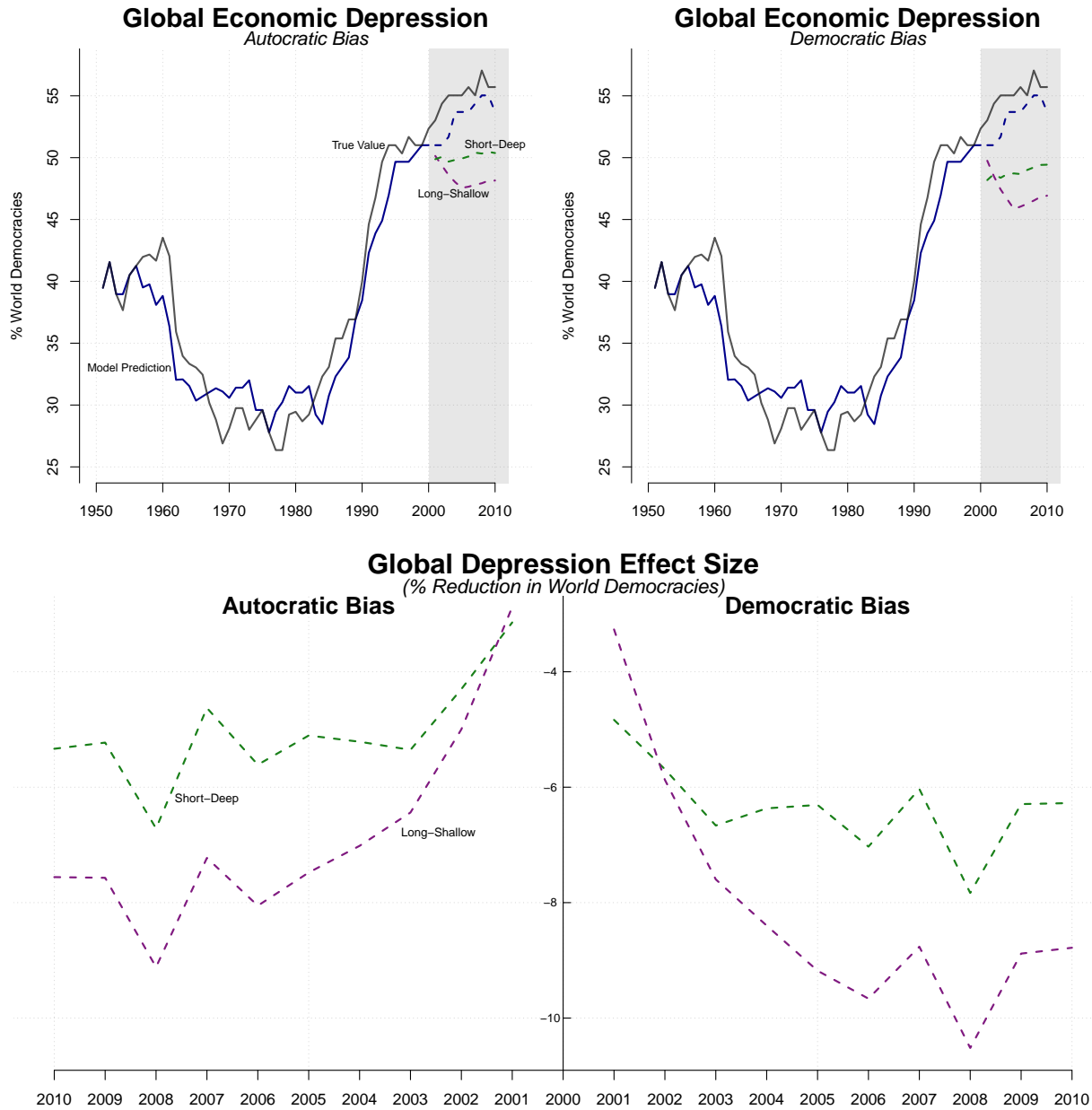


Figure 7: Second Great Depression Counterfactuals

Notes. This figure reports results from two counterfactual exercises. First, we present predictions following a short-deep crisis of a single year (2000) of 6% average worldwide economic contraction (green). Second, we present predictions based on a long-shallow crisis of 5 years (2000-2004) of 2% average contraction (purple). In the top half of the figure, we plot the true percentage of world democracies (gray) as well as baseline model predictions (blue). We concentrate the contraction in autocracies in the left-hand panel, and in democracies in the right-hand panel. The lower half of the figure gives effect-size estimates for each counterfactual equal to the difference between the true proportion of world democracies and our counterfactual estimates.

4.2 The Third Wave

A number of studies highlight the influence of external actors on the prospects for democracy.²⁸ Especially for the early “third-wave” democratization events in Greece, Portugal, and Spain, the potential for accession to the European Community (EC) has been put forward as a crucial determinant of their respective transitions.²⁹ Brussels’ requirement that community members maintain a form of government consistent with liberal democracy coupled with the economic benefits of access to the common market generated an incentive to democratize. In this section we show that, rather than serving as an institutional target, much of the impact the EC had upon third-wave democratization operated through its constituent states’ economic performance, which affected beliefs about the economic effects of democracy in potential member states.

To show this, we construct a counterfactual wherein we generate a recession in the EC’s three largest economies, Britain, France, and Germany, of 2% average annual contraction for the two years preceding the first transition of Greece in 1974.³⁰ Then, to obtain an estimate of this counterfactual recession’s impact on the transitions of Greece, Portugal, and Spain, we compare our predicted transitions in this counterfactual world to the truth. Our results are given in Figure 8.

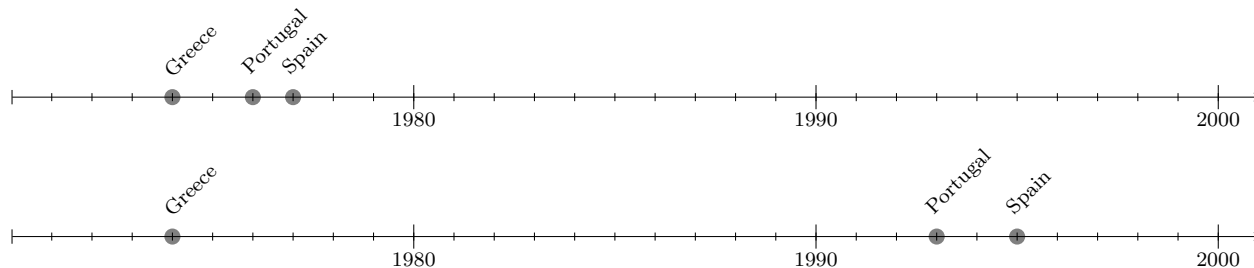


Figure 8: Counterfactual Third Wave

Notes. The top panel shows the true timeline of democratization of Greece, Portugal, and Spain. In the bottom panel, we present this timeline under our counterfactual scenario of a two-year recession (1972 and 1973) in Britain, France, and Germany of 2% average annual contraction.

For Portugal and Spain, the effect of lower growth in Britain, France, and Germany is considerable, delaying their transitions to democracy by almost 20 years. In contrast, we find no effect of this recession on Greece’s transition. The reason for this becomes apparent once we compare the evolution of beliefs in our three example countries. We plot in Figure 9, for each of these cases, our estimates of beliefs under the observed economic conditions (solid)

²⁸See, for example, Pevehouse (2002, 2005).

²⁹Pridham (1991); Whitehead (1996); Powell (1996)

³⁰Recessions in our sample last 2 years on average, with an average 2% drop in per capita GDP.

and under our counterfactual timeline (dashed). Note that for Portugal and Spain there is a substantial divergence in beliefs between the observed and counterfactual timelines. Following our counterfactual recession, Portuguese and Spanish beliefs become markedly less favorable towards the potential for growth under democracy. Notably, in both of these cases, beliefs also become increasingly uncertain (right-hand panel of Figure 9). In contrast, for Greece this is not the case: there is no substantial difference in their beliefs about the economic effects of democracy, and uncertainty in their beliefs continues to decline over time. The reason for this, according to our model, is that Greek elites pay little attention to the large, relatively distant Western economies we used to construct our counterfactual scenario. Britain, France, and Germany are simply too different from Greece to be used as a reference for learning.

Ultimately, our counterfactual experiment suggests that the distinguishing characteristic of Portugal and Spain in contrast to, for example, Brazil or Mexico is not their underlying propensity for democracy. Indeed, in terms of their estimated baseline political cost of democracy, Portugal and Brazil and Spain and Mexico are very close.³¹ Rather, Portugal and Spain democratized early because they learned from Western Europe, benefiting from proximity to successful liberal democracies.

4.3 Chinese (and North Korean) Democracy

With an eye to contemporary politics, we evaluate the stability of a pair of geopolitically important autocracies. We explore conditions under which our model predicts that China and North Korea would democratize. We focus first on the Chinese case. Here, we look for the minimal average growth rate among China’s democratic neighbors, in a five-year economic expansion, that would result in transitions of two types.³² First, we find the rate of growth that would result in at least a single year of democracy. Second, we establish the growth rate that would deliver a permanent transition to democracy.

To obtain a predicted single-year Chinese transition to democracy, we estimate that it would take five years (2000-2004) of 6.5% average annual growth in China’s democratic neighbors. After this single year of democracy (2005), our model predicts an immediate reversal to autocracy. To obtain a “permanent” transition—that is, a prediction of democracy until the end of our sample—we estimate that China’s democratic neighbors would have to grow at an average annual rate of 11% between 2000 and 2004. In contrast, under the same

³¹The baseline political cost of democracy for Portugal and Brazil is estimated at 0.01 and -0.14, and in terms of rank order they are 71 and 39, respectively. Estimates for Spain and Mexico are 0.06 and 0.12, yielding a rank order of 83 and 93, respectively.

³²Economic expansions in our sample last five years on average.

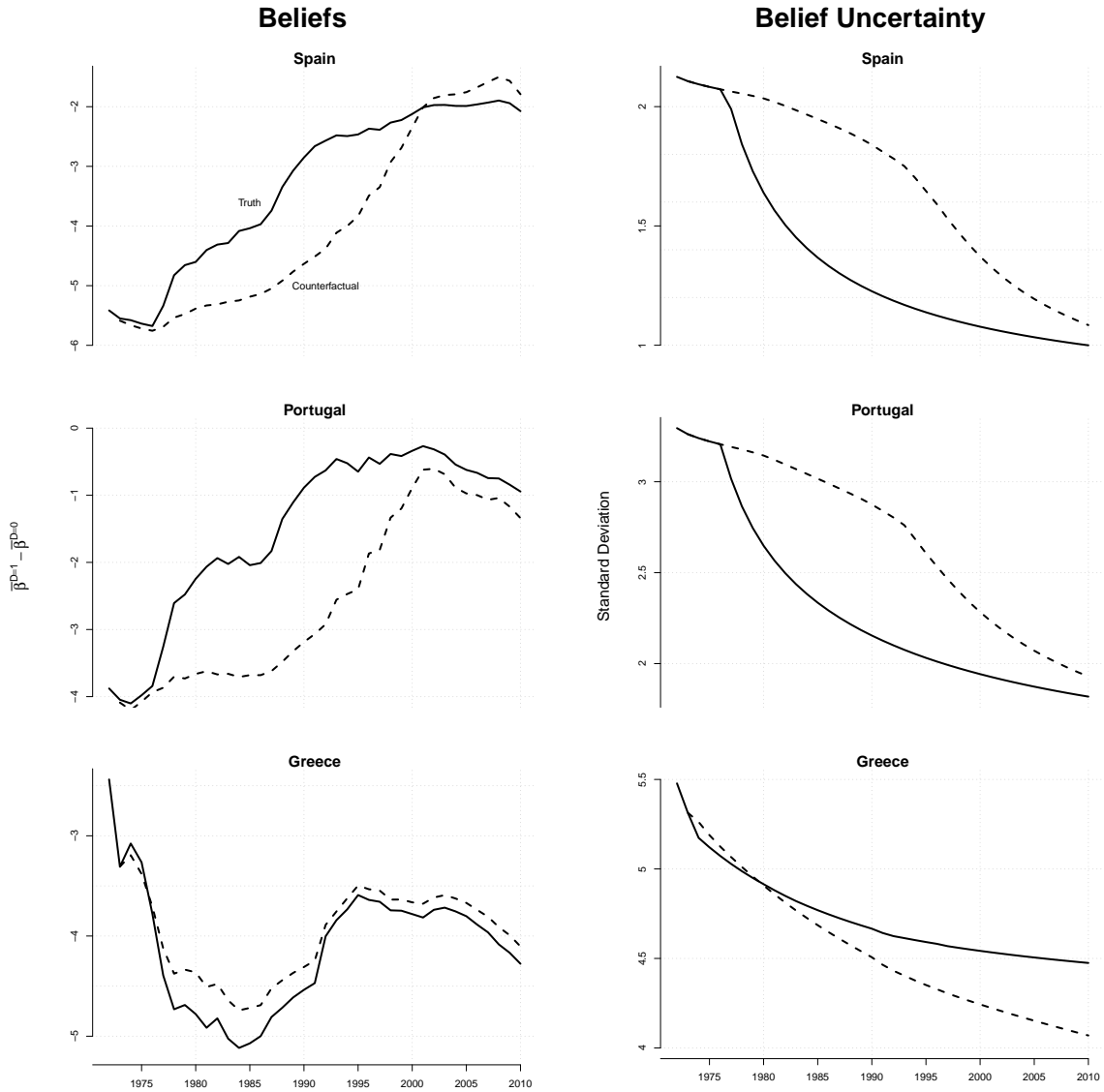


Figure 9: Counterfactual Third Wave Beliefs

Notes In the left-hand portion of this figure, we plot estimates of Spain, Portugal, and Greece's mean beliefs under the true growth rates (solid) and under a counterfactual two-year recession (1973 and 1974) in Britain, France, and Germany of 2% average contraction (dashed). In the right hand panel, we plot changes in each country's belief uncertainty under the same true and counterfactual conditions.

set of counterfactual conditions, North Korea does not democratize for any period. North Korea would democratize for a single year following five years of 16.5% average growth in its democratic neighbors, and permanently following five years of 20.5% average growth.

To highlight the differences in learning, in Figure 10 we plot Chinese and North Korean beliefs under the “Chinese democracy” scenarios of 6.5% and 11% average growth in their democratic neighbors. For China, our intervention increases the perception that democracy outperforms autocracy. In contrast, for North Korea, we observe a substantially smaller change in beliefs about the economic effects of democracy. Furthermore, in the lower panel we see that both counterfactual expansions, though slightly, reduce uncertainty in Chinese beliefs. In North Korea, there is virtually no change in the precision of beliefs.

These results suggest that the prospects for Chinese and North Korean democracy are limited. It would take a strikingly large economic boom in Asian democracies for China to democratize, and an even larger and implausible boom to generate the same outcome in North Korea.

5 Conclusion

In this paper, we propose and estimate a learning model of democratization. We show that worldwide learning about the economic consequences of democracy is crucial to explaining observed patterns of economic growth and democracy adoption. In particular, our model jointly rationalizes the cross-sectional correlation between income and democracy and the clustering of democratic transitions.

Our parameter estimates indicate that, for office-holding elites, growth is stabilizing in democracies and destabilizing in autocracies. As such, autocracy entails an incentive to suppress economic growth, and democracy the converse. When democratic incumbents come to believe that democracy does not produce sufficient growth to win a fair election, they will subvert popular rule. Conversely, autocratic incumbents will liberalize if they expect sufficiently fast subsequent economic development.

Furthermore, we show that learning is highly circumscribed geographically. Our model indicates that countries only update their beliefs using their most proximate neighbors as references. Learning dissipates rapidly outside of a 5,000 kilometer radius. As a consequence, we provide a structural explanation for the regional clustering of transitions to and from democracy.

In combination, these features of our model allow us to successfully predict, both in (1951-2000) and out of sample (2001-2010), much of the observed variation in democracy adoption. Moreover, our approach allows to evaluate the prospects for democracy. Rather

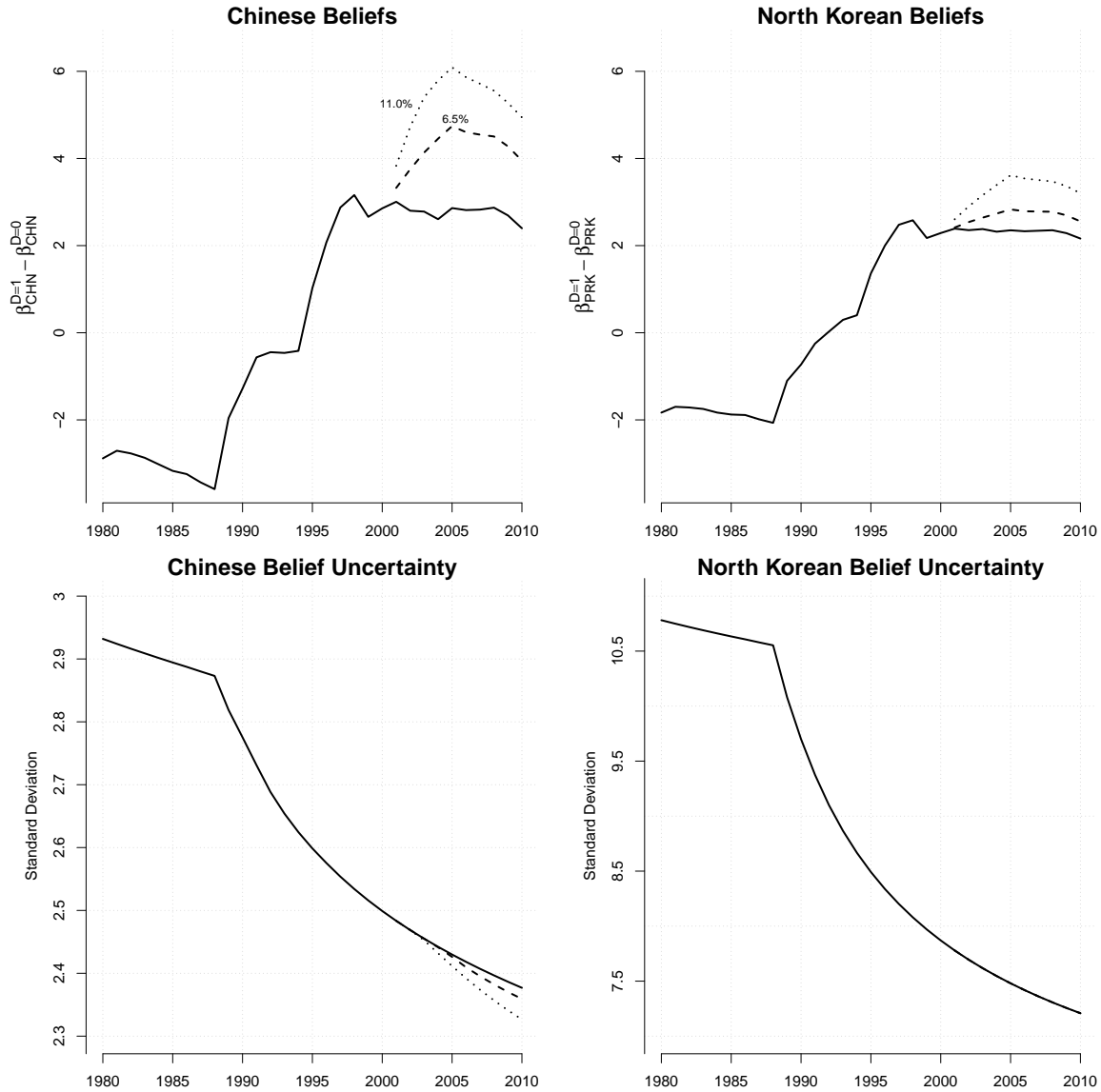


Figure 10: Chinese and North Korean Beliefs

Notes. In the top portion of this figure, we plot estimates of Chinese and North Korean beliefs about the economic effects of democracy under observed growth (solid) and under two counterfactual five-year expansions (2000-2004) in their democratic neighbors of 6.5% (dashed) and 11% (dots) average annual growth. In the lower panel, we plot changes in each country's belief uncertainty under the same true and counterfactual conditions.

than an “end of history,” we show that democracy is only as resilient as the economic performance it engenders. Even in the 1990s, when the success of democratic systems made such proclamations seem reasonable, we find substantial variation in the worldwide distribution of beliefs about democracy’s impact on growth and, moreover, substantial uncertainty in these beliefs. As recent history suggests, and as our counterfactual experiments demonstrate, systemic economic crises, particularly those concentrated in democracies, have the potential to generate waves of autocratic reversals.

Appendix

A1 Data

<i>Variable</i>	<i>Source</i>
Per Capita GDP	Maddison (2010)
Democracy	Boix et al. (2013)
Executive Faction	Goemans et al. (2009)
Trade Share	Gleditsch (2002)

A2 Prior

We set our prior over the model parameters in Table 1 as follows. We assume that

$$\begin{aligned}\alpha_i &\stackrel{\text{i.i.d.}}{\sim} N(\underline{\alpha}, \omega_\alpha^2), \\ \theta^{D=0,1} &\stackrel{\text{i.i.d.}}{\sim} N(\underline{\theta}, \omega_\theta^2), \\ \bar{\beta}_{i,0}^{D=0} &\stackrel{\text{i.i.d.}}{\sim} N(\underline{\beta}_0^{D=0}, \omega_\beta^2), \\ \bar{\beta}_{i,0}^{D=1} &\stackrel{\text{i.i.d.}}{\sim} N(\underline{\beta}_0^{D=1}, \omega_\beta^2), \\ v_i &\stackrel{\text{i.i.d.}}{\sim} \text{IG}(s_v, d_v), \\ f_i &\stackrel{\text{i.i.d.}}{\sim} N(\underline{f}, \omega_f^2), \\ \varsigma_i &\stackrel{\text{i.i.d.}}{\sim} \text{IG}(s_\varsigma, d_\varsigma), \\ \gamma &\sim \text{Uniform}, \\ \xi &\sim \text{Uniform},\end{aligned}$$

where $IG(s, d)$ denotes the Inverse-Gamma distribution with shape parameter s and rate parameter d . We calibrate our prior using pre-sample data from 1875-1950 (excluding the two world wars):

- We set $\underline{\beta}_0^{D=0} = 0.0180$ and $\underline{\beta}_0^{D=1} = 0.0218$, the average annual growth rates among autocracies and democracies, respectively, in the pre-sample period. We then set $\omega_\beta = 0.02$, allowing for considerable uncertainty about the mean of initial beliefs.
- We select $s_v = 3$ and $d_v = 0.7423$ so that the prior mean and standard deviation of $v_i\sigma_i$ equal the standard deviation of average growth rates, \bar{y}_i , in the pre-sample period. A pre-sample estimate of the mean of σ_i (equal to 0.0531) is obtained from the residuals of a regression of GDP growth on country and time fixed effects. We then set the prior mean of v_i equal to $\sqrt{\text{Var}(\bar{y}_i)}/0.0531 = \sqrt{0.0004}/0.0531 = 0.3711$.
- We set $\underline{\theta} = 0$ to adopt an agnostic starting point about whether GDP growth has a stabilizing or destabilizing effect on elite turnover across forms of government, and we normalize $\omega_\theta = 1$.
- To adopt an agnostic starting point regarding the political cost of democracy, we set $\underline{f} = 0$. We describe our choice of ω_f below.
- To ensure prior correlations of beliefs between 0 and 1, we adopt a flat (improper) prior over $\gamma \geq 0$. For the political cost of democracy, we center the variables in $X_{i,t}$ around their sample means so that $K_{i,t}$ has an expected value of zero (in line with our agnostic view of f_i), and we adopt a flat (improper) prior over ξ .
- Letting $\bar{\theta y}_i$ and \overline{KD}_i denote the within-country means of $\theta^{D=D_{i,t}}y_{i,t}$ and $K_{i,t}D_{i,t}$, respectively, and noting that $\alpha_i + \bar{\theta y}_i - \overline{KD}_i$ approximately equals the log-odds of staying in power in country i , we select $\underline{\alpha}$, ω_α , and ω_f to match the first two moments of these log-odds across countries in the pre-sample period. Since $E(\alpha_i + \bar{\theta y}_i - \overline{KD}_i) = \underline{\alpha}$, we

set $\underline{\alpha} = 1.8432$, the average log-odds in the pre-sample period.³³ Noting that the variance of the log-odds among autocracies is approximately equal to $\omega_\alpha^2 + \text{Var}(\bar{y}_i)$, while the variance among democracies is approximately equal to $\omega_\alpha^2 + \text{Var}(\bar{y}_i) + \omega_f^2$, we set $\omega_\alpha^2 + 0.0005 = 0.722$ and $\omega_\alpha^2 + 0.0007 + \omega_f^2 = 1.0802$, so $\omega_\alpha = 0.8494$ and $\omega_f = 0.5984$.

- Finally, to discourage the model from fitting the data with large (absolute) realizations of the unobserved political cost shock $\kappa_{i,t}$, we set $s_\varsigma = 3$ and $d_\varsigma = 0.2992$ so that ς_i has a prior mean and standard deviation of $\omega_f/4 = 0.1496$.

A3 Computing $\bar{\kappa}_{i,t}(\cdot)$: A Fixed-Point Algorithm

We provide an algorithm for computing $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ from (6). To simplify notation, let $V_0(\beta_i^{D=0}, \epsilon_{i,t}) \equiv \alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t})$ and $V_1(\beta_i^{D=1}, \epsilon_{i,t}) \equiv \alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t}) - f_i - X'_{i,t}\xi$.

Define $g : \mathbb{R} \rightarrow \mathbb{R}$ by

$$g(\kappa) \equiv E_{i,t-1} \left[\frac{\exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)}{1 + \exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)} \right].$$

Differentiating under the integral, we have

$$g'(\kappa) = -E_{i,t-1} \left[\frac{\exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)}{1 + \exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)} \left(1 - \frac{\exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)}{1 + \exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)} \right) \right] < 0$$

and $-g'(\kappa) < g(\kappa)$ as $\frac{\exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)}{1 + \exp(V_1(\beta_i^{D=1}, \epsilon_{i,t}) - \kappa)} \in (0, 1)$. Now, define $h : \mathbb{R} \rightarrow \mathbb{R}$ by

$$h(\kappa) \equiv \kappa + \log(g(\kappa)) - \log(U_0),$$

³³For countries that experienced no elite turnover in the pre-sample period, we limit the probability of staying in power to equal the maximum among countries with turnover (0.95).

where $U_0 \equiv E_{i,t-1} \left[\frac{\exp(V_0(\beta_i^{D=0}, \epsilon_{i,t}))}{1 + \exp(V_0(\beta_i^{D=0}, \epsilon_{i,t}))} \right]$. Notice that $h'(\kappa) = 1 + \frac{g'(\kappa)}{g(\kappa)} \in (0, 1)$, which implies that, for any $\kappa', \kappa \in \mathbb{R}$,

$$|h(\kappa') - h(\kappa)| = \left| \int_{\kappa}^{\kappa'} h'(k) dk \right| < |\kappa' - \kappa|.$$

Thus, h is a contraction. By the Contraction Mapping Theorem, h has a unique fixed point, which by construction is equal to $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ and can be computed iteratively as follows:

Step 0. Initialize $\kappa_0 = -f_i - X'_{i,t}\xi$ and set $l = 1$.

Step l . Set $\kappa_l = h(\kappa_{l-1})$. If $\kappa_l = \kappa_{l-1}$, stop; otherwise, continue to step $l + 1$.

As logarithms are computationally more expensive than exponentials, $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ can be computed more efficiently using an exponential version of the algorithm.³⁴

³⁴Iteratively evaluate $h_e(w) \equiv w \left(\frac{g_e(w)}{U_0} \right)$, where $g_e(w) \equiv E_{i,t-1} \left[\frac{\exp(V_1(\beta_i^{D=1}, \epsilon_{i,t})/w)}{1 + \exp(V_1(\beta_i^{D=1}, \epsilon_{i,t})/w)} \right]$, and compute $\bar{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \boldsymbol{\theta})$ as the natural logarithm of the unique fixed point of h_e . The starting point for the algorithm in this case should be constrained to satisfy $g_e(w_0)^2 < U_0$.

References

- Acemoglu, Daron, Simon Johnson, James A Robinson, and Pierre Yared**, “Income and democracy,” *The American Economic Review*, 2008, 98 (3), 808–842.
- , – , – , and – , “Reevaluating the modernization hypothesis,” *Journal of Monetary Economics*, 2009, 56 (8), 1043–1058.
- Ahlquist, John S and Erik Wibbels**, “Riding the Wave: World Trade and Factor-Based Models of Democratization,” *American Journal of Political Science*, 2012, 56 (2), 447–464.
- Alesina, Alberto and Nicola Fuchs-Schündeln**, “Good-Bye Lenin (or Not?): The Effect of Communism on People’s Preferences,” *American Economic Review*, 2007, 97 (4), 1507–1528.
- , **Nouriel Roubini, and Gerald D Cohen**, *Political cycles and the macroeconomy*, MIT press, 1997.
- Almond, Gabriel Abraham and Sidney Verba**, *The civic culture: Political attitudes and democracy in five nations*, Princeton University Press, 1963.
- Armingeon, Klaus and Kai Guthmann**, “Democracy in crisis? The declining support for national democracy in European countries, 2007–2011,” *European Journal of Political Research*, 2014, 53 (3), 423–442.
- Barro, Robert J**, “Democracy and growth,” *Journal of economic growth*, 1996, 1 (1), 1–27.
- , “Determinants of democracy,” *Journal of Political economy*, 1999, 107 (S6), S158–S183.
- Bartels, Larry and Nancy Bermeo**, *Mass politics in tough times: Opinions, votes and protest in the great recession*, Oxford University Press, 2013.
- Bartels, Larry M**, “Political effects of the Great Recession,” *The Annals of the American Academy of Political and Social Science*, 2013, 650 (1), 47–76.
- Basora, Adrian A**, “Must Democracy Continue to Retreat in Postcommunist Europe and Eurasia?,” *Orbis*, 2008, 52 (1), 3–24.
- Berg-Schlosser, Dirk and Jeremy Mitchell**, *The Conditions of Democracy in Europe 1919-39: Systematic Case Studies*, Springer, 2000.
- and – , *Authoritarianism and democracy in Europe, 1919-39: comparative analyses*, Springer, 2002.
- Besley, Timothy J. and Anne Case**, “Diffusion as a Learning Process: Evidence from HYV Cotton,” Research Program in Development Studies Discussion Paper 174, Woodrow Wilson School of Public and International Affairs 1994.
- Boix, Carles**, “Democracy, development, and the international system,” *American Political Science Review*, 2011, 105 (04), 809–828.

- and Susan C Stokes, “Endogenous democratization,” *World politics*, 2003, 55 (04), 517–549.
- , Michael Miller, and Sebastian Rosato, “A complete data set of political regimes, 1800–2007,” *Comparative Political Studies*, 2013, 46 (12), 1523–1554.
- Bolton, Patrick and Christopher Harris**, “Strategic Experimentation,” *Econometrica*, 1999, 67 (2), 349–374.
- Bramoulleé, Yann, Rachel Kranton, and Martin D’Amours**, “Strategic Interaction and Networks,” *American Economic Review*, 2014, 104 (3), 898–930.
- Brender, Adi and Allan Drazen**, “How do budget deficits and economic growth affect reelection prospects? Evidence from a large panel of countries,” *The American Economic Review*, 2008, 98 (5), 2203–2220.
- Brinks, Daniel and Michael Coppedge**, “Diffusion is no illusion: Neighbor emulation in the third wave of democracy,” *Comparative Political Studies*, 2006, 39 (4), 463–489.
- Bromhead, Alan De, Barry Eichengreen, and Kevin H O’Rourke**, “Political extremism in the 1920s and 1930s: Do German lessons generalize?,” *The Journal of Economic History*, 2013, 73 (2), 371–406.
- Buera, Francisco J., Alexander Monge-Naranjo, and Giorgio E. Primiceri**, “Learning the Wealth of Nations,” *Econometrica*, 2011, 79 (1), 1–45.
- Chen, Jie and Chunlong Lu**, “Democratization and the middle class in China: The middle class’s attitudes toward democracy,” *Political Research Quarterly*, 2011, 64 (3), 705–719.
- Conley, Timothy G. and Christopher R. Udry**, “Learning about a New Technology: Pineapple in Ghana,” *American Economic Review*, 2010, 100 (1), 35–69.
- Dahl, Robert Alan**, *Polyarchy: Participation and opposition*, Yale University Press, 1973.
- , *On Democracy*, Yale University Press, 1998.
- Diamond, Larry**, “Why democracies survive,” *Journal of Democracy*, 2011, 22 (1), 17–30.
- Douthat, Ross**, “Recession and Revolution,” *New York Times*, 6-15-2009.
- Drazen, Allan**, “The political business cycle after 25 years,” *NBER macroeconomics annual*, 2000, 15, 75–117.
- Foster, Andrew D. and Mark R. Rosenzweig**, “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture,” *Journal of Political Economy*, 1995, 103 (6), 1176–1209.
- Frey, Bruno S and Hannelore Weck**, “A statistical study of the effect of the Great Depression on elections: The Weimar Republic, 1930–1933,” *Political Behavior*, 1983, 5 (4), 403–420.

- Friedman, Thomas**, “The democratic recession,” *New York Times*, 5-7-2008.
- García-Jimeno, Camilo**, “The Political Economy of Moral Conflict: An Empirical Study of Learning and Law Enforcement under Prohibition,” *Econometrica*, 2016, *84* (2), 511–570.
- Gleditsch, Kristian Skrede**, “Expanded Trade and GDP Data (version 6.0),” *Journal of Conflict Resolution*, 2002, *46* (5), 712–724.
- **and Michael D Ward**, “Diffusion and the international context of democratization,” *International organization*, 2006, *60* (04), 911–933.
- Goemans, Henk E, Kristian Skrede Gleditsch, and Giacomo Chiozza**, “Introducing Archigos: A dataset of political leaders,” *Journal of Peace research*, 2009, *46* (2), 269–283.
- Goldstone, Jack A**, *Revolution and rebellion in the early modern world*, Univ of California Press, 1991.
- Gunitsky, Seva**, “From shocks to waves: Hegemonic transitions and democratization in the twentieth century,” *International Organization*, 2014, *68* (03), 561–597.
- Hibbs, Douglas A**, “Political parties and macroeconomic policy,” *American political science review*, 1977, *71* (4), 1467–1487.
- Hirschman, Albert O and Michael Rothschild**, “The Changing Tolerance for Income Inequality in the Course of Economic Development* With A Mathematical Appendix,” *The Quarterly Journal of Economics*, 1973, *87* (4), 544–566.
- Houle, Christian, Mark A Kayser, and Jun Xiang**, “Diffusion or Confusion? Clustered Shocks and the Conditional Diffusion of Democracy,” *International Organization*, 2016, *70* (4), 687.
- Huntington, Samuel P**, *Political order in changing societies*, Yale University Press, 1968.
- , *The third wave: Democratization in the late twentieth century*, Vol. 4, University of Oklahoma press, 1993.
- Jackson, Nicole J**, “The role of external factors in advancing non-liberal democratic forms of political rule: a case study of Russia’s influence on Central Asian regimes,” *Contemporary Politics*, 2010, *16* (1), 101–118.
- Kadera, Kelly M, Mark JC Crescenzi, and Megan L Shannon**, “Democratic survival, peace, and war in the international system,” *American Journal of Political Science*, 2003, *47* (2), 234–247.
- King, Gary, Ori Rosen, Martin Tanner, and Alexander F Wagner**, “Ordinary economic voting behavior in the extraordinary election of Adolf Hitler,” *The Journal of Economic History*, 2008, *68* (4), 951–996.
- Krugman, Paul**, “Depression and democracy,” *New York Times*, 12-11-2011, *12*.

- Kurlantzick, Joshua and Perry Link**, “China: Resilient, sophisticated authoritarianism,” *Undermining democracy*, 2009, pp. 13–28.
- Landier, Augustin, David Thesmar, and Mathias Thoenig**, “Investigating Capitalism Aversion,” in G. de Ménéil, R. Portes, and H.-W. Sinn, eds., *Economic Policy 55*, Oxford, UK: Blackwell Publishing Ltd., 2008, pp. 465–497.
- Levitsky, Steven and Lucan Way**, “International linkage and democratization,” *Journal of democracy*, 2005, 16 (3), 20–34.
- Li, Quan and Rafael Reuveny**, “Economic globalization and democracy: An empirical analysis,” *British Journal of Political Science*, 2003, 33 (01), 29–54.
- Lipset, Seymour Martin**, “Some social requisites of democracy: Economic development and political legitimacy,” *American political science review*, 1959, 53 (01), 69–105.
- Londregan, John B and Keith T Poole**, “Poverty, the coup trap, and the seizure of executive power,” *World politics*, 1990, 42 (2), 151–183.
- and —, “Does high income promote democracy?,” *World politics*, 1996, 49 (1), 1–30.
- Maddison, Angus**, “Statistics on world population, GDP and per capita GDP, 1-2008 AD,” *Historical Statistics*, 2010, pp. 1–36.
- Mansfield, Edward D, Helen V Milner, and B Peter Rosendorff**, “Free to trade: Democracies, autocracies, and international trade,” *American Political Science Review*, 2000, 94 (02), 305–321.
- Matérn, Bertil**, “Spatial variation: Stochastic models and their application to some problems in forest surveys and other sampling investigations,” *Meddelanden från Statens Skogsforskningsinstitut*, 1960, 49 (5), 1–144.
- Mossel, Elchanan, Allan Sly, and Omer Tamuz**, “Strategic Learning and the Topology of Social Networks,” *Econometrica*, 2015, 83 (5), 1755–1794.
- Nordhaus, William D**, “The political business cycle,” *The review of economic studies*, 1975, 42 (2), 169–190.
- Norris, Pippa**, *Critical citizens: Global support for democratic government*, Oxford University Press, 1999.
- Olson, Mancur**, “Rapid growth as a destabilizing force,” *The Journal of Economic History*, 1963, 23 (4), 529–552.
- Payne, Stanley G**, *A history of fascism, 1914–1945*, University of Wisconsin Press, 1996.
- Pevehouse, Jon C**, “Democracy from the outside-in? International organizations and democratization,” *International organization*, 2002, 56 (03), 515–549.

- , *Democracy from above: Regional organizations and democratization*, Cambridge University Press, 2005.
- Powell, Charles**, “International aspects of democratization: The case of Spain,” in Laurence Whitehead, ed., *The international dimensions of democratization: Europe and the Americas*, Oxford: Oxford University Press, 1996, chapter 11, pp. 285–314.
- Pridham, Geoffrey**, *Encouraging democracy: The international context of regime transition in Southern Europe*, Burns & Oates, 1991.
- Primiceri, Giorgio E.**, “Why Inflation Rose and Fell: Policymakers’ Beliefs and US Post-war Stabilization Policy,” *Quarterly Journal of Economics*, 2006, 121 (3), 867–901.
- Przeworski, Adam and Fernando Limongi**, “Modernization: Theories and facts,” *World politics*, 1997, 49 (02), 155–183.
- , **Michael Alvarez, Jose Cheibub, and Fernando Limongi**, “Democracy and Development: Political Institutions and Well-Being in the World, 1950-1990 (Cambridge Studies in the Theory of Democracy),” 2000.
- Rodrik, Dani**, “Democracies Pay Higher Wages,” *Quarterly Journal of Economics*, 1999, 114 (3), 707–738.
- Rueschemeyer, Dietrich, Evelyne Huber Stephens, and John D Stephens**, “Capitalist development and democracy,” *Cambridge, UK*, 1992.
- Sargent, Thomas J.**, *The Conquest of American Inflation*, Princeton, NJ: Princeton University Press, 1999.
- Skocpol, Theda**, *States and social revolutions: A comparative analysis of France, Russia and China*, Cambridge University Press, 1979.
- Spolaore, Enrico and Romain Wacziarg**, “The Diffusion of Development,” *The Quarterly Journal of Economics*, 2009, 124 (2), 469–592.
- Su, Che-Lin and Kenneth L Judd**, “Constrained Optimization Approaches to Estimation of Structural Models,” *Econometrica*, 2012, 80 (5), 2213–2230.
- Torfason, Magnus Thor and Paul Ingram**, “The global rise of democracy: A network account,” *American Sociological Review*, 2010, 75 (3), 355–377.
- tzung Chang, Yu, Yunhan Zhu, and Chong min Pak**, “Authoritarian nostalgia in Asia,” *Journal of Democracy*, 2007, 18 (3), 66–80.
- Wejnert, Barbara**, “Diffusion, development, and democracy, 1800-1999,” *American Sociological Review*, 2005, 70 (1), 53–81.
- Whitehead, Laurence**, *The international dimensions of democratization: Europe and the Americas*, Oxford University Press, 1996.