



Electoral Contestation: A Comprehensive Polity-Level Analysis

John Gerring, Allen Hicken, Daniel Weitzel, and Lee Cojocaru

August 2018

Working Paper

SERIES 2018:73

THE VARIETIES OF DEMOCRACY INSTITUTE



UNIVERSITY OF GOTHENBURG
DEPT OF POLITICAL SCIENCE

Varieties of Democracy (V-Dem) is a new approach to conceptualization and measurement of democracy. The headquarters – the V-Dem Institute – is based at the University of Gothenburg with 17 staff, and a project team across the world with 6 Principal Investigators, 14 Project Managers, 30 Regional Managers, 170 Country Coordinators, Research Assistants, and 3,000 Country Experts, the V-Dem project is one of the largest ever social science research-oriented data collection programs.

Please address comments and/or queries for information to:

V-Dem Institute

Department of Political Science

University of Gothenburg

Sprängkullsgatan 19, PO Box 711

SE 40530 Gothenburg

Sweden

E-mail: contact@v-dem.net

V-Dem Working Papers are available in electronic format at www.v-dem.net.

Copyright © 2018 by authors. All rights reserved.

Electoral Contestation: A Comprehensive Polity-Level Analysis*

John Gerring

Department of Government
University of Texas at Austin

Allen Hicken

Department of Political Science
University of Michigan

Daniel Weitzel

Department of Government
University of Texas at Austin

Lee Cojocaru

Department of Political Science
Boston University

* This research project was supported by Riksbankens Jubileumsfond, Grant M13-0559:1, PI: Staffan I. Lindberg, V-Dem Institute, University of Gothenburg, Sweden; by Knut and Alice Wallenberg Foundation to Wallenberg Academy Fellow Staffan I. Lindberg, Grant 2013.0166, V-Dem Institute, University of Gothenburg, Sweden; as well as by internal grants from the Vice-Chancellor's office, the Dean of the College of Social Sciences, and the Department of Political Science at University of Gothenburg. We performed simulations and other computational tasks using resources provided by the Notre Dame Center for Research Computing (CRC) through the High Performance Computing section and the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre in Sweden, SNIC 2017/1-407 and 2017/1-68. We specifically acknowledge the assistance of In-Saeng Suh at CRC and Johan Raber at SNIC in facilitating our use of their respective systems.

Abstract

The study of electoral contestation generally focuses on districts or regions rather than polities. We present a new dataset that measures electoral contestation through historical records of elections in sovereign and semi-sovereign polities throughout the world from 1789 to the present. We also offer a new index of contestation intended to capture multiple dimensions of this complex concept. Our second objective is to explain variation across polities and through time in electoral contestation. We argue that the degree of contestation in a polity is affected by demography, with larger polities fostering greater electoral contestation. This hypothesis is tested with a series of cross-national regression tests that employ a variety of specifications and estimators –cross-sectional, fixed-effect, and instrumental variable. We find a robust association between population and contestation extending throughout the modern era.

Introduction

Electoral contestation (aka competition) is widely regarded as a central element of democracy (Becker 1958; Dahl 1956, 1971; Sartori 1976: 217; Schumpeter 1942/1950; Strom 1992). Studies indicate that enhanced competitiveness in votes or seats, or frequent turnover, is associated with more programmatic politics (Keefer 2006; Keefer & Vlaicu 2008; Lachat 2011), greater activity on the part of representatives (Konisky & Ueda 2011), greater responsiveness or accountability (Ansolabehere et al. 2001; Beer & Mitchell 2004; Gordon & Huber 2007; Griffin 2006; Jones 2013; Powell 2000; but see Brunell 2008; Cleary 2007; Fiorina 1973), greater efficiency and lower political rents (Barro 1973; Stigler 1972; Wittman 1989, 1995), political reform (Borges 2008; Geddes 1991, 1994; Grzymala-Busse 2007; Heller, Kyriacou & Roca-Sagalés 2011; Ting et al. 2013), lower corruption (Weitz-Shapiro 2012), lower levels of political protest (Arce & Mangonnet 2013), and stronger growth performance (Berkowitz & Clay 2012; Besley, Persson & Sturm 2010; Padovano & Ricciuti 2009).¹

Although a good deal of work has accumulated on these subjects, the study of contestation is limited in two respects. First, it generally focuses on districts or regions rather than polities, meaning that we know a lot less about contestation at national levels than at subnational levels. Second, extant work generally focuses on the probable effects of contestation, rather than its causes. To the extent that contestation affects outcomes we care about, we should also be interested in its determinants.

This study makes several contributions. First, we measure electoral contestation through historical records of elections in sovereign and semi-sovereign polities throughout the world from 1789 to the present, producing a new dataset with ~36,000 observations. Second, we offer a new measure of contestation intended to capture multiple dimensions of this complex concept. Third, we explain variation across polities and through time. We argue that the degree of contestation in a polity is affected by demography, with larger polities fostering greater electoral contestation.

The first section of the paper lays out our explanatory framework. The second section discusses issues of conceptualization and measurement, introduces the data, illustrates various

¹ Of course, contestation can have deleterious effects if parties utilize para-military groups to round up supporters and intimidate opponents, as occurred in Weimar Germany and contemporary Jamaica (Clarke 2006). We should also bear in mind that in winner-take-all elections competition may entail a decline in responsiveness and trust (Brunell 2008; Brunell & Buchler 2009; Buchler 2005, 2011). Since only one candidate or party can win, all those supporting the losing candidate or party are not likely to be well-represented and are likely to be less trusting of the candidate who wins. In a highly competitive district, this is nearly half of the voters. (The same dynamic does not apply in an electoral system in which votes are allocated proportionally to candidates and there are multiple seats in a single district.) There is also work that finds that high levels of competitiveness can exacerbate clientelism and corruption (Lindberg and Morrison 2008; Nyblade and Reed 2008; Golden and Tiwari 2009). These caveats notwithstanding, contestation seems to be associated with more good things than bad. For a wide-ranging discussion, see Bardhan & Yang (2004).

aggregation techniques, and compares our index with alternatives. The third section displays historical patterns in electoral contestation throughout the world. The fourth section tests the argument with a series of cross-national regression tests that employ a variety of specifications and estimators – cross-sectional, fixed-effect, and instrumental-variable. We find a robust association between population and contestation extending throughout the modern era.

I. Explanatory Framework

Many factors assuredly contribute to electoral contestation; like democracy, it is a complex outcome sensitive to innumerable inputs (Coppedge 2012). Here, we focus on a distal feature – demography – that we believe to be sizeable and persistent, perhaps more than any other causal factor.

We begin with a number of core assumptions, which we regard as plausible even if not entirely provable. We assume that leaders (the individual or group that controls the executive) would prefer to monopolize power but also value personal security (for their life and property, and the lives and property of their supporters), revenue, and territory. While citizens may have no *ex ante* preference for a particular constitutional structure, we assume that they have a strong preference to be governed by “one of their own,” i.e., someone whom they trust, who hails from the social group with which they identify, and who is assumed to represent the interests of that group.²

Now, let us consider the impact of demography. As the size of a community grows, the challenges of governance become more complex. This is a product of sheer numbers and also of the greater social diversity that usually accompanies greater size.

By diversity, we mean the *number of viable social groups* in a polity, as defined by ethnicity, religion, language, ideology, social class, and/or region. This should be distinguished from fractionalization indices, which measure the population *distribution* across social groups. The first is strongly correlated with polity size, while the second is only weakly correlated (AUTHOR FORTHCOMING). Note also that social groups may be defined in many different ways. In some polities, religion is paramount, in others language or ethnicity – or some other factor not captured in any standard measure. Regional identity may be especially important as parties often have a regional base, which means that a greater number of regional identities translates into a larger party system (*ceteris paribus*). If social groups are regionally defined one would expect this to foster greater contestation even if the regions do not have highly distinctive ethnic, religious, or linguistic

² This follows from studies of in-group trust (Foddy et al. 2009)

characteristics (as in South Korea). So, diversity refers here to *any* sort of social identity, a matter that we expect to vary across polities and through time.

Size and diversity, in turn, generate a greater number of political conflicts (Raleigh & Hegre 2009: 224), greater social distrust (Putnam 2007), and greater political distrust (Denters 2002; Hansen 2013; Matsubayashi 2007; Rahn & Rudolph 2005).

It follows that in a *small* society leaders may satisfy everyone's preferences without resort to elections if the leadership group is viewed as belonging to the dominant social group. And, if multi-party elections are allowed the leadership group is likely to claim a vast majority of the votes. Because elections are mute (either they are not held at all or they are monopolized by a single party), political demands will be handled through mechanisms of consultation and patronage.

By contrast, in a *large* society the leadership group will find it more difficult to satisfy citizens' desire to be governed by one of their own. While leaders may pretend to represent everyone, this claim is unlikely to be very convincing in a diverse society. Likewise, mechanisms of consultation and patronage are unlikely to operate effectively in a society of tens of millions, or hundreds of millions. Private agreements, informal understandings, and personal relationships will be harder to monitor, harder to enforce, and less legitimate.

Thus, while electoral contestation may not be anyone's first choice, it may provide an optimal solution in a large (heterogeneous) society because it promises representation for multiple groups, satisfying the core demand that citizens be ruled by one of their own. Even those not embraced by the current ruler or ruling party may enjoy a share of power in the legislature and may reasonably hope to govern someday in their own right. Importantly, the same demographic factors that augur for an election-based system of leadership selection also augur for a fragmented electoral field in which no single party is likely to gain a majority of votes or seats and in which there is likely to be frequent rotation among leadership groups. Demography thus drives all three dimensions of contestation.

Granted, size also poses coordination problems for citizens. In a model of democratization in which citizens must work together in order to achieve their preference for democracy, this is problematic. However, we do not assume a preference for democracy or the narrower goal of electoral contestation, and our framework does not depend upon coordination among various out-groups. Note that many forms of dissent can be exercised unilaterally. A single social group may refuse to pay taxes, agitate for independence, wage a guerrilla war, or sponsor political assassinations if their desire to be governed by one of their own is violated (i.e., if they feel excluded from power). Insofar as rulers value security, revenue, and territory (core assumptions in our model), these are costly dissents. If multi-party elections can solve these problems, while still

offering the leader some degree of control, this may be seen as an acceptable solution. Thus, while leaders are unlikely to look favorably upon electoral contestation they may be more inclined to acquiesce in a large society than in a small society, where monopolization of power is easier to maintain.

By way of conclusion, it is important to emphasize that this is a theory of contestation, not of democratization. Indeed, size poses problems for democracy, which may account for why studies have found no consistent relationship between the population of a polity and its regime-type (author forthcoming). We pointed out that a larger polity is likely to foster a greater number of conflicts. This may negatively impact the quality of elections as well as relationships among political institutions (legislature, executive, judiciary, independent agencies, the press), inhibiting the independence of each body. Insofar as social peace and consensus on core policies are required for successful democratic consolidation large societies are probably at a disadvantage. It is easier to respect civil liberty if no social group is agitating for separation or threatening to overthrow the state. Thus, we make no claims about the relationship of demography to democratization, writ large. Our claims are limited to one component of democracy, electoral contestation.

II. Measuring Electoral Contestation

As with any abstract concept, electoral contestation invites many approaches to operationalization (Bartolini 1999, 2000; Strom 1989). In this section, we discuss data sources, introduce a new index, discuss alternate measures, and compare our index with indices of democracy.

Data

Our goal is to measure contestation across all polities in the modern era, i.e., from 1789 to the present. Polities include formally sovereign countries and also defunct countries (e.g., DDR), semisovereign territories (e.g., Bermuda), and colonies (e.g., Gold Coast).

Fortunately, national elections are high-profile events and election results are recorded in a number of places. It is important to bear in mind that these are *official* results, as posted by the government or electoral management body. They may, or may not, reflect the intentions of voters, not to mention the intentions of citizens who did not vote. Access to the ballot may be limited, access to polling places may be restricted, and election fraud and coercion may affect the officially posted results. Electoral contestation refers to the degree of contestation that governments are willing to permit, and to report. A government that suppresses contestation receives a lower score on our indices, which seems appropriate given that electoral contestation is unlikely to have any impact on electoral accountability if the government does not recognize it.

Core data sources include Caramani (2000), Nohlen (2005), Nohlen, Grotz & Harmann (2002), Nohlen, Krennerich & Thibaut (1999), and Nohlen & Stover (2010). These are supplemented, as needed, by Wikipedia entries and the homepages of national parliaments. For some national elections, district-level results are also available. However, the largest district-level databases – the Constituency-Level Elections Archive [CLEA] (Kollman et al. 2011) and the Multi-level Elections Archive [MLEA] (Gerring et al. 2015) – do not have broad polity and historical coverage. Consequently, we cannot use these district-level databases to construct a national-level database.

Sources generally do not record election results for all parties. We make a strategic decision to collect results for the top three parties in each election (though even here data is not entirely complete). This data is collected for national elections to the lower chamber or unicameral chamber of the legislature and the presidency (if the head of state is directly elected). For each, we record the number of votes obtained by the top three contestants, and – for legislative elections – the number of seats. Contestants are identified by party if there is a party affiliation; otherwise, they are noted as independent. We also record the *total* number of votes and seats so that vote- and seat-shares can be calculated. With this data, we undertake to measure the level of contestation in legislative elections (based on votes or seats) and presidential elections (based on votes).

A New Index

Our understanding of contestation (aka competition) is guided by the goal of electoral accountability – the capacity of electors to monitor elected officials and, subsequently, to reward or punish them through the electoral process.³ Accordingly, parties rather than individual candidates are regarded as vehicles of contestation, as the former cannot be held accountable across various branches of government and over the longer-term. Indeed, where elections are fought among independent (non-party affiliated) candidates this is often a conscious ploy on the part of elites to fragment political opposition.

Electoral contestation thus refers to the degree of party-based competition within a polity. Where contestation is high, no single party is able to dominate the electoral field. Monopolization of power is prevented, and elected officials are under continual threat of losing their jobs.⁴ Uncertainty reigns.

³ Contestation affects other aspects of politics as well. For example, high contestation generally enhances turnout, and for this purpose a somewhat different operationalization of the concept may be required (Grofman & Selb 2009).

⁴ We will largely focus on political parties as the relevant actor of interest, rather than individual candidates, since these organizations are generally more suitable as vehicles for collective accountability. However, in the absence of (strong) parties, many of the same dynamics we identify would apply to individual politicians.

So defined, electoral contestation may be operationalized according to three empirical dimensions. There is, first of all, (a) the existence of elections. A polity with no elections has zero contestation, by definition. Where elections occur, the degree of contestation in a polity may be measured by (b) the closeness of the outcome (“competitiveness”) and (c) the occurrence of turnover (“alternation”). We assume that turnover is essential for obtaining electoral accountability, and hence an integral empirical dimension of the concept.

To integrate these dimensions into a single index we propose an *incumbent-challenger* formula. This is calculated as the incumbent share (of votes or seats) minus the share of the largest challenger, subtracted from 100. “Incumbent” is defined as the plurality winner in the previous election. “Challenger” is defined as the runner-up in the current election. When the incumbent falls into second place, or further down, we judge that “turnover” has occurred. The resulting measure varies, in principle, from 0 (no contestation) to 200 (the incumbent party receives no votes or seats), with 100 signaling the point at which turnover occurs.

In a polity’s first election, or the first election after an interregnum (caused by a seizure of power, a discontinuation of elections, a new constitution, foreign occupation, or some other mishap), there is no incumbent. Here, the largest party is treated as the incumbent. “First” elections cannot receive a score above 100 because turnover is (by definition) impossible.

Where the incumbent party drops out entirely, or falls below third place, we view this as an instance of party failure rather than electoral contestation. Sometimes, party failure is a product of political unrest or persecution. Sometimes, it is a product of political reorganization (indeed, a new party may contain many of the members of the old party). And sometimes, it is a product of a party that served as a vehicle for a particular candidate, who drops out leaving the party bereft of supporters. None of these scenarios conform to theoretical models of electoral accountability, where a degree of continuity among major parties is assumed. Thus, in those instances where the incumbent party vote share falls below the vote share of the third largest party, we treat the latter as the “incumbent” in our calculations.

Where multiple elections for a single office (president or parliament) occur in a single year we record the *last* of these elections as the value for that year. (Multiple elections are an infrequent occurrence. Taking the *average* value across elections within that year has virtually no effect on our index or on any of the results reported below.)

We code all years *prior* to the first election or during an *electoral interregnum* (when elections are interrupted or the elective body is prorogued) as zero (0), signaling an absence of electoral contestation. This coding is based on the Electoral regime (“elecrec”) variables, calculated for

legislative and presidential elections respectively, in the Varieties of Democracy dataset (Coppedge et al. 2018).

Years *in between* elections are filled in with results from the previous election unless there is an electoral interregnum, as discussed. The assumption is that results from the last election characterize the state of electoral contestation until the next election.

This set of coding procedures may be applied to any elective body – collective (legislative) or unitary (presidential). Election results may be measured in votes (for legislative or presidential elections) or seats (for legislative elections).

Since votes are informative across all elective bodies, we choose this outcome for our core index, which combines results from legislative and presidential elections. In parliamentary systems, the value of the composite index is solely a product of legislative elections. In presidential systems, we count elections to both branches separately and combine them by taking the mean value. The assumption of equal weighting presumes that both elections are highly salient (even if not equally consequential), and thus relevant for understanding a polity's overall level of electoral contestation. We do not presume to judge the relative power of different branches, a feature that depends upon both formal powers and informal practices and is likely to change over time – a difficult measurement issue that lies orthogonal to our theoretical concerns (Shugart & Carey 1992).

The resulting index refers to a party's *electoral* status, measured in votes or seats. This may or may not translate into control over the legislature or the executive. The largest party in the legislature may be shut out of power due to coalitional politics; very occasionally, the largest vote-getter in a presidential election does not obtain office (as in the US election in 2016). Our index registers electoral power, not policymaking power, although we assume the two concepts are highly correlated.

Finally, it is important to emphasize that while electoral contestation is a component of electoral democracy it is not an adequate proxy for that broader concept, as discussed in Appendix C.

Largest-party and Top-two Indices

The incumbent-challenger formula may be contrasted with other approaches to measuring electoral contestation. Foremost among these are formulas based on the largest party and the top-two parties.

The *largest-party* formula focuses on the vote- or seat-share of the largest party, subtracted from 100. This index varies, in principle, from 0 (monopolization of votes or seats by one party

or candidate) to 99 (where the largest party or candidate wins an infinitesimal share of votes or seats).

This approach has the benefit of broad coverage and ease of coding and interpretation. However, it takes no notice of the performance of other parties/candidates. If a party wins 51% of the votes or seats the contestation score is 49, regardless of how votes or seats are distributed among the challengers. An election in which the most successful challenger earns 49% of the votes is coded the same as an election in which ten challengers divide up the 49%. This is problematic insofar as a fragmented opposition makes it more difficult to solve coordination dilemmas, necessary to provide an effective counterweight to the dominant party.

The *top-two* formula is calculated as 100 minus the difference between the first and second place parties. The resulting measure ranges from 0 (where one party takes all the votes or seats) to 100 (where two parties tie). This approach allows one to distinguish between a unified and divided opposition.

However, neither of these traditional formulas is capable of distinguishing instances of turnover from instances of single-party dominance. A polity where a single party wins every election receives the same score as a polity where different parties alternate in power (so long as the vote shares of the winning parties in both polities are the same). Likewise, a polity where parties alternate in power, each winning by large margins, receives a *lower* score than a polity in which a single party wins every election by a close margin.

Table 1 illustrates how the largest-party and top-two formulas compare with the incumbent-challenger formula by looking at a set of hypothetical electoral outcomes and the scores each measure would produce. All three formulas identify profile #1 as the least competitive. By contrast, the profiles with the highest contestation are different across the three measures. Under the largest-party measure, profile #5 has the highest level of contestation. Profiles #2 and #3 are the most competitive under the top two measure, while #5 has the highest level of contestation under our incumbent-challenger measure. Another advantage of our incumbent-challenger measure is apparent when we compare profiles #2 and #3. The first two measures give us no way to distinguish between these two profiles. However, incumbent-challenger takes into account that there is turnover in which party is the largest vote-getter in profiles 3 & 5, reflecting scores above 100.

Each of these three formulas (incumbent-challenger, largest-party, top-two) may be applied to different outcomes (legislative and presidential votes, legislative votes, and legislative seats), following the data and coding procedures described in the previous section. This generates a typology with nine possible indices of contestation, as shown in panel (a) of Table 2. Our

preferred index – the incumbent-challenger index applied to legislative and presidential votes – is in cell #1. (We omit the purely presidential index as it applies only to presidential systems, a small and unrepresentative sub-sample of the world.)

Table 1: Three Contestation Formulas Illustrated

<i>Election results</i>			Largest-Party 100-A	Top-Two 100-(A-B)	Incumbent-Challenger 100-(I-Ch)
1.	A (I)	80			
	B (Ch)	15	20	35	35
	C	05			
2.	A (I)	51	49	98	98
	B (Ch)	49			
3.	A (Ch)	51	49	98	102
	B (I)	49			
4.	A (I)	51			
	B (Ch)	25	49	74	74
	C	24			
5.	A (Ch)	35			
	B (I)	30	65	95	105
	C	25			

Note: Hypothetical election results, showing vote shares for the top parties and the coding they would receive according to three contestation formulas. *I* = incumbent party, *Ch* = largest challenger party.

Panel (b) displays descriptive statistics for each of these indices. It will be seen that slightly more data is available for seats than for votes, and more data is available for the largest-party formula than the other formulas, which require information for two parties rather than just one. All measures contain a mode at zero, representing a non-electoral period (before elections are established or during which elections are suspended) or an election in which the largest party or incumbent wins all the votes or seats.

Importantly, these six measures of contestation provide a more extensive sample than is provided by any extant measure of politics or political institutions. By way of comparison, the democracy index with the most extensive country and historical coverage offers ~17,000 country-years (Skaaning et al. 2015). Our samples extend to 192-94 polities, ~3,000 elections and ~36,000 polity-year observations. This is primarily because we are able to measure contestation in semisovereign units such as colonies, while traditional indices are limited to sovereign units. A full list of countries and years in our dataset is contained in Table A4. (Readers should bear in mind that samples are smaller in many of the regression analyses due to missing data on relevant covariates.

Table 2: Nine Contestation Indices

	Legislative/ presidential votes	Legislative votes	Legislative seats
Incumbent-challenger	1*	2	3
Largest-party	4	5	6
Top-two	7	8	9

(a) Typology: three formulas applied to three electoral outcomes

	<i>Polities</i>	<i>Elections</i>	<i>Observations</i>	<i>Mean</i>	<i>Std dev</i>	<i>Zeros (%)</i>	<i>Max</i>
1	194	2,921	35,814	23	39	70	190
2	192	2,834	34,812	22	39	73	190
3	194	3,179	36,585	23	39	70	193
4	194	2,978	36,198	14	24	70	125
5	192	2,868	35,005	14	25	73	94
6	194	3,224	36,871	14	24	70	100
7	194	2,960	36,038	22	36	69	125
8	192	2,868	35,005	22	37	71	125
9	194	3,185	36,660	22	35	67	100

(b) Descriptive statistics (numbers rounded to nearest integer)

	1	2	3	4	5	6	7	8	9
1	1								
2	0.98	1							
3	0.90	0.93	1						
4	0.91	0.89	0.78	1					
5	0.90	0.91	0.80	0.97	1				
6	0.81	0.83	0.85	0.84	0.89	1			
7	0.95	0.92	0.81	0.95	0.94	0.81	1		
8	0.93	0.93	0.83	0.93	0.95	0.84	0.98	1	
9	0.85	0.86	0.90	0.84	0.88	0.95	0.87	0.90	1

(c) Intercorrelations (Pearson's r)

	1	2	3	4	5	6	7	8	9
1	1								
2	0.95	1							
3	0.78	0.85	1						
4	0.72	0.59	0.44	1					
5	0.61	0.62	0.49	0.90	1				
6	0.42	0.47	0.64	0.60	0.76	1			
7	0.83	0.72	0.51	0.85	0.74	0.51	1		
8	0.75	0.76	0.56	0.76	0.81	0.61	0.94	1	
9	0.53	0.56	0.73	0.57	0.64	0.89	0.64	0.72	1

(d) Intercorrelations (Pearson's r) – zero values excluded

Panel (c) displays a correlation table, showing how closely the nine indices track each other. A principal component analysis, shown in Table A3, demonstrates that nearly all of the variance (0.95) is contained in the first component, suggesting that these nine indices are measuring the

same underlying latent trait. However, when zero scores are excluded (i.e., multi-party elections are allowed), scores differ appreciably across the indices, as shown in panel (d).

Other Alternatives

Having introduced nine indices, we turn to several other alternatives that are more complex – and, for that reason, harder to measure. We argue that our preferred measure – the incumbent-challenger index – is a more appropriate choice as an overall measure of electoral contestation at national levels.

Competition (Vanhanen). The only extant global dataset that measures contestation at national levels was produced several decades ago by Tatu Vanhanen (2000: 253-55). Vanhanen adopts the largest-party formula with several additional considerations that set it apart from options illustrated in Table 1. First, where information on vote shares are unavailable, seat shares are substituted. Second, if competitors in legislative elections are independent candidates rather than organized parties the largest party is assigned a score of 30%. Third, if the vote (or seat) share garnered by the largest party falls below 30% it is nonetheless assigned a score of 30%, under the assumption that any further diminution is a product of electoral system laws and is irrelevant to the quality of democracy. Fourth, if the executive is unelected (e.g., a monarch, military leader, revolutionary group), the largest party is assumed to have won 100% of the vote. Fifth, in polities with a separately elected chief executive, results for presidential and legislative elections are combined based on a judgment of how dominant each branch is. If branches are co-equal, each is assigned a weight of 50%; if the executive is more powerful it receives a weighting of 75%; and so forth. It is not possible to tell from the data and supporting materials how many observations, or which observations, these five ad hoc coding principles affect.

Vanhanen’s special coding rules may be justifiable in light of his theoretical goal to construct a composite measure of democracy, which he derives from a multiplication of two indices: Competition (above) and Participation (turnout in each election). Our goal, however, is focused narrowly on contestation (aka competition). For this purpose, we need a more finely-grained set of measures.

Party-system measures. A rather different approach to measuring contestation encompasses the entire party system via indices of “fractionalization,” “effective parties,” or the entry and exit of new parties (Harmel & Robertson 1985). This approach draws on a market model of politics, where the dispersion of providers (parties or firms) is thought to provide a reasonable test of how open the market is and hence how efficient is it likely to be (Becker 1958). In a polity, however, goods are not provided by individual parties but rather by the party, or parties, who control the

executive. Since only a few – generally large – parties are in a position to occupy that policymaking role, smaller parties matter only insofar as they affect the behavior of those larger actors. Arguably, a measure of contestation should focus on the parties that matter most – for voters, for setting policy, and for achieving accountability. As a practical matter, sources do not consistently provide vote and seat totals for very small parties, so we cannot implement the party system approach in a comprehensive fashion.

Effective competition (Altman/Pérez-Liñán). A more nuanced approach incorporates all parties but with concern for the relative coherence of government and opposition forces. In this vein, Altman & Pérez-Liñán (2002) develop an index of effective competition that is applied to eighteen Latin American polities from 1978 to 1996. Unfortunately, data is not available to extend this measure to a large sample of polity-years.

Volatility. Volatility is understood generally as the shift in votes or seats from one election to the next. High levels of volatility might conceivably strike fear in the hearts of politicians, increasing accountability. This aspect of contestation is captured to some extent in our incumbent-challenger formula, but only in relation to the major parties. In any case, from the perspective of electoral accountability, volatility is an imperfect measure of contestation. At high levels of volatility, where party systems are constantly churning – with new parties arising and old parties falling in every electoral cycle – voters will find it difficult to exercise collective accountability and party leaders will tend to adopt very short time-horizons (Hicken 2018). Thus, we do not view volatility as an adequate measure of contestation.

Comparisons. In Table 3, we examine correlations between our incumbent-challenger formula and these alternatives. The latter are drawn directly from the cited studies and cover varying samples, dependent upon data coverage. Because some extant measures focus on votes and others on seats we include two indices based on the incumbent-challenger formula – the first focused on legislative and presidential votes and the second focused solely on legislative seats.

It will be seen that Vanhanen’s measure of competition and Altman & Pérez-Liñán’s measure of effective competition are fairly highly correlated with the incumbent-challenger indices, indicating some degree of convergent validity. However, indices of effective parties and volatility are not well correlated, suggesting that these indices are measuring something quite different.

Table 3: Alternate Indices of Electoral Contestation

	Incumbent-Challenger formula					
	<i>Legislative & Pres. votes</i>			<i>Legislative seats</i>		
	<i>Pearson's r</i>	<i>Polities</i>	<i>Elections</i>	<i>Pearson's r</i>	<i>Polities</i>	<i>Elections</i>
Competition – votes (Vanhanen 2000)	0.80	169	2,343			
Effective parties – seats (Borman & Golder 2013)				0.10	124	1,168
Effective competition – seats (Altman & Pérez-Liñán 2002)				0.62	21	178
Volatility – seats (Mainwaring et al. 2017)				0.09	142	1,054

Note: Pearson's r correlation between the incumbent-challenger indices of contestation based on (a) legislative/presidential votes and (b) legislative seats and four alternate indices of electoral contestation.

III. Patterns of Contestation

Having chosen a benchmark index – the incumbent-challenger index applied to legislative and presidential votes – we turn to the empirical record. What patterns of contestation obtain across polities and through time?

Of particular interest is the dispersion of election results during periods in which national elections are on course, i.e., excluding polities that have held no national elections as well as years prior to the first election in a polity and years in which the electoral system was interrupted (e.g., by a coup or foreign occupation). Polity-year data is displayed as a histogram in Figure 1, where the Y axis is percent and the X axis is our index of electoral contestation, divided into 5-point increments.

The resulting curve is strongly bimodal. The first mode at 0-5, comprising roughly 9% of all polity-years, represents a setting in which the incumbent party wins all, or nearly all, of the votes. This may be regarded as strong *prima facie* evidence of autocracy.

To the right of the first mode we find a substantial dropoff in frequency. It is uncommon for the incumbent-challenger differential to fall between 40 and 95% of the vote (5-60 on our index of electoral contestation). Where multi-party competition is allowed, huge wins for the incumbent are rare.

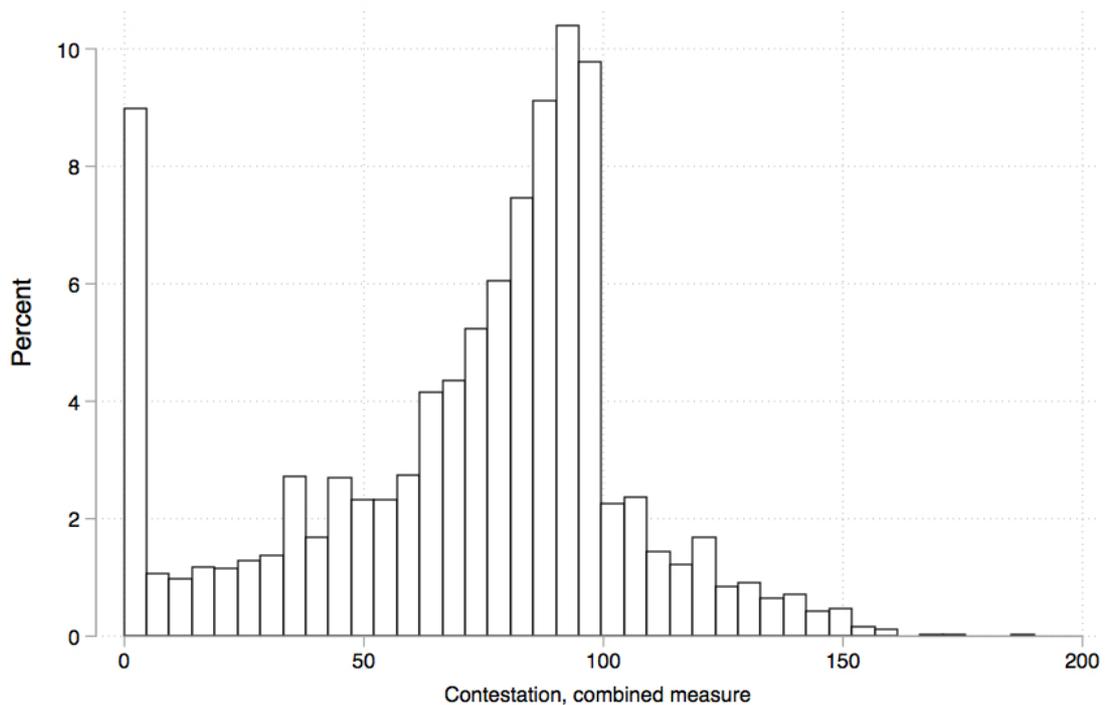
Most of the electoral outcomes fall between 60 and 100 on our index, representing situations where the incumbent-challenger differential is between 0 and 40. The three highest bars, including nearly 30% of all the data points, represent situations in which the incumbent party narrowly retains its plurality status, winning up to 15% more votes than the next largest challenger.

After 100, we find a dramatic fall in frequency. Note that a score of 100-110 represents a narrow loss for the incumbent party and a probable shift in control over the executive. There are two possible interpretations of the non-symmetrical distribution around 100. One might surmise

that incumbent parties are catering to their core constituencies, following a minimal-winning strategy (Riker 1962) and thereby eking out narrow victories. Alternatively, or additionally, one might surmise that incumbent parties are manipulating the electoral process just enough to stay in power. In any case, it bears emphasis that narrow wins for the incumbent are *much* more common than narrow losses.

A final aspect of the histogram that deserves emphasis is the extremely thin right tail. Big losses for the incumbent are the rarest of all possible outcomes. This, too, is open to varying interpretations. It might be an indication of the stability of party ties. Since incumbents are likely to be entrenched in the electorate, it would be surprising if their support collapsed all of a sudden (from one election to the next). It might also be an indication of the advantages of incumbency, which serves as ballast even in contrary electoral tides.

Figure 1: Histogram of Contestation

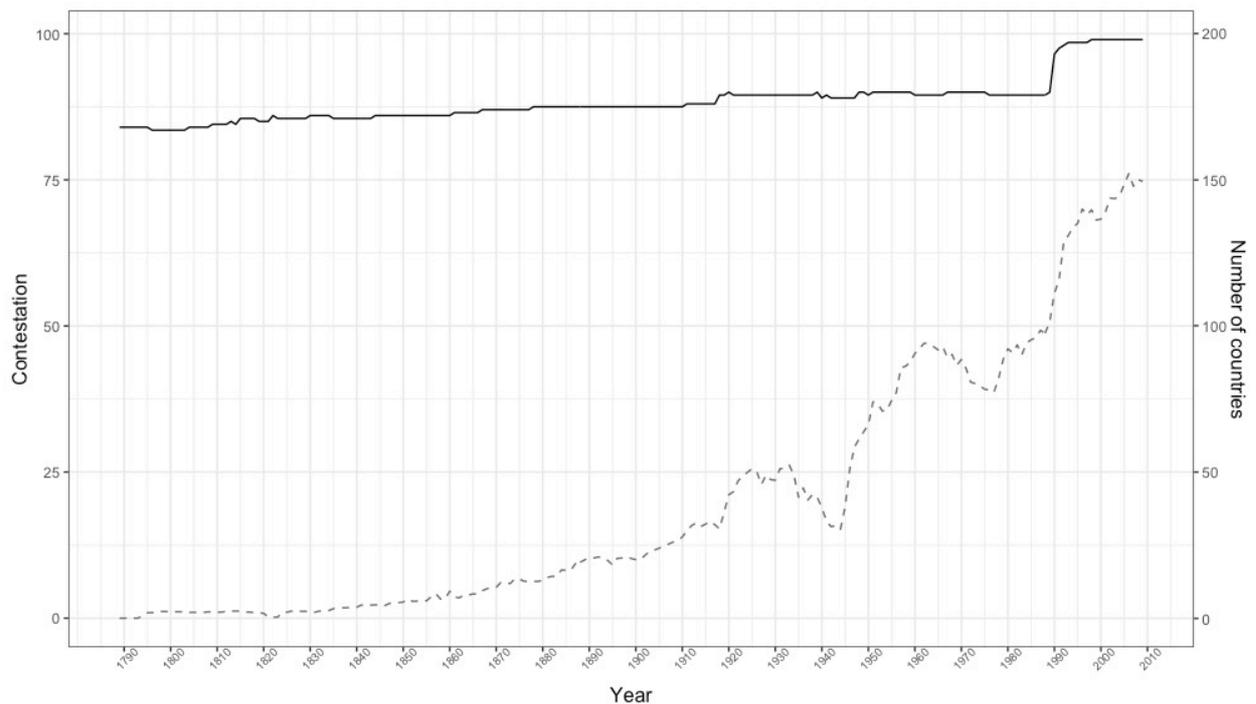


Note: Histogram of contestation (incumbent-challenger), excluding polity-years prior to a polity's first election or interregnums when elections are discontinued.

To get a sense of how contestation varies over time, we graph the mean value of our index across all available polity-years from 1789 to the present. The long-term pattern depicted in Figure 2 demonstrates that contestation has increased dramatically from the turn of the nineteenth century to the turn of the twenty-first century. Only one significant reversal is apparent, associated

with the fascist era in the 1940s. This lasts only a few years, after which contestation continues its ascent. A plateau appears from 1960-90 and again in the first decade of the twenty-first century. The latter may reflect the achievement of an equilibrium, or it may be simply a pause in a longer secular-historical development.

Figure 2: Contestation through Time



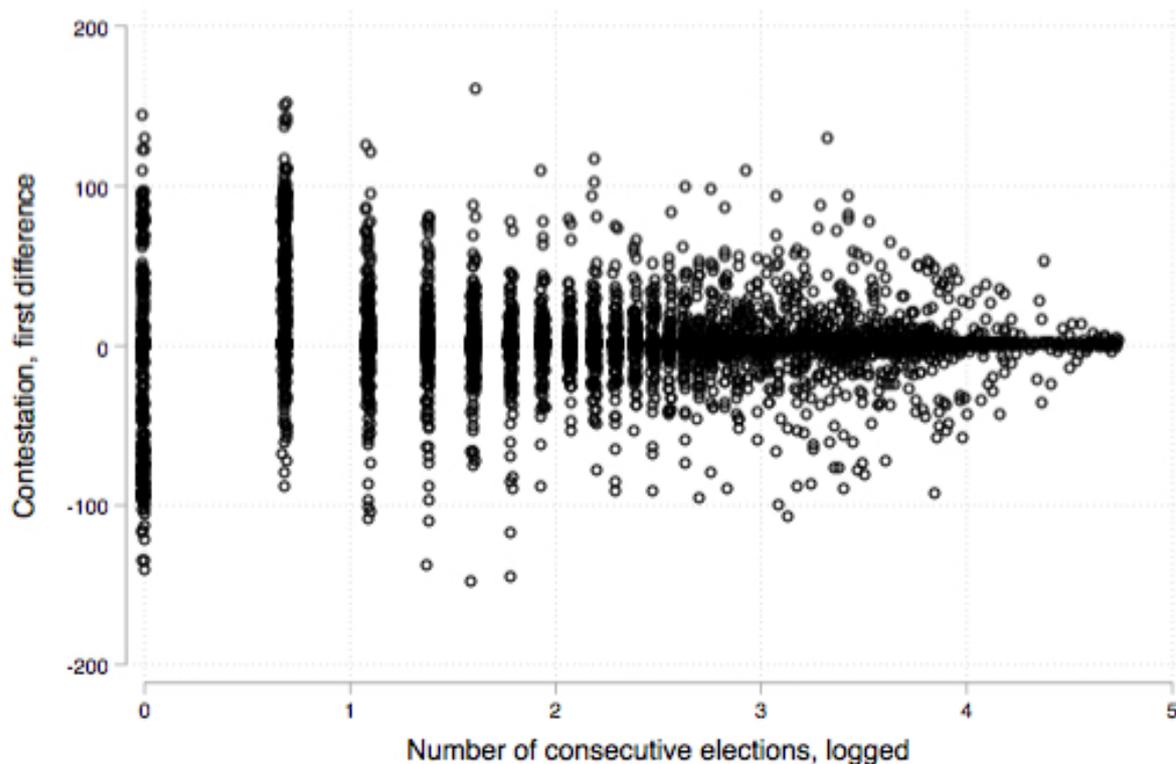
Note: The solid line shows the number of countries for each year in the data set. The dashed line shows contestation (incumbent-challenger) through time, calculated as the polity-year mean across all polities for which data is available.

Leaving aside global patterns, we turn to polity-specific patterns. One may suspect that for any given polity there exists an equilibrium – a level of contestation that is normal for that polity, given its structural endowments (whatever relatively fixed factors affect contestation). Accordingly, over time, as a polity gains experience with elections, one would expect election-to-election variation in contestation to moderate. Players (both masses and elites) should learn the rules of the game and solve their coordination problems. Likewise, one would expect the rules of the game, including electoral system laws, to stabilize.

To test this proposition, we graph year-to-year volatility in contestation (the first-difference of contestation) against the number of years a polity has held uninterrupted elections, transformed by the natural logarithm. Figure 3 supports the hypothesis of increasing institutionalization, with polities converging on a polity-level mean. This is confirmed by a statistical analysis in which the

first-difference of contestation is regressed against the number of consecutive elections (log) along with contestation (lagged), polity dummies, and year dummies, shown in Table B1.

Figure 3: Volatility across Elections



Note: Y axis: volatility, measured as the first-difference of contestation (incumbent-challenger). X axis: number of consecutive (uninterrupted) legislative elections, transformed by the natural logarithm.

IV. Analysis

A positive relationship between population size and competitiveness has been found at district levels (Gerring et al. 2015). However, it is not at all clear that district-level findings can be extrapolated to national levels. At issue is not simply an aggregation problem but also very different samples and populations. Extant district-level analyses focus on a thin slice of contemporary nation-states and a handful of (highly unrepresentative) nation-states offering long-time series. They are also restricted to polities where a modicum of party competition exists. This study, by contrast, encompasses both the fact of contestation (the existence of multi-party elections) and the degree of contestation (the extent to which a single party or candidate is able to monopolize votes and seats, and prevent turnover). Consequently, the population of theoretical interest

includes all polities, democratic or autocratic. From this population, we draw a global sample of polities that is orders of magnitude larger (in polity-year coverage) than any district-level analysis.

To test the proposition that size enhances contestation we regress our incumbent-challenger index against population (transformed by the natural logarithm) in a global sample of polities, extending from 1789 to 2009. Definitions, sources, and descriptive statistics for all variables are contained in Appendix A. Because the outcome is censored at 0, a tobit estimator is employed in benchmark models (Long & Freese 2014). Decade dummies are included to de-trend the data (year dummies are not tractable with a tobit estimator). Standard errors are clustered by polity in order to mitigate the serial correlation of errors.

Because contestation is not a well-studied topic there is no standard statistical or theoretical model that identifies relevant covariates. Plausibly, the same factors that favor democracy might also foster contestation. Accordingly, specification tests in Table 4 draw on the substantial literature on democratization (Coppedge 2012).

Model 1 includes per capita GDP (log), reflecting the modernization hypothesis (Knutsen et al. 2018) and a dummy for English colonial status (contemporary or former), reflecting the assumption that British colonies were more likely to foster elective assemblies (Bernhard et al. 2004; Lange, Mahoney & Vom Hau 2006; Olsson 2009). Both factors may be regarded as exogenous, both have prima facie plausibility as causes, and both perform fairly well in subsequent tests. We regard this as the benchmark model.

Model 2 includes the regressor of theoretical interest along with dummies for each region of the world – West Europe, East Europe, Central Asia, Latin America, MENA, Sub-Saharan Africa, North America, East Asia, Southeast Asia, South Asia, Pacific, and the Caribbean. Regional fixed effects, although not theoretically informed, represent the possibility that contestation might be affected by historical and cultural factors specific to different regions and avoids any possibility of post-treatment confounding.

Model 3 includes urbanization, a correlate of modernization, and a factor that may be correlated with population (in fact, the correlation is modest). Urbanization shows a strong relationship to contestation but is not robust in later tests.

Model 4 includes dummy variables representing four major electoral system types: (a) majoritarian (the excluded category), (b) proportional, (c) mixed (majoritarian and proportional), and (d) other (single non-transferable voting, limited voting, et al). Bear in mind that although our outcome is votes, rather than seats, one may assume that the rules of the game condition choices by elites and voters, prompting both to act strategically (Cox 1997). Some electoral systems are apt to foster greater contestation than others. Indeed, proportional electoral systems foster higher

contestation than majoritarian electoral systems, as expected. We do not include this in the benchmark because of possible endogeneity relative to the factor of theoretical interest.

Model 5 includes a measure of democracy, the Lexical index of electoral democracy. This is an important covariate insofar as we want to distinguish contestation from democratization. On the other hand, the two concepts are interwoven (as discussed), and it is by no means clear that democracy should be regarded as a “cause” of contestation. Thus, we do not retain democracy in the benchmark model.

Model 6 includes measures of ethnic, religious, and linguistic fractionalization, factors often regarded as causes of democracy (Gerring et al. 2018). Results are mixed, with ethnic fractionalization showing a negative relationship to contestation, religious fractionalization showing a positive relationship, and linguistic fractionalization showing no consistent relationship. Bear in mind that our theory rests on the *number* of different social identities within a society, while fractionalization indices measure their *distribution*. Our theory is also agnostic on what sorts of identities might be relevant for politics. They might be ethnic, religious, linguistic, ideological, social class, region, or some mixture of the above, and the features that matter will surely vary from time to time and place to place. Thus, the mixed results for these common predictors of democracy does not speak directly on our theory. Fractionalization indices might be regarded as background covariates, but not causal mechanisms. Because they are also likely to be endogenous to the aggregate size of a society we exclude them from the benchmark specification.

Model 7, a kitchen-sink specification, includes all of the foregoing factors. By virtue of list-wise deletion, this sample includes only sovereign countries, alleviating concerns that our results might be driven by the inclusion of colonies and dependencies.

The final test, shown in Model 8, addresses potential problems of causal inference through an instrumental-variable design. We identify two instruments that strongly affect the population of a polity – territorial size (log) and agricultural suitability. The critical assumption is that these factors have no direct effect on contestation conditional on observed covariates – which, we note, includes per capita GDP. We regard this as a plausible, though not unimpeachable, assumption. In any case, the coefficient estimate is very close to the benchmark model.

Results across the eight models presented in Table 4 reveal estimates for population that are remarkably stable, suggesting that the observed relationship is not a product of idiosyncratic choices in model specification. By contrast, estimates for background covariates are not especially robust, though generally signed in the expected direction.

Table 4: Tobit Models

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	6.047*** (1.680)	5.946*** (1.407)	3.981** (1.899)	8.378*** (1.779)	2.901*** (0.904)	2.848* (1.551)	5.835*** (1.196)		7.797** (3.301)
GDP pc (log)	21.913*** (2.445)		17.274*** (3.405)	24.59*** (2.339)	4.311*** (1.542)	16.79*** (2.903)	2.153 (2.636)	-0.0690 (0.085)	22.758*** (2.585)
English colony	18.731** (7.642)		19.148*** (6.300)	25.03*** (8.116)	6.718* (3.524)	11.01 (7.406)	13.88*** (4.688)	0.038 (0.234)	13.408* (8.076)
Urbanization			49.204*** (18.215)				5.889 (10.28)		
Latitude (log)							-5.199** (2.181)		
Oil income pc							9.029 (5.659)		
Years indep (log)							-0.00449** (0.00184)		
Protestant							-0.230 (1.081)		
Muslim							0.0946 (0.0602)		
Island							0.0368 (0.0728)		
<i>Fractionalization</i>									
Ethnic							-40.15*** (14.75)	-25.50** (10.96)	
Religious							34.20** (14.71)	17.89** (9.061)	
Linguistic							-7.787 (13.28)	6.782 (10.84)	
Lexical index					17.09*** (0.864)		13.63*** (0.844)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				16.38** (7.435)			8.886** (4.043)		
Mixed				13.94* (7.450)			3.951 (5.870)		
Other				-10.69 (24.91)			-37.43*** (12.61)		
<i>Instruments</i>									
Area (log)								0.627*** (0.061)	
Agricultural suitability								2.192*** (0.392)	
Region FE		✓					✓		
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	57.16*** (2.189)	50.12*** (1.662)	50.74*** (2.331)	52.75*** (1.612)	36.28*** (1.156)	50.69*** (2.333)	33.91*** (1.323)		
<i>Polities</i>	197	197	163	166	191	177	142		164
<i>Years</i>	226	227	111	215	215	120	107		215
<i>Observations</i>	21,068	21,068	11,502	18,796	14,351	12,944	7,947		18,601
<i>Pseudo R2</i>	0.088	0.122	0.046	0.099	0.128	0.045	0.109		
<i>Chi2</i>									421.42

Note. Outcome: contestation (incumbent-challenger). *Estimator:* tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

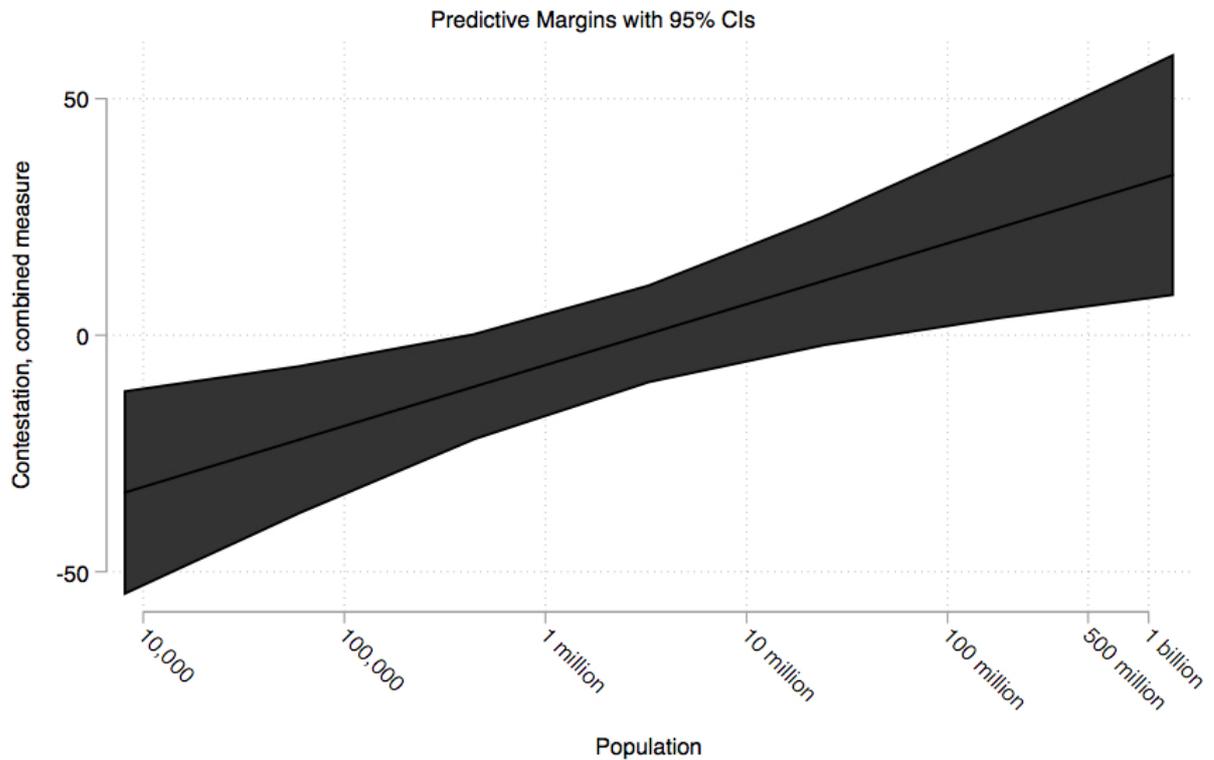
In Appendix B, analyses performed in Table 4 for the incumbent-challenger index are replicated for the five other indices of contestation presented in Table 2. These analyses show very similar results, demonstrating that the relationship between population and contestation is robust with various formulas and using votes or seats as measures of party performance. Across all these tests population is a robust predictor of contestation, while other factors are less consistent.

Taken together, the tests shown in Table 4 and Appendix B suggest that among the theoretically plausible causes of contestation, population is the only strong and consistent predictor. Of course, there may be other causal factors that we have not managed to identify or to properly measure. And there may be causal relationships among the chosen regressors that appear only in certain contexts – delimited by cultural, historical, political, economic scope-conditions that further study may identify. We do not wish to over-interpret these results. Still, it is remarkable that many of the purported causes of democracy do not apply – at least, not consistently or strongly – to electoral contestation.

To gain a sense for the significance of this relationship, Figure 4 graphs predicted values for contestation as population varies, based on Model 1 in Table 4, with covariates set to their sample means. Note that values below zero refer to the values that (according to the tobit estimator) would have been realized if the scale were not truncated. Note also that because of the logged scale, the impact of population on contestation is much greater at lower levels of population than at higher levels. For example, an increase in population from one hundred thousand to one million is associated with a (roughly) 20-point increase in contestation. This is equivalent to an increase from 10,000 to 100,000 (on the low end) or 10 million to 100 million (on the high end). In any case, the effect is non-trivial. Larger polities generate considerably higher levels of party competition.

To test this hypothesis in a more disaggregated fashion we explore a series of sample restrictions in Table 5. Model 1 focuses on the nineteenth century while Model 2 focuses on the twentieth century. Strikingly, the relationship between population and contestation appears to be much stronger in the nineteenth century than in the twentieth century. However, the historical sample is much smaller than the contemporary sample and there is less variation in the variables of theoretical interest, so it is difficult to know whether it provides adequate grounds for generalization.

Figure 4: Predicted Values



Note: Predicted values for contestation (incumbent-challenger) as population changes, based on Model 1, Table 2, holding other variables at their means, surrounded by a 95% confidence interval.

Model 3 excludes polity-years for which contestation is zero, presenting an analysis based only on positive values (as displayed in Figure 1). Model 4 transforms the contestation index into a binary index, coded as 0 (no contestation) or 1 (some contestation). In this fashion, we disaggregate the two elements of the index. Both analyses show a positive relationship to population, though it is considerably weaker when disaggregated in this fashion.

In Table B2, we test the sensitivity of the benchmark model by excluding specific regions of the world. Here, we find relatively little variation in the performance of population, for which estimates vary only by a few points across models. Apparently, our analysis is not subject to influential regions.

Table 5: Sample Restrictions

<i>Estimator</i>	Tobit	Tobit	OLS	OLS
<i>Outcome</i>	Contestation	Contestation	Contestation	Contestation(0/1)
<i>Sample</i>	Year<1900	Year>1900	Contestation(>0)	All
	1	2	3	4
Population (log)	34.758*** (7.719)	3.397** (1.315)	2.676*** (0.764)	0.024** (0.011)
<i>Polities</i>	92	197	156	186
<i>Years</i>	111	114	215	226
<i>Obs</i>	6,842	14,152	9,221	18,228
R2	0.110	0.038	0.201	0.355

Note: *Outcome:* contestation (incumbent-challenger). Restricted to positive values in Model 3 and re-coded as a binary variable in Model 4. All models include per capita GDP (log), English colony, and decade dummies. Model 3 also includes electoral system dummies. Coefficients and standard errors (clustered by polity) shown only for the variable of theoretical interest. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

In Table 6, we approach the question with time-series models, using a linear (ordinary least squares) estimator along with polity and year fixed effects. Errors are clustered by polity, as previously. Model 1 follows the benchmark specification (excluding English colony, which is time-invariant). Model 2 includes only the regressor of theoretical interest, population. Model 3 substitutes urbanization for GDP. Model 4 returns to the benchmark specification, this time with a ten-year (rather than one-year) lag between right- and left-side variables. Model 5 introduces a lagged dependent variable.

All models show a robust association between population and contestation. Note also that the estimate from Model 1 tracks the estimate from the benchmark analysis (Model 1, Table 4) very closely. Importantly, the fixed-effect framework alleviates concerns about confounders stemming from static features of each polity (e.g., geography, culture, pre-modern history). The robustness of the relationship with different lag structures, as well as with a lagged dependent variable, suggests that simultaneity (between right- and left-sides of the model) is unlikely to pose a problem.

Table 6: Fixed-effect Models

<i>Right-side lag (years)</i>	1	1	1	10	1
	1	2	3	4	5
Population (log)	6.443** (2.530)	4.373* (2.364)	13.337*** (3.491)	6.831** (2.631)	0.964** (0.408)
GDP pc (log)	5.846*** (1.706)			4.443** (1.475)	0.496* (0.256)
Dependent variable (lagged)					0.861*** (0.008)
Urbanization			-1.767 (18.233)		
<i>Polities</i>	197	197	163	196	196
<i>Years</i>	226	227	109	219	225
<i>Obs.</i>	21,083	21,083	11,545	19,949	20,802
R2	0.336	0.399	0.076	0.312	0.904

Note: *Outcome:* contestation (incumbent-challenger). *Estimator:* ordinary least squares with polity and year fixed effects. Standard errors clustered by polity in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Constant omitted.

V. Discussion

In this study, we introduce a new index of electoral contestation. The proposed formula encompasses several dimensions of contestation – the size of the incumbent party, the size of the largest challenger, as well as turnover in the pole position (the largest party) – and thus provides a more complete measure of this complex concept than existing indices. This measure is applied to polities from 1789 (or year of independence) to the present, producing a comprehensive dataset of electoral contestation extending to nearly 36,000 polity-year observations.

We also put forth an explanation for why some polities achieve higher contestation than others. This explanation rests on demography. Larger polities, we argue, are more likely to adopt multi-party elections as a mechanism for choosing rulers and the electoral arena is more likely to be highly contested.

We reiterate that contestation is not a proxy for democracy but nonetheless constitutes an important dimension of that nebulous concept. Readers may wonder what other functions an index of contestation performs. A wealth of studies, referenced at the outset, suggest that contestation fosters better governance. However, these studies focus almost exclusively at subnational levels – at regions or electoral districts. It remains to be seen whether electoral contestation has a positive impact on governance at national levels. We view this as an important agenda for future research.

VI. References

- Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8:2. 155-194.
- Alvarez, Michael, Jose A. Cheibub, Fernando Limongi, Adam Przeworski. 1996. "Classifying Political Regimes." *Studies in Comparative International Development* 31:2, 3-36.
- Altman, David, Anibal Pérez-Liñán. 2002. "Assessing the Quality of Democracy: Freedom, Competitiveness and Participation in Eighteen Latin American Countries." *Democratization* 9:2, 85-100, DOI: 10.1080/714000256
- Ansolabehere, Stephen, James M. Snyder, Jr., Charles Stewart, III. 2001. "Candidate Positioning in U.S. House Elections." *American Journal of Political Science* 45:1 (January) 136-59.
- Arce, Moisés, Jorge Mangonnet. 2013. "Competitiveness, Partisanship, and Subnational Protest in Argentina." *Comparative Political Studies* 46, 895-919.
- Bardhan, Pranab, Tsung-Tao Yang. 2004. "Political Competition in Economic Perspective." BREAD Working Paper No. 078, Bureau for Research in Economic Analysis of Development.
- Barro, Robert J. 1973. "The Control of Politicians: An Economic Model." *Public Choice*, 14, 19-42.
- Bartolini, Stefano. 1999. "Collusion, competition and democracy: Part I." *Journal of theoretical politics* 11.4: 435-470.
- Bartolini, Stefano. 2000. "Collusion, competition and democracy: Part II." *Journal of theoretical politics* 12.1: 33-65.
- Becker, Gary S. 1958. "Competition and Democracy." *Journal of Law and Economics* 1, 105-9.
- Beer, Caroline, Neil J. Mitchell. 2004. "Democracy and Human Rights in the Mexican States: Elections or Social Capital?" *International Studies Quarterly* 4:8 (June) 293- 312.
- Berkowitz, Daniel, Karen B. Clay. 2012. *The Evolution of a Nation: How Geography and Law Shaped the American States*. Princeton: Princeton University Press.
- Bernhard, Michael, Christopher Reenock, Timothy Nordstrom. 2004. "The Legacy of Western Overseas Colonialism on Democratic Survival." *International Studies Quarterly* 48, 225-50.

- Besley, Timothy, Torsten Persson, Daniel M. Sturm. 2010. "Political Competition and Economic Performance: Evidence from the United States." *Review of Economic Studies* 77:3, 1329-52.
- Bormann, Nils-Christian, Matt Golder. 2013. "Democratic Electoral Systems Around the World, 1946-2011." *Electoral Studies* 32: 360-369.
- Borges, André. 2008. "State Government, Political Competition and Education Reform: Comparative Lessons from Brazil." *Bulletin of Latin American Research* 27:2, 235-54.
- Brunell, Thomas L. 2008. *Redistricting and Representation: Why Competitive Elections are Bad for America*. New York: Routledge.
- Brunell, Thomas L., Justin Buchler. 2009. "Ideological representation and competitive congressional elections." *Electoral Studies* 28.3: 448-457.
- Buchler, Justin. 2005. "Competition, representation and redistricting: the case against competitive congressional districts." *Journal of Theoretical Politics* 17.4: 431-463.
- Buchler, Justin. 2011. *Hiring and firing public officials: Rethinking the purpose of elections*. Oxford University Press.
- Caramani, Daniele. 2000. *Elections in Western Europe since 1815: electoral results by constituencies*. Macmillan.
- Clarke, Colin. 2006. "Politics, Violence and Drugs in Kingston, Jamaica." *Bulletin of Latin American Research* 25:3, 420-440.
- Cleary, Matthew R. 2007. "Electoral Competition, Participation, and Government Responsiveness in Mexico." *American Journal of Political Science* 51:2 (April) 283-99.
- Coppedge, Michael. 2012. *Democratization and research methods*. Cambridge: Cambridge University Press.
- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yiting Wang, Tore Wig, and Daniel Ziblatt. 2018. "V-Dem Codebook v8" Varieties of Democracy (V-Dem) Project.
- Cox, Gary W. 1997. *Making Votes Count*. Cambridge: Cambridge University Press.

- Dahl, Robert A. 1971. *Polyarchy: Participation and Opposition*. New Haven: Yale University Press.
- Denters, Bas. 2002. "Size and Political Trust: Evidence from Denmark, the Netherlands, Norway, and the United Kingdom." *Environment and Planning C: Government and Policy* 20, 793-812.
- Fariss, C. J., C. D. Crabtree, T. Anders, Z. M. Jones, F. J. Linder, J. N. Markowitz. 2017. "Latent Estimation of GDP, GDP per capita, and Population from Historic and Contemporary Sources." URL: <http://arxiv.org/abs/1706.01099>
- Fiorina, Morris P. 1973. "Electoral Margins, Constituency Influence, and Policy Moderation: A Critical Assessment." *American Politics Quarterly* 1:4, 479-98.
- Foddy, Margaret, Michael J. Platow, Toshio Yamagishi. 2009. "Group-based trust in strangers: The role of stereotypes and expectations." *Psychological Science* 20.4: 419-422.
- Freedom House. 2007. "Methodology," Freedom in the World 2007. New York. (http://www.freedomhouse.org/template.cfm?page=351&ana_page=333&year=2007), accessed September 5, 2007.
- Geddes, B. 1991. "A Game Theoretical Model of Reform in Latin American Democracies." *American Political Science Review* 85: 371-92.
- Geddes, Barbara. 1994. *Politician's Dilemma: Building State Capacity in Latin America*. Berkeley: University of California Press.
- Gerring, John, Michael Hoffman, Dominic Zarecki. 2018. "The Diverse Effects of Diversity on Democracy." *British Journal of Political Science* (April).
- Gerring, John, Maxwell Palmer, Jan Teorell, Dominic Zarecki. 2015. "Demography and Democracy: A Global, District-level Analysis of Electoral Contestation." *American Political Science Review* 109:3 (August) 574-91.
- Golden Miriam A and Devesh Tiwari. 2009. "Criminality and malfeasance among national legislators in contemporary India." Presented at Annual Meeting American Political Science Association, Toronto, Sep. 3-6.
- Gordon, Sanford C., Gregory A. Huber. 2007. "The Effect of Electoral Competitiveness on Incumbent Behavior." *Quarterly Journal of Political Science* 2: 107-38.
- Griffin, John D. 2006. "Electoral Competition and Democratic Responsiveness: A Defense of the Marginality Hypothesis." *Journal of Politics* 68:4 (November) 911-21.

- Grofman, Bernard, Peter Selb. 2009. "A fully general index of political competition." *Electoral Studies* 28.2 (2009): 291-296.
- Grzymala-Busse, Anna. 2007. *Rebuilding Leviathan: Party Competition and State Exploitation in Post-Communist Democracies*. Cambridge: Cambridge University Press.
- Hansen, Sune W. 2013. "Country Size and Local Political Trust: A Quasi-experiment Using Municipal Mergers in Denmark." *Scandinavian Political Studies* 36:1, 43-66.
- Harmel, Robert, John D. Robertson. 1985. "Formation and success of new parties: A cross-national analysis." *International political science review* 6.4: 501-523.
- Heller, William B., Andreas P. Kyriacou, Oriol Roca-Sagalés. 2011. "Party Competition and Government Quality: The Politics of Turning a Blind Eye to Poor Governance." Presented at the annual meetings of the Midwest Political Science Association, Chicago, IL.
- Hicken, Allen. 2018. "Party Systems and the Politics of Development." In *Handbook on the Politics of Development*, Oxford University Press. Carol Lancaster and Nicolas van de Walle, eds.
- Jones, Philip Edward. 2013. "The Effect of Political Competition on Democratic Accountability." *Political Behavior* 35:481–515.
- Keefer, Philip. 2007. "Clientelism, credibility, and the policy choices of young democracies." *American Journal of Political Science* 51(4):804--21.
- Keefer, Philip, R. Vlaicu. 2008. "Democracy, credibility and clientelism." *Journal of Law, Economics, and Organization* 24(2):371—406.
- Knutsen, Carl Henrik, John Gerring, Svend-Erik Skaaning, Jan Teorell, Matthew Maguire, Michael Coppedge, Staffan Lindberg. 2018. "Economic Development and Democracy: An Electoral Connection." *European Journal of Political Research* (forthcoming).
- Kollman, Ken, Allen Hicken, Daniele Caramani, David Backer. 2011. "Constituency-Level Elections Archive." (CLEA, www.electiondataarchive.org), September 14, 2011 [dataset]. Ann Arbor, MI: University of Michigan, Center for Political Studies.
- Konisky, David M., Michiko Ueda. 2011. "The Effects of Uncontested Elections on Legislator Performance." *Legislative Studies Quarterly* 36:2, 199-229.
- Lachat, Romain. 2011. "Electoral Competitiveness and Issue Voting." *Political Behavior* 33, 645–63.

- La Porta, Rafael, Florencio Lopez-de-Silanes, Andrei Shleifer, Robert W. Vishny. 1999. "The Quality of Government." *Journal of Law, Economics, and Organization* 15(1): 222-79.
- Lindberg Staffan I, Minion K.C. Morrison 2008. "Are African Voters Really Ethnic or Clientelistic?: Survey Evidence from Ghana." *Political Science Quarterly* 123(1): 95-122.
- Long, J. Scott, Jeremy Freese. 2014. *Regression Models for Categorical Dependent Variables Using Stata, 3rd ed.* Stata Press.
- Mainwaring, Scott, Carlos Gervasoni, Annabella España Najera. 2017. "Extra- and Within-System Electoral Volatility." *Party Politics* 23:6 (November): 623-635.
- Matsubayashi, Tetsuya. 2007. "Population Size, Local Autonomy, and Support for the Political System." *Social Science Quarterly* 88:3 (September) 830-49.
- Nohlen, Dieter (ed). 2005. *Elections in the Americas: A Data Handbook, vols 1-2.* New York: Oxford University Press.
- Nohlen, Dieter; Florian Grotz; Christof Harmann (eds). 2002. *Elections in Asia and the Pacific: A Data Handbook, vols 1-2.* New York: Oxford University Press.
- Nohlen, Dieter; Michael Krennerich; Bernhard Thibaut (eds). 1999. *Elections in Africa: A Data Handbook.* Oxford: Oxford University Press.
- Nohlen, Dieter; Philip Stover (eds). 2010. *Elections in Europe: A Data Handbook.* Nomos Verlagsgesellschaft.
- Nyblade, Benjamin and Steven R. Reed. 2008. "Who cheats? Who loots? Political competition and corruption in Japan, 1947--1993." *American Journal of Political Science* 54(4):926—41.
- Olsson, Ola. 2009. "On the Democratic Legacy of Colonialism." *Journal of Comparative Economics* 37:4, 534-51.
- Padovano, Fabio, Roberto Ricciuti. 2009. "Political Competition and Economic Performance: Evidence from the Italian Regions." *Public Choice* 138, 263—77.
- Powell, G. Bingham, Jr. 2000. *Elections as Instruments of Democracy: Majoritarian and Proportional Visions.* New Haven and London: Yale University Press.
- Putnam, Robert D. 2007. "E pluribus unum: Diversity and community in the twenty-first century. The 2006 Johan Skytte Prize lecture." *Scandinavian Political Studies* 30, 137e174.

- R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rahn, Wendy M., Thomas J. Rudolph. 2005. "A tale of political trust in American cities." *Public Opinion Quarterly* 69.4: 530-560.
- Raleigh, Clionadh, Håvard Hegre. 2009. "Population size, concentration, and civil war. A geographically disaggregated analysis." *Political geography* 28.4: 224-238.
- Riker, William H. 1962. *The theory of political coalitions*. New Haven: Yale University Press.
- Sartori, Giovanni. 1976. *Parties and Party Systems*. Cambridge: Cambridge University Press.
- Schumpeter, Joseph A. 1942/1950. *Capitalism, Socialism and Democracy*. New York: Harper & Bros.
- Shugart, Matthew Soberg, John M. Carey. 1992. *Presidents and assemblies: Constitutional design and electoral dynamics*. Cambridge: Cambridge University Press.
- Skaaning, Svend-Erik, John Gerring, Henrikas Bartusevičius. 2015. "A Lexical Index of Electoral Democracy." *Comparative Political Studies* 48:12 (October) 1491-1525.
- StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC
- Stigler, George J. 1972. "Economic Competition and Political Competition." *Public Choice* 13 (September) 91-106.
- Strom, Kaare. 1989. "Inter-party competition in advanced democracies." *Journal of Theoretical Politics* 1.3: 277-300.
- Strom, Kaare. 1992. "Democracy as political competition." *American Behavioral Scientist* 35.4-5: 375-396.
- Ting, Michael M., James M Snyder, Jr., Shigeo Hirano, Olle Folke. 2013. "Elections and Reform: The Adoption of Civil Service Systems in the U.S. States." *Journal of Theoretical Politics* 25, 363-87.
- Vanhanen, Tatu. 2000. "A New Dataset for Measuring Democracy, 1810-1998." *Journal of Peace Research* 37, 251-65.
- Weitz-Shapiro, Rebecca. 2012. "What Wins Votes: Why Some Politicians Opt Out of Clientelism." *American Journal of Political Science* 56:3 (July) 568-83.
- Wittman, Donald. 1989. "Why Democracies Produce Efficient Results." *Journal of Political Economy*, 97(6), pp. 1395-1424.

Wittman, Donald. 1995. *The Myth of Democratic Failure: Why Political Institutions are Efficient*.
Chicago: University of Chicago Press.

Online Appendix A: Data Description

Table A1: Variable Definitions

Left-side Variables

Incumbent-challenger, votes, parl. 100 – (vote share of the incumbent in legislative elections – vote share of the challenger in legislative elections). The incumbent is the largest party in the previous elections. If it is the first democratic election of a polity or the incumbent is not among the top three parties of the most recent election the largest party of the current election is included in the formula instead. The challenger is the second largest party if the incumbent is again the largest party or if it is the first democratic election in a polity or the incumbent is not among the top three parties anymore. If the incumbent is the second or third largest party in the election then the challenger is the largest party. *Source:* authors. *Scale:* interval. `contestation_vote`

Incumbent-challenger, seats, parl. 100 – (seat share of the incumbent – seat share of the challenger). The incumbent is the largest party in the previous elections. If it is the first democratic election of a polity or the incumbent is not among the top three parties of the most recent election the largest party of the current election is included in the formula instead. The challenger is the second largest party if the incumbent is again the largest party or if it is the first democratic election in a polity or the incumbent is not among the top three parties anymore. If the incumbent is the second or third largest party in the election then the challenger is the largest party. *Source:* authors. *Scale:* interval. `contestation_seat2`

Incumbent-challenger, votes, pres. 100 – (vote share of the incumbent in presidential elections – vote share of the challenger in presidential elections). The incumbent is the largest party in the previous elections. If it is the first democratic election of a polity or the incumbent is not among the top three parties of the most recent election the largest party of the current election is included in the formula instead. The challenger is the second largest party if the incumbent is again the largest party or if it is the first democratic election in a polity or the incumbent is not among the top three parties anymore. If the incumbent is the second or third largest party in the election then the challenger is the largest party. *Source:* authors. *Scale:* interval. `pres_contestation_vote`

Incumbent-challenger, combined. This is the benchmark variable in our paper. It is a combined version of Incumbent-challenger, votes, parl. and Incumbent-challenger, votes, pres. The variable takes the average of presidential and legislative contestation when both are taking place in a country.

Largest-party, votes, parl. 100 – vote share of the largest party in a legislative election. *Source:* authors. *Scale:* interval. `v2ellovtlg_100`

Largest-party, seats, parl. 100 – seat share of the largest party in a legislative election. *Source:* authors. *Scale:* interval. `v2ellostsl_100`

Largest-party, votes, pres. 100 – vote share of the largest party in a presidential election. *Source:* authors. *Scale:* interval. `v2elvotlrg_100`

Top two parties, votes, parl. 100 – (vote share of the largest party – vote share of the second largest party) in a legislative election. *Source:* authors. *Scale:* interval. `contestation_top2`

Top two parties, seats, parl. 100 – (seat share of the largest party – seat share of the second largest party) in a legislative election. *Source:* authors. *Scale:* interval. `contestation_seat_top2`

Top two parties, votes, pres. 100 – (vote share of the largest party – vote share of the second largest party) in a presidential election. *Source:* authors. *Scale:* interval. `pres_contestation_top2`

Right-side Variables

Agricultural suitability. Geographic endowments favoring agricultural production including climate, soil, and terrain. *Source:* Agro-Ecological Zones system (GAEZ), developed by the Food and Agriculture Organization of the United Nations (FAO), downloaded (October 2017) from <http://gaez.fao.org/Main.html#>. *Scale:* logarithmic. `suita_GAEZ_ln`

- English colony** Former English colony. Source: Authors. Scale: binary. *English_legal_origin*
- Fractionalization, Ethnic.** Herfindahl index of fractionalization. Specifically, the probability of two randomly chosen people belonging to the same ethnic group. Source: Alesina et al. (2003). Scale: interval. *al_ethnic*
- Fractionalization, Religious.** Herfindahl index of fractionalization. Specifically, the probability of two randomly chosen people belonging to the same religious group. Source: Alesina et al. (2003). Scale: interval. *al_religion*
- Fractionalization, Linguistic.** Herfindahl index of fractionalization. Specifically, the probability of two randomly chosen people belonging to the same linguistic group. Source: Alesina et al. (2003). Scale: interval. *al_language*
- GDP per cap.** Gross domestic product per capita in constant 1990 dollars, based on data from the Maddison Project (Bolt & van Zanden 2014), supplemented by estimates from Bairoch (1976), Broadberry (2015), Broadberry/Klein (2012), Gleditsch (2002), and the WDI (World Bank 2016), which are combined in a dynamic, three-dimensional latent trait model. Source: Fariss et al. (2017). Scale: logarithmic. *Maddison_gdppc_1990_estimate_ln*
- Land area.** Land area of polity. Source: Agro-Ecological Zones system (GAEZ), developed by the Food and Agriculture Organization of the United Nations (FAO), downloaded (October 2017) from <http://gaez.fao.org/Main.html#>. Extra data linearly imputed with data from WDI (World Bank 2016). Scale: logarithmic. *area_GAEZ_ln_imp*
- Latitude.** Distance from equator. Source: QoG (Teorell et al. 2016). Scale: logarithmic. *Latitude_ln*
- Lexical index of electoral democracy.** An ordinal index measuring the electoral components of democracy in a cumulative fashion. That is, to qualify for a given level (0-6) all previous conditions must be satisfied. 0 = No elections. (Elections are not held for any policymaking offices. This includes situations in which elections are postponed indefinitely or the constitutional timing of elections is violated in a more than marginal fashion.) 1 = Elections with no parties or only one party. (There are regular elections but they are non-partisan or only a single party or party grouping is allowed to participate.) 2 = Multi-party elections for legislature. (Opposition parties are allowed to participate in legislative elections and to take office.) 3 = Multi-party elections for executive. (The executive is chosen directly or indirectly – by an elected legislature – through elections. 4 = Minimally competitive elections for both executive and legislature. (The chief executive offices and the seats in the effective legislative body are – directly or indirectly – filled by elections characterized by uncertainty, meaning that the elections are, in principle, sufficiently free to enable the opposition to win government power.) 5 = Male or female suffrage. (Virtually all adult male or female citizens are allowed to vote in elections.) 6 = Universal suffrage. (Virtually all adult citizens are allowed to vote in elections.) Source: Skaaning, Gerring & Bartusevičius (2015). Scale: ordinal. *lexical_index*
- Muslim.** Percentage of population that claims to be Muslim in 1980. Source: La Porta et al. (1999). Scale: binary. *Muslim*
- Oil wealth.** The aggregated real value of a polity's petroleum production, as a share of total population. Source: Haber & Menaldo (2011). Scale: interval. *e_Total_Oil_Income_PC*
- Polyarchy.** Electoral democracy index. Source: V-Dem (Coppedge et al. 2018; Teorell et al. 2016). Scale: interval. *v2x_polyarchy*
- Population.** Official population of a polity, counting only those acknowledged as citizens. This is based on data from the Maddison Project (Bolt & van Zanden 2014), supplemented by estimates from Broadberry/Klein (2012), Gleditsch (2002), Singer et al. (1972), and WDI (World Bank 2016), which are combined in a dynamic, three-dimensional latent trait model. Source: Fariss et al. (2017). Scale: logarithmic. *Maddison_pop_estimate_ln*
- Protestant.** Percentage of population that claims to be part of a Protestant denomination in 1980. Source: La Porta et al. (1999). *lp_protmg80*
- Regions.** A vector of dummies: Eastern Europe and Central Asia (including Mongolia), Latin America, Middle East & North Africa, Sub-Saharan Africa, Western Europe and North America, East Asia,

South-East Asia, South Asia, the Pacific, and the Caribbean. *Source:* QoG (Teorell et al. 2013). *Scale:* nominal. *e_regionpol*

Urbanization. Share of population living in urban areas (%). *Source:* V-Dem (Coppedge et al. 2018). *Scale:* interval. *e_miurbani*

Years independent (log). Years since formal independence (logged). *Source:* authors. *Scale:* logarithmic. *indep_yrs_since_ln*

Table A2: Descriptive Statistics

	<i>Observations</i>	<i>Mean</i>	<i>Median</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
(1) 100 – (Inc. – chall.), votes, combined	35,814	22.66	0	38.55	0	190
(2) 100 – (Inc. – chall.), votes, parl.	34,812	22.34	0	39.41	0	190
(3) 100 – (Inc. – chall.), seats, parl.	36,585	23.03	0	38.85	0	193.4
(4) 100 – largest party, votes, combined	36,198	13.65	0	23.73	0	124.79
(5) 100 – largest party, votes, parl.	35,005	14.21	0	25.16	0	93.50
(6) 100 – largest party, seats, parl.	36,871	14.08	0	24.03	0	99.53
(7) 100 – Top Two, votes, combined	36,038	21.95	0	36.34	0	124.7
(8) 100 – Top Two, votes, parl.	35,005	21.77	0	36.90	0	124.7
(9) 100 – Top Two, seats, parl.	36,660	21.56	0	34.94	0	100
Urbanization	15,562	0.349	0.296	0.247	0.00787	1
Ethnic fractionalization	19,514	0.437	0.431	0.258	0	0.930
Religious fractionalization	43,319	0.438	0.463	0.232	0.00229	0.860
GDP pc (log)	26,479	7.624	7.397	1.155	3.868	14.40
Lexical index	17,154	2.915	3	2.352	0	6
Region	47,838	4.433	4	2.602	1	10
Agricultural Suitability	35,260	0.419	0.438	0.266	0	0.965
Linguistic fractionalization	19,891	0.387	0.357	0.284	0.00211	0.923
Protestant	44,499	14.29	2.800	21.70	0	97.80
Years indep (log)	43,296	1.564	0	2.256	0	7.574
English colony	47,838	0.348	0	0.476	0	1
Muslim	42,863	23.42	1.200	36.04	0	99.90
Area (log)	44,255	11.22	11.69	2.789	2.227	18.36
Population (log)	26,284	15.04	15.11	1.931	7.507	21.38
Latitude (log)	42,728	-1.591	-1.496	0.938	-4.500	-0.325
Oil income pc	14,399	344.0	0	2,666	0	78,589
Electoral System	36,624	0.475	0	0.725	0	3

Table A3: Principal components (eigenvectors), all measures

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	8.55033	8.35346	0.95	0.95
Comp2	0.19687	0.0426093	0.0219	0.9719
Comp3	0.15426	0.100744	0.0171	0.9891
Comp4	0.053516	0.0332166	0.0059	0.995
Comp5	0.0202994	0.00754942	0.0023	0.9973
Comp6	0.01275	0.00618389	0.0014	0.9987
Comp7	0.00656607	0.00290478	0.0007	0.9994
Comp8	0.00366129	0.00191608	0.0004	0.9998
Comp9	0.00174521	.	0.0002	1

Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Comp8	Comp9
1	0.3357	-0.2542	0.3355	0.0766	-0.2158	-0.5759	-0.0736	-0.4301	0.3683
2	0.3352	-0.2198	0.3872	0.1605	-0.3747	-0.1749	-0.3058	0.5401	0.3276
3	0.3257	0.3967	0.6053	0.1561	0.1390	0.4862	0.2816	-0.1037	-0.0068
4	0.3347	-0.2235	-0.3305	0.3567	0.6028	0.2563	-0.1806	0.3338	-0.1689
5	0.3356	-0.1647	-0.3325	0.3999	-0.4479	0.2529	-0.0719	-0.5357	0.1819
6	0.3268	0.5612	-0.3435	0.2139	-0.1096	-0.4594	0.3779	0.2177	-0.0335
7	0.3367	-0.2784	-0.0946	-0.4418	0.2791	-0.0592	0.3979	-0.0862	0.5982
8	0.3372	-0.2368	-0.0975	-0.4463	-0.3678	0.2355	0.2360	0.2022	-0.5781
9	0.3321	0.4518	-0.1217	-0.4656	0.0674	0.0409	-0.6535	-0.1359	0.0480

Number of obs. = 34,697, Number of comp. = 9; Trace = 9; Rotation: (unrotated = principal); Rho = 1.0000

Table A4: Full Dataset

Country	First Year	Last Year	Country	First Year	Last Year
Afghanistan	1789	2014	Chile	1789	2013
Albania	1789	2017	China	1789	2017
Algeria	1789	2017	Colombia	1789	2014
Andorra	1993	2015	Comoros	1789	2016
Angola	1789	2012	Congo DR	1789	2011
Antigua and Barbuda	1789	2014	Congo, Republic of	1789	2016
Argentina	1789	2015	Costa Rica	1789	2014
Armenia	1919	2017	Croatia	1941	2016
Australia	1789	2016	Cuba	1789	2013
Austria	1789	2016	Cyprus	1789	2016
Azerbaijan	1995	2015	Czech Republic	1918	2013
Bahamas	1956	2017	Denmark	1789	2015
Bahrain	1789	2001	Djibouti	1789	2016
Bangladesh	1789	2014	Dominica	1789	2014
Barbados	1951	2013	Dominican Republic	1789	2016
Belarus	1994	2016	Ecuador	1789	2013
Belgium	1789	2014	Egypt	1789	2015
Belize	1789	2015	El Salvador	1789	2015
Benin	1789	2016	Equatorial Guinea	1789	2016
Bhutan	1789	2013	Eritrea	1789	2017
Bolivia	1789	2014	Estonia	1918	2016
Bosnia and Herzegovina	1990	2014	Ethiopia	1789	2015
Botswana	1789	2014	Fiji	1789	2014
Brazil	1789	2014	Finland	1809	2015
Brunei	1789	2017	France	1795	2017
Bulgaria	1878	2016	Gabon	1789	2016
Burkina Faso	1789	2015	Gambia	1789	2017
Burma	1789	2015	Georgia	1919	2016
Burundi	1789	2015	German Democratic Republic	1945	1990
Cambodia	1789	2013	Germany	1789	2013
Cameroon	1789	2013	Ghana	1789	2016
Canada	1789	2015	Greece	1822	2015
Cape Verde	1789	2016	Grenada	1789	2013
Central African Republic	1789	2016	Guatemala	1789	2015
Chad	1789	2016	Guinea	1789	2015

Country	First Year	Last Year	Country	First Year	Last Year
Guinea Bissau	1789	2014	Malaysia	1789	2013
Guyana	1789	2015	Maldives	1789	2014
Haiti	1789	2016	Mali	1789	2013
Honduras	1789	2013	Malta	1789	2017
Hong Kong	1789	2016	Marshall Islands	1789	1976
Hungary	1789	2014	Mauritania	1789	2014
Iceland	1789	2016	Mauritius	1789	2014
India	1789	2014	Mexico	1789	2015
Indonesia	1789	2014	Micronesia	1789	1978
Iran	1789	2016	Moldova	1994	2014
Iraq	1789	2014	Monaco	1789	2013
Ireland	1918	2016	Mongolia	1789	2016
Israel	1948	2015	Montenegro	1789	2016
Italy	1861	2013	Morocco	1789	2016
Ivory Coast	1789	2016	Mozambique	1789	2014
Jamaica	1865	2016	Namibia	1789	2014
Japan	1789	2014	Nauru	1789	1950
Jordan	1789	2006	Nepal	1789	2013
Kazakhstan	1999	2016	Netherlands	1789	2017
Kenya	1789	2013	New Zealand	1789	2014
Kiribati	1789	2016	Nicaragua	1789	2016
Kosovo	1789	2014	Niger	1789	2016
Kuwait	1789	1989	Nigeria	1789	2015
Kyrgyzstan	1995	2015	North Korea	1945	2014
Laos	1789	1988	Norway	1789	2013
Latvia	1920	2014	Oman	1789	1999
Lebanon	1789	1942	Pakistan	1789	2013
Lesotho	1789	2015	Palau	1789	2016
Liberia	1789	2011	Palestine (West Bank)	1948	2017
Libya	1789	2017	Panama	1789	2014
Liechtenstein	1789	2017	Papua New Guinea	1789	2012
Lithuania	1918	2016	Paraguay	1811	2013
Luxembourg	1815	2013	Peru	1789	2016
Macedonia	1990	2016	Philippines	1789	2016
Madagascar	1789	2013	Poland	1789	2015
Malawi	1789	2014	Portugal	1789	2016

Country	First Year	Last Year	Country	First Year	Last Year
Qatar	1789	2017	Sweden	1887	2014
Romania	1789	2016	Switzerland	1789	2015
Russia	1789	2016	Syria	1789	2017
Rwanda	1789	2013	Taiwan	1789	2016
Saint Vincent and the Grenadines	1789	2015	Tajikistan	1991	2015
Samoa	1789	2016	Tanzania	1789	1960
San Marino	1789	2016	Thailand	1789	2014
Sao Tome and Principe	1789	2016	Timor Leste	1789	2017
Saudi Arabia	1789	2017	Togo	1789	2015
Senegal	1789	2012	Tonga	1789	2014
Serbia	1804	2016	Trinidad and Tobago	1789	2015
Seychelles	1789	2016	Tunisia	1789	2014
Sierra Leone	1789	2012	Turkey	1789	2015
Singapore	1867	2015	Turkmenistan	1994	2012
Slovakia	1938	2016	Tuvalu	1789	1976
Slovenia	1989	2014	Uganda	1789	2016
Solomon Islands	1789	2014	Ukraine	1991	2014
Somalia	1789	2016	United Arab Emirates	1789	2005
Somaliland	1789	2010	United Kingdom	1832	2015
South Africa	1789	2014	United States	1794	2016
South Korea	1789	2016	Uruguay	1789	2014
South Sudan	1789	2017	Uzbekistan	1789	2016
South Yemen	1789	2017	Vanuatu	1789	2016
Spain	1789	2016	Venezuela	1789	2015
Sri Lanka	1789	2015	Vietnam	1789	2016
St Kitts and Nevis	1789	2015	Vietnam, Democratic Republic of	1789	1975
St Lucia	1789	2016	Yemen	1789	2017
Sudan	1789	2015	Zambia	1789	2016
Suriname	1789	2015	Zimbabwe	1789	2013
Swaziland	1789	2013			

Online Appendix B: Robustness Tests

Table B1: Volatility across Elections

	1
Consecutive elections (log)	-10.249*** (1.544)
Lagged DV	0.989*** (0.20)
<i>Polities</i>	153
<i>Years</i>	221
<i>Obs.</i>	2,359
R2	0.434

Outcome: contestation (incumbent-challenger), differenced. *Estimator:* ordinary least squares with polity and year fixed effects. Standard errors clustered by polity in parentheses. *** p<0.01, ** p<0.05, * p<0.1 Constant omitted.

Table B2: Sample Restriction by Region

<i>Excluded region</i>	West Europe	East Europe Central Asia	Latin America	MENA	Sub-Saharan Africa	North America	East Asia	Southeast Asia	South Asia	Pacific	Caribbean
	1	2	3	4	5	6	7	8	9	10	11
Population (log)	4.490** (1.932)	6.375*** (1.794)	6.656*** (1.767)	4.821*** (1.544)	6.113*** (1.884)	5.551*** (1.828)	6.796*** (1.574)	5.823*** (1.684)	6.271*** (1.818)	6.692*** (1.788)	7.170*** (1.816)
<i>Polities</i>	174	185	177	176	148	194	191	191	189	184	185
<i>Years</i>	226	226	226	226	226	226	226	226	226	226	226
<i>Obs</i>	17,688	18,953	17,921	18,216	17,145	20,044	20,111	19,441	20,147	20,722	20,263
R2	0.095	0.087	0.091	0.087	0.087	0.090	0.087	0.085	0.086	0.089	0.090

Outcome: contestation (incumbent-challenger). All models include per capita GDP (log) and decade dummies. *Estimator:* tobit regression, coefficients and standard errors (clustered by polity) shown only for the variable of theoretical interest. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B3: Incumbent-challenger formula applied to legislative seats

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	7.577*** (1.648)	6.624*** (1.390)	4.768*** (1.828)	9.454*** (1.686)	3.980*** (1.052)	4.218*** (1.510)	4.605*** (1.380)		7.631** (3.196)
GDP pc (log)	21.547*** (2.236)		16.889*** (3.316)	23.62*** (2.229)	5.152*** (1.555)	16.11*** (2.498)	-0.233 (2.855)	-0.069 (0.084)	21.819*** (2.401)
English colony	9.518 (7.680)		9.394 (6.154)	15.25** (7.755)	-4.053 (3.994)	0.299 (6.652)	-0.114 (3.697)	0.039 (0.244)	6.060 (8.076)
Urbanization			38.459** (18.942)				-5.172 (11.44)		
Latitude (log)							-3.852 (2.530)		
Island							-4.102 (6.115)		
Oil income pc							0.000550 (0.000798)		
Years indep (log)							-0.615 (0.937)		
Protestant							0.00166 (0.0489)		
Muslim							-0.0292 (0.0817)		
<i>Fractionalization</i>									
Ethnic						-41.67*** (12.57)	-4.607 (9.328)		
Religious						36.08*** (12.72)	39.26*** (9.677)		
Linguistic						-2.717 (10.80)	-7.093 (7.987)		
Lexical index					16.31*** (0.849)		13.65*** (0.789)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				13.22* (6.767)			2.634 (4.089)		
Mixed				6.579 (6.468)			-2.851 (4.978)		
Other				-9.210 (19.57)			-5.466 (18.15)		
<i>Instruments</i>									
Area (log)								0.627*** (0.061)	
Agricultural Suitability								2.192*** (0.392)	
Region FE		✓					✓		
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	55.94*** (1.854)	49.43*** (1.494)	50.88*** (1.971)	52.53*** (1.493)	38.78*** (1.156)	50.43*** (1.878)	36.19*** (1.251)		
<i>Polities</i>	197	197	163	166	192	177	143		164
<i>Years</i>	227	226	111	215	215	120	107		215
<i>Observations</i>	21,721	21,721	11,817	19,320	14,536	13,425	8,082		19,001
<i>Pseudo R2</i>	0.080	0.111	0.040	0.089	0.107	0.040	0.088		
<i>Chi2</i>									468.35

Replicates Table 4 with a different measure of contestation – the incumbent-challenger formula applied to seats won in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B4: Largest-party formula applied to legislative votes

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	5.244*** (1.110)	4.653*** (0.937)	3.622*** (1.386)	6.740*** (1.203)	2.921*** (0.652)	3.015*** (1.099)	5.075*** (1.018)		4.697** (2.161)
GDP pc (log)	15.94*** (1.711)		13.16*** (2.524)	17.13*** (1.688)	3.129*** (1.145)	12.29*** (2.140)	1.171 (2.161)	-0.073 (0.085)	16.097*** (1.830)
English colony	10.81** (4.647)		10.24** (4.330)	17.47*** (5.296)	-0.244 (2.118)	5.169 (5.092)	8.679** (3.757)	0.000 (0.236)	7.304 (5.121)
Urbanization			29.79** (12.53)				0.378 (8.775)		
Latitude (log)							-4.376** (1.897)		
Island							5.227 (4.803)		
Oil income pc							-0.00279* (0.00143)		
Years indep (log)							-1.054 (0.864)		
Protestant							0.193*** (0.0480)		
Muslim							0.0856 (0.0692)		
<i>Fractionalization</i>									
Ethnic						-38.11*** (9.820)	-16.12 (10.07)		
Religious						17.34* (10.33)	6.840 (7.311)		
Linguistic						7.730 (8.932)	11.81 (8.626)		
Lexical index					11.76*** (0.606)		9.289*** (0.593)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				15.80*** (5.439)			7.999** (3.306)		
Mixed				12.66** (5.448)			4.306 (4.245)		
Other				-4.775 (18.81)			-22.41*** (8.124)		
<i>Instruments</i>									
Area (log)								0.631*** (0.061)	
Agricultural suitability								2.198*** (0.396)	
Region FE		✓						✓	
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	39.06*** (1.639)	34.38*** (1.320)	35.65*** (1.765)	36.09*** (1.320)	24.80*** (1.137)	35.11*** (1.688)	22.49*** (1.252)		
<i>Polities</i>	196	196	162	165	188	176	141		163
<i>Years</i>	226	226	111	215	215	120	107		215
<i>Observations</i>	20,277	20,277	10,904	18,037	13,511	12,364	7,465		17,852
<i>Pseudo R2</i>	0.103	0.139	0.055	0.119	0.147	0.053	0.132		
<i>Chi2</i>									349.24

Replicates Table 4 with a different measure of contestation – the largest-party formula applied to votes in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B5: Largest-party formula applied to legislative seats

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	5.220*** (0.990)	4.596*** (0.866)	2.975** (1.203)	6.183*** (1.008)	2.705*** (0.671)	2.934*** (0.958)	2.599** (1.082)		4.007** (1.901)
GDP pc (log)	14.11*** (1.483)		10.59*** (2.373)	15.29*** (1.572)	3.652*** (1.103)	10.59*** (1.707)	-2.807 (2.231)	-0.62 (0.085)	14.441*** (1.561)
English colony	3.797 (4.399)		3.577 (4.070)	8.162* (4.689)	-7.486*** (2.191)	-2.371 (4.354)	-3.000 (3.852)	0.053 (0.244)	0.892 (4.714)
Urbanization			32.36** (12.86)				7.883 (9.135)		
Latitude (log)							-1.868 (2.020)		
Oil income pc							-2.540 (4.173)		
Years indep (log)							8.45e-05 (0.000611)		
Protestant							-1.438* (0.770)		
Muslim							0.116*** (0.0417)		
							-0.0574 (0.0691)		
<i>Fractionalization</i>							2.824		
Ethnic						-29.77*** (8.444)	2.824 (8.263)		
Religious						18.22** (8.597)	22.17** (9.812)		
Linguistic						4.391 (7.587)	-0.162 (6.521)		
Lexical index					10.68*** (0.576)		8.810*** (0.561)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				11.55** (4.521)			4.503 (3.275)		
Mixed				5.649 (4.432)			0.589 (3.214)		
Other				3.194 (13.70)			4.910 (11.57)		
<i>Instruments</i>									
Area (log)								0.628*** (0.060)	
Agricultural suitability								2.170*** (0.381)	
Region FE		✓							
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	37.46*** (1.335)	33.53*** (1.247)	34.52*** (1.533)	35.48*** (1.206)	26.36*** (1.099)	33.77*** (1.424)	24.08*** (1.221)		
<i>Polities</i>	197	197	163	166	192	177	143		164
<i>Years</i>	227	227	111	215	215	120	107		215
<i>Observations</i>	21,925	21,925	12,014	19,517	14,677	13,607	8,192		19,181
<i>Pseudo R2</i>	0.083	0.114	0.046	0.094	0.116	0.043	0.104		
<i>Chi2</i>									390.29

Replicates Table 4 with a different measure of contestation – the largest-party formula applied to seats won in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B6: Top-two formula applied to legislative votes

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	6.529*** (1.693)	5.896*** (1.358)	4.293** (1.820)	8.810*** (1.789)	3.038*** (0.942)	3.127** (1.533)	6.297*** (1.294)		7.192** (3.257)
GDP pc (log)	22.01*** (2.377)		17.25*** (3.265)	24.07*** (2.239)	4.520*** (1.507)	17.17*** (2.887)	3.401 (2.611)	-0.0725 (0.085)	21.993*** (2.586)
English colony	18.20** (7.417)		17.25*** (5.979)	25.13*** (7.838)	5.256 (3.634)	8.027 (7.023)	9.647** (4.263)	0.003 (0.235)	14.067* (7.947)
Urbanization			40.42** (17.11)				2.847 (9.916)		
Latitude (log)							-8.269*** (2.044)		
Oil income pc							10.12** (4.304)		
Years indep (log)							-0.00238* (0.00134)		
Protestant							-1.150 (0.993)		
Muslim							0.168*** (0.0536)		
							0.0513 (0.0713)		
<i>Fractionalization</i>									
Ethnic							-42.66*** (13.60)	-19.45* (10.07)	
Religious							35.43** (14.10)	15.00* (8.709)	
Linguistic							5.574 (12.39)	14.39 (9.867)	
Lexical index					15.69*** (0.870)		12.29*** (0.762)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				15.55** (7.444)			5.039 (3.999)		
Mixed				12.78* (7.170)			1.734 (5.302)		
Other				-16.83 (24.91)			-39.20*** (10.92)		
<i>Instruments</i>									
Area (log)								0.630*** (0.061)	
Agricultural suitability								2.220*** (0.401)	
Region FE		✓					✓		
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	54.31*** (2.367)	47.15*** (1.719)	46.80*** (2.386)	49.49*** (1.744)	33.21*** (1.302)	47.40*** (2.393)	28.94*** (1.500)		
<i>Polities</i>	196	196	162	165	188	176	141		163
<i>Years</i>	226	226	111	215	215	120	107		215
<i>Observations</i>	20,277	20,277	10,904	18,037	13,511	12,364	7,465		17,852
<i>Pseudo R2</i>	0.097	0.131	0.049	0.111	0.136	0.048	0.119		
<i>Chi2</i>									398.36

Replicates Table 4 with a different measure of contestation – the top-two formula applied to votes won in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B7: Top-two formula applied to legislative seats

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	6.724*** (1.455)	5.901*** (1.172)	4.037** (1.614)	8.408*** (1.468)	3.225*** (0.924)	3.466*** (1.344)	3.809*** (1.244)		6.351** (2.848)
GDP pc (log)	19.64*** (2.012)		15.20*** (3.010)	21.55*** (2.004)	5.131*** (1.377)	14.81*** (2.261)	-0.0289 (2.532)	-0.0671 (0.084)	19.868*** (2.181)
English colony	6.578 (6.697)		6.321 (5.339)	11.63* (6.754)	-5.902* (3.298)	-2.570 (5.894)	-2.620 (3.694)	0.040 (0.243)	3.321 (7.084)
Urbanization			36.94** (16.84)				0.795 (10.20)		
Latitude (log)							-3.327 (2.236)		
Oil income pc							-1.993 (4.974)		
Years indep (log)							0.000253 (0.000695)		
Protestant							-1.528* (0.860)		
Muslim							0.0603 (0.0472)		
							-0.0659 (0.0720)		
<i>Fractionalization</i>									
Ethnic							-35.87*** (11.59)	2.497 (8.079)	
Religious							32.57*** (11.46)	29.90*** (9.331)	
Linguistic							0.629 (10.05)	-3.145 (6.503)	
Lexical index					14.28*** (0.744)			11.52*** (0.651)	
<i>Electoral systems</i>									
Majoritarian				[omitted]				[omitted]	
Proportional				11.99** (6.065)				2.004 (3.713)	
Mixed				5.090 (5.544)				-2.081 (3.899)	
Other				-6.804 (18.14)				0.497 (15.86)	
<i>Instruments</i>									
Area (log)								0.629*** (0.060)	
Agricultural suitability								2.151*** (0.384)	
Region FE		✓						✓	
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	49.75*** (1.709)	42.23*** (1.347)	43.77*** (1.889)	46.48*** (1.372)	32.66*** (1.074)	43.49*** (1.789)	28.92*** (1.227)		
<i>Polities</i>	197	197	163	166	192	177	143		164
<i>Years</i>	227	227	111	215	215	120	107		215
<i>Observations</i>	21,782	21,782	11,889	19,385	14,585	13,466	8,120		19,059
<i>Pseudo R2</i>	0.082	0.116	0.043	0.092	0.115	0.042	0.101		
<i>Chi2</i>									471.75

Replicates Table 4 with a different measure of contestation – the top-two formula applied to seats won in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B8: Incumbent-challenger formula applied to legislative votes

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	7.812*** (1.826)	6.970*** (1.539)	5.369*** (2.055)	10.34*** (1.972)	3.920*** (1.011)	4.027** (1.700)	7.287*** (1.449)		8.038** (3.474)
GDP pc (log)	25.24*** (2.584)		20.10*** (3.700)	26.97*** (2.485)	4.814*** (1.683)	18.93*** (3.192)	1.322 (2.971)	-0.0779 (0.0850)	25.12*** (2.798)
English colony	21.84*** (8.015)		21.95*** (6.716)	31.49*** (8.828)	6.814* (3.750)	12.68 (7.922)	16.37*** (5.583)	0.00330 (0.236)	17.13** (8.622)
Urbanization			46.77** (19.36)				0.495 (12.42)		
Latitude (log)							-4.573 (2.818)		
Oil income pc							8.336 (6.672)		
Years indep (log)							-0.00428* (0.00222)		
Protestant							-0.253 (1.131)		
Muslim							0.115* (0.0652)		
							0.129 (0.0870)		
<i>Fractionalization</i>									
Ethnic							-52.27*** (14.62)	-29.87** (12.05)	
Religious							35.16** (15.68)	23.12** (10.00)	
Linguistic							-0.705 (13.55)	7.532 (11.85)	
Lexical index					18.42*** (0.956)			14.74*** (0.884)	
<i>Electoral systems</i>									
Majoritarian				[omitted]				[omitted]	
Proportional				22.32*** (8.615)				12.41** (5.070)	
Mixed				17.05** (8.328)				3.744 (6.486)	
Other				-10.19 (26.79)				-43.23*** (12.58)	
<i>Instruments</i>									
Area (log)								0.632*** (0.0612)	
Agricultural suitability								2.209*** (0.401)	
Region FE		✓						✓	
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	61.06*** (2.455)	53.52*** (1.873)	54.86*** (2.641)	56.19*** (1.792)	38.71*** (1.358)	54.13*** (2.598)	35.83*** (1.558)		
<i>Polities</i>	196	196	162	165	188	176	141		163
<i>Years</i>	226	226	111	215	215	120	107		215
<i>Observations</i>	20,114	20,114	10,846	17,914	13,413	12,262	7,415		17,729
<i>Pseudo R2</i>	0.096	0.128	0.051	0.109	0.136	0.050	0.117		
<i>Chi2</i>									407.35

Replicates Table 4 with a different measure of contestation – the top-two formula applied to votes won in elections to the lower (or unicameral) chamber of the national legislature. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B9: Top-two formula applied to combined measure

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	5.177*** (1.567)	5.092*** (1.253)	3.351* (1.725)	7.432*** (1.638)	2.360*** (0.840)	2.182 (1.425)	5.021*** (1.095)		7.048** (3.119)
GDP pc (log)	19.84*** (2.253)		15.81*** (3.054)	22.56*** (2.116)	4.100*** (1.358)	15.54*** (2.679)	3.310 (2.167)	-0.0692 (0.0845)	20.62*** (2.407)
English colony	16.40** (7.129)		16.26*** (5.694)	21.42*** (7.381)	5.503* (3.325)	8.239 (6.738)	8.928** (3.559)	0.0318 (0.234)	11.82 (7.532)
Urbanization			42.20** (16.40)				6.424 (8.970)		
Latitude (log)							-6.219*** (1.654)		
Oil income pc							10.24** (4.137)		
Years indep (log)							-0.00328** (0.00135)		
Protestant							-0.848 (0.953)		
Muslim							0.129*** (0.0494)		
							0.000449 (0.0599)		
<i>Fractionalization</i>									
Ethnic							-34.54** (13.97)	-17.12* (9.772)	
Religious							33.67** (13.60)	13.37* (7.807)	
Linguistic							-2.745 (12.58)	9.856 (9.849)	
Lexical index					15.27*** (0.791)		12.16*** (0.748)		
<i>Electoral systems</i>									
Majoritarian				[omitted]			[omitted]		
Proportional				12.69* (6.705)			3.824 (3.351)		
Mixed				11.37 (6.943)			1.938 (5.321)		
Other				-14.62 (23.25)			-32.96*** (10.76)		
<i>Instruments</i>									
Area (log)								0.627*** (0.0611)	
Agricultural suitability								2.180*** (0.392)	
Region FE		✓							
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	52.25*** (2.111)	45.44*** (1.543)	44.94*** (2.172)	47.85*** (1.565)	32.16*** (1.136)	45.76*** (2.206)	28.64*** (1.347)		
<i>Polities</i>	197	197	163	166	191	177	142		164
<i>Years</i>	227	227	111	215	215	120	107		215
<i>Observations</i>	21,264	21,264	11,576	18,946	14,476	13,038	7,996		18,751
<i>Pseudo R2</i>	0.089	0.124	0.046	0.100	0.129	0.043	0.114		
<i>Chi2</i>									398.52

Replicates Table 4 with a different measure of contestation – the top-two formula applied to votes won in elections to the lower (or unicameral) chamber of the national legislature and presidential elections. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Table B10: Largest party formula applied to combined measure

<i>Estimator</i>	Tobit 1	Tobit 2	Tobit 3	Tobit 4	Tobit 5	Tobit 6	Tobit 7	Tobit, IV	
								8a	8b
Population (log)	3.973*** (0.941)	3.646*** (0.771)	2.713** (1.203)	5.325*** (0.978)	2.174*** (0.520)	2.017** (0.951)	3.328*** (0.880)		3.934** (1.983)
GDP pc (log)	13.93*** (1.559)		11.41*** (2.225)	15.53*** (1.523)	3.224*** (0.989)	10.89*** (1.923)	0.621 (1.832)	-0.0737 (0.0844)	14.48*** (1.652)
English colony	9.973** (4.313)		10.06*** (3.903)	14.62*** (4.607)	1.341 (1.887)	5.468 (4.644)	6.894** (3.091)	0.0220 (0.235)	6.479 (4.751)
Urbanization			30.19*** (11.40)				8.214 (7.997)		
Latitude (log)							-3.477** (1.615)		
Oil income pc							5.254 (4.016)		
Years indep (log)							-0.00235** (0.00116)		
Protestant							-0.736 (0.843)		
Muslim							0.142*** (0.0484)		
							0.0156 (0.0545)		
<i>Fractionalization</i>									
Ethnic							-33.54*** (9.376)	-14.50* (8.115)	
Religious							16.58* (9.432)	4.509 (6.877)	
Linguistic							5.423 (8.504)	9.838 (7.204)	
Lexical index					10.66*** (0.509)		8.637*** (0.520)		
<i>Electoral systems</i>									
Majoritarian				[omitted]				[omitted]	
Proportional				12.15*** (4.465)			6.414** (2.613)		
Mixed				9.513** (4.671)			2.803 (3.679)		
Other				-2.785 (16.75)			-14.34** (7.291)		
<i>Instruments</i>									
Area (log)								0.627*** (0.0609)	
Agricultural suitability								2.103*** (0.390)	
Region FE		✓						✓	
Decade FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sigma	35.35*** (1.351)	31.21*** (1.128)	31.84*** (1.444)	32.61*** (1.077)	22.42*** (0.989)	31.77*** (1.422)	20.69*** (1.096)		
<i>Polities</i>	197	197	163	166	191	177	142		164
<i>Years</i>	227	227	111	215	215	120	107		215
<i>Observations</i>	21,424	21,424	11,710	19,106	14,633	13,134	8,077		18,871
<i>Pseudo R2</i>	0.097	0.133	0.052	0.110	0.142	0.050	0.127		
<i>Chi2</i>									376.03

Replicates Table 4 with a different measure of contestation – the largest-party formula applied to votes won in elections to the lower (or unicameral) chamber of the national legislature and in presidential elections. *Estimator*: tobit, left-censored at 0. Standard errors in parentheses, clustered by polity. ***p<0.01 **p<0.05 *p<0.10 (two-tailed) Constant omitted.

Online Appendix C: Relationship to Democracy

Our ambition is to measure contestation – not the larger, more diffuse concept of electoral democracy (not to mention the even larger and more diffuse concept of democracy, *tout court*). Nonetheless, the two concepts are closely related, so it is worth reviewing this conceptual and empirical relationship before continuing.

Arguably, indices of contestation offer a *reflection* of the freeness/fairness of an election. If the incumbent party buys votes on a massive scale, imprisons opposition leaders, monopolizes sources of campaign finance and access to the media, or simply prevents other parties from appearing on the ballot, that election is likely to receive a lower contestation score, as pointed out. Note that it is usually the incumbent party that violates democratic norms – tilting the playing field in its favor – for it is by construction the largest party and is in a position to utilize the apparatus of the state. As such, the incumbent-challenger index – and, to some extent, the other indices introduced above – may be regarded as outcome-based measures of electoral democracy (Altman & Pérez-Liñán 2002).

However, they are imperfect measures. Sometimes, incumbents win big even though the rules of the game are fair (e.g., Lyndon Johnson in the US presidential election of 1964). Sometimes, incumbents lose even though the rules of the game are unfair (e.g., Malaysia in 2018). Likewise, our measures take no notice of suffrage rights or ways in which electoral participation is informally restricted. Needless to say, our indices do not reflect dimensions of democracy that are non-electoral in character, e.g., judicial independence, rule of law, civil liberty, horizontal accountability, the strength of civil society, et al.

To shed light on the empirical relationship between electoral contestation and democracy we present a series of descriptive statistics in Table C1. The first set correlates the incumbent-challenger index with prominent indices of electoral democracy – the Polyarchy (aka Electoral democracy) index from the Varieties of Democracy project (Teorell et al. 2016), the Polity2 index from the Polity IV project (Marshall et al. 2014), the Lexical index of electoral democracy (Skaaning et al. 2015), and the BMR index from Boix, Miller & Rosato (2013). The second set focus on more specific components often associated with democracy – clean elections, party institutionalization, liberal freedoms, participation, and equality – as measured by the Variety of Democracy dataset (Coppedge et al. 2018).

Although correlations between contestation and various measures of democracy are fairly strong in the unrestricted sample (column 1), they drop considerably in the sample that excludes 0 scores (column 1). This suggests that a judgment of no contestation accords with a judgment of no (or very little) democracy: multi-party elections, after all, are commonly viewed as a *sine qua non* of democracy. However, where multi-party elections are allowed, degrees of contestation do not accord very closely to degrees of electoral democracy as measured by standard indices of the latter.

Table C1: Contestation and Democracy

	Contestation	
	Full	θ^*
<i>Electoral democracy indices</i>		
Polyarchy (V-Dem)	0.80	0.48
Polity2 (Polity IV)	0.73	0.54
Lexical (Skaaning et al.)	0.78	0.47
BMR (Boix et al.)	0.69	0.50
<i>Democracy components</i>		
Clean elections (V-Dem)	0.84	0.51
Party institutionalization (V-Dem)	0.53	0.40
Liberal (V-Dem)	0.66	0.52
Participatory (V-Dem)	0.74	0.46
Egalitarian (V-Dem)	0.59	0.37

Note: Pearson's r correlation between contestation (*incumbent-challenger*) and various democracy indices and democracy components, tested in a full sample and a sample excluding zero contestation scores.

Online Appendix C: References

- Boix, Carles, Michael Miller, Sebastian Rosato. 2013. "A complete data set of political regimes, 1800–2007." *Comparative Political Studies* 46.12: 1523-1554.
- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerløw, Adam Glynn, Allen Hicken, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Aksel Sundtröm, Eitan Tzelgov, Luca Uberti, Yi-ting Wang, Tore Wig, and Daniel Ziblatt. 2018. "V-Dem Codebook v8" Varieties of Democracy (V-Dem) Project.
- Marshall, Monty, Ted Gurr, Keith Jagers. 2014. *Polity IV project: Dataset users' manual*. <http://www.systemicpeace.org/inscr/p4manualv2013.pdf>
- Skaaning, Svend-Erik, John Gerring, Henrikas Bartusevičius. 2015. "A Lexical Index of Electoral Democracy." *Comparative Political Studies* 48:12 (October) 1491-1525.
- Teorell, Jan, Michael Coppedge, Svend-Erik Skaaning, Staffan I. Lindberg. 2016. "Measuring Electoral Democracy with V-Dem Data: Introducing a New Polyarchy Index." Working Paper Series 2016:25, Varieties of Democracy Institute.