Democratic Values and Institutions[†]

By TIMOTHY BESLEY AND TORSTEN PERSSON*

This paper builds a model of the two-way interaction between democratic values and institutions to bridge sociological research, focusing on values, with economics research, which studies strategic decisions. Some citizens hold values that make them protest to preserve democracy with the share of such citizens evolving endogenously over time. There is then a natural complementarity between values and institutions creating persistence without assuming any form of commitment. The approach unifies ideas in the literature, explains observed patterns in the data on democratic values and political institutions, and suggests new insights into sources of heterogeneity in values. (*JEL* D02, D72)

> (*I*)*f a political system is not characterized by a value system allowing the peaceful "play" of power … there can be no stable democracy.*
> —Lipset (1959, p. 71)

> *During the nineteenth century most Western societies extended voting rights, ... these political reforms can be viewed as strategic decisions by the political elite to prevent widespread social unrest and revolution.*
> —Acemoglu and Robinson (2000, p. 1167)

Looking across today's world and its history, the heterogeneity of democratic experiences is striking. Some polities have made secure transitions into democracy, and these institutions are accepted pretty much by everybody. Others have never secured democracy. A third group occupies a middle ground with a history of institutional reversals with occasional transitions to the stable groups.

Understanding what drives democratic reforms is important intrinsically, as well as instrumentally—a body of research gives political institutions a central role in explaining cross-country differences in economic growth and development (e.g., North 1990).

The initial quotes illustrate two approaches to democratic reform. Recent research in economics argues that democratic institutions and reforms are the result of strategic, forward-looking decisions by dominant groups. An older body of research in political science and sociology holds that democratic values are key in inducing and supporting democratic institutions. Although both approaches highlight important drivers, few have investigated whether joining them together generates new insights.

This paper models the drivers of democratic reforms with dynamic democratic values and strategic choices—including decisions to fight—by prospective winners and losers. Neither institutions nor values have an upper hand in democratic change; the two evolve jointly and interdependently.

The now standard model of institutional change from Acemoglu and Robinson (2000, 2006) assumes that decision-makers can commit institutions one or more periods ahead. We dispense with any commitment assumption: institutional reforms are sustainable only if they are incentive compatible for the *current* incumbent. Democratic values is the single slow-moving state variable which generates persistent change. The model allows us to interpret broad patterns of democratic reforms and values found in Polity IV (PIV) and World Value Survey (WVS) data. It also generates new predictions, including the effects on values of foreign occupations, via colonialism or the Cold War. We present some within-country correlations from the WVS consistent with these auxiliary predictions.

The next section overviews different approaches to democratic institutions and provides background facts about the dispersion of democratic institutions and values over countries and time. Section II sketches a simple model of the interplay between democratic institutions and democratic values. Section III shows how this model helps to interpret patterns of institutional dynamics and values, unifying ideas in the existing literature, and pinpointing auxiliary predictions, which are consistent with the data. Section IV concludes. An online Appendix collects supporting materials.

I. Background

A. *Related Ideas*

Cultural, value-based arguments for democracy go back to Aristotle. But the locus classicus is Montesquieu (1748), who spells out how geography and climate interact with culture to shape how alternative political institutions work. In modern political science, Lipset (1959) and Almond and Verba (1963) pioneer the argument that political culture and values are vital prerequisites for democracy.

These ideas have influenced the measurement of values and attitudes (Inglehart 1997). Drivers and consequences of values are subjects of an evolving literature, which argues that mass attitude as measured in the WVS, gauge the demand for democratic change (Inglehart and Welzel 2005) and demonstrate the willingness to struggle for democracy (Welzel 2007). Fuchs-Schündeln and Schündeln (2015) show that experience with democracy raises support for it, while Neundorf (2010) exploits Eastern European political attitudes to show that such support is considerably weaker for individuals who grew up during the Cold War. Gorodnichenko and Roland (2015) emphasize why individualistic rather than collectivist cultures are more likely to underpin democratization.

Almond and Verba (1963, p. 367) discuss how civic culture is shaped by social-ization, which "includes training in many social institutions—family, peer group, school, work place, as well as in the political system itself." Our approach builds on models of cultural evolution beginning with Cavalli-Sforza and Feldman (1981) and Boyd and Richerson (1985).

Research on culture, individual behavior, and institutions has increased among economists in recent years (Bisin and Verdier 2011). We model cultural change through the dynamics of preferences or values (rather than behavior or beliefs) fol-lowing the indirect evolutionary approach of Güth and Yaari (1992).

Acemoglu and Robinson (2000, 2006) suppose that an elite uses the franchise as a commitment device to guarantee the masses more favorable policy treatment. On top of the case studies in these works, Aidt and Jensen (2014) and Aidt and Franck (2015) provide supportive econometric evidence. Our approach follows Acemoglu and Robinson (2000) except in one key dimension. In their model, political institu-tions are a state variable causing persistence, on the argument that they are harder to change than economic policies. In our model, democratic values are the only state variable, on the argument that they move more slowly than institutions.

Closest to our approach is Ticchi, Verdier, and Vindigni (2013) who model the interaction between value formation and political reforms, giving an explicit role to education. Their model has two state variables and assumes, in common with the earlier literature, that political institutions can be committed one period ahead. Studying the coevolution of institutions and culture, Bisin and Verdier (2017) also make this assumption.

Our approach is also akin to Weingast (1997) who shows how rights can emerge as a self-enforcing equilibrium and Lagunoff (2001) who shows how greater polit-ical turnover raises support for civil liberties. However, neither has a role for dem-ocratic values.

B. *Motivating Facts*

The model links two sets of facts: heterogeneity in country-level democratic his-tories and covariation of these histories with democratic values.

We gauge each country's democratic history from the PIV, classifying it as dem-ocratic if the *polity2* variable—measured on a 20-step scale from −10 to +10—is greater than zero. When documenting the patterns of democratic reforms, we con-fine ourselves to the 50 countries that appear in the PIV data in each year from 1875 onward. We summarize the heterogeneity of country dynamics as follows:

Institutions: Histories of democratic reforms come in three broad forms: always nondemocratic, permanent transition to democracy, or churning between the two, with the churning group the most prevalent one.

Table 1 illustrates these facts, classifying each country according to its history. The left-most column shows that three of the 50 countries have never been dem-ocratic. The top of the right-most column shows a striking institutional longevity in countries with democracy from the outset (or from 1800), although transitions to democracy are more recent in countries at the bottom of the right-most column,

Table 1—Classification of Countries by Democratic History

Weak	Mixed		Strong
Always nondemocratic	Multiple changes (number upward, number downward)		Always democratic
Afghanistan	Argentina*† (7,6)	Haiti (4,4)	Canada*†
Morocco*†	Austria (3,2)	Honduras (3,2)	New Zealand†
Oman	Belgium (3,2)	Iran*† (1,1)	Switzerland*
	Bolivia (2,1)	Japan*† (2,1)	United States*†
	Brazil*† (2,1)	Liberia (1,1)	
Permanent switch to non-democracy (year of switch)	Chile*† (3,2)	Nepal (3,2)	Permanent switch to democracy (year of switch)
	China*† (1,1)	Netherlands*† (2,1)	
	Colombia*† (3,2)	Norway* (2,1)	
	Denmark (3,2)	Peru*† (8,7)	Costa Rica (1841)
	Dominican Republic (2,1)	Portugal (3,2)	El Salvador (1982)
	Ecuador† (3,2)	Paraguay (2,1)	Hungary* (1989)
	Ethiopia*† (1,1)	Serbia* (4,3)	Italy* (1945)
	France* (3,2)	Spain*† (4,3)	Mexico*† (1994)
	Germany*† (2,1)	Thailand*† (5,4)	Nicaragua (1990)
	Greece (5,4)	Turkey*† (3,2)	Romania (1990)
	Guatemala (6,5)	Venezuela (1,1)	Russia*† (1992)
			Sweden*† (1910)
			United Kingdom*† (1837)
			Uruguay* (1910)

Notes: Sample is 50 countries which appear in the Polity IV database as independent countries in 1875. The dataset covers the period 1800 to 2011 and Table 3 displays when each country first entered the data. Data for Germany are for unified Germany; West Germany had strong executive constraints from 1950 onward. A * denotes a country in wave 5 and a † denotes a country in wave 6 of the World Values Survey.

except Costa Rica and Sweden. Countries with transitions in both directions, in the middle column, are the largest group.

If we extend this table to all PIV countries, all columns have more entries. A few countries, like South Korea and Taiwan, have made single transitions to democracy while others, like Gambia or Somalia, have made single transitions in the other direction. However, as in Table 1, most countries fall into the mixed category.

To study democratic values, we use data from WVS waves 5 and 6. V. 140 asks people to rate the importance of democracy on a 10-point scale. We adopt a binary indicator: someone has (strong) democratic values if she rates democracy strictly above 8. This variable, with a global mean of about 0.6, reveals the following:

Values: Support for democracy varies across individuals and countries, with strongest (weakest) support in countries with long (short) histories of democracy.

To illustrate these facts, panel A in Figure 1 shows a positive relation between a country's share of people with democratic values (relative to the global mean) and its fraction of democratic years. Panel B shows a similar relation, when conditioning on individual gender, education, age, and income (see figure notes). Panel C shows that democratic support is about 25 percent higher in countries with a once-and-for-all entry into democracy (right column of Table 1) rather than a mixed history (left and middle columns).[1]

[1] Fuchs-Schündeln and Schündeln (2015) use a country-fixed-effects regression with WVS data to show that eight more years of exposure to democracy raises individual support for democracy by the equivalent of secondary (rather than primary) school education.
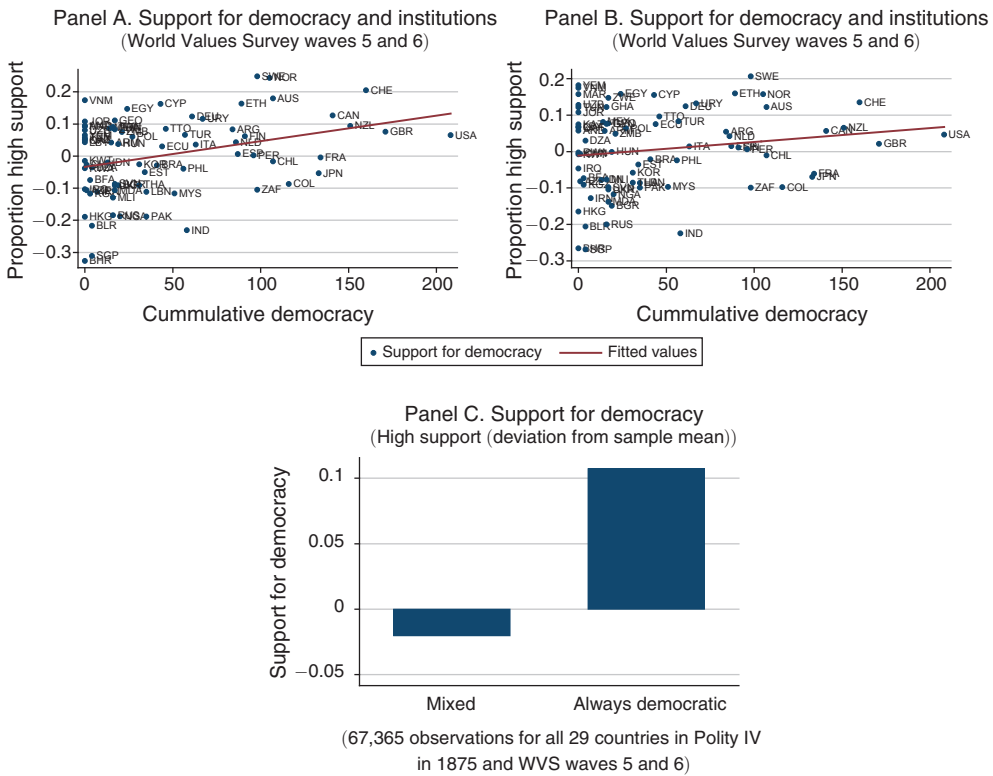
FIGURE 1. DEMOCRATIC VALUES AND DEMOCRATIC HISTORY

Notes: The data on institutions come from the Polity IV website (http://www.systemicpeace.org/polityproject. html). For democracy, we use the variable *polity2* (on a -10, $+10$ scale) to create a dummy variable which is equal to 1 if *polity2* takes a positive value in a given country-year. The horizontal axes in panels A and B display the number of years for which a country has had a 1 for this democracy dummy. Support for democracy is an individual dummy variable from the World Values Survey (http://www.worldvaluessurvey.org/wvs.jsp) waves 5 and 6 which equals 1 if the individual expresses support for democracy (on a 10-point scale) at 9 or 10. The vertical axis gives the average value of the dummy variable for each country across both waves. Panel A plots the raw data. Panel B holds constant each individual's gender, education, age, and income: we estimate an individual-level linear probability model with the dummy for democratic support on the left-hand side including on the right-hand side controls for gender, ten dummies for income groups, three for education groups, and three age bands. To construct the figure, we average the residuals at the country level. Panel C compares the values in countries (in the top right panel of Table 1 along with Sweden, the United Kingdom, and Uruguay) that have one long-standing transition into democracy with those with a recent, multiple, or no transition into democracy (in the left and middle panels of Table 1 along with Hungary, Italy, Mexico, and Russia).

II. Model

Our framework highlights a conflict of interest over democratic institutions between an incumbent group (a "political elite") and its opposition. In each period, the incumbent chooses whether to install a democracy or an autocracy, without being able to commit to future institutions. The *only* state variable is the proportion of individuals with democratic values, who may fight for democracy against autocracy.

Groups and Payoffs.—There are two groups of equal size, each normalized to measure 1. Their roles may shift across periods, as indicated by $G \in \{I, O\}$

with I denoting the incumbent and O the opposition.[2] Institutions are denoted by $D_t \in \{0, 1\}$ where $D_t = 1$ is democracy and $D_t = 0$ autocracy. Payoffs depend on this institutional indicator and the realization of a random variable $x_t \in [\underline{x}, \overline{x}]$, with distribution function $H(\cdot)$.

At realization x and institution D, group G's material payoff is denoted by $u^G(x, D)$, which we assume is (weakly) increasing in x. We make the following assumptions:

$$u^I(x, 0) - u^I(x, 1) = \Gamma(x) > 0 \quad \text{is increasing in } x \text{ for all } x \in [\underline{x}, \overline{x}]$$

and

$$(1) \quad u^O(x, 1) - u^O(x, 0) = \gamma(x) > 0 \quad \text{is increasing in } x \text{ for all } x \in [\underline{x}, \overline{x}].$$

A higher value of x implies a greater incentive for the incumbent to maintain $D_t = 0$ and a greater value to the opposition of $D_t = 1$.

Institutional Interpretation.—Why is this a plausible reduced-form model of democracy? Crucially, D_t captures a basic conflict of interest over the private material payoffs under alternative political institutions: incumbents prefer autocracy while oppositions prefer democracy. Our online Appendix sketches two examples that provide microfoundations and capture a core element of democratic institutions. The first highlights constraints on executive power—here, x_t represents some (resource) rents to be split between the two groups at t. The second example highlights open access to executive power where x_t represents the incumbent's current unpopularity—the probability that the opposition would win an electoral contest at t. However, a similar framework could be used to model the sustainability of any institutional arrangement favoring one group over another.

Types, Democratic Values, and Fairness.—Citizens are of two types, whose shares are endogenous. Fraction $1 - \mu_t$ are *passive* (type P)—if they protest, this is only due to private gains. Their date-t utility is $u^O(x_t, D_t)$. The remaining fraction , μ_t are *concerned* (type C)—a prospective civil society willing to support democracy—who care about the payoffs of society at large.[3] Concerned-citizen payoffs are $u^O(x_t, D_t) + s(x_t, D_t)$ with

$$(2) \quad s(x_t, D_t) = \begin{cases} \gamma(x) & \text{if } D_t = 1 \\ -\chi\gamma(x) & \text{if } D_t = 0 \end{cases},$$

where (2) gives a positive payoff if $D_t = 1$, a negative one if $D_t = 0$, and parameter $\chi \geq 1$ represents loss aversion by concerned citizens. These reference-dependent social preferences (Kahneman and Tversky 1979) capture how citizens value political rights. As discussed in the online Appendix, they can be microfounded

---

[2] The assumption of two groups with equal size is for analytical convenience. Other assumptions—e.g., allowing for multiple groups, or letting the incumbent elite have negligible size—would produce similar qualitative results.

[3] Democratic values are universal rather than particularistic. The complementarity of institutions and values we emphasize below would be stronger still if concerned citizens had "tribal preferences," i.e., cared only about the payoffs of other concerned citizens.

by concerned citizens judging the outcome as a gain or loss relative to their pre-ferred institution.[4] The formulation makes democratic values distinct from standard preferences, as in the distinction between acquisition utility and transactions utility, which can also reflect a sense of justice (Thaler 1999).

We assume that concerned citizens are equally distributed across the two groups. Democratic values serve two roles. They can motivate concerned citizens to pro-test. They also affect the "psychological fitness" of such citizens relative to passive citizens, because—beyond material payoffs—concerned citizens rejoice when they have democratic rights, but despair otherwise.

Concerned Citizens and Incumbent Fighting.—A successful protest can impose democracy via a successful coup or social pressure.

If a protest involves a fraction ϕ_t of citizens in period t, then the probability of success is $\phi_t p(f_t)$. Here, f_t are the resources that the incumbent devotes to preventing or fighting the protest, at a cost of wf_t.[5] This is consistent with a complementarity in collective action with a greater return to protesting when more citizens join in.[6]

Protests have a random binary cost, which is common to all individuals and denoted by $c_t \in \{\underline{c}, \overline{c}\}$ where ρ is the probability of low protest costs $c_t = \underline{c}$. Draws of c_t are iid over time. Assume that

$$(3) \qquad \gamma(x) \; < \; \underline{c} \; < \; [2 + \chi]\,\gamma(x)\,p(f) \; < \; \overline{c} \quad \text{for all } x \,\in\, [\underline{x}, \overline{x}] \text{ and } f \geq 0,$$

so that material gains are never sufficient to induce protest while democratic values can be. We assume that concerned citizens in the incumbent group never protest in support of democracy.[7] Also, function $p(\,\cdot\,)$ is decreasing and log convex, with $p(0) = 1$ and $\lim_{f \to 0} p'(0) = -\infty$ so that it is always worth devoting some resources to fighting a citizen-protest.

Democratic Values Transmission.—Over time, values follow an evolutionary dynamic based on a revision protocol (Sandholm 2010). Formally, the protocol is a continuous function $\varsigma^{I,J}(\Delta, \mu_t) \,\in\, [0, 1]$, which specifies a conditional switching rate from type I to J. Sandholm (2010) suggests a general class of dynamics that yield

$$(4) \qquad\qquad \mu_{t+1} - \mu_t \; = \; (1 - \mu_t)\,\varsigma^{P,C} - \mu_t \varsigma^{C,P},$$

where

$$\varsigma^{P,C} \; > \; 0 \;\; \Leftrightarrow \; \Delta \; > \; 0 \quad \text{and} \quad \varsigma^{C,P} \; > \; 0 \Leftrightarrow \Delta \; < \; 0.$$

---

[4] Our formulation follows Loomes and Sugden (1982) where an individual experiences either regret or rejoices depending on her reference point. This formulation is related to Passarelli and Tabellini (2017), who consider how values underpin citizens' willingness to protest against policies they regard as unfair.

[5] We do not allow the incumbent to buy off protesters, although this would lead to similar trade-offs.

[6] There could be a further complementarity if the cost of protest (per concerned citizen) would decrease with the number of participants.

[7] This could be rationalized by supposing there is a higher protest cost for such citizens due to within-group peer pressure.

We call Δ the *relative* (*psychological*) *fitness*—the *expected gain or loss*—of being a concerned citizen. The evolution of values has a "Darwinian" element: if concerned citizens have strictly higher (lower) payoffs than passive citizens, their share in the population increases (decreases) over time. The sign of $\Delta_\mu(\mu)$ affects the equilibrium dynamics (see further below).

The online Appendix shows that (4) can be given microfoundations where parents socialize their children (strategically or non-strategically). It can also be derived from a replicator-dynamic where the young are influenced by "cultural parents" and/or imitate more successful types.[8]

Timing.—The timing within a generation has four steps:

Step 1: A leader in generation t is selected from incumbent group I, and x_t is realized.

Step 2: This leader chooses D_t and f_t.

Step 3: Under democracy $D_t = 1$, the payoffs are $u^G(x_t, 1)$ for $G \in \{I, O\}$. Under autocracy $D_t = 0$, c_t is realized and citizens decide whether to protest. With an unsuccessful protest, payoffs are $u^G(x_t, 0)$ for $G \in \{I, O\}$. A successful protest imposes $D_t = 1$ and payoffs $u^G(x_t, 1)$ for $G \in \{I, O\}$.

Step 4: Payoffs are realized, a new generation is born and socialized, changing μ_t to μ_{t+1}. A non-unseated incumbent stays until period $t + 1$. With an unseated incumbent (successful protest), the opposition at t becomes the new incumbent at $t + 1$.

Preliminaries.—The online Appendix analyzes optimal fighting and protesting at stages 2 and 3. Based on these choices, we define two functions $V(x_t, \mu_t)$ and $U(x_t)$ for the incumbent's equilibrium payoffs under autocracy and democracy, respectively, and a survival function $\lambda(x, \mu)$, for the expected probability of successfully enforcing $D_t = 0$ with optimal fighting on both sides. We show that for all $\mu \in [0, 1]$ and $x \in [\underline{x}, \overline{x}]$, a higher x increases $\lambda(x, \mu)$ and $V(x, \mu) - U(x)$. That is, a higher x raises the incumbent group's gain from remaining in office and its benefit to fighting—it thus makes autocracy more attractive. A larger share of concerned citizens μ has the opposite effect: it decreases expected survival $\lambda(x, \mu)$ and the equilibrium gain from autocracy $V(x, \mu) - U(x)$.

For Proposition 1, we also need the following assumption.

ASSUMPTION 1: (*i*) *The payoff functions satisfy* $V(x, 1) - U(x) < 0$, *and* (*ii*) *there exists* $\underline{\mu} > 0$ *such that* $V(\underline{x}, \underline{\mu}) - U(\underline{x}) = 0$.

In this assumption, (i) says that it is never worthwhile to maintain autocracy if all citizens are concerned, while (ii) says that μ has a lower bound, which makes the

---

[8]Depending on the exact model, relative fitness can depend either on tomorrow's share of concerned citizens, $\Delta(\mu_{t+1})$, or today's share, $\Delta(\mu_t)$. However, the steady states of the model do not depend on this detail.

incumbent indifferent between autocracy and democracy at the lowest realization of x. A necessary condition for (ii) is that concerned citizens do protest at $(\underline{x}, \underline{\mu})$.

Equilibrium Institutions.—To choose D_t at step 2, the incumbent compares $V(x_t, \mu_t)$ with $U(x_t)$, given realized x_t, and the share of concerned citizens μ_t. Define value $\hat{x}(\mu)$ that makes the incumbent indifferent between the two: $V(\hat{x}(\mu), \mu) = U(\hat{x}(\mu))$. Then, the choice of democracy D_t satisfies the following.[9]

PROPOSITION 1: *Under Assumption* 1, *there are two values* $\mu^L < \mu^H$, *such that for*

(i) $\mu \leq \mu^L, D(\mu, x) = 0$ *for all* $x \in [\underline{x}, \overline{x}]$;

(ii) $\mu \geq \mu^H, D(\mu, x) = 1$ *for all* $x \in [\underline{x}, \overline{x}]$; *and*

(iii) $\mu \in [\mu^L, \mu^H]$ *there exists* $\hat{x}(\mu) \in [\underline{x}, \overline{x}]$ *such that* $D(\mu, x) = 0$ *if and only if* $x \geq \hat{x}(\mu)$.

The result is intuitive. With weak democratic values (low μ), protesters are unlikely to win and the incumbent leader can safely choose autocracy $D_t = 0$ and spend little on fighting. When democratic values are strong, incumbent loss is instead likely and, as fighting is costly, citizens get democracy. These polar cases hold independently of x_t. However, for intermediate democratic values, institutions depend on the realization of x_t—at high (low) x, the leader stays with autocracy (installs democracy).

Evolving Values.—Evolving democratic values reflect the relative fitness of being concerned versus passive, as determined by expected utilities at date $t + 1$ (or t). As the material payoffs of passive and concerned citizens are the same, they cancel out. Hence, only (2), the society-wide component of utility for concerned citizens, matters. This leads to the following cultural dynamics.

From (4), $\mu_{t+1} - \mu_t$ is positive (negative) whenever $\Delta(\mu_t)$ is positive (negative). Using (2) and Proposition 1, and recalling that x has cdf H, we can write the expression for $\Delta(\mu_t)$ as

$$(5) \quad \Delta(\mu) = \begin{cases} \int_{\underline{x}}^{\overline{x}} \gamma(x)\, dH(x) & \mu \geq \mu^H \\ \int_{\underline{x}}^{\hat{x}(\mu)} \gamma(x)\, dH(x) - \int_{\hat{x}(\mu)}^{\overline{x}} L(x, \lambda(x, \mu))\, dH(x) & \mu \in [\mu^L, \mu^H], \\ -\int_{\underline{x}}^{\overline{x}} L(x, \mu)\, dH(x) & \mu \leq \mu^L \end{cases}$$

---

[9] We prove this proposition in the online Appendix.

where $L(x, \lambda) = \left[\chi - \rho(1 - \lambda)(1 + \chi) \right] \gamma(x) + \rho\underline{c}$ is the loss from $D_t = 0$, which is increasing in λ. We focus on the case where $L(x, \lambda) > 0$ for all x, λ, which always holds with sufficient loss aversion χ.

There are three regions for μ. When $\mu \geq \mu^H$, democratic values have reached a point where incumbents always choose democracy $D_t = 1$ and no protests occur. The concerned have an intrinsic gain from this institution, so their share is growing. When $\mu \leq \mu^L$, the incumbent group get its preferred autocracy $D_t = 0$ for any realization of x and the few concerned individuals feel a perpetual sense of injustice, which gives them an intrinsic loss. Hence, democratic values are shrinking. In an intermediate range for democratic values, realized x determines the incumbent's institutional choice. From Proposition 1 and (2), a gain $(\Delta(\mu) > 0)$ only occurs if $D_t = 1$ which requires $x \leq \hat{x}(\mu)$. Otherwise, incumbents choose $D_t = 0$, which leads to losses as defined in (2). Democratic values grow (shrink) when expected gains exceed (fall below) expected losses, which in turn requires expected x to fall below (above) threshold $\hat{x}(\mu)$, according to distribution H. As we show in the online Appendix, $\partial\hat{x}(\mu)/\partial\mu > 0$, which implies $\Delta_\mu(\mu) \geq 0$ for all $\mu \in [0, 1]$.

From (2), the loss from being a concerned citizen is higher when x is high and the probability of a protest unseating the incumbent is low, which happens when μ is low, since the survival function $\lambda(x, \mu)$ is then close to one. At the other extreme, the loss is low when the incumbent almost surely loses a rebellion, as $\lambda(x, \mu)$ is close to zero.

Steady States and Inertia.—The possible steady states are described as follows.

PROPOSITION 2: *There exists a critical value $\hat{\mu}$ defined by*

$$\int_{\underline{x}}^{\hat{x}(\hat{\mu})} \gamma(x)\, dH(x) = \int_{\hat{x}(\hat{\mu})}^{\overline{x}} L\big(x, \lambda(x, \hat{\mu})\big)\, dH(x).$$

Whenever $\mu_0 \geq \hat{\mu}$, the polity converges to $\mu = 1$. However, for $\mu < \hat{\mu}$, the polity converges to $\mu = 0$.

To see why this is true, note that $\Delta(0) < 0$ and $\Delta(1) > 0$. Because $\Delta(\mu)$ is (weakly) monotonically increasing, there must exist a unique level $\hat{\mu}$ such that $\Delta(\hat{\mu}) = 0$. Moreover, this interior point is unstable, meaning that the dynamics described in (4) will converge slowly to either of two extremes (see the online Appendix for further discussion).

This convergence is associated with a specific path of democratic institutions. Once democratic values on an upward path reach region $\mu \geq \mu^H$, democracy becomes permanently chosen. Equally, once democratic values on a downward path reach the region where $\mu \leq \mu^L$, autocracy becomes permanent. The intermediate region for μ can have reforms in both directions depending on x_t.

To summarize, democratic institutions are persistent without assuming any form of institutional commitments. Institutional inertia reflects slow-moving democratic values which feed back to democratic reform. Democratic institutions also feed back to democratic values.

III. Insights

The model is consistent with the two motivating facts in Section IB. Its predictions encompass a range of findings discussed in existing research. Moreover, beyond reproducing the two motivating facts, the model makes some auxiliary predictions on democratic values that we may confront with data.

A. *Motivating Facts Redux*

Institutions.—Table 1 documented three groups of country histories: permanent transitions into democracy, into autocracy, and flip-flopping between the two. These correspond neatly to the predictions from Propositions 1 and 2: an upper and lower region for democratic values where democracy and autocracy become absorbing states, and an intermediate range where reforms occur in both directions due to country-specific shocks. The model predicts heterogeneous institutional responses for *temporary* shocks to x, depending on the value of μ. This, together with separate starting values μ_0, implies that countries follow their own paths which reflect an evolving state variable rather than multiple equilibria.

Values.—Figure 1 documented that people in societies that have never or rarely transitioned into democratic institutions value democracy less than people in long-consolidated democracies. Our model underpins this fact: (4) and (5), together with the complementarity between $D = 1$ and μ, imply that we should observe a larger share of citizens with high democratic values—a higher μ—today, the longer in history their society had positive and high values of Δ. This, in turn, is associated with more time spent with democratic institutions.

B. *Relationship to Existing Ideas*

Persistence.—Our model suggests a mechanism behind a long-lived effect of historical political institutions, like the colonial-origins hypothesis of Acemoglu, Johnson, and Robinson (2001). However, it also suggests why cumulated values—like social or democratic capital—may consolidate change, as in Putnam (1993) and Persson and Tabellini (2009). Even though incumbents are free to reform in any period, political institutions become sticky in equilibrium due to slow-moving democratic values.

Varieties of Reform.—The model allows different types of political reforms: "defensive," when ruling elites voluntarily relinquish political control (Acemoglu and Robinson 2000, 2006), and "offensive," when citizens force ruling elites to implement institutional change (Marx and Engels 1848, Kuran 1995).

Critical Junctures.—Except shedding light on the effect of temporary shocks, x_t, and conflicts of interest between ruling elites and opposition groups, the model also shows how *permanent* shocks might matter. Specifically, it underpins how *critical junctures* may shape long-run outcomes, as stressed by Acemoglu and Robinson (2012). Two otherwise similar countries with democratic values just above and below $\hat{\mu}$, the country-specific threshold for the dynamics, can have radically different

trajectories. Moreover, a permanent shock to the *distribution* of x around $\hat{\mu}$, can flip a country to the opposite side of $\hat{\mu}$. Propositions 1 and 2 suggest that such shifts could have long-run consequences for democratic values and institutions. For example, interpreting x_t as resource rents, resource discoveries could affect the trajectory of democratic values. This merits further investigation, especially since WVS data show a negative correlation between support for democracy and contemporaneous natural-resource intensity.

Initial Conditions.—The model also highlights the importance of historical processes that change μ or function $Q(\Delta)$. One example is the transformation of political views when the ideas of Locke (1690), Montesquieu (1748), and Paine (1776) influenced the US Founding Fathers, and challenged ruling elites elsewhere. Christian teaching and institutions may also have changed exposure to liberal thought. Our model predicts that once the democratic genie is out of the bottle and μ exceeds $\hat{\mu}$, democratic reform will be sustained.

Reversing this logic, democratic institutions installed before democratic values are built may be hard to sustain. Some postcolonial African states—Nigeria, Sudan, Somalia, and Uganda—began with European-style democratic (parliamentary) regimes, but these broke down within a decade. This could be because lacking democratic values made it hard to support defense of democracy.

Economic Growth.—The model suggests how economic development may sustain democracy. As development raises wages w, the opportunity cost of fighting rises, making it less likely that incumbents will resist democratic rights. If the costs of protests also rise with economic growth, however, this pulls in the opposite direction. But the complementarity at the heart of the model also suggests a coevolution of democratic values and the economy, capturing the predictions of modernization theorists such as Lipset (1959).

Autocracy Traps.—Our model suggests how weak democratic values may create an "autocracy trap." Russia's short democratic history (in PIV) and low democratic values (in WVS) is a case in point. Previous Soviet repression (high f) weakened democratic values and thus undermined later reform attempts, like that by Boris Yeltsin (upon a low c)—giving democracy little chance of becoming permanent. Changing Russia's trajectory would require different fundamentals or a favorable shock to values μ. Examples could be a weaker repression capacity (raising the influence of given democratic values) or lower resource rents x (cutting the additional rents to power from autocracy).

Democratic Capital.—Section II showed democratic support to be strongest in countries that made once-and-for-all democratic transitions. Persson and Tabellini (2009) interpreted institutional persistence in terms of "democratic capital." This is a classic case where state dependence and unobserved heterogeneity provide competing interpretations. Our model suggests that democratic capital may reflect an unobserved omitted variable—democratic values—rather than state dependence, i.e., past experience with democracy directly causing future democracy. Moreover, our model suggests that causality runs both ways.

C. *Auxiliary Predictions and Data*

The model makes some auxiliary predictions about values.

Foreign Occupation.—World history is replete with examples, such as colonization or Soviet occupation of Eastern Europe, where foreign powers dictate domestic political institutions. Our framework can interpret these as foreign imposition of institutions $D_t = 0$ via repressive use of force f_t.

Such historical episodes should have *persistent* effects via evolving democratic values. The dynamic complementarity between institutions and values implies that a state whose democracy is interrupted by foreign-imposed autocracy may have weaker democratic values in *future* periods.

What if foreign occupation simply replaces an existing domestic autocracy? Under the plausible assumption that a major power is more likely to enforce autocracy through repression than a domestic autocrat, an occupied country will have lower future democratic values compared to spending the same amount of time in homegrown autocracy. To see why, let $\Lambda(x, \mu)$ be the probability that autocracy persists under foreign occupation and, as before, $\lambda(x, \mu)$ the same probability under domestic autocracy. If $\Lambda(x, \mu) > \lambda(x, \mu)$, (4) and (5) imply that today's μ must be lower in an occupied country, ceteris paribus, for the same number of years spent in autocracy.[10]

Colonialism.—Colonial powers mostly established autocratic regimes, though some colonies—e.g., Australia, Canada, New Zealand, and South Africa—got elements of democracy. Acemoglu, Johnson, and Robinson (2001) distinguish extractive and inclusive institutions, which we could portray as different values of D. The empirical findings in Acemoglu, Johnson, and Robinson (2001) are then readily interpretable in our model. Maintaining $D_t = 0$ $(D_t = 1)$, colonialism may have permanently affected postcolonial democratic institutions by inhibiting (promoting) emerging democratic values.[11] Countries with repressed values would then face long-run effects of colonialism, beyond any initial efforts to bring in democratic reforms.

To shed light on this prediction, we exploit *within-country* cross-cohort variation. Taken literally, the model's generational structure translates the predicted variation in values across time into variation across cohorts. Empirically, this requires that democratic values are formed relatively early and become sticky over an individual's lifetime. Then, the model predicts individuals with their formative years under colonization to have lower democratic values than those growing up post independence. We check this against WVS data in postcolonial countries, comparing individuals who had, or had not, turned 16 (results are similar for other cutoffs) by the country-specific independence year. Thus we follow a similar approach as earlier studies of age-dependent political preferences (Alesina and Fuchs-Schündeln 2007, Kaplan and Mukand 2014).

---

[10] This follows since loss function $L(x, \cdot)$ is increasing.

[11] Acemoglu, Johnson, and Robinson (2001) use strong executive constraints—a component of the *polity2* democracy index—as a dependent variable in the postcolonial era.

Specifically, we estimate the following linear probability model:

$$(6) \qquad v_{b,c,w} = \alpha_b + \alpha_c + \alpha_w + \delta_{b,c,w} + \gamma x_{b,c,w} + \varepsilon_{b,c,w},$$

where $v_{b,c,w}$ is a dummy variable for democratic support in the WVS (as in Section IB), for an individual born in year b in country c answering the question in survey wave w. We include a full set of birth-year, wave, and country dummies $\{\alpha_b, \alpha_w, \alpha_c\}$, as well as a set of individual controls $x_{b,c,w}$ as detailed in the note to Table 2 (results are similar with 10-year cohort dummies replacing birth-year dummies). The individual treatment variable $\delta_{b,c,w}$ is a binary indicator set equal to one if the individual was 16 or older at the end of colonialism.

Table 2, column 1, shows that a smaller share of cohorts with early-life exposure to colonialism holds strong democratic values. The cross-cohort difference is about 10 percent of the overall (world) sample mean. Moreover, column 2 shows that the result holds up when we estimate the same regression on the subsample of ever colonies. This adds further credibility to the idea that democratic values reflect past political regimes as posited by the theoretical model.

Communism.—We can apply a similar logic to Cold War occupation, when the USSR absorbed some independent countries—such as the Baltic ones—and made others satellites. Among countries with WVS data, we code 16 (see the Table 2 note) as subject to Soviet occupation. The population proportion that nowadays strongly supports democracy in these countries is 0.54, versus 0.61 in non-USSR influenced countries.

Column 3 estimates a version of (6) where the treatment, $\delta_{b,c,w} = 1$, now applies to those who turn 16 before the end of USSR occupation, set at 1990 in all countries. Like in columns 1 and 2, we thus only exploit within-country cross-cohort variation in values. We find a negative and significant correlation between democratic values and formative years under Soviet influence—the same effect as for colonialism both qualitatively and quantitatively. This result echoes the finding of Neundorf (2010) how within-country intergenerational preferences for democracy in ten Eastern European countries depend on Soviet influence. Column 4 estimates this on the subsample of countries, which were ever subject to Soviet influence. Although the point estimate is the same as in column 3, the lower power in a much smaller sample makes the coefficient statistically insignificant.

IV. Conclusion

We model the two-way interaction between democratic values and institutions with a single state variable: the proportion of citizens holding strong enough values to defend democracy. Rejoicing or despair about political institutions among these citizens helps propagate democratic values via a dynamic complementarity. Institutional change becomes a gradual process, not because incumbents can commit future incumbents, but because these pay close attention to gradually evolving democratic values. Shocks along this path create the kinds of episodic change seen in the data.

TABLE 2—EXTERNAL INFLUENCE ON INDIVIDUAL DEMOCRATIC VALUES

	(1)	(2)	(3)	(4)
Colonial rule at 16	−0.062	−0.058		
	(0.015)	(0.016)		
USSR occupation at 16			−0.069	−0.067
			(0.018)	(0.088)
Individual controls	Yes	Yes	Yes	Yes
Birth-year dummies	Yes	Yes	Yes	Yes
Country dummies	Yes	Yes	Yes	Yes
Countries in sample	All	Past colonies	All	Post-USSR block
Number of countries				
Observations	140,311	103,776	140,311	25,952
R^2	0.075	0.074	0.075	0.056

Notes: All the estimates in Table 2 come from individual-level, linear-probability models, where the left-hand side variable is our dummy for a score of 9 or 10 of democratic support. We control for a wave dummy, country dummies, dummies for birth year, gender, ten dummies for income, three dummies for education, and three age bands. For the end of colonialism, we use the following list of countries from the WVS with their dates of decolonization in parentheses: Algeria (1963), Argentina (1853), Australia (1901), Bahrain (1971), Brazil (1822), Burkina Faso (1960), Canada (1867), Chile (1818), Colombia (1810), Cyprus (1960), Ecuador (1822), Egypt (1922), Finland (1917), Ghana (1957), India (1947), Indonesia (1949), Iraq (1932), Jordan (1946), South Korea (1948), Kuwait (1962), Lebanon (1941), Libya (1951), Malaysia (1957), Mali (1958), Mexico (1810), Morocco (1955), New Zealand (1907), Nigeria (1960), Norway (1905), Pakistan (1947), Peru (1821), Philippines (1898), Qatar (1971), Rwanda (1962), Singapore (1965), South Africa (1910), Taiwan (1949), Trinidad and Tobago (1962), Tunisia (1956), Uruguay (1825), United States (1776), Vietnam (1945), Yemen (1967), Zambia (1964), Zimbabwe (1980). For Soviet influence, we use data for Armenia, Azerbaijan, Bulgaria, Belarus, Estonia, Georgia, Hungary, Kazakhstan, Kyrgyzstan, Moldova, Poland, Romania, Russia, Slovenia, Serbia, Ukraine, and Uzbekistan. In columns 1 and 2, Colonial rule at 16 is a dummy variable equal to one if the individual was aged 16 or older in the year her country gained independence. In columns 3 and 4, USSR occupation at 16 is a dummy variable equal to one if the individual was 16 or older when Soviet occupation ended, which we set to 1990 for all countries. Standard errors are adjusted for clustering at the country level.

Our model bridges the cultural and strategic approaches to institutional change: democratic values and democratic reforms reinforce each other. These joint dynamics help us better understand persistence and change in political institutions across countries and time. The model can cast light on the heterogeneous country experience with democratic reform—it also allows us to be precise about critical junctures and the role of initial conditions. Finally, we present some within-country correlations consistent with the model's auxiliary predictions for the effect of foreign occupation on domestic democratic values.

The paper suggests a wider agenda. On the empirical side, our model has a number of implications, which could be explored beyond simple correlations. On the theoretical side, little research has been devoted to the codetermination of values and institutional rules. Models like ours can be deployed to study related phenomena, such as the joint dynamics of organizational cultures and organizational designs (Besley and Persson 2018).

REFERENCES

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91 (5): 1369–1401.

Acemoglu, Daron, and James A. Robinson. 2000. "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective." *Quarterly Journal of Economics* 115 (4): 1167–99.

Acemoglu, Daron, and James A. Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. New York: Cambridge University Press.

Acemoglu, Daron, and James A. Robinson. 2012. *Why Nations Fail*. New York: Crown Publishers.

Aidt, Toke S., and Raphaël Franck. 2015. "Democratization Under the Threat of Revolution: Evidence From the Great Reform Act of 1832." *Econometrica* 83 (2): 505–47.

Aidt, Toke S., and Peter S. Jensen. 2014. "Workers of the World, Unite! Franchise Extensions and the Threat of Revolution in Europe, 1820–1938." *European Economic Review* 72: 52–75.

Alesina, Alberto, and Nicola Fuchs-Schündeln. 2007. "Goodbye Lenin (or Not?): The Effect of Communism on People's Preferences." *American Economic Review* 97 (4): 1507–28.

Almond, Gabriel A., and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.

Bandiera, Oriana, Myra Mohnen, Imran Rasul, and Martina Viarengo. 2016. "Nation-Building through Compulsory Schooling during the Age of Mass Migration." Unpublished.

Besley, Timothy, and Torsten Persson. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." *American Economic Review* 99 (4): 1218–44.

Besley, Timothy, and Torsten Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton: Princeton University Press.

Besley, Timothy, and Torsten Persson. 2018. "Organizational Dynamics: Culture, Design and Performance." Unpublished.

Besley, Timothy, and Torsten Persson. 2019. "Democratic Values and Institutions: Dataset." *AER: Insights.* https://doi.org/10.1257/aeri.20180248.

Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory* 97 (2): 298–319.

Bisin, Alberto, and Thierry Verdier. 2011. "The Economics of Cultural Transmission and Socialization." In *Handbook of Social Economics*, Vol. 1, edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, 339–416. Amsterdam: Elsevier B. V.

Bisin, Alberto, and Thierry Verdier. 2017. "On the Joint Evolution of Culture and Institutions." National Bureau of Economic Research Working Paper 23375.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Cavalli-Sforza, Luigi Luca, and Marcus W. Feldman. 1981. *Cultural Transmission and Evolution*. Princeton: Princeton University Press.

Fuchs-Schündeln, Nicola, and Matthias Schündeln. 2015. "On the Endogeneity of Political Preferences: Evidence from Individual Experience with Democracy." *Science* 347 (6226): 1145–48.

Gorodnichenko, Yuriy, and Gerard Roland. 2015. "Culture, Institutions, and Democratization." Unpublished.

Güth, Werner, and Menahem E. Yaari. 1992. "Explaining Reciprocal Behavior in Simple Strategic Games: An Evolutionary Approach." In *Explaining Process and Change: Approaches to Evolutionary Economics*, edited by Ulrich Witt, 23–34. Ann Arbor: University of Michigan Press.

Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton: Princeton University Press.

Inglehart, Ronald, and Christian Welzel. 2005. *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge, UK: Cambridge University Press.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–91.

Kaplan, Ethan, and Sharun Mukand. 2014. "The Persistence of Political Partisanship: Evidence from 9/11." Unpublished.

Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *Quarterly Journal of Economics* 121 (4): 1133–65.

Kuran, Timur. 1995. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.

Lagunoff, Roger. 2001. "A Theory of Constitutional Standards and Civil Liberty." *Review of Economic Studies* 68 (1): 109–32.

Lipset, Seymour Martin. 1959. "Some Social Requisites of Democracy: Economic Development and Political Legitimacy." *American Political Science Review* 53 (1): 69–105.

Locke, John. 1690. *Two Treatises of Government*. Cambridge Texts in the History of Political Thought. Cambridge, UK: Cambridge University Press, 1988.

Loomes, Graham, and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty." *Economic Journal* 92 (36): 805–24.

Marshall, Monty G., and Keith Jaggers. 2005. "POLITY IV Project: Political Regime Characteristics and Transitions, 1800–2004." http://www.columbia.edu/acis/eds/data_search/1080.html.

Marx, Karl, and Friedrich Engels. 1848. *Manifesto of the Communist Party*. Marxist Internet Archive, 2004. www.marxists.org/archive/marx/works/1848/communist-manifesto/index.htm.

Montesquieu, Charles de. 1748. *The Spirit of the Laws*. Cambridge, UK: Cambridge University Press, 1989.

Neundorf, Anja. 2010. "Democracy in Transition: A Micro Perspective on System Change in Post-Socialist Societies." *Journal of Politics* 72 (4): 1096–1108.

North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. Cambridge, UK: Cambridge University Press.

Paine, Thomas. 1776. *Common Sense, the Rights of Man, and Other Essential Writings*. St. Louis: Turtleback Books, 2003.

Passarelli, Francesco, and Guido Tabellini. 2017. "Emotions and Political Unrest." *Journal of Political Economy* 125 (3): 903–46.

Persson, Torsten, and Guido Tabellini. 2008. "The Growth Effect of Democracy: Is it Heterogeneous and How Can It Be Estimated?" In *Institutions and Economic Performance*, edited by Elhanan Helpman, 544–85. Cambridge, MA: Harvard University Press.

Persson, Torsten, and Guido Tabellini. 2009. "Democratic Capital: The Nexus of Political and Economic Change." *American Economic Journal: Macroeconomics* 1 (2): 88–126.

Putnam, Robert D., Robert Leonardi, and Raffaella Y. Nanetti. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton: Princeton University Press.

Sandholm, William H. 2010. *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press.

Siedentop, Larry. 2014. *Inventing the Individual: The Origins of Western Liberalism*. Cambridge, MA: Harvard University Press.

Thaler, Richard H. 1999. "Mental Accounting Matters." *Journal of Behavioral Decision Making* 12 (3): 183–206.

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

Weber, Max. 1922. *Economy and Society*. Berkeley, CA: University of California Press, 1968.

Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *American Political Science Review* 91 (2): 245–63.

Welzel, Christian. 2007. "Are Levels of Democracy Affected by Mass Attitudes? Testing Attainment and Sustainment Effects on Democracy." *International Political Science Review* 28 (4): 397–424.