Online Appendix for *Learning about Growth and Democracy*[*]

Scott F Abramson[†] Sergio Montero[‡]

[*]This version: April 17, 2020.

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

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

A1 List of Countries

Table A1 lists all countries in our main sample used to estimate the baseline specification of our model.

Table A1: Main Sample

Country	Years	Notes	Country	Years	Notes
United States	1951-2000		Canada	1951-2000	
Cuba	1951-2000		Dominican Republic	1951-2000	
Jamaica	1964-2000		Trinidad and Tobago	1964-2000	
Mexico	1951-2000		Guatemala	1951-2000	
Honduras	1951-2000		El Salvador	1951-2000	
Nicaragua	1951-2000		Costa Rica	1951-2000	
Panama	1951-2000		Colombia	1951-2000	
Venezuela	1951-2000		Ecuador	1951-2000	
Peru	1951-2000		Brazil	1951-2000	
Bolivia	1951-2000		Paraguay	1951-2000	
Chile	1951-2000		Argentina	1951-2000	
Uruguay	1951-2000		United Kingdom	1951-2000	
Ireland	1951-2000		Netherlands	1951-2000	
Belgium	1951-2000		France	1951-2000	
Switzerland	1951-2000		Spain	1951-2000	
Portugal	1951-2000		Germany	1951-2000	Federal Republic of Germany from 1951-1990.
Poland	1951-2000		Austria	1951-2000	
Hungary	1951-2000		Czech Republic	1951-2000	Czechoslovakia from 1951-1992.
Slovakia	1994-2000		Italy	1951-2000	
Albania	1951-2000		Macedonia	1993-2000	
Croatia	1992-2000		Yugoslavia	1951-1991	
Bosnia and Herzegovina	1993-2000		Slovenia	1992-2000	
Greece	1951-2000		Bulgaria	1951-2000	
Moldova	1992-2000		Romania	1951-2000	
Russia	1951-2000	U.S.S.R. from 1951-1991.	Estonia	1993-2000	
Latvia	1992-2000		Lithuania	1992-2000	
Ukraine	1992-2000		Belarus	1993-2000	
Armenia	1993-2000		Georgia	1993-2000	
Azerbaijan	1993-2000		Finland	1951-2000	
Sweden	1951-2000		Norway	1951-2000	
Denmark	1951-2000		Republic of Guinea-Bissau	1976-2000	
Equatorial Guinea	1969-2000		Gambia	1967-2000	
Mali	1962-2000		Senegal	1962-2000	
Benin	1961-2000		Mauritania	1962-2000	
Niger	1962-2000		Côte D'Ivoire	1962-2000	
Republic of Guinea	1960-2000		BurkinaFaso	1962-2000	
Liberia	1951-2000		Sierra Leone	1963-2000	
Ghana	1958-2000		Togo	1962-2000	
Cameroon	1961-2000		Nigeria	1962-2000	
Gabon	1962-2000		Central African Republic	1962-2000	
Chad	1962-2000		Republic of the Congo	1962-2000	
Zaire	1962-2000	Congo-Léopoldville from 1962-1965, Democratic Republic of Congo from 1997-2000.	Uganda	1964-2000	
Kenya	1965-2000		Tanzania	1963-2000	
Burundi	1964-2000		Rwanda	1963-2000	
Somalia	1962-1991		Djibouti	1979-2000	
Ethiopia	1994-2000		Angola	1977-2000	
Mozambique	1977-2000		Zambia	1966-2000	
Zimbabwe	1967-2000		Malawi	1966-2000	
South Africa	1951-2000		Namibia	1992-2000	
Lesotho	1968-2000		Botswana	1968-2000	
Swaziland	1970-2000		Madagascar	1962-2000	
Comoro Islands	1977-2000		Mauritius	1970-2000	
Morocco	1958-2000		Algeria	1964-2000	
Tunisia	1951-2000		Libya	1953-2000	
Sudan	1957-2000		Iran	1951-2000	
Turkey	1951-2000		Iraq	1951-2000	
Egypt	1951-2000		Syria	1951-2000	
Lebanon	1951-2000		Jordan	1951-2000	
Israel	1951-2000		Saudi Arabia	1951-2000	
Kuwait	1952-2000		Bahrain	1973-2000	
Qatar	1973-2000		United Arab Emirates	1973-2000	
Oman	1951-2000		Afghanistan	1951-2000	
Turkmenistan	1992-2000		Tajikistan	1993-2000	
Kyrgyzstan	1992-2000		Uzbekistan	1992-2000	
Kazakhstan	1992-2000		China	1951-2000	
Mongolia	1951-2000		Taiwan	1951-2000	
North Korea	1951-2000		South Korea	1951-2000	
Japan	1951-2000		India	1951-2000	
Pakistan	1951-2000		Bangladesh	1973-2000	
Burma	1951-2000		Sri Lanka	1951-2000	
Nepal	1951-2000		Thailand	1951-2000	
Cambodia	1956-2000		Laos	1955-2000	
Democratic Republic of Vietnam	1951-2000	Socialist Republic of Vietnam from 1976-2000.	Malaysia	1959-2000	
Singapore	1961-2000		Philippines	1951-2000	
Indonesia	1951-2000		Australia	1951-2000	
New Zealand	1951-2000				

A2 Preliminary Evidence

To motivate our model, we estimate a pair of reduced-form regressions. The purpose is to demonstrate two empirical patterns underpinning our structural approach: (i) democracy adoption systematically covaries with observed differences in economic performance between neighboring democracies and autocracies—which is suggestive of the learning process we model—and (ii) economic growth has a differential impact on elite turnover under democracy versus autocracy.

First, we estimate the following linear probability model via ordinary least squares:

$$D_{i,t} = \phi_0 + \phi_1 D_{i,t-1} + \phi_2 \left[\bar{y}^{D=1}_{i,t-1} - \bar{y}^{D=0}_{i,t-1}\right] + \phi_3 \bar{D}_{i,t-1} + u_{i,t}, \tag{A1}$$

where $\bar{y}^{D=1}_{i,t-1}$ and $\bar{y}^{D=0}_{i,t-1}$ are weighted averages of past GDP per capita growth rates among country i's democratic and autocratic neighbors, respectively.[1] Thus, $\bar{y}^{D=1}_{i,t-1} - \bar{y}^{D=0}_{i,t-1}$ proxies for beliefs in country i about the impact of democracy on economic growth. We vary the size of i's effective neighborhood by setting the decay parameter equal to the estimated value from our structural model ($\gamma = 0.4234$) as well as twice (2γ) and half this value ($\frac{1}{2}\gamma$). Furthermore, we control for $\bar{D}_{i,t-1}$, the weighted proportion of democracies in i's neighborhood.

If $\phi_2 \neq 0$, then, consistent with our model, democracy adoption systematically covaries with observed growth differences in democratic versus autocratic neighbors. We evaluate this hypothesis in Table A2, which shows that, across specifications and spatial weights, the reduced-form evidence is indeed consistent with our structural estimates: as $\phi_2 > 0$, the likelihood of democracy increases with superior economic performance in democracies.

Second, to motivate incumbents' objective function in our model, we show that the effect of GDP growth on their likelihood of retaining power in the subsequent period is heterogenous

[1] Formally, as in Buera, Monge-Naranjo and Primiceri (2011), we compute $\bar{y}^{D=1}_{i,t-1} - \bar{y}^{D=0}_{i,t-1} = \sum_{\tau=t-3}^{t-1} \left(\frac{\sum_{j\neq i} \exp(-\gamma d_{i,j}) y_{j\tau} D_{j\tau}}{\sum_{j\neq i} \exp(-\gamma d_{i,j}) D_{j\tau}} - \frac{\sum_{j\neq i} \exp(-\gamma d_{i,j}) y_{j\tau} (1 - D_{j\tau})}{\sum_{j\neq i} \exp(-\gamma d_{i,j})(1 - D_{j\tau})} \right)$, where $d_{i,j}$ is the distance between countries i and j.

Table A2: Reduced-Form Evidence of Effect of Learning about Growth on Democracy

	Dependent variable: $D_{i,t}$											
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)	(XI)	(XII)
Decay:	$\frac{1}{2}\gamma$				γ				2γ			
ϕ_1	0.854***	0.851***	0.849***	0.849***	0.854***	0.851***	0.849***	0.849***	0.853***	0.851***	0.849***	0.850***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
ϕ_2	0.801***	0.527**	0.556**	0.527**	0.419***	0.279*	0.289*	0.258*	0.214**	0.163*	0.161*	0.141
	(0.202)	(0.217)	(0.217)	(0.221)	(0.149)	(0.154)	(0.154)	(0.156)	(0.097)	(0.098)	(0.098)	(0.098)
ϕ_3	0.121***	0.128***	0.128***	0.133***	0.117***	0.114***	0.115***	0.115***	0.104***	0.096***	0.097***	0.096***
	(0.025)	(0.025)	(0.025)	(0.026)	(0.021)	(0.021)	(0.021)	(0.022)	(0.018)	(0.018)	(0.018)	(0.018)
Controls:												
$\log(Y_{i,t-1})$	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$\text{Trade}_{i,t-1}/Y_{i,t-1}$	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes
Time in Power	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes
Observations	5,623	5,623	5,623	5,579	5,623	5,623	5,623	5,579	5,623	5,623	5,623	5,579

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

Notes. Ordinary least squares estimates from regression (A1). All models include country fixed effects. Standard errors in parentheses.

across democracies and autocracies, with its impact in democracies being consistently larger. Specifically, we estimate the following linear probability model via least squares:

$$Pr(RetainPower_{i,t+1}|y_{i,t}, D_{i,t}) = \lambda_0 + \lambda_1 y_{i,t} + \lambda_2 D_{i,t} + \lambda_3 y_{i,t} \times D_{i,t}. \quad \text{(A2)}$$

Of course, given the selection into or out of democracy we model, estimates of λ will be inconsistent. To address this without imposing additional structure, we use $\bar{y}^{D=1}_{i,t-1}$, $\bar{y}^{D=0}_{i,t-1}$, and $\bar{D}_{i,t-1}$ from (A1) to build instruments for the regressors in (A2).[2] That is, in line with our assumption that beliefs have no direct impact on outcomes (other than through institutional choices), if we correctly measured beliefs, we could obtain consistent estimates of λ via two-stage least squares. Since we view $\bar{y}^{D=1}_{i,t-1}$ and $\bar{y}^{D=0}_{i,t-1}$ only as rough proxies for beliefs, however, we take the instrumental-variables estimates of λ_1 and λ_3 with a grain of salt.

Estimates of λ_1 and $\lambda_1 + \lambda_3$ measure the effects of economic growth on elite survival under autocracy and democracy, respectively, while λ_2 gives the difference in baseline survival under democracy (the average of $-f_i$ in our structural model). Ordinary and two-stage least

[2] We set $\gamma = 0.4234$.

Table A3: Heterogenous Impact of Growth on Incumbent Stability

	Dependent variable: $RetainPower_{i,t+1}$							
	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)
	OLS				2SLS			
λ_1	0.336^{***}	0.350^{***}	0.353^{***}	0.378^{***}	-26.391	-11.995	-13.179	-9.583
	(0.073)	(0.073)	(0.074)	(0.074)	(109.797)	(22.195)	(26.917)	(15.707)
λ_2	-0.081^{***}	-0.082^{***}	-0.083^{***}	-0.080^{***}	-1.161	-0.500	-0.533	-0.276
	(0.014)	(0.014)	(0.014)	(0.014)	(4.343)	(0.731)	(0.869)	(0.291)
λ_3	0.289^{*}	0.288^{*}	0.286^{*}	0.245	91.411	37.182	41.071	27.540
	(0.168)	(0.170)	(0.170)	(0.179)	(378.318)	(68.632)	(83.795)	(45.160)
Controls:								
$\log(Y_{i,t-1})$	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$\text{Trade}_{i,t-1}/Y_{i,t-1}$	No	No	Yes	Yes	No	No	Yes	Yes
Time in Power	No	No	No	Yes	No	No	No	Yes
Observations	5,966	5,924	5,919	5,845	5,622	5,622	5,622	5,578

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

Notes. Ordinary (OLS) and two-stage least squares (2SLS) estimates from regression (A2). All models include country fixed effects. Standard errors in parentheses.

squares estimates are presented in Table A3. Across specifications, we find that democracies consistently reward incumbents for high rates of growth more than autocracies. Moreover, the two-stage least squares results are consistent with our structural estimates in that growth is stabilizing for incumbents under democracy but not under autocracy. These estimates are imprecise, however, given that the instruments are only rough proxies for the learning process we model.

A2.1 Growth in Democracies Versus Autocracies

Figure A1 illustrates the variation in the data that drives our main results by plotting the evolution of the difference in growth rates in democracies versus autocracies. The plot shows that, following the oil crisis, democracies experienced a markedly superior recovery. This underlies the sharp revision of beliefs and subsequent transitions to democracy that characterize the most consequential period in our sample.

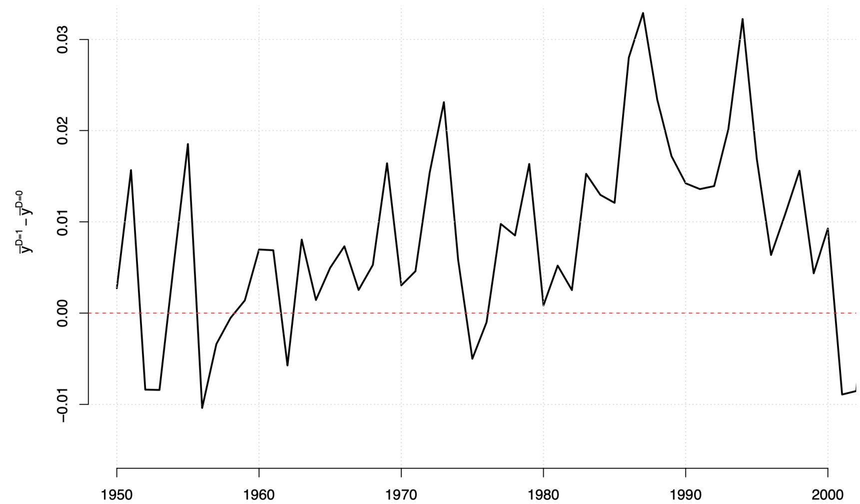

Figure A1: Difference in Average Growth Rates in Democracies Versus Autocracies

A3 Likelihood of the Data

Recall that $W^T \equiv \{I_t, y_t, D_t, X_t\}_{t=1}^T$ denotes the set of all data available up to period T. With a slight abuse of notation—using $\mathcal{L}$ to denote arbitrary densities of the data—the likelihood function can be written as

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

where $W_t \equiv \{I_t, y_t, D_t, X_t\}$ collects the data generated in period t. As discussed in the paper, 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

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

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

To compute $\mathcal{L}(D_t|X_t, W^{t-1},\varphi)$, notice from (5) in the paper that, given $(X_{i,t}, W^{t-1},\varphi)$, 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 $\overline{\kappa}_{i,t}(X_{i,t}, W^{t-1},\varphi)$, is defined implicitly by

$$\begin{aligned}&E_{i,t-1}\left[\frac{\exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t}) - f_i - X'_{i,t}\xi - \overline{\kappa}_{i,t}(X_{i,t}, W^{t-1},\varphi))}{1+\exp(\alpha_i + \theta^{D=1}(\beta_i^{D=1} + \epsilon_{i,t}) - f_i - X'_{i,t}\xi - \overline{\kappa}_{i,t}(X_{i,t}, W^{t-1},\varphi))}\right]\\ &\qquad\qquad = E_{i,t-1}\left[\frac{\exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))}{1+\exp(\alpha_i + \theta^{D=0}(\beta_i^{D=0} + \epsilon_{i,t}))}\right].\end{aligned} \tag{A3}$$

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

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

where

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

and Φ denotes the standard Normal cumulative distribution function.

A4 Prior

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

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

where $\text{IG}(s, d)$ denotes the Inverse-Gamma distribution with shape parameter s and scale 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 systems 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 belief correlations 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 $\overline{\theta y}_i$ 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 $\alpha_i + \overline{\theta y}_i - \overline{DK}_i$ 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 $E(\alpha_i + \overline{\theta y}_i - \overline{DK}_i) = \underline{\alpha}$, we set $\underline{\alpha} = 1.8432$, the average log-odds in the pre-sample period.[3] 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 $\omega_f/4 = 0.1496$.

A5 Maximum-A-Posteriori Estimator: MPEC Approach

As noted in the paper, calculating $\overline{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \varphi)$ from (A3) to evaluate the likelihood of the data is computationally expensive. To avoid this burden, we follow the Mathematical

[3] For 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 (95%).

Programming with Equilibrium Constraints (MPEC) approach of Su and Judd (2012) to compute our maximum-a-posteriori (MAP) estimator of φ. The idea behind this approach is simple: instead of calculating $\overline{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \varphi)$ at every trial value of φ, one treats each $\overline{\kappa}_{i,t}$ as an auxiliary parameter and imposes (A3)—the optimality (or equilibrium) condition of the model—as a feasibility constraint on the log-posterior maximization program. Accordingly, we estimate φ by solving $\max_{\varphi,\overline{\boldsymbol{\kappa}}} \log\left(p(\varphi, \overline{\boldsymbol{\kappa}}|W^T)\right)$ subject to the constraint that $\overline{\kappa}_{i,t}$ satisfies (A3) for all i and t.

As shown by Su and Judd (2012), MPEC and the standard approach of directly maximizing $\log\left(p(\varphi|W^T)\right)$ yield theoretically-identical estimates of φ. 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 (A3) is satisfied exactly only once rather than at every trial value of φ. Moreover, for this reason, MPEC is robust to sensitivity issues that may arise from not setting a sufficiently stringent convergence criterion when computing $\overline{\kappa}_{i,t}(X_{i,t}, W^{t-1}, \varphi)$ (Dubé, Fox and Su, 2012). A potential disadvantage is that, by introducing $\overline{\boldsymbol{\kappa}}$ 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 $\overline{\kappa}_{i,t}$ entering a single constraint.

To reap the computational benefits of the MPEC approach, it is essential to employ optimization software tailor-made to handle large-scale problems—with thousands of variables and nonlinear constraints. Accordingly, we implement our MPEC-MAP estimator using the industry-leading software Knitro.[4] Due to memory and computational constraints—our baseline model with no covariates features 8,459 variables—our implementation relies on Knitro's Interior/Direct algorithm with their limited-memory quasi-Newton BFGS approximation of the Hessian of the Lagrangian. Nevertheless, we provide exact first derivatives of the log-posterior and constraints.[5] With a 3.0 GHz machine, it takes about 5-6 days to

[4]https://www.artelys.com/en/optimization-tools/knitro

[5]Knitro offers a derivative-check option—which our implementation passes—to test the code for exact derivatives against finite-difference approximations.

estimate our model once.

To mitigate concerns about potential local maxima, for each model specification we randomly draw 5 sets of starting values for the optimization algorithm from the prior distribution of the model parameters described in Appendix A4. We then select, for each specification, the solution that achieves the highest log-posterior value. Reassuringly, there is very little divergence in solutions across starting values.

Standard errors for our parameter estimates are parametrically bootstrapped (Davison and Hinkley, 1997). Due to the considerable computational cost of estimating our model, we only compute standard errors for our baseline specification with no covariates. This has the added advantage that only estimates from the true DGP described in Footnote 36 are necessary to generate bootstrap samples.

A final notable computational challenge is that the integrals in (A3) have no closed-form solution. Rather than employing a Monte Carlo approximation, which would require independent draws across all i and t to prevent simulation error from propagating, we rely on sparse-grid integration as implemented by Heiss and Winschel (2008). This approach is much more efficient and delivers virtually exact integral computations for integrands that are well approximated by polynomials—as is the case for the integrals in (A3).[6]

A6 Coefficient Estimates: Baseline Model

Table 1 lists and describes all the parameters of our model. In our baseline specification with no covariates, the vector ξ is empty. We report estimates of θ and γ in Footnotes 31 and 34, respectively, and estimates of f_i for each country in Figure 3. Table A4 presents all remaining coefficient estimates for our baseline specification (standard errors in parentheses).

[6] Our implementation computes exact integrals of fifteenth-degree polynomials.

Table A4: Estimates of Country-Specific Parameters α_i, $\beta_{i,0}^{D=0}$, $\beta_{i,0}^{D=1}$, v_i, and ς_i

Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i	Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i
USA	1.8147	0.0189	0.0284	0.1902	0.0728	CAN	1.8000	0.0191	0.0249	0.1823	0.0723
	(0.0487)	(0.0198)	(0.0176)	(0.1953)	(0.3791)		(0.0521)	(0.0209)	(0.0193)	(0.1956)	(0.3637)
CUB	1.7449	0.0422	0.0431	0.5696	0.0469	DOM	1.9302	0.0020	0.0005	0.6220	0.0442
	(0.0743)	(0.0327)	(0.0320)	(0.2930)	(0.2320)		(0.1091)	(0.0436)	(0.0359)	(0.3591)	(0.4087)
JAM	1.8361	0.0185	0.0313	0.5088	0.0737	TRI	1.7974	0.0192	0.0165	0.1946	0.0722
	(0.0877)	(0.0400)	(0.0346)	(0.2522)	(0.3667)		(0.0194)	(0.0268)	(0.0247)	(0.1860)	(0.3887)
MEX	1.8773	0.0126	0.0206	0.1913	0.0705	GUA	1.9023	0.0453	0.0386	0.2184	0.0598
	(0.0705)	(0.0331)	(0.0285)	(0.1876)	(0.1898)		(0.0645)	(0.0759)	(0.0649)	(0.1662)	(0.1646)
HON	1.8249	0.0008	0.0082	0.2591	0.0585	SAL	1.8631	-0.0085	-0.0154	0.4512	0.0712
	(0.0672)	(0.0346)	(0.0312)	(0.1910)	(0.1484)		(0.3185)	(0.0687)	(0.0635)	(0.2657)	(0.2856)
NIC	1.8652	0.0154	0.0264	0.1166	0.0278	COS	1.7512	0.0197	0.0142	0.2037	0.0715
	(0.0408)	(0.0398)	(0.0353)	(0.1565)	(0.1964)		(0.0748)	(0.0699)	(0.0576)	(0.1898)	(0.3795)
PAN	1.8437	0.0426	0.0089	0.1317	0.1102	COL	2.5650	0.0072	0.0046	1.4453	0.0723
	(0.0118)	(0.0285)	(0.0277)	(0.1404)	(0.3525)		(0.4678)	(0.1286)	(0.1070)	(0.4873)	(0.4519)
VEN	2.3228	0.0123	0.0106	1.4937	0.0531	ECU	1.4737	0.0385	0.0295	0.4143	0.0423
	(0.3398)	(0.0835)	(0.0641)	(0.4937)	(0.4106)		(0.0809)	(0.1467)	(0.1283)	(0.2001)	(0.1520)
PER	1.8327	0.0300	0.0037	0.2171	0.0917	BRA	1.8211	0.0214	0.0030	0.1191	0.1059
	(0.0450)	(0.0428)	(0.0365)	(0.1901)	(0.2221)		(0.0328)	(0.0305)	(0.0295)	(0.1557)	(0.2135)
BOL	1.9907	0.0068	0.0099	1.6876	0.0612	PAR	1.8948	0.0362	0.0208	0.2246	0.0735
	(2.3220)	(0.4624)	(0.3649)	(0.3540)	(0.7489)		(0.0795)	(0.0757)	(0.0645)	(0.1882)	(0.3755)
CHL	1.4128	0.0289	0.0285	0.5013	0.0357	ARG	1.8439	0.0057	0.0356	0.0735	0.1003
	(0.0817)	(0.0996)	(0.0840)	(0.2114)	(0.2287)		(0.0497)	(0.0545)	(0.0481)	(0.1578)	(0.2731)
URU	1.8164	0.0322	0.0077	0.2035	0.0365	UKG	1.7922	0.0191	0.0244	0.1782	0.0724
	(0.1215)	(0.0647)	(0.0553)	(0.1432)	(0.1799)		(0.0523)	(0.0182)	(0.0141)	(0.1726)	(0.3912)
IRE	1.8277	0.0186	0.0243	0.1729	0.0734	NTH	1.6578	0.0202	0.0242	0.1883	0.0711
	(0.0603)	(0.0134)	(0.0113)	(0.1723)	(0.3845)		(0.0928)	(0.0376)	(0.0304)	(0.2439)	(0.3671)
BEL	1.8362	0.0185	0.0203	0.1847	0.0736	FRN	1.7776	0.0192	0.0260	0.1702	0.0724
	(0.0610)	(0.0259)	(0.0218)	(0.2549)	(0.3996)		(0.0623)	(0.0282)	(0.0232)	(0.1777)	(0.3995)
SWZ	1.8035	0.0190	0.0250	0.1680	0.0727	SPN	1.8729	0.0161	-0.0132	0.4770	0.0422
	(0.0418)	(0.0208)	(0.0164)	(0.1565)	(0.4014)		(0.1007)	(0.0575)	(0.0466)	(0.3326)	(0.3714)
POR	1.8777	0.0124	-0.0022	0.3360	0.0505	GMY	1.8419	0.0182	0.0285	0.1457	0.0745
	(0.1023)	(0.0748)	(0.0608)	(0.2820)	(0.3306)		(0.0362)	(0.0376)	(0.0328)	(0.2375)	(0.3980)
POL	1.9991	0.0175	0.0093	1.0647	0.0546	AUS	1.8432	0.0180	0.0740	0.1156	0.0748
	(1.5958)	(0.4168)	(0.3400)	(0.3578)	(0.5581)		(0.0139)	(0.0748)	(0.0641)	(0.2030)	(0.3932)
HUN	2.3984	0.0179	0.0195	2.0474	0.0505	CZR	1.8709	-0.0005	-0.0157	0.3155	0.0421
	(0.4215)	(0.1945)	(0.1690)	(0.7520)	(0.2212)		(0.0882)	(0.0173)	(0.0183)	(0.2906)	(0.1772)
SLO	1.7086	0.0197	0.0244	0.2122	0.0725	ITA	1.8411	0.0182	0.0272	0.1694	0.0743
	(0.0224)	(0.0158)	(0.0131)	(0.2786)	(0.4015)		(0.0218)	(0.0096)	(0.0090)	(0.1807)	(0.3986)
ALB	1.7750	0.0006	0.0240	0.2592	0.0593	MAC	1.7758	0.0193	0.0222	0.2043	0.0724
	(0.0518)	(0.0172)	(0.0132)	(0.2334)	(0.1494)		(0.0143)	(0.0089)	(0.0088)	(0.2129)	(0.4087)
CRO	1.8592	0.0180	0.0191	0.2071	0.0644	YUG	1.9095	0.0319	0.0188	0.2142	0.0712
	(0.0130)	(0.0179)	(0.0161)	(0.2801)	(0.2775)		(0.0211)	(0.0361)	(0.0281)	(0.2360)	(0.4035)
BOS	1.8004	0.0190	0.0230	0.3225	0.0730	SLV	1.5727	0.0205	0.0249	0.2008	0.0718
	(0.0187)	(0.0255)	(0.0205)	(0.2473)	(0.4074)		(0.0410)	(0.0134)	(0.0107)	(0.1868)	(0.4094)
GRC	1.5500	0.0468	0.0367	0.6486	0.0410	BUL	2.2611	0.0074	0.0222	1.0768	0.0579
	(0.0769)	(0.0774)	(0.0720)	(0.2902)	(0.3054)		(0.3696)	(0.3163)	(0.2659)	(0.3649)	(0.2674)
MLD	1.7984	0.0191	0.0236	0.1817	0.0727	RUM	2.4243	0.0186	0.0149	6.6647	0.0651
	(0.0172)	(0.0507)	(0.0414)	(0.1988)	(0.3884)		(3.2608)	(0.1847)	(0.0889)	(2.0319)	(0.6177)
RUS	1.8535	0.0186	0.0146	0.1078	0.1104	EST	1.8239	0.0186	0.0231	0.1796	0.0735
	(0.0115)	(0.0256)	(0.0204)	(0.1631)	(0.2625)		(0.0068)	(0.0102)	(0.0109)	(0.2519)	(0.4018)

Table A4 (continued)

Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i	Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i
LAT	1.8201	0.0184	0.0204	0.2166	0.0655	LIT	1.7210	0.0196	0.0254	0.2098	0.0722
	(0.0207)	(0.0094)	(0.0084)	(0.2520)	(0.2467)		(0.0281)	(0.0124)	(0.0123)	(0.2735)	(0.3954)
UKR	1.7982	0.0191	0.0240	0.1907	0.0728	BLR	1.8559	0.0187	0.0204	0.1672	0.0750
	(0.0149)	(0.0122)	(0.0118)	(0.2818)	(0.4079)		(0.0058)	(0.0185)	(0.0162)	(0.2341)	(0.2946)
ARM	1.8847	0.0165	0.0203	0.1667	0.0730	GRG	1.8432	0.0183	0.0218	0.1925	0.0748
	(0.0096)	(0.0064)	(0.0062)	(0.2125)	(0.3988)		(0.0258)	(0.0211)	(0.0154)	(0.2378)	(0.4042)
AZE	1.8470	0.0183	0.0217	0.2031	0.0747	FIN	1.7190	0.0197	0.0279	0.1797	0.0713
	(0.0237)	(0.0194)	(0.0138)	(0.2389)	(0.4047)		(0.0298)	(0.0241)	(0.0182)	(0.2173)	(0.3992)
SWD	1.8431	0.0181	0.0348	0.1434	0.0748	NOR	1.8031	0.0190	0.0222	0.1954	0.0726
	(0.0191)	(0.0102)	(0.0099)	(0.1754)	(0.3970)		(0.0701)	(0.0256)	(0.0208)	(0.1995)	(0.3838)
DEN	1.6931	0.0200	0.0242	0.2011	0.0713	GNB	1.8194	-0.0055	-0.0042	0.5045	0.0472
	(0.0786)	(0.0494)	(0.0404)	(0.2254)	(0.3822)		(0.0177)	(0.0130)	(0.0144)	(0.2605)	(0.2337)
EQG	1.9776	0.0191	0.0187	0.2287	0.0722	GAM	1.5020	0.0155	0.0217	0.6286	0.0586
	(0.1098)	(0.0786)	(0.0633)	(0.2426)	(0.4031)		(0.0847)	(0.0472)	(0.0391)	(0.2995)	(0.1793)
MLI	1.8795	0.0060	0.0008	0.5336	0.0453	SEN	1.9353	0.0363	0.0207	0.1905	0.1028
	(0.0524)	(0.0145)	(0.0136)	(0.2888)	(0.1921)		(0.0557)	(0.0161)	(0.0110)	(0.1466)	(0.2668)
BEN	1.8362	0.0025	-0.0126	0.3807	0.0561	MAA	1.9058	0.0188	0.0195	0.1744	0.0720
	(0.0667)	(0.0559)	(0.0472)	(0.2814)	(0.1693)		(0.0336)	(0.0238)	(0.0183)	(0.1920)	(0.4051)
NIR	1.8202	0.0196	0.0266	0.1989	0.0727	CDI	1.8757	0.0190	0.0182	0.1511	0.0708
	(0.0626)	(0.0709)	(0.0587)	(0.2127)	(0.2148)		(0.0402)	(0.0483)	(0.0389)	(0.1858)	(0.3991)
GUI	1.8766	0.0162	0.0188	0.1989	0.0714	BFO	1.8553	0.0160	0.0175	0.2194	0.0704
	(0.0273)	(0.0144)	(0.0118)	(0.1909)	(0.3986)		(0.0269)	(0.0140)	(0.0094)	(0.2435)	(0.3962)
LBR	1.8596	0.0028	0.0185	0.1369	0.0709	SIE	1.8233	0.0373	0.0505	0.4757	0.0479
	(0.0295)	(0.0261)	(0.0214)	(0.2099)	(0.4058)		(0.1108)	(0.0493)	(0.0399)	(0.2757)	(0.2199)
GHA	1.8585	0.0139	0.0084	0.1360	0.0512	TOG	1.9611	0.0276	0.0189	0.1611	0.0721
	(0.0289)	(0.0179)	(0.0135)	(0.1645)	(0.2214)		(0.1153)	(0.0640)	(0.0552)	(0.1741)	(0.3625)
CAO	1.8599	0.0171	0.0177	0.1792	0.0704	NIG	1.8737	0.0237	0.0192	0.0982	0.1973
	(0.0206)	(0.0184)	(0.0154)	(0.2056)	(0.4004)		(0.0525)	(0.0230)	(0.0201)	(0.1822)	(0.2475)
GAB	1.9154	0.0159	0.0203	0.1530	0.0732	CEN	1.7973	0.0129	0.0257	0.4532	0.0517
	(0.0657)	(0.0711)	(0.0589)	(0.1964)	(0.3965)		(0.1195)	(0.0652)	(0.0528)	(0.2711)	(0.2460)
CHA	1.8823	0.0189	0.0212	0.1889	0.0740	CON	1.8664	0.0243	0.0212	0.1445	0.0976
	(0.0629)	(0.0340)	(0.0272)	(0.1768)	(0.3713)		(0.0409)	(0.0342)	(0.0288)	(0.1896)	(0.1751)
DRC	1.8770	0.0176	0.0185	0.1853	0.0710	UGA	1.8465	0.0165	0.0206	0.1186	0.0982
	(0.0428)	(0.0226)	(0.0170)	(0.2403)	(0.4070)		(0.0343)	(0.0188)	(0.0148)	(0.1935)	(0.2448)
KEN	1.8534	0.0046	0.0181	0.1584	0.0706	TAZ	1.9018	0.0176	0.0204	0.2016	0.0732
	(0.0065)	(0.0101)	(0.0076)	(0.2181)	(0.4083)		(0.0418)	(0.0177)	(0.0130)	(0.2054)	(0.3987)
BUI	1.9025	0.0173	0.0193	0.1830	0.0720	RWA	1.8749	0.0185	0.0213	0.2281	0.0741
	(0.0273)	(0.0189)	(0.0144)	(0.2132)	(0.4052)		(0.0514)	(0.0516)	(0.0420)	(0.2277)	(0.4004)
SOM	1.8143	0.0390	0.0384	0.4018	0.0447	DJI	1.8634	0.0164	0.0183	0.1944	0.0709
	(0.0661)	(0.0447)	(0.0351)	(0.2292)	(0.1870)		(0.0228)	(0.0089)	(0.0078)	(0.2686)	(0.3988)
ETH	1.8853	0.0175	0.0205	0.1796	0.0733	ANG	1.8744	0.0174	0.0213	0.1840	0.0742
	(0.0108)	(0.0065)	(0.0051)	(0.2619)	(0.4026)		(0.0438)	(0.0322)	(0.0251)	(0.1923)	(0.3917)
MZM	1.8505	-0.0069	-0.0098	0.2624	0.0446	ZAM	1.9325	0.0172	0.0195	0.1916	0.0727
	(0.0272)	(0.0494)	(0.0418)	(0.2579)	(0.1588)		(0.0525)	(0.0288)	(0.0229)	(0.1848)	(0.4029)
ZIM	1.9480	0.0205	0.0192	0.1595	0.0724	MAW	1.8417	0.0161	0.0196	0.3374	0.0565
	(0.0624)	(0.0286)	(0.0236)	(0.1429)	(0.3932)		(0.0879)	(0.0912)	(0.0756)	(0.2880)	(0.2650)
SAF	1.8835	0.0143	-0.0226	0.9136	0.0475	NAM	1.9146	0.0191	0.0204	0.1937	0.0734
	(0.0427)	(0.0511)	(0.0448)	(0.3972)	(0.2298)		(0.0398)	(0.0261)	(0.0206)	(0.2140)	(0.4067)
LES	2.0190	0.0179	0.0187	0.1879	0.0722	BOT	1.8295	0.0189	0.0053	0.1126	0.0727
	(0.1155)	(0.0834)	(0.0691)	(0.2240)	(0.4024)		(0.0359)	(0.0209)	(0.0218)	(0.1892)	(0.4041)
SWA	2.0060	0.0293	0.0192	0.1458	0.0726	MAG	1.8638	0.0151	0.0024	0.3599	0.0522
	(0.1293)	(0.0495)	(0.0437)	(0.1803)	(0.4026)		(0.0818)	(0.0482)	(0.0395)	(0.2580)	(0.1413)
COM	1.8525	0.0464	0.0167	0.1309	0.0695	MAS	1.8395	0.0183	0.0282	0.1617	0.0742
	(0.0265)	(0.0417)	(0.0341)	(0.1620)	(0.4078)		(0.0504)	(0.0157)	(0.0172)	(0.1598)	(0.4021)
MOR	1.9467	0.0251	0.0179	0.1848	0.0709	ALG	1.8899	0.0245	0.0172	0.1966	0.0702
	(0.0897)	(0.0433)	(0.0347)	(0.2236)	(0.4046)		(0.0492)	(0.0448)	(0.0365)	(0.1958)	(0.4080)

Table A4 (continued)

Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i	Country	α_i	$\beta_{i,0}^{D=0}$	$\beta_{i,0}^{D=1}$	v_i	ς_i
TUN	1.8613	0.0144	0.0170	0.2899	0.0702	LIB	1.8824	0.0112	0.0185	0.0907	0.0709
	(0.0253)	(0.0323)	(0.0258)	(0.2818)	(0.4061)		(0.0239)	(0.0120)	(0.0095)	(0.1662)	(0.4058)
SUD	1.8564	0.0204	0.0221	0.1877	0.0834	IRN	1.8658	0.0214	0.0177	0.2123	0.0702
	(0.0278)	(0.0201)	(0.0164)	(0.2124)	(0.1882)		(0.0157)	(0.0155)	(0.0115)	(0.2701)	(0.3650)
TUR	1.8914	0.0019	-0.0004	0.2253	0.0524	IRQ	1.9082	0.0198	0.0189	0.1784	0.0714
	(0.0894)	(0.0825)	(0.0664)	(0.2256)	(0.2125)		(0.0533)	(0.0229)	(0.0176)	(0.2391)	(0.4042)
EGY	1.8520	0.0023	0.0172	0.2901	0.0698	SYR	2.1083	0.0087	0.0170	0.3711	0.0707
	(0.0465)	(0.0345)	(0.0287)	(0.2560)	(0.4075)		(0.2113)	(0.0841)	(0.0717)	(0.3321)	(0.3848)
LEB	1.8962	0.0266	0.0236	0.2089	0.0402	JOR	1.9433	0.0198	0.0191	0.1241	0.0716
	(0.0939)	(0.0622)	(0.0552)	(0.2392)	(0.2457)		(0.0733)	(0.0393)	(0.0329)	(0.1834)	(0.4037)
ISR	1.8391	0.0183	0.0182	0.1804	0.0740	SAU	1.8854	0.0180	0.0209	0.1885	0.0736
	(0.0429)	(0.0134)	(0.0152)	(0.1730)	(0.3986)		(0.0808)	(0.0625)	(0.0502)	(0.2255)	(0.3956)
KUW	1.8888	0.0182	0.0192	0.1997	0.0718	BAH	1.8733	0.0063	0.0196	0.1443	0.0723
	(0.0980)	(0.0414)	(0.0321)	(0.2272)	(0.4032)		(0.0156)	(0.0364)	(0.0259)	(0.1757)	(0.4062)
QAT	1.8708	0.0232	0.0212	0.2552	0.0741	UAE	1.8746	0.0185	0.0189	0.2131	0.0718
	(0.1227)	(0.0518)	(0.0411)	(0.3129)	(0.3987)		(0.0445)	(0.0203)	(0.0133)	(0.2240)	(0.3665)
OMA	1.8797	0.0193	0.0182	0.1693	0.0707	AFG	1.9012	0.0273	0.0202	0.1375	0.0726
	(0.0282)	(0.0162)	(0.0146)	(0.2570)	(0.4062)		(0.0563)	(0.0269)	(0.0213)	(0.1821)	(0.3988)
TKM	1.8841	0.0191	0.0199	0.2080	0.0729	TAJ	1.8431	0.0228	0.0218	0.3066	0.0748
	(0.0212)	(0.0156)	(0.0113)	(0.2607)	(0.4040)		(0.0487)	(0.0321)	(0.0247)	(0.2275)	(0.4002)
KYR	1.8431	0.0171	0.0218	0.1955	0.0748	UZB	1.8608	0.0129	0.0214	0.1608	0.0743
	(0.0367)	(0.0404)	(0.0330)	(0.2160)	(0.4039)		(0.0210)	(0.0183)	(0.0134)	(0.2176)	(0.4055)
KZK	1.8563	0.0186	0.0215	0.1883	0.0744	CHN	2.0790	0.0168	0.0152	0.1761	0.0698
	(0.0329)	(0.0188)	(0.0139)	(0.2559)	(0.4052)		(0.3876)	(0.2475)	(0.2102)	(0.2163)	(0.3799)
MON	1.8330	-0.0068	-0.0246	0.1608	0.0279	TAW	1.8480	0.0198	0.0117	0.1921	0.0395
	(0.0277)	(0.0198)	(0.0251)	(0.2343)	(0.1869)		(0.0725)	(0.0762)	(0.0659)	(0.2445)	(0.1506)
PRK	1.9355	0.0179	0.0161	0.1068	0.0701	ROK	1.8577	0.0225	0.0037	0.2863	0.0733
	(0.2099)	(0.0544)	(0.0449)	(0.1801)	(0.4006)		(0.2518)	(0.0955)	(0.0823)	(0.2341)	(0.2046)
JPN	1.7171	0.0185	0.0284	0.1619	0.0624	IND	1.6707	0.0206	0.0040	0.1502	0.0699
	(0.2023)	(0.1375)	(0.1192)	(0.1430)	(0.5419)		(0.0310)	(0.0443)	(0.0362)	(0.1571)	(0.3696)
PAK	1.8169	0.0067	0.0321	0.3833	0.0444	BNG	1.8673	0.0297	-0.0072	1.0305	0.0455
	(0.0765)	(0.0490)	(0.0404)	(0.2426)	(0.1674)		(0.0367)	(0.0723)	(0.0652)	(0.3952)	(0.1721)
MYA	1.6948	0.0551	0.0344	0.3119	0.0479	SRI	1.7513	0.0412	0.0328	0.2532	0.0290
	(0.1321)	(0.1018)	(0.0826)	(0.2006)	(0.1879)		(0.1899)	(0.1046)	(0.0813)	(0.1753)	(0.2451)
NEP	1.9384	0.0090	0.0035	0.5470	0.0396	THI	1.8232	-0.0573	0.0006	0.1695	0.0413
	(0.0592)	(0.0718)	(0.0582)	(0.2928)	(0.1939)		(0.0237)	(0.1572)	(0.1165)	(0.3165)	(0.1885)
CAM	1.8808	0.0237	0.0186	0.3581	0.0716	LAO	1.6121	0.0362	0.0316	0.3543	0.0511
	(0.2869)	(0.1764)	(0.1528)	(0.2957)	(0.3984)		(0.6854)	(0.3354)	(0.2872)	(0.2662)	(0.2295)
DRV	1.8845	0.0136	0.0204	0.2178	0.0732	MAL	1.8437	0.0210	0.0218	0.2600	0.0748
	(0.3411)	(0.2784)	(0.2450)	(0.1317)	(0.3981)		(0.7790)	(0.3452)	(0.2911)	(0.2960)	(0.3965)
SIN	1.8514	0.0040	0.0180	0.1237	0.0706	PHI	1.6377	0.0166	0.0257	0.2015	0.0379
	(0.0352)	(0.0632)	(0.0491)	(0.1237)	(0.4075)		(0.1145)	(0.1574)	(0.1436)	(0.1361)	(0.2380)
INS	1.8459	0.0241	0.0185	0.2275	0.0587	AUL	1.8423	0.0183	0.0212	0.1803	0.0740
	(0.0790)	(0.0538)	(0.0504)	(0.2990)	(0.2023)		(0.0404)	(0.0722)	(0.0734)	(0.2334)	(0.3995)
NEW	1.5256	0.0210	0.0246	0.1843	0.0696						
	(0.0624)	(0.1167)	(0.0986)	(0.1256)	(0.3372)						

A7 Additional Results

In this appendix, we describe in detail various additional results mentioned in the paper.

Model fit by world region. Figure A2 presents goodness-of-fit and out-of-sample prediction results disaggregated by four regions of the world: the Americas, Europe, Africa, and Asia-Oceania. As in Figure 1, we compare the true proportion of democracies (gray) in each region with 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, and the bottom panel, from our model with two covariates. Notably, both specifications perform well at this, or indeed any, level of geographic aggregation. And, while not included in Figure A2 to avoid clutter, our learning model still significantly outperforms any alternative that ignores the role of learning.

Direct diffusion of democracy. To evaluate the possibility that our model's empirical success is simply an artifact of some alternative process of democratic diffusion that is indirectly picked up by our model's spatial and temporal flexibility, we construct a distance-weighted measure of how democratic each country's neighborhood is over time, and we reestimate our baseline model using this measure as a control for direct diffusion effects on the political cost of democracy. Specifically, we include in $X_{i,t}$ the weighted average

$$\bar{D}_{i,t-1} = \frac{\sum_{j \neq i} \exp(-\delta d_{i,j}) D_{j,t-1}}{\sum_{j \neq i} \exp(-\delta di, j)},$$

where $d_{i,j}$ denotes the distance between i and j's capitals.

To reduce the computational burden, instead of estimating δ—the parameter determining the size of each country's effective neighborhood—we consider five scenarios. For our "medium neighborhood" scenario, we set δ equal to the estimated value of the parameter governing the spatial decay of learning in our baseline specification. We also consider a "smaller" and "smallest" neighborhood scenarios, where we increase the value of δ two-fold

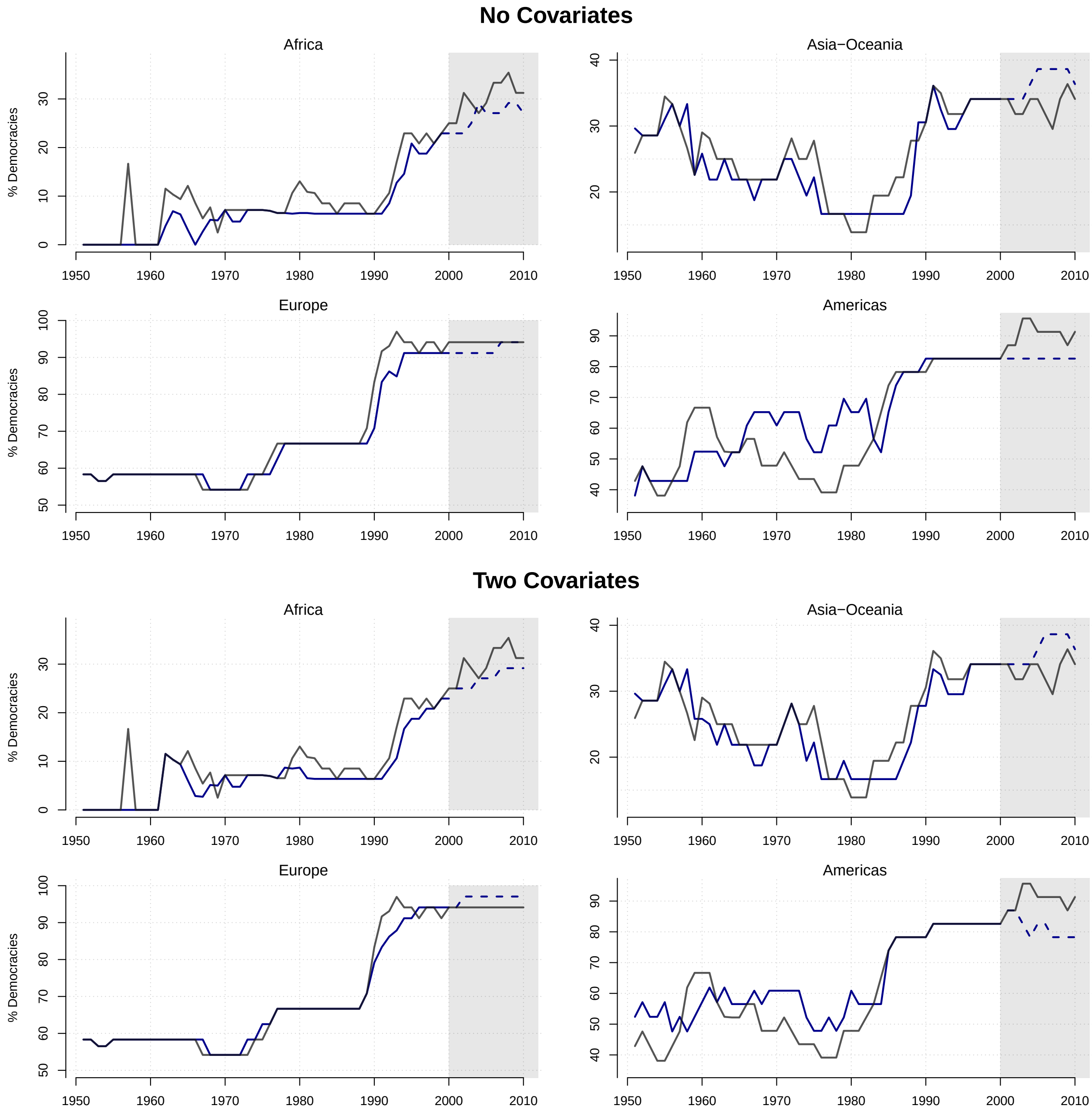

Figure A2: Observed versus Predicted Prevalence of Democracy by World Region

Notes. By world region, this figure compares the true proportion of democracies (gray) with 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.

and five-fold, respectively, and a "larger" and "largest" neighborhood scenarios, where we divide δ by two and five, respectively. Table A5 presents the results of this exercise, following the format of Table 2. We find that controlling for direct diffusion provides little increase in predictive power, which indicates that it is our proposed mechanism of learning about the economic effects of democracy—and not some alternative process of diffusion—what drives our results.

Table A5: Direct Diffusion of Democracy

	Smallest Neighborhood		Smaller Neighborhood		Medium Neighborhood		Larger Neighborhood		Largest Neighborhood	
	Learning	No Learning	Learning	No Learning	Learning	No Learning	Learning	No Learning	Learning	No Learning
Choices (% correct)	96.3	92.1	96.0	92.4	96.3	92.3	95.9	92.6	96.7	92.7
Transitions (% correct)										
±0 years	11.6	5.4	12.4	3.9	13.2	4.7	20.2	4.7	18.6	9.3
±2 years	41.9	18.6	54.3	20.9	54.3	25.6	50.4	22.5	56.6	22.5
Log-likelihood	-536.6	-1,019.1	-539.7	-964.3	-511.4	-931.7	-553.6	-939.3	-481.5	-946.2
Observations	5,925	5,925	5,925	5,925	5,925	5,925	5,925	5,925	5,925	5,925

Notes. From left to right, respectively, models in each "neighborhood" scenario control for direct diffusion effects on the political cost of democracy using distance weights $\delta = 2.5$, $\delta = 1$, $\delta = 0.5$, $\delta = 0.25$, and $\delta = 0.1$. 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.

Robustness to alternative measure of democracy and time frame. Next, we explore the sensitivity of our results to (i) our preferred measure of democracy and (ii) the time frame.

First, as noted in the paper, various alternative measures of democracy have been employed in the literature. To test the robustness of our results to this feature of the data, we take Acemoglu et al.'s (2019) preferred measure and reestimate our baseline model.[7] Since the BMR coding is more comprehensive than this alternative measure, to keep the results as comparable as possible—in particular, to avoid having to modify the time span of the sample—we use the BMR coding to fill any gaps in Acemoglu et al.'s (2019) data.

Second, to address potential concerns about the myopia of incumbents in our model and whether an annual time frame is appropriate to study changes in systems of government, we

[7]To that end, we also reestimate the true DGP (see Footnote 36) to obtain a new estimate of Σ.

estimate a five-year panel version of our baseline model (and true DGP). Following Acemoglu et al. (2008), we take the observation of democracy every fifth year as our measure for the five-year panel.

Table A6 summarizes the results of these robustness exercises, following the format of Table 2. Our findings are virtually unchanged.

Table A6: Robustness to Democracy Measure and Time Frame

	Alternative Democracy Measure		Five-Year Panel	
	Learning	No Learning	Learning	No Learning
Choices (% correct)	95.4	87.5	95.2	85.5
Transitions (% correct)				
± 0 years	12.5	0.0		
± 2 years	46.1	0.0	46.6	0.0
Log-likelihood	-647.9	-1,555.4	-160.4	-388.1
Observations	5,925	5,925	1,437	1,437

Notes. Models in the first and second columns are estimated using Acemoglu et al.'s (2019) preferred measure of democracy with no covariates. Models in the last two columns are estimated using a five-year panel version of our data with no covariates. For each model, we report the percentage of correctly predicted in-sample system of government choices (first row) and 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.

Robustness to alternative measures of similarity. While geographic distance is highly correlated with various dimensions of similarity across countries, we present in Table A7 results from an alternative version of our baseline specification in which we allow the cross-country correlation in initial beliefs to also depend on genetic distance—as measured by Spolaore and Wacziarg (2009)—and on economic distance in terms of initial levels of development. For the latter, we use a linear trend to impute values of GDP per capita in 1950 for countries without such observations. We then compute the economic distance between countries i and j as $(Y_{i0} - Y_{j0})^2/\mathrm{Var}(Y_0)$. Our results are identical.

Table A7: Robustness to Similarity Measures

	Learning
Choices (% correct)	95.2
Transitions (% correct)	
± 0 years	9.3
± 2 years	41.1
Log-likelihood	-581.4
Observations	5,925

Notes. Results are from a specification with no covariates for the political cost of democracy but three measures of similarity to capture cross-country correlation in initial beliefs: geographic distance, genetic distance, and economic distance. We report the percentage of correctly predicted in-sample system of government choices (first row) and 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.

Income and leader turnover. While we conceive elites broadly as the political party or faction in power in each country rather than as individual leaders, previous work has highlighted the influence individual leader exit can have on democratic transitions. To explore this, we estimate an alternative specification of our model that uses, as in Treisman (2015), (lagged) log-GDP per capita, (lagged) leader exit, and their interaction to characterize the political cost of democracy. As shown in Table A8, and consistent with our main results, accounting for learning dwarfs the explanatory benefits from such a specification.

Heterogeneous effects of growth on elite turnover. Finally, we explore whether there is heterogeneity across countries in $\theta = (\theta^{D=0}, \theta^{D=1})$, the effects of GDP growth on elite turnover. Specifically, we estimate a version of our model where we allow θ to vary by region of the world as in Figure A2. Table A9 reports the corresponding point estimates, and Table A10 summarizes goodness of fit. Our results provide little evidence of substantively important heterogeneity.

Table A8: Income, Leader Turnover, and Democratization

	Learning	No Learning
Choices (% correct)	95.8	90.5
Transitions (% correct)		
± 0 years	13.3	7.0
± 2 years	50.0	21.1
Log-likelihood	-557.6	-1,150.9
Observations	5,882	5,882

Notes. Results are from a specification that controls for (lagged) log-GDP per capita, (lagged) leader turnover, and their interaction as in Treisman (2015). We report the percentage of correctly predicted in-sample system of government choices (first row) and 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.

Table A9: Effects of Economic Growth on Elite Turnover by World Region

Region	$\theta^{D=0}$	$\theta^{D=1}$
Africa	-1.6859	5.0174
Asia-Oceania	-4.0284	5.7846
Europe	-1.7962	4.6478
Americas	-1.9665	5.6278

Table A10: Heterogeneous Effects of GDP Growth on Elite Turnover

	Learning	No Learning
Choices (% correct)	95.3	88.9
Transitions (% correct)		
± 0 years	7.0	0.0
± 2 years	36.4	0.0
Log-likelihood	-614.3	-1,390.8
Observations	5,925	5,925

Notes. Results are from a specification where we allow the effects of GDP growth on elite turnover to vary by world region (Africa, Asia-Oceania, Europe, Americas). We report the percentage of correctly predicted in-sample system of government choices (first row) and 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.

References

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

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A. Robinson. 2019. "Democracy Does Cause Growth." *Journal of Political Economy* 127(1):47–100.

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

Davison, Anthony C. and David V. Hinkley. 1997. Bootstrap Methods and their Application. In *Cambridge Series in Statistical and Probabilistic Mathematics*. New York: Cambridge University Press.

Dubé, Jean-Pierre, Jeremy T. Fox and Che-Lin Su. 2012. "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation." *Econometrica* 80(5):2231–2267.

Heiss, Florian and Viktor Winschel. 2008. "Likelihood approximation by numerical integration on sparse grids." *Journal of Econometrics* 144(1):62–80.

Spolaore, Enrico and Romain Wacziarg. 2009. "The Diffusion of Development." *Quarterly Journal of Economics* 124(2):469–529.

Su, Che-Lin and Kenneth L Judd. 2012. "Constrained Optimization Approaches to Estimation of Structural Models." *Econometrica* 80(5):2213–2230.

Treisman, Daniel. 2015. "Income, Democracy, and Leader Turnover." *American Journal of Political Science* 59(4):927–942.