Introducing the Historical Varieties of Democracy Dataset: Political Institutions in the Long 19th Century

Carl Henrik Knutsen, Jan Teorell, Agnes Cornell, John Gerring, Haakon Gjerløw, Svend-Erik Skaaning, Tore Wig, Daniel Ziblatt, Kyle L. Marquardt, Dan Pemstein, Brigitte Seim

April 2018

Working Paper

SERIES 2018:65

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.

Introducing the Historical Varieties of Democracy Dataset: Political Institutions in the Long 19th Century*

Carl Henrik Knutsen, University of Oslo
Jan Teorell, Lund University
Agnes Cornell, Lund University
John Gerring, University of Texas at Austin
Haakon Gjerløw, University of Oslo
Svend-Erik Skaaning, Aarhus University
Tore Wig, University of Oslo
Daniel Ziblatt, Harvard University
Kyle L. Marquardt, University of Gothenburg
Dan Pemstein, North Dakota State University
Brigitte Seim, University of North Carolina at Chapel Hill

* We gratefully acknowledge coding efforts and other research assistance provided by Solveig Bjørkholt, Ben Chatterton, Vlad Ciobanu, Lee Cojocaru, Vilde Lunnan Djuve, Kristian Frederiksen, Sune Orloff Hellegaard, Bernardo Isola, Sindre Haugen, Haakon Haugevik Jernsletten, Claudia Maier, Swaantje Marten, Selemon Negash, Moa Olin, Konstantinos Skenteri, and Katharina Sibbers; help with constructing vignettes by Amanda Haraldsson, Kersti Hazell and Alexander Kuhn; assistance with implementing the measurement model by Joshua Krusell and Johannes von Römer; and help with creating expert surveys, managing the data, coordinating, discussing and resolving conceptual and technical issues, etc., by numerous people at the V-Dem Institute in Gothenburg, including Frida Andersson, Staffan I. Lindberg, Valeriya Mechkova, Moa Olin, Josefine Pernes, Laura Saxer, and Natalia Stepanova. We also thank our country experts and numerous scholars (who are too many to mention), both inside and outside the wider V-Dem team, for inputs at various stages in the process. Finally, we acknowledge funding from various larger and smaller grants for the data collection for Historical V-Dem (see V-Dem Organization and Management document for details). The two largest sources of funding were Swedish Research Council Grant 421-2014-1283, PI: Jan Teorell, Department of Political Science, Lund University and Norwegian Research Council Grant pnr 240505, PI: Carl Henrik Knutsen, Department of Political Science, University of Oslo. Another main funding source was Innovationsfonden Grant 4110-00002B, PI: Svend-Erik Skaaning, Department of Political Science, Aarhus University. Further, the V-Dem data collection 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 Historical Varieties of Democracy Dataset (Historical V-Dem) is a new dataset containing about 260 indicators, both factual and evaluative, describing various aspects of political regimes and state institutions. The dataset covers 91 polities globally – including most large, sovereign states, as well as some semi-sovereign entities and large colonies – from 1789 to 1920 for many cases. The majority of the indicators are also included in the Varieties of Democracy dataset, which covers the period from 1900 to the present – and together these two datasets cover the bulk of "modern history". Historical V-Dem also includes several new indicators, covering features that are pertinent for 19th century polities. We describe the data, the process of coding, and the different strategies employed in Historical V-Dem to cope with issues of reliability and validity and ensure inter-temporal- and cross-country comparability. To illustrate the potential uses of the dataset we provide a descriptive account of patterns of democratization in the "long 19th century." Finally, we perform an empirical investigation of how inter-state war relates to subsequent democratization.

1. Introduction

Although many datasets describe political institutions in countries across the world (see Coppedge et al. 2017a), the Varieties of Democracy (V-Dem) dataset (Coppedge et al. 2017b,c) is the most wide-ranging – including several hundred indicators and indices. While country coverage is impressive, historical coverage begins in 1900, omitting half of the period commonly included under the rubric of "modern history". This omission poses a hindrance to systematic comparative description of institutional and political developments during this era, but also implies that several theories of political development (in the given time period and more generally) lack the requisite data for testing.

To remedy this situation, we introduce the Historical Varieties of Democracy (Historical V-Dem) dataset. Historical V-Dem spans all major countries and several other polities in the world between 1789 and 1920, encapsulating what Hobsbawm (1962; 1975; 1987) has labeled the "Long 19th Century". These data mean that most of the indicators contained in V-Dem now extend back to 1789, offering a continuous time series across more than 225 years for many polities. Historical V-Dem also provides several new indicators, many of them focused on features of state institutions and state capacity and on the type of coalitions that supported political regimes.

In this article, we describe the dataset and the data collection process. Next, we address issues of reliability, validity, and inter-temporal- and cross-country comparability, and describe our strategy for dealing with them. Finally, we illustrate the potential uses of the data with two empirical applications. First, we map global patterns of democratization across the "long 19th century" using several measures from Historical V-Dem and comparing these patterns with those displayed by Polity2, one of the most widely used existing measures. Second, we analyze the relationship between international war and subsequent regime change along different dimensions. A key finding is that war participation correlates positively with indicators related to the electoral dimension, such as clean elections and suffrage, but not with other aspects of democracy.

2. Historical V-Dem and extant datasets

Despite the proliferation of high-quality datasets describing 20th and 21st century political institutions, there is a dearth of data for the 19th century. Of the widely used indices, only a few (e.g., Marshall et al. 2015; Boix et al 2013) extend back to 1800. Moreover, the quality and level of detail for the 19th century coding in the historical time series of measures such as Polity2 are sometimes questionable, as we detail below. Further, they cover a limited range of institutional features, grounded in a specific conception of democracy.

One obvious issue stemming from the dearth of systematically compiled and comparable cross- country data on historical institutions relates to lacking description of institutional features and developments in the long 19th century. Key descriptive questions in comparative politics are thus left open. For example, did the long, first wave of democratization stretch back to the beginning of the 19th century (Huntington 1991) or erupt only after WWI (Doorenspleet 2005)? Were there separate sub-waves of democratization after the 1848 revolutions (Weyland 2014)? Was the movement towards democracy across the long 19thcentury discontinuous or gradual, and was it monotonic or characterized by reversals (Congleton 2011; Ziblatt, 2017)?

The lack of data also means that scholars more generally have been unable to satisfactorily address key questions pertaining to the causes and consequences of institutional development. Note that the link between institutions and most outcomes of interest (e.g., economic growth) is difficult to parse because of limited variation and because of the sluggish nature of institutions (and many outcomes). Only with a suitably long time-series can one hope to disentangle cause and effect (Knutsen, Møller and Skaaning 2016). Historical V-Dem thus opens up new opportunities for social scientists studying the historical trajectories of political-institutional developments – including sequences of institutional reforms in different areas – as well as the causes and effects of political-institutional developments.

3. What does Historical V-Dem cover?

Historical V-Dem is divided into 10 surveys, covering different areas of political life: Elections; Parties; Executive; Legislature; Judiciary; Civil Liberties; State; Civil Society; Media; and Political Equality. There are two types of indicators: factual indicators coded by RAs ("*A* indicators") and evaluative indicators coded by country experts ("*C* indicators"). *A* indicators involve features such as election dates, names of local government entities, the legal status of slavery, and the

existence of entities such as statistical agencies or national banks. *C* indicators pertain to features such as the extent of election violence, the relative power of elected and non-elected offices at the local level, *de facto* freedom from forced labor, and the extent to which recruitment to the bureaucracy is merit-based.

In total, there are 149 *C* indicators and 110 *A* indicators in Historical V-Dem. Appendix II provides condensed lists of all indicators (full details in the V-Dem codebook). 129 C indicators are adopted from V-Dem, whereas there 20 are new C indicators. About 50 of the 110 A indicators are new to Historical V-Dem. Many of the new indicators are of special relevance for the 19th century.

The 19th century was an era of state building, and Historical V-Dem contains several new indicators pertaining to the development of state bureaucracies, armed forces, and various other agencies relevant for the capacity of states to gather information, monitor citizens and project power. For example, Historical V-Dem includes several indicators focused on how bureaucrats (and army officers) are recruited and remunerated – capturing important dimensions of a "Weberian" bureaucracy. These variables will, e.g., allow for systematic, empirical studies of processes of modern state formation – a core area of political science where most empirical contributions to date have been based on lengthy case narratives (e.g., Fukuyama 2014).

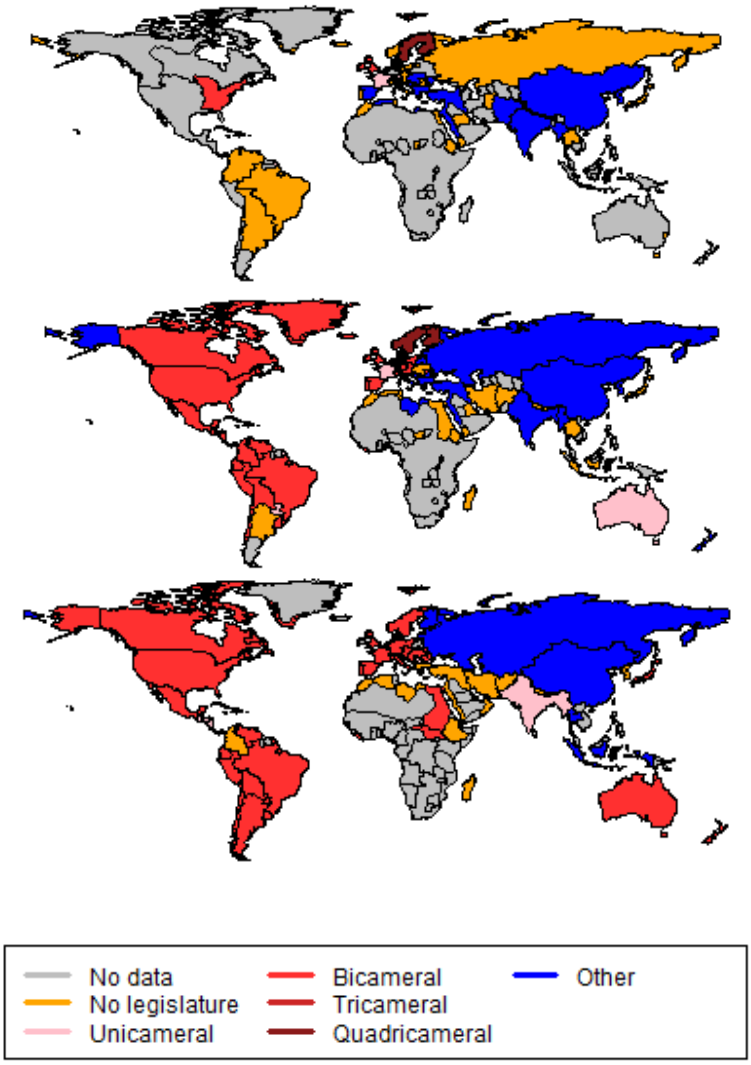
Second, Historical V-Dem includes new indicators pertaining to "regimes" – understood here as a set of formal and/or informal rules that govern the choice of political leaders and their exercise of power. For instance, indicators capture when and how a particular regime ended, the size of regime support coalitions, and which social groups are included in that coalition. These data will allow for empirical testing of arguments pertaining to particular social groups, e.g., agrarian elites or urban middle class, and their relevance for regime stability and change (see, e.g., Moore 1966; Ansell and Samuels 2015). Likewise, they allow for test9ng whether size of regime support coalition has implications for policymaking in foreign and domestic policy arenas (Bueno de Mesquita et al. 2003).

Polities included in Historical V-Dem are a) sizeable (>250000 inhabitants); b) sovereign during an extended time period between 1789 and 1900, either in a formal-juridical or *de facto* sense; and, c) are continuous with present-day states.[1] The resulting sample includes 91 polities – 14 from Africa and the Middle East, 21 from the Americas, 14 from Asia and the Pacific, and 42

[1] Regarding b), we include data for 1789-1920 even if a unit was not independent during the entire period, given that the area is not directly covered by the coding of another polity. To exemplify, this means that Brazil is coded from 1789-1920, even if Brazil became truly independent from Portugal only in 1822.

from Europe – whereof 71 are listed by Gleditsch & Ward (1999). Details on the sample are presented in Appendix I.

Figure 1: Existence and type of legislature (*v3lgbicam)* for 1790 (top), 1850 (middle), and 1899 (bottom).



Note: The maps are produced for Historical V-Dem by digitizing and editing online map sources (see anonymized reference). Correspondence with Historical V-Dem units is, however, still not entirely accurate, especially for non-sovereign entities.

The modal time series is 1789-1920. More specifically, 41 polities are coded for this interval (66 start in 1789). However, some polities cease to exist as independent entities well before 1920, such as Bavaria (coded 1789-1871) with the creation of the German Empire. Others cease to exist for a period of time before they reappear. Tuscany, for example, is not

coded between 1807 and 1814 as it was annexed by France under Napoleon. The rule is that a particular area should not be coded for more than one political entity in a single year, and we have carefully gone through the history of border changes and specified the entities for the different parts of the time period (see the V-Dem countries document for details). Figure 1 maps the polities included in Historical V-Dem in 1790, 1850 and 1900, respectively, coloring countries by the existence and chamber structure of the legislature (*v3lgbicam*), one of the indicators that cover all 91 polities.

4. How was Historical V-Dem constructed?

Constructing this dataset required significant human and financial resources. Planning started in 2013, using as our point of departure the contemporary V-Dem codebook (Coppedge et al. 2017b). Successive rounds of deliberation were required in order to identify contemporary V-Dem questions to a) omit, b) adjust (in order to fit the historical context), or c) create anew. Pilot surveys were conducted on Denmark and Colombia in 2014, after which we received comments and identified potentially problematic questions that needed to be dropped or revised.

Although V-Dem coding for the contemporary era (1900–) rests on a group of coders (generally about five per country), it was not feasible to achieve the same complement for the historical era. Detailed historical knowledge of political affairs is much rarer than knowledge of contemporary political affairs, especially with respect to small and understudied countries. Under these circumstances, only a few experts around the world would be able to code Bavaria, Madagascar or Oman in 1800. Thus, we followed a narrow strategy of recruitment, seeking to identify one or two highly qualified experts for each historical case. We also compensated experts for their time in a fairly generous fashion (1250 to 2000 Euro per country, depending on estimated workload), with the understanding that they would need to consult sources in order to answer many of the questions – a time-consuming process.

Team members and research assistants compiled long lists of potential country experts, employing scholarly networks and web- and literature searches. Ideal experts should have an academic track record working on the political history of the country. Experts with identifiable competencies in a broad range of political-institutional features were prioritized, and, everything else equal, experts with comparative knowledge of other countries were also prioritized (see V-Dem Organization and Management document for details). In the end, most experts were historians or historically oriented political scientists. A few experts were asked to code more than

one polity if they had comprehensive knowledge of different polities (for example, the expert for Baden also coded Würtemberg). The coding was conducted through a web-platform constructed for V-Dem and customized for Historical V-Dem. Experts had the opportunity to contact the team with questions of clarification and information about potential issues with the pre-coded data on, e.g., election dates or heads of state and government. These issues were then discussed by the team, and identified errors were corrected before the expert ensued coding.

Country-expert coding (including updated coding for the pilot countries) started in December 2015 and is still ongoing, currently with a special focus on double-coding using a second country expert. (The ambition is to have a high ratio of double coded polities for updated versions of Historical V-Dem within the next couple of years.) Research assistants, located at several universities, were involved in coding the *A* variables. Thereafter, codings would be checked by a team member or another RA for validation (and possible adjustments).[2]

5. Methodological problems and solutions

The specificity of most indicators in the Historical V-Dem dataset ameliorates the fuzziness of questions in other datasets, which often pertain to diffuse topics as "executive constraints" or the "competitiveness of executive recruitment" (Polity IV). However, this specificity also places a tremendous burden on coders to ascertain the facts of a historical case, e.g., to pin down the extent of vote fraud in an election. Most experts agreed to be publically acknowledged for their work on a particular country, ensuring full transparency and offering an additional incentive to provide accurate coding.

As with contemporary V-Dem, we faced a challenge in achieving equivalence across countries and experts. We want to ensure that when, e.g., scores between France and Russia in 1880 differ, this is because the situation in these two countries differ and not simply because our expert on France is more or less "conservative" than the Russian expert. We therefore employ a latent variable model to generate estimates based on various sources of information, described below, anchoring scores across time and space to a common scale. Point estimates in this dataset are accompanied by uncertainty estimates (Pemstein et al. 2017) to reflect measurement error; for additional information regarding uncertainty, experts also rate their own subjective certainty

[2] For the pre-unification German and Italian states, we employed a separate German RA and Italian RA, respectively, for many A questions. This reflects the demanding source situation for these small, no longer existing states, and the importance of identifying and reading native-language sources. These RAs were provided with similar instructions and coding templates as the "global" RAs for each question.

(from 0-100) for each observation (typically a variable-country-year). Issues of uncertainty are perhaps even more pertinent for the historical period than more recent years, due to a dearth of sources and fewer scholars that specialize in the political institutions of this period. Accordingly, uncertainty about historical point estimates is generally higher than in contemporary V-Dem.

Incorporating historical ratings into the V-Dem modeling framework required the team to implement several model refinements.[3] Regarding key sources of information fed into the measurement model, historical experts were encouraged with an additional monetary incentive to code three extra countries for a single year. Participants in this additional coding selected the three countries from a list of six (USA, UK, France, Mexico, China, and Russia), and coded all variables for the first year after 1900 with an election for each selected country. This procedure provides us with one source of information for assessing how historical experts differ in their understanding of the question scale.

Second, *all* historical experts coded an identical set of indicator-specific anchoring vignettes (King & Wand, 2007) prior to coding their cases. Vignettes provide a powerful tool for addressing differences in ordinal scale perception ("Differential Item Functioning") by allowing us to compare coders who do not share expertise across cases. In our case, vignettes represent hypothetical cases specific to each indicator that have two plausible scores on the question scale (see Appendix III). Experts' ratings of the hypothetical cases provide information about differences in how each expert translates concrete aspects of cases into ordinal ratings.

Third, experts also coded an overlap period with contemporary V-Dem of about twenty years, typically 1900-1920, for either the polity that they coded prior to 1900 or that country's successor state (e.g. Italy for Modena). Overlap years thus include data from historical *and* contemporary experts. By comparing an historical expert's scores during this period to those of her contemporary colleagues, the measurement model algorithm is able to assess both her reliability and the degree to which she systematically codes different ordinal categories than her peers. Because those within-country peers are also bridged – through coding of additional cases and through vignettes – to the rest of the contemporary coders, this overlap period helps to anchor historical coders to the contemporary period.

Though these methods could have in principle been sufficient to ensure cross-temporal and cross-national comparability, preliminary analyses indicated that there were too few overlapping observations for the original measurement model to adequately adjust for

[3] See Pemstein et. al. (2017) for a full technical description of V-Dem's latent modeling framework. In particular, section 2.7 provides an in-depth description of issues related to Historical V-Dem.

differences in expert scale perception. Specifically, in initial runs of the measurement model we discovered substantial disjunctures between the pre- and post-1900 periods. An inspection of raw coder scores indicated that this disjuncture is due to historical experts systematically diverging in their codings from their contemporary V-Dem counterparts. Intuitively, experts might adjust their scales to the range of institutional quality that they observe across the observations that they consider – with historical experts applying more favorable judgments to the quality of democracy in the 19th century, presumably because they are implicitly "historicizing" their subject matter. To compensate for this effect, we have adjusted the measurement model to include country-specific offsets into the prior values for the years that historical experts coded.[4]

6. Patterns of democratization in the early part of modern history

Historical V-Dem includes data for 91 countries; however, coverage varies across questions. We focus here on 68 polities that have data for all indicators entering the V-Dem Polyarchy ("Electoral democracy") index (Teorell et al. 2018). (In subsequent editions of the dataset we hope to rectify missingness so that close to the full complement of 91 countries can be included.) We start by considering the average trend in Polyarchy from 1789 to 1945. This period includes Huntington's (1991) "first wave of democratization" but also the "first reverse wave" in the inter-war years.

[4] Specifically, we model our prior belief about the value of a historical observation as the sum of the ordinal value provided by the expert for that observation and the average difference between her yearly codings during the overlap period (typically 1900-1920) and the average yearly codings of the contemporary experts, restricted such that the value does not go beyond the range of the ordinal scale. This sum is normalized across all country-years (contemporary and historical) to calculate the prior.

Figure 1. The First Wave, 1789-1944

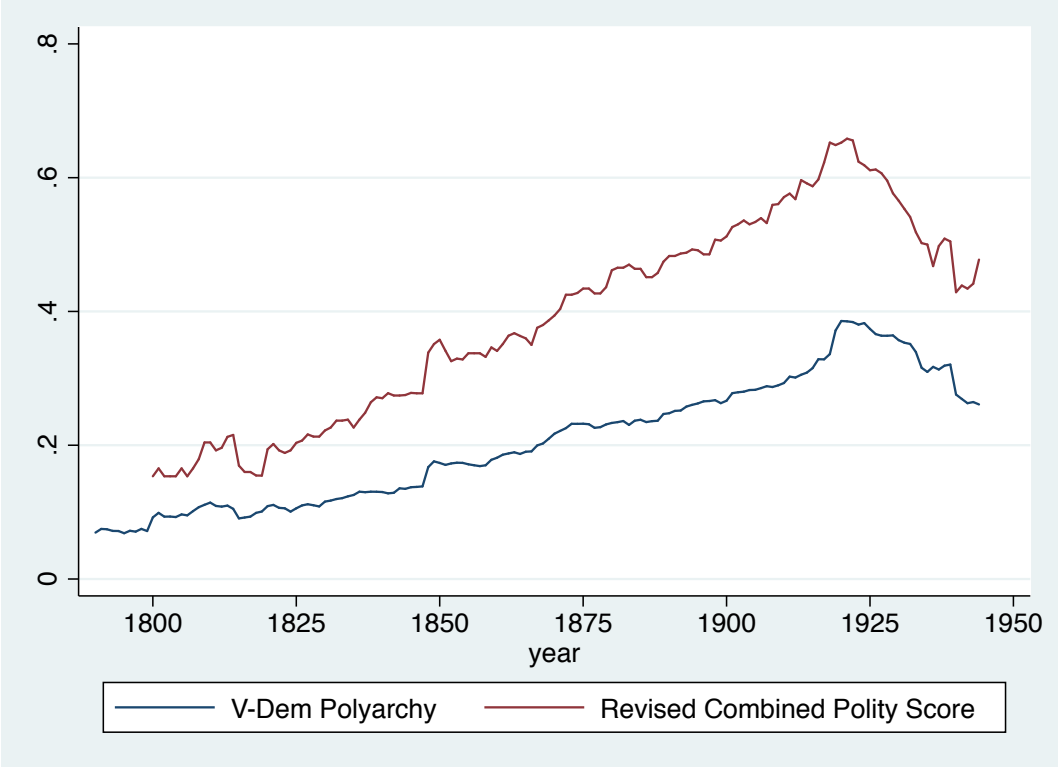
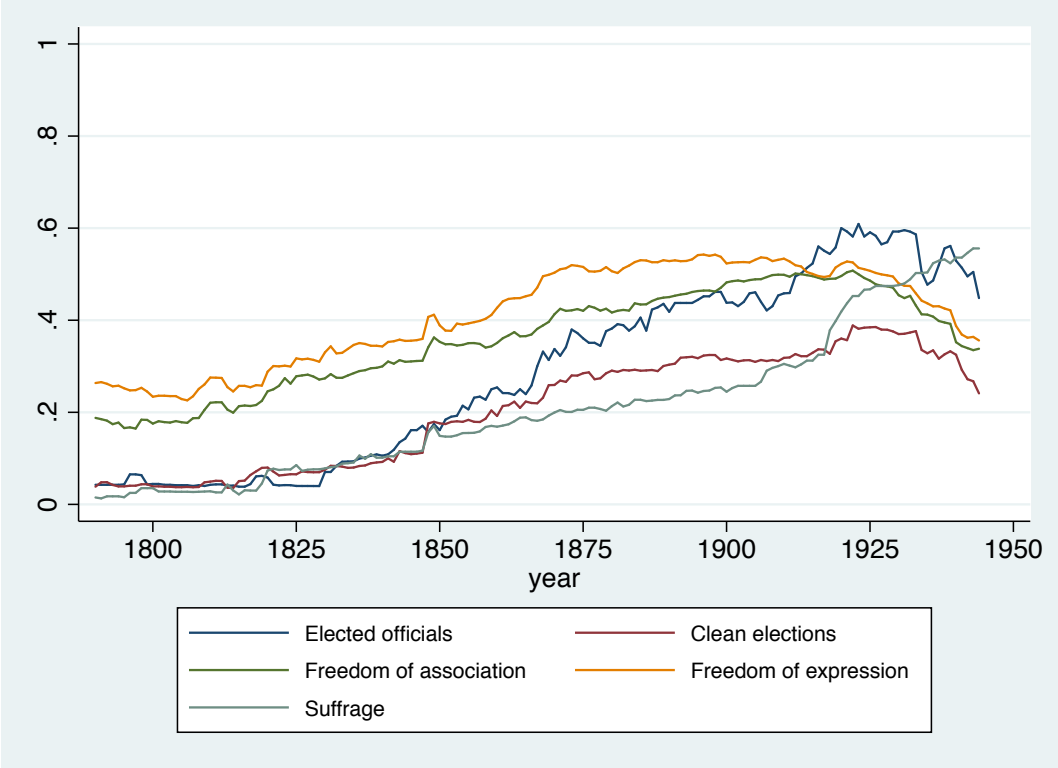


Figure 1 reveals that the upward trend in Polyarchy from 1789 to WWI, i.e., the long 19th century, is gradual. There is a brief dent in the steady upward slope around the revolutionary year of 1848, but overall, as argued by Weyland (2014), several of the revolutionary events were largely contained within the respective countries and did not ripple across either Europe or other continents. Only with the truly international event of WWI do we see a large spike in Polyarchy. Overall, the shape of the trend is in line with Congleton's (2011) description of the 19th century as an era of multiple, minor, liberal reforms. The first wave was not only a long wave; it was also a slow wave.

This aggregate pattern is fairly similar according to the Polity2 index, as Figure 1 shows. However, these data sources are quite different in other respects. First, since the Polyarchy index combines information from a number of underlying indicators, we are able to drill down to view the evolution of its constituent parts. (Polity2 also offers opportunities for disaggregation. However, there are just a few components of this index, and these components are themselves highly aggregated and therefore not as informative.) In Figure 2, we show the trajectories of all five of Dahl's (1998) institutional guarantees (the components of Polyarchy): elected officials,

free and fair elections, freedom of association, freedom of expression, and suffrage.[5] With few exceptions, they trend upwards throughout the long 19th century, but they also reveal some hitherto unexplored patterns.

Figure 2. Polyarchy Components, 1789-1944



To begin with, freedom of expression actually declined after the French Revolution and during the Napoleonic wars. Moreover, the "freedom" components are the one in highest observance, whereas the more strictly political ones, concerning the electability of executives and legislatures, the fairness (or even holding) of elections, and suffrage extension, display much lower average scores throughout most of the 19th century. This is markedly different from the 20th century after WWII, where suffrage and elected officials are the clearly highest-ranking components of Polyarchy. Finally, Figure 2 shows that suffrage is the aspect of Polyarchy that had the lowest average scores, at least from 1850 to WWI, which might explain why universal suffrage has often been treated as the "crowning event" of democratization during the first wave.

[5] These components, in turn, draw on, respectively, 15, 8, 6, 8, and 1 Historical V-Dem indicators. Freedom of expression is the only part of the index construction that differs (though only slightly) from contemporary V-Dem: One media indicator (v2mecenefm) was not included in the historical survey. The Bayesian Factor Analysis index on freedom of expression is therefore run without this indicator.

The second conspicuous difference between Polity2 and Polyarchy is that Polity2 offers a more lenient standard of democracy, signaled by Polyarchy's consistently lower values in Figure 1. To show this difference more precisely, Figure 3 plots the Polyarchy scores against Polity2 scores (re-scaled 0-1), averaged across the 1789–1945 period, for the 56 countries covered by both measures. The diagonal line marks no average differences (which might mask yearly differences that cancel each other out), so countries above the line have larger Polyarchy scores, and countries below have larger Polity2 scores. Consistent with the over-time trends, few countries, on average, have higher Polity2- than Polyarchy scores. We have highlighted the three top countries in the former group (Denmark, Bavaria and Bulgaria), and the ten countries falling furthest below the line in the latter.

Figure 3. Comparing V-Dem Polyarchy to Polity2, 1800-1944

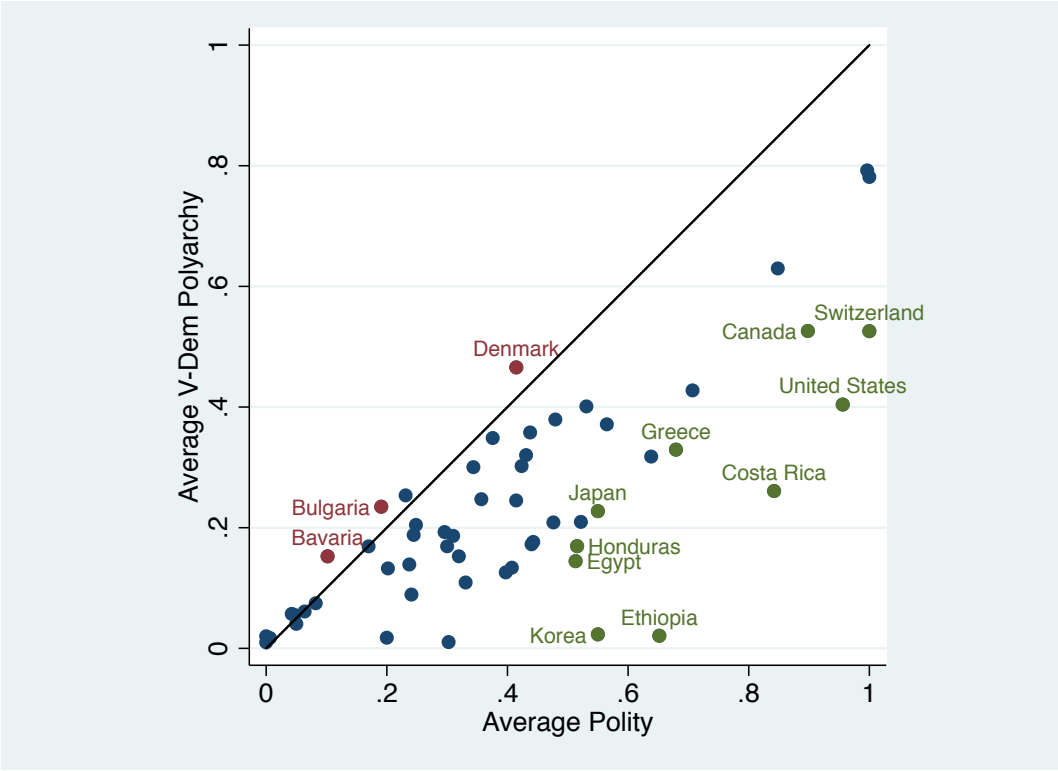
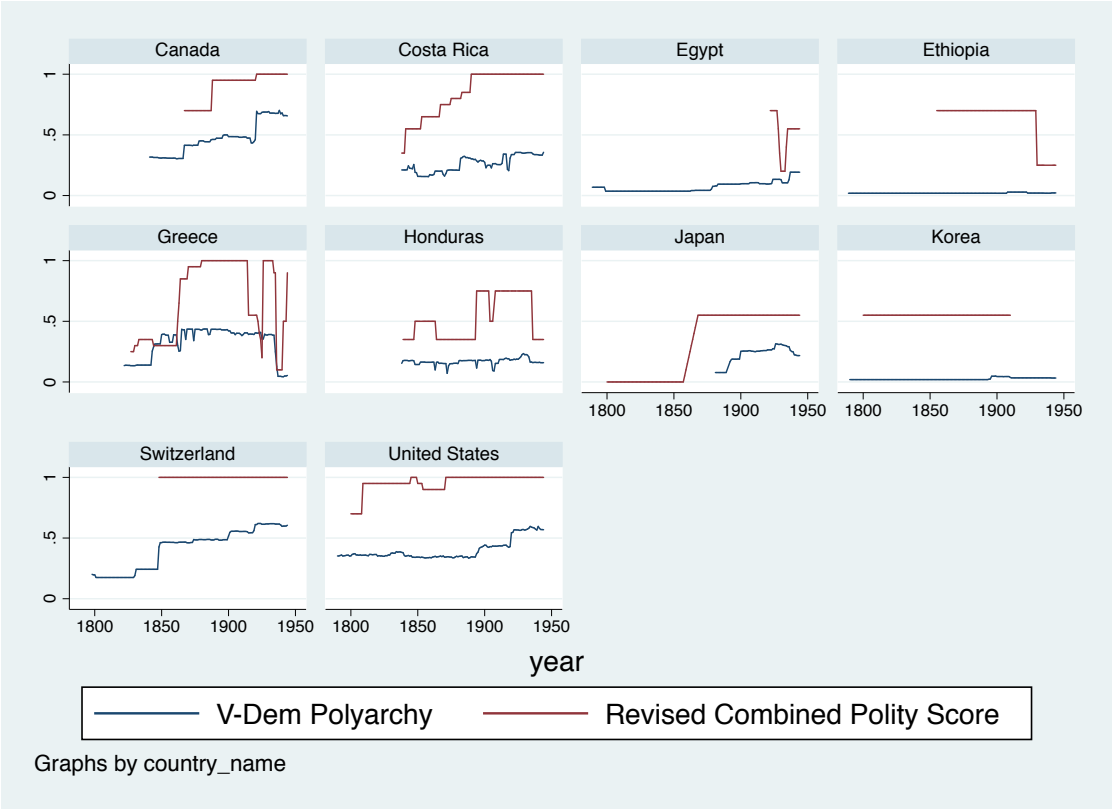


Figure 4 plots the latter "top ten" countries over time. The differences are quite substantial. Polity2 scores the US at its maximum already in 1871, and does not pick up any subsequent change in democracy, despite, for example, *de jure* and *de facto* restrictions on voting rights for large parts of the population, including women and African-Americans (especially) in the South. Similarly, Polity2 ignores suffrage restrictions in Canada, Costa Rica, Greece and

Switzerland. Polity2 also has a surprisingly high appraisal of democracy in Ethiopia and Korea, despite these polities never holding elections and, with the partial exception of the Great Korean Empire from 1897 until Japanese annexation in 1910, severely restricted freedoms of expression and association.

Figure 4. Ten Largest Country Discrepancies in Polyarchy vs. Polity2, 1789-1944



7. The role of war in democratization across modern history

Finally, we employ Historical V-Dem data to investigate a potential determinant of democracy: international conflict. Key criticisms of the democratic peace literature have argued that the latter is "putting the cart before the horse" (Thompson, 1996): war affects regime type and not (just) vice versa. Gibler (2012) provides a recent, comprehensive empirical treatment, arguing that (territorial) war breeds autocracy. According to Gibler, wars create larger armies, which, in turn, can be used for internal repression. Wars also induce political centralization, which can lead to

dictatorship. Further, populations facing external threats supposedly turn more willing to defer to ascendant autocrats.

Yet, other scholars have argued that war can favor subsequent democratization. Summarizing the record in Europe after the two world wars, Therborn (1977, 19) proposed that "democracy is largely a martial accomplishment". Regimes ruling countries that lose in inter-state wars are sometimes toppled through external intervention (Bueno de Mesquita and Downs 2006; Pickering and Peceny 2006; Grimm 2008). War, and especially loss in war, can also alter the relative power of key domestic groups, sometimes undermining entrenched autocrats and strengthening domestic constituencies favoring regime change.

Empirical studies suggest that the evidence is mixed. There are some indications that war hinders democratization (see, e.g., Reiter, 2001; Gibler, 2012; Mitchell, Gates and Hegre, 1999), whereas other studies yield null findings (e.g., Oneal and Russett, 2000; Mousseau and Shi, 1999; Mansfield and Snyder 2010). Evidently, a careful assessment of how war affects regime type requires data with long time series that also capture detailed institutional features. This is especially important given (a) the paucity of inter-state wars; (b) the possibility of temporal heterogeneity in the relationship, given changes to the international system and power structure (see Boix 2011); and (c) the possibility that war might affect some aspects of democracy, but not others. For example, suffrage expansions are often viewed as concessions in return for mass-conscription (for men) and female labor force participation during times of warfare (e.g., Ticchi and Vindigni, 2008).

We employ Correlates of War (COW) data on inter-state war (Sarkees et al. 2010) for 1817-2007. To capture the impact of war, we register the number of years a country has experienced a war between t-1 and t-5.[6] Since an ongoing war may have different implications for current regime type than past war exposure, we control for war ongoing at t. For democracy, we focus on the discussed Polyarchy measure, but contrast results with Polity2 to investigate whether estimates hinge on the measurement of democracy. Our baseline specification is intentionally sparse, controlling only for GDP per capita, population, and year-fixed effects. We mostly find similar results in models that add country-fixed effects (see Appendix IV).

[6] We find very similar results when using logged number of years.

Table 1. Regressing interstate war on Polity2 and Polyarchy

	(1) Polity2 1817-2006 LDV	(2) Polyarchy 1817-2006 LDV	(3) Polity2 1817-1918 LDV	(4) Polyarchy 1817-1918 LDV	(5) Polyarchy 1918-2006 LDV
War past 5 years	-0.098*	0.004**	0.028	0.001	0.005**
	(-1.87)	(2.47)	(0.40)	(1.21)	(2.15)
Ongoing war	0.112	-0.001	0.145	0.000	-0.002
	(1.48)	(-0.28)	(1.26)	(0.16)	(-0.64)
Ln(GDPpc)	0.097***	0.002***	0.065***	0.001***	0.002**
	(5.49)	(3.03)	(3.55)	(3.19)	(2.52)
Ln(population)	0.027***	0.000	0.020**	0.000	-0.000
	(2.61)	(0.39)	(2.20)	(1.05)	(-0.05)
Year-FE	Y	Y	Y	Y	Y
N	14116	16202	4786	5094	11030
R^2	0.955	0.978	0.977	0.983	0.975

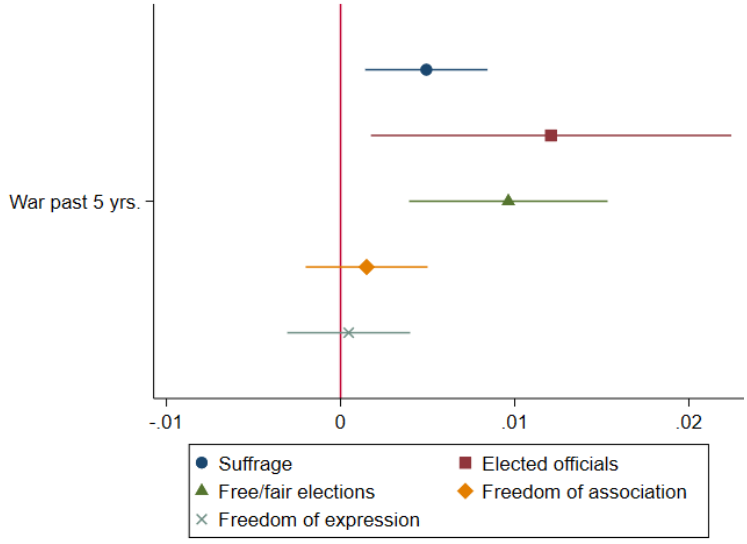
Notes: ***p<0.01; **p<0.05; *p<0.1. All models are OLS with errors clustered by country. T-values reported in parentheses.

We start out, in Column 1, Table 1, by considering Polity2 for all observations with available data (1817–2006). The war experience (past 5 years) coefficient is weakly significant and negative, indicating that downturns in Polity2 often follow wars (the negative coefficient is further weakened when adding country-fixed effects). Column 2 reports a model (full sample) using Polyarchy. In stark contrast to the Polity2 result, Polyarchy is positive and more precisely estimated, suggesting that interstate war correlates with subsequent democratization. When restricting the sample to the long 19th century, however, ongoing war is statistically insignificant both when using Polity2 and Polyarchy. Columns 3-4 re-estimate Columns 1-2, but restricted to 1817-1918. There is no clear evidence for a relationship in this period, independent of democracy measure used. In contrast, when we only employ post-WWI data and use Polyarchy (Column 5), we find a clear, positive relationship.

To probe deeper into what might be driving the relationship between war and democracy in the full sample, we disaggregate Polyarchy into its subcomponents, and use them as dependent variables in our benchmark specification. These results (see Figure 5) show that freedom of association and freedom of expression are not significantly related to past war exposure. In contrast, the suffrage-, elected officials-, and free and fair elections indices are all positively correlated with past war exposure. Thus, the positive relationship between war and democracy seems primarily to work through the electoral channel. This is consistent with the notion that participation in free and fair elections (suffrage) is widened by experiences with interstate conflict,

perhaps due to dynamics relating to mass mobilization and subsequent bargaining with elites (see Ticchi and Vindigni, 2008).

Figure 5. Coefficient estimates with 95% confidence intervals for "War past 5 years" from benchmark (full time series) run on Polyarchy's five subcomponents



In sum, our results based on Polyarchy contrast with arguments on the democracy-hampering effect of war, such as those proposed by Reiter (2001) and Gibler (2012). When employing our data, we find that war exposure correlates positively with democracy, and particularly when focusing on electoral components such as suffrage extension and cleanness of elections. The analysis also leads to two other key observations: First, when checking for heterogeneous effects across time, there are indications that the relationship between war and democratization has evolved throughout the course of modern history. Second, the choice of democracy measure matters for estimates of the relationship between war and democratization. For instance, utilizing the Polyarchy measure generates a clear positive association between prior war exposure and democratization in the post WWI period, whereas this relationship is different when using Polity2. This seems, at least partly, to stem from differences in components included, as Polity2, for example, basically ignores suffrage, a vital component in Polyarchy (and most other common notions of democracy).

8. Conclusion

We have laid out the general features and content of Historical V-Dem, and described how it addresses issues of reliability, validity, inter-temporal- and cross-country comparability. When combined with contemporary V-Dem, the about 260 indicators contained in Historical V-Dem open up new possibilities for drawing on historical information from the entirety of "modern history" to inform the study of democracy and related phenomena. Here, we have shown how the detailed nature of V-Dem data can be used to identify trends in democracy and to explore the relationship between interstate war and democratization. Subsequent research can use these data to delve more closely into potential determinants and effects of different varieties of democracy, as well as effects of more specific political institutions.

References

Anonymous. 2018. "Ports and Democracy". Working paper.

Ansell, Ben & David Samuels. 2015. *Inequality and Democratization: An Elite-Competition Approach.* New York: Cambridge University Press.

Boix, Carles. 2011. Democracy, Development, and the International System. *American Political Science Review* 105(4): 809-828.

Boix, Carles, Michael Miller and Sebastian Rosato. 2013. A Complete Dataset of Political Regimes, 1800-2007, *Comparative Political* Studies 46(12): 1523-1554

Bueno de Mesquita, Bruce, Alastair Smith, Randolph M. Siverson & James D. Morrow. 2003. *The Logic of Political Survival.* Cambridge, MA.: The MIT Press.

Bueno de Mesquita, Bruce, and George Downs. 2006. "Intervention and Democracy." *International Organization* 60(3): 627-49

Congleton, Roger. 2011. *Perfecting Parliament. Constitutional Reform, Liberalism, and the Rise of Western Democracy*. Cambridge: Cambridge University Press.

Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning & Jan Teorell. 2017a. "V-Dem Comparisons and Contrasts with Other Measurement Projects." V-Dem Working Paper 2017:45.

Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Pamela Paxton, Daniel Pemstein, Laura Saxer, Brigitte Seim, Rachel Sigman & Jeffrey Staton. 2017b. "V-Dem Codebook v7." Varieties of Democracy (V-Dem) Project.

Coppedge, Michael, John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Joshua Krusell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa Olin, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Constanza Sanhueza Petrarca, Johannes von Römer, Laura Saxer, Brigitte Seim, Rachel Sigman, Jeffrey Staton, Natalia Stepanova & Steven Wilson. 2017c. "V-Dem [Country-Year/Country-Date] Dataset v7.1" Varieties of Democracy (V-Dem) Project.

Dahl, Robert. 1971. *Polyarchy*. New Haven & London: Yale University Press.

Dahl, Robert. 1998. *On Democracy*. New Haven & London: Yale University Press.

Doorenspleet, Renske. 2005. *Democratic Transitions: Exploring the Structural Sources of the Fourth Wave*. Boulder and London: Lynne Rienner

Fukuyama, Francis. 2014. *Political order and political decay : from the industrial revolution to the globalization of democracy*: New York : Farrar, Straus and Giroux, 2014

Gibler, Douglas M. 2012. *The territorial peace: Borders, state development, and international conflict*. Cambridge University Press.

Gleditsch, Kristian S. & Michael D. Ward. 1999. "Interstate System Membership: A Revised List of the Independent States since 1816." *International Interactions* 25(4): 393-413.

Grimm, Sonja. 2009. "External Democratization after War: Success and Failure." *Democratization* 15(3): 525-49

Hobsbawm, Eric. 1962. The Age of Revolution 1789–1848. London: Weidenfeld and Nicolson.

Hobsbawm, Eric. 1975. The Age of Capital 1848–1875. London: Weidenfeld and Nicolson.

Hobsbawm, Eric. 1987. The Age of Empire 1875–1914. London: Weidenfeld and Nicolson.

Huntington, Samuel. 1991. *The Third Wave: Democratization in the Late Twentieth Century*. Norman and London: University of Oklahoma Press

King, Gary & Jonathan Wand. 2007. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes." *Political Analysis* 15(1):46-66.

Knutsen, Carl Henrik, Jørgen Møller & Svend-Erik Skaaning. 2016. "Going historical: Measuring democraticness before the age of mass democracy." *International Political Science Review, 37*(5): 679-689.

Marshall, Monty G., Keith Jaggers & Tedd R. Gurr. 2015. Polity IV project: Political regime characteristics and transitions, 1800-2013. Dataset users' manual, version 2014. Fort Collins: Colorado State University.

Mansfield, Edward and James Snyder. 2010. "Does War Influence Democratization?", in Kier, E. & R Krebs, eds. *In War's Wake. International Conflict and the Fate of Liberal Democracy*. Cambridge: Cambridge University Press.

Mitchell, Sara McLaughlin, Scott Gates, and Håvard Hegre. 1999. "Evolution in democracy-war dynamics." *Journal of Conflict Resolution* 43(6): 771-792.

Moore, Barrington. 1966. *Social origins of dictatorship and democracy : lord and peasant in the making of the modern world*. Boston: Beacon.

Mousseau, Michael, and Yuhang Shi. 1999. "A Test for Reverse Causality in the Democratic Peace Relationship." Journal of Peace Research 36(6): 639-663.

Oneal, John, and Bruce Russett. 2000. "Why 'An Identified Systemic Model of the Democracy-Peace Nexus' Does Not Persuade." Defence and Peace Economics 11(2): 197-214.

Pemstein, Dan, Kyle L. Marquardt, Eitan Tzelgov, Yi-ting Wang, Joshua Krussel & Farhad Miri. 2018. "The V-Dem Measurement Model: Latent Variable Analysis for Cross-National and Cross-Temporal Expert-Coded Data." V-Dem Working Paper 21 [Updated version, April 2018].

Pickering, Jeffrey and Mark Peceny. 2006. "Forging Democracy at Gunpoint." *International Studies Quarterly* 50: 539-59

Reiter, Dan. 2001. "Does peace nurture democracy?" *Journal of Politics* 63(3): 935-948.

Teorell, Jan, 2018, Michael Coppedge, Svend-Erik Skaaning & Stafan I. Lindberg. "Measuring Polyarchy across the Globe, 1900-2016." *Studies in Comparative International Development*. Forthcoming.

Therborn, Göran. 1977. "The Rule of Capital and the Rise of Democracy." *New Left Review* I/103: 3-41

Thompson, William R. 1996. "Democracy and peace: putting the cart before the horse?" *International Organization* 50(1): 141-174.

Weyland, Kurt. 2014. *Making Waves: Democratic Contention in Europe and Latin America since the Revolutions of 1848*. Cambridge: Cambridge University Press.

Ziblatt, Daniel. 2017. *Conservative Parties and the Birth of Democracy*. Cambridge: Cambridge University Press

Appendix I: Sample

The authoritative list by Gleditsch & Ward (1999) served as the point of departure for delimiting the current sample.[i] More specifically, the main criteria for including polities is that they are a) fairly sizeable (>250000 inhabitants); b) sovereign during an extended time period between 1789 and 1900, either in a formal-juridical or *de facto* sense; and, that they c) match present-day state units.

Historical V-Dem includes another 20 polities not covered by Gleditsch & Ward (1999). After a careful mapping of potential polities to be included these are polities that corresponds to a contemporary state and that, despite the lack of international sovereignty, wielded sufficient *de facto* domestic sovereignty (over an extended period prior to 1900) for being considered as at least semi-sovereign. This means that we included Australia, Finland, Hungary, Kuwait, Norway, New Zealand, Poland and Yemen as well as two "precursor" polities of contemporary states where borders do not quite fit the latter (Nejd/Saudi Arabia, Bukhara/Uzbekistan). In addition, we included a selection of colonies/protectorates, including the two most populous, namely British India and the Dutch West Indies (Indonesia), plus three smaller, namely Cuba, Singapore and Zanzibar. Finally, due to a particular extra grant, we included five additional pre-unification German principalities below the 250,000 population threshold (Brunswick, Hamburg, Oldenburg, Nassau and Saxe-Weimar).

Table A.I lists the time series for each polity included in Historical V-Dem.

[i] Gleditsch and Ward (1999) identify 75 independent polities pre-1900. Two polities from this list, which do not neatly map onto borders of a contemporary state entity (Orange Free State and Transvaal), plus one short-lived polity (Algeria prior to the French conquest), are currently not included in Historical V-Dem. Since we also treat Colombia and Gran Colombia, as well as Guatemala and the United Provinces of Central America, as one case each, but Piedmont-Sardinia as separate from pre-unification Italy, we end up with 91 polities after adding 20 extra polities (75-3-2+1+20=91).

Table A.I: Countries and years covered in Historical V-Dem

Polity	Years covered	Polity	Years covered
Afghanistan	1789-1920	Mecklenburg Schwerin	1789-1867
Argentina	1789-1920	Mexico	1789-1920
Australia	1789-1920	Modena	1789-1797; 1814-1859
Austria	1789-1938	Montenegro	1789-1918
Baden	1789-1871	Morocco	1789-1920
Bavaria	1789-1871	Nassau	1806-1866; 1900-1920
Belgium	1789-1795; 1830-1920	Nepal	1789-1920
Bolivia	1825-1920	Netherlands	1789-1810; 1813-1920
Brazil	1789-1920	New Zealand	1841-1920
Brunswick	1789-1807; 1813-1867	Nicaragua	1838-1920
Bulgaria	1878-1920	Norway	1789-1920
Burma/Myanmar	1789-1920	Oldenburg	1789-1810; 1813-1867
Canada	1841-1920	Oman	1789-1920
Chile	1789-1920	Orange Free State	1854-1910
China	1789-1920	Papal States	1789-1809; 1814-1870
Colombia	1789-1920	Paraguay	1811-1920
Costa Rica	1838-1920	Parma	1789-1802; 1814-1859
Cuba	1789-1920	Peru	1789-1920
Denmark	1789-1920	Piedmont-Sardinia	1789-1861
Dominican Republic	1789-1822; 1844-1920	Poland	1789-1795; 1807-1867; 1918-1938
Ecuador	1830-1920	Portugal	1789-1920
Egypt	1789-1920	Romania	1789-1920
El Salvador	1838-1920	Russia	1789-1920
Ethiopia	1789-1920	Saudi Arabia	1789-1818; 1822-1952
Finland	1809-1920	Saxe-Weimar-Eisenach	1809-1867
France	1789-1920	Saxony	1789-1867
Germany	1789-1920	Serbia	1804-1813; 1815-1918
Greece	1822-1920	Singapore	1867-1920
Guatemala	1789-1822; 1823-1920	Spain	1789-1920
Haiti	1789-1920	Sweden	1789-1920
Hamburg	1789-1810; 1813-1867	Switzerland	1798-1920
Hanover	1789-1810; 1813-1866	Thailand	1789-1920
Hesse-Darmstadt	1789-1871	Transvaal	1852-1910
Hesse-Kassel	1789-1866	Tunisia	1789-1920
Honduras	1838-1920	Turkey	1789-1920
Hungary	1789-1938	Tuscany	1789-1807; 1814-1861
India	1789-1920	Two Sicilies	1789-1860
Indonesia	1800-1920	United Kingdom	1789-1920
Iran	1789-1920	United States of America	1789-1920
Italy	1861-1920	Uruguay	1825-1920
Japan	1789-1920	Uzbekistan	1789-1920; 1990-2010
Korea, South	1789-1920	Venezuela	1789-1819; 1830-1920
Kuwait	1789-1920	Vietnam, Republic of	1802-1922
Liberia	1821-1920	Wurtemberg	1789-1871
Libya	1789-1834; 1911-1933; 1952-1972	Yemen	1789-1850; 1918-1938
Luxembourg	1815-1920	Zanzibar	1856-1920
Madagascar	1817-1920		

Note: This is the maximum coverage in the dataset, pertaining to some of the included (A) variables. Coverage varies between variables.

Appendix II: Variables included

Table A.II: "A variables" included in Historical V-Dem.

Variable tag	Variable name
v3canagelc	Minimum candidate age parliament/lower chamber
v3canageuc	Minimum candidate age upper chamber
v3clslavery	Slavery
v3elage	Minimum voting age parliament/lower chamber
v3elagepr	Minimum voting age presidency
v3elageuc	Minimum voting age upper chamber
v3eldirelc	Direct parliamentary/lower chamber elections
v3eldirepr	Direct presidential elections
v3eldireuc	Direct upper chamber elections
v3elfemrst	Female suffrage restricted
v3elloelsy	Lower chamber electoral system, fine-grained
v3elloseat	Lower chamber election seats
v3ellostlg	Lower chamber election seats won by largest party
v3ellostsl	Lower chamber election seat share won by largest party
v3ellostsm	Lower chamber election seats won by second largest party
v3ellostss	Lower chamber election seat share won by second largest party
v3ellosttm	Lower chamber election seats won by third largest party
v3ellostts	Lower chamber election seat share won by third largest party
v3ellovtlg	Lower chamber election vote share of largest vote-getter
v3ellovtsm	Lower chamber election vote share of second-largest vote-getter
v3ellovttm	Lower chamber election vote share of third-largest vote-getter
v3elncbmaj	Minority or majority government
v3elncbpr	Effective number of cabinet parties
v3elparlel	Lower chamber electoral system
v3elrstrlc	Candidate exclusions (*de jure*) parliament/lower chamber
v3elrstrpr	Candidate exclusions (*de jure*) presidential elections
v3elrstrup	Candidate exclusions (*de jure*) upper chamber
v3elsec	(*De jure*) ballot secrecy
v3elsuffrage	Percentage of population with suffrage
v3eltrnout	Election turnout
v3eltvrig	Lower chamber election turnover
v3eltvriguc	Upper chamber election turnover
v3elupseat	Upper chamber election seats
v3elupstsl	Upper chamber election seats won by largest party
v3elupstsm	Upper chamber election seats won by second largest party
v3elupvtlg	Upper chamber election vote share of largest vote-getter
v3elupvtsm	Upper chamber election vote share of second-largest vote-getter
v3elvotlrg	Presidential election vote share of largest vote-getter
v3elvotsml	Presidential election vote share of second-largest vote-getter
v3elvstrlc	Suffrage exclusions (*de jure*) parliament/lower chamber
v3elvstrpr	Suffrage exclusions (*de jure*) presidential elections
v3elvstruc	Suffrage exclusions (*de jure*) upper chamber

v3elwomcab	Election women in the cabinet
v3exagehog	HOG age
v3exagehos	HOS age
v3exaphogp	HOG selection by legislature in practice
v3exaphos	HOS selection by legislature in practice
v3exapup	Chief executive appointment by upper chamber
v3exapupap	Chief executive appointment by upper chamber implicit approval
v3exdeathog	HOG year of death
v3exdeathos	HOS year of death
v3exothhgl	HOG other appointing body in practice
v3exothhs	HOS other appointing body in practice
v3expathhg	HOG appointment in practice
v3expathhs	HOS appointment in practice
v3lgamend	Legislature amends constitution
v3lgamnsty	Legislature amnesties
v3lgcamoth	Legislature other than uni- or bicameral
v3lgelecup	Upper chamber elected
v3lgello	Lower chamber elected
v3lginello	Lower chamber indirectly elected
v3lginelup	Upper chamber indirectly elected
v3lgintblo	Lower chamber introduces bills
v3lgintbup	Upper chamber introduces bills
v3lglegllo	Lower chamber legislates by law
v3lgleglup	Upper chamber legislates by law
v3lgqumin	Lower chamber quota for social groups
v3lgtreaty	Legislature approval of treaties by law
v3lgwarlaw	Legislature declares war by law
v3lpname	Name of largest party
v3pechilabl	Child labor laws
v3peminwage	Minimum wage
v3peminwagerestr	Minimum wage provision
v3psagefirst	Party age largest
v3psagepm	Party age executive
v3psagesecond	Party age second largest
v3psagethird	Party age third largest
v3regendtypems	Regime end type
v3regint	Regime interregnum
v3serfdeju	Serfdom
v3slpname	Name of second largest party
v3stcensus	Census
v3stcitlaw	Citizenship laws
v3stflag	Flag
v3stnatant	National anthem
v3stnatbank	National bank
v3ststatag	Statistical agency
v3ststybcov	Statistical yearbook covered
v3ststybpub	Statistical yearbook published
v3tlpname	Name of third largest party

v3ttlvote	Total votes

Note: A variables are coded by research assistants. See V-Dem v.8 codebook for specifics on questions, clarifications, and answer categories. HOG=Head of Government. HOS=Head of State.

Table A.III: "A* variables" included in Historical V-Dem (see V-Dem v.8 codebook for specifics).

Variable tag	Variable name
v3ellocelc	Local government elected
v3ellocgov	Local government exists
v3ellocnam	Local government name
v3elreggov	Regional government exists
v3elregnam	Regional government name
v3elsrgel	Regional government elected
v3eltype	Election type
v3exhoshog	HOS = HOG
v3exnamhog	HOG name
v3exnamhos	HOS name
v3extithog	Title of HOG
v3extithos	HOS title
v3juhcname	High court name
v3juhcourt	High court existence
v3lgbicam	Legislature bicameral
v3lgnamelo	Lower chamber legislature name
v3lgnameup	Upper chamber name
v3regendtype	Regime end type
v3reginfo	Regime information

Note: A* variables are pre-coded by research assistants and are entered as relevant information in the expert surveys. These variables may be adjusted based on expert feedback. See V-Dem v.8 codebook for specifics on questions, clarifications, and answer categories. HOG=Head of Government. HOS=Head of State.

Table A.IV: "C variables" included in Historical V-Dem

Variable tag	Variable name
v3clacfree	Freedom of academic and cultural expression
v3clacjstm	Access to justice for men
v3clacjstw	Access to justice for women
v3clacjust	Social class equality in respect for civil liberty
v3cldiscm	Freedom of discussion for men
v3cldiscw	Freedom of discussion for women
v3cldmovem	Freedom of domestic movement for men
v3cldmovew	Freedom of domestic movement for women
v3clfmove	Freedom of foreign movement
v3clkill	Freedom from political killings
v3cllabrig	Labor rights
v3clprptym	Property rights for men
v3clprptyw	Property rights for women
v3clrelig	Freedom of religion
v3clrgunev	Regional unevenness in respect for civil liberties
v3clrspct	Rigorous and impartial public administration
v3clslavef	Freedom from forced labor for women
v3clslavem	Freedom from forced labor for men
v3clsocgrp	Social group equality in respect for civil liberties
v3clstown	State ownership of economy
v3cltort	Freedom from torture
v3cltrnslw	Transparent laws with predictable enforcement
v3csanmvch	CSO anti-system movement character
v3csantimv	CSO anti-system movements
v3cscnsult	CSO consultation
v3cseeorgs	CSO entry and exit
v3csgender	CSO women's participation
v3csprtcpt	CSO participatory environment
v3csreprss	CSO repression
v3csrlgcon	Religious organization consultation
v3csrlgrep	Religious organization repression
v3csstruc	CSO structure
v3dlconslt	Range of consultation
v3dlencmps	Particularistic or public goods
v3dlengage	Engaged society
v3elaccept	Election losers accept results
v3elasmoff	Election assume office
v3elbalpap	Voting, voice or ballot
v3elbalstat	Ballot printing
v3elboycot	Election boycotts
v3elcomvot	Compulsory voting
v3eldonate	Disclosure of campaign donations
v3elecsedf	Secret ballot, *de facto*
v3elembaut	EMB autonomy
v3elembcap	EMB capacity
v3elffelr	Subnational elections free and fair

v3elfrfair	Election free and fair
v3elintim	Election government intimidation
v3elirreg	Election other voting irregularities
v3ellocpwr	Local offices relative power
v3elmalalc	Malapportionment legislature/lower chamber
v3elmalauc	Malapportionment upper chamber
v3elmalsuf	Election male suffrage in practice
v3elmulpar	Elections multiparty
v3elpeace	Election other electoral violence
v3elpubfin	Public campaign finance
v3elreapplc	Reapportionment legislature/lower chamber
v3elreappuc	Reapportionment upper chamber
v3elrgpwr	Regional offices relative power
v3elrgstry	Election voter registry
v3elsnlsff	Subnational election unevenness
v3elvotbuy	Election vote buying
v3equavolc	Equal vote legislature/lower chamber
v3equavouc	Equal vote upper chamber
v3exbribe	Executive bribery and corrupt exchanges
v3excrptps	Public sector corrupt exchanges
v3exctlhg	HOG control over
v3exctlhog	HOG other body controls
v3exctlhos	HOS other body controls
v3exctlhs	HOS control over
v3exdfcbhs	HOS appoints cabinet in practice
v3exdfdmhs	HOS dismisses ministers in practice
v3exdfdshg	HOG dismisses ministers in practice
v3exdfdshs	HOS dissolution in practice
v3exdfpphg	HOG proposes legislation in practice
v3exdfpphs	HOS proposes legislation in practice
v3exdfvthg	HOG veto power in practice
v3exdfvths	HOS veto power in practice
v3exdjcbhg	HOG appoints cabinet in practice
v3exdjdshg	HOG dissolution in practice
v3exembez	Executive embezzlement and theft
v3exremhog	HOG removal by legislature in practice
v3exremhsp	HOS removal by legislature in practice
v3exrescon	Executive respects constitution
v3exrmhgnp	HOG removal by other in practice
v3exrmhgop	HOG other body remove HOG in practice
v3exrmhsnl	HOS other body removes in practice
v3exrmhsol	HOS removal by other in practice
v3exthftps	Public sector theft
v3juaccnt	Judicial accountability
v3jucomp	Compliance with judiciary
v3jucorrdc	Judicial corruption decision
v3juhccomp	Compliance with high court
v3juhcind	High court independence

v3juncind	Lower court independence
v3jureview	Judicial review
v3lgbudglo	Lower chamber budget
v3lgbudgup	Upper chamber budget
v3lgcomslo	Lower chamber committees
v3lgcrrpt	Legislature corrupt activities
v3lgdomchm	Legislature dominant chamber
v3lgfunds	Legislature controls resources
v3lginses	Lower chamber in session
v3lginsesup	Upper chamber in session
v3lginvstp	Legislature investigates in practice
v3lglegplo	Lower chamber legislates in practice
v3lglegpup	Upper chamber legislates in practice
v3lgoppart	Legislature opposition parties
v3lgotovst	Executive oversight
v3lgqstexp	Legislature questions officials in practice
v3lgsrvlo	Lower chamber members serve in government
v3meaccess	Media access
v3mebias	Media bias
v3mecrit	Print/broadcast media critical
v3meharjrn	Harassment of journalists
v3merange	Print/broadcast media perspectives
v3meslfcen	Media self-censorship
v3partyid	Party identification
v3pepwrgen	Power distributed by gender
v3pepwrses	Power distributed by socioeconomic position
v3pepwrsoc	Power distributed by social group
v3psbantar	Party ban target
v3psbars	Barriers to parties
v3pscnslnl	Candidate selection---national/local
v3pscohesv	Legislative party cohesion
v3pscomprg	Party competition across regions
v3psoppaut	Opposition parties autonomy
v3psorgs	Party organizations
v3psparban	Party ban
v3psplats	Distinct party platforms
v3psprbrch	Party branches
v3psprlnks	Party linkages
v3pssunpar	Subnational party control
v3psswitch	Party switching
v3regimpgroup	Regime most important support group
v3regsupgroups	Regime support groups
v3regsupgroupssize	Regime support groups size
v3regsuploc	Regime support location
v3stcritapparm	Criteria for appointment decisions in the armed forces
v3stcritrecadm	Criteria for appointment decisions in the state administration
v3stfisccap	State fiscal capacity
v3strenadm	Bureaucratic remuneration

v3strenarm	Remuneration in the Armed Forces
v3struinvadm	Rulers' involvement in the state administration
v3ststeecap	State steering capacity
v3svdomaut	Domestic autonomy
v3svinlaut	International autonomy
v3svstpop	State authority over population
v3svstterr	State authority over territory

Note: C variables are coded by country experts, and scores are subsequently adjusted in the V-Dem measurement model to achieve cross-country and inter-temporal comparability. See V-Dem v.8 codebook for specifics on questions, clarifications, and answer categories. HOG=Head of Government. HOS=Head of State. EMB=Election Monitoring Board. CSO=Civil Society Organization.

Appendix III: Anchoring vignettes

Historical V-Dem makes extensive use of anchoring vignettes to improve cross-country comparability (King et al., 2004; King and Wand, 2007; Bakker et al., 2014). Anchoring vignettes are descriptions of specific, but hypothetical – or at least unnamed – cases that provide the information required to answer a certain question. In the context of V-Dem, they are descriptions of hypothetical country-years that focus on describing the country's status specific to one V-Dem indicator. Coders' ratings of the hypothetical cases, once combined, provide information about differences in how they translate concrete aspects of cases into ordinal ratings. There are several reasons vignettes are a powerful and efficient tool for addressing differential item functioning (DIF) in V-Dem ratings.

- Raters have all the information about the case in question at their fingertips and coding vignettes, therefore, requires substantially less coder effort than evaluating actual cases. This makes vignettes substantially less costly for coders than bridge or lateral coding and raters can provide more vignette responses in a given set of time.
- Vignettes require no case knowledge, so everyone can do them, even experts who are not qualified to rate multiple countries.
- Vignettes provide perfect overlap, because every rater answers the same questions.
- Vignettes provide high threshold variability, because we control their content and strive to maximize that variability.
- Because we know that every rater considers the same information when they rate a vignette, we can assume potentially low random error in the rating process and treat all cross-coder variation as evidence of threshold differences.
- In asking all coders to code vignettes, we address potential selection bias introduced by having only those who opt in to bridge and lateral coding - i.e., those who are either most knowledgeable about the world or those who just think they are - provide data to adjust for cross-country comparability.

Ultimately, the vignettes exercise within Historical V-Dem included vignettes for the vast majority of its expert coded (C) questions, with multiple vignettes for each of these questions. This section of the appendix describes our approach to this exercise and explains the choices we made.

Where possible, the V-Dem vignettes approach followed the prevailing best practices according to the literature. However, the literature provides little guidance on a number of problems that are specific to V-Dem. In particular, the raw size of the V-Dem survey, and our reliance on a limited pool of expert coders, introduces a number of logistical constraints. First, we had to decide which questions should be vignetted. The literature tells us to vignette any question that could be subject to cross-country coding differences. However, vignetting all questions that would fall in this category was not feasible. We were therefore in the position of attempting to define what would constitute "high priority" vignette questions. We evaluated each question on 73 characteristics that we thought would affect coding difficulty and, in turn, increase cross-country rating differences (e.g., length of the question, whether or not relative terms of degree differentiate the answer categories, and whether the question was getting at one or more underlying concepts). Using these characteristics, we selected the questions we thought were vulnerable to DIF, for a total of 117 questions, out of 149 C-questions in total.

The next task was to determine how we would construct vignettes for these selected questions. The literature suggests constructing many vignettes for each question, so as to generate as much information as possible about each coder's thresholds, and to maximize the probability of producing vignettes that are discriminating and effectively span the latent scale in question (Hopkins and King, 2010). But, mainly because we did not want to ask too much of our coders – and also because time constraints made producing sufficient quantities of vignettes difficult – we fielded only a few vignettes for each question. We attempted to field vignettes that would give us the greatest DIF information for least coder time. Since thresholds are effectively a coder's border between two answer categories, we decided to attempt to construct vignettes that were right on these thresholds, designed to be challenging to code as they appeared to straddle two answer categories. This would give us information about coder's thresholds in that we would learn whether a coder tended to code these border vignettes as belonging in the higher category or the lower category. If a question has k answer categories, then it has $k-1$ borders between answer categories.

However, given how the vignettes were embedded in the data collection tool within Historical V-Dem (described below), we were concerned about the potential for coders – either consciously or subconsciously – to order the vignettes as they coded them. This would mean that the data obtained from the vignettes would not be a clear signal of DIF, but instead would be tainted by the degree to which a logical order was apparent across the set of vignettes for a given question. To address this, we decided to include "decoy" vignettes for some questions – extra

vignettes for one or two border categories to preclude - or at least obstruct - the ordering of the vignettes by the coders. Out of the 117 questions to receive vignettes in historical V-Dem, we randomly assigned questions to receive either *k-1* vignettes (one for each border category), *k* vignettes (one for each border category and one decoy), or *k+1* vignettes (one for each border category and two decoys).

We then had to design a process to write the vignettes. We did not want to require Project Managers (PMs) and Principal Investigators (PIs) to write many paragraphs about fake countries, but we nonetheless wanted them to vet the vignettes. To strike a balance, a group of masters students at the University of Gothenburg wrote the vignettes (editing each other's work in an iterated process), and then the V-Dem Project Manager in charge of the question edited the vignettes for that question. The overarching guidelines for writing the vignettes were as follows:

- ON ONE HAND: Each vignette should be as specific as possible. It should include details about all aspects of the answer categories.
- ON THE OTHER HAND: Vignettes should NOT include details that are not part of the answer categories. Even the shortest word (e.g. adding the word "small" or "European" to describe the country) can add an irrelevant case detail that affects coding.
- Vignettes should be 1-2 paragraphs long.
- Vignettes should not obviously map to a question category. One way to ensure this is not the case is to vary the order of pieces of information in the vignette compared to pieces of information in the answer categories. For example, if an answer category talks about the geographic spread of fraud and then the kind of fraud, either reverse the order in the vignette or find a way to talk about both together.
- Vignettes should not obviously map to real cases. If coders can discern the case on which a vignette is based, then they could be influenced by the irrelevant details of that case. In other words, drawing content from a real case is fine, but please make sure it is not obvious.

An example of question wording and corresponding border vignettes appears in Table A.V.[ii]

[ii] Note that the election vote buying represents a rather difficult question to vignette. The question requires a clarification and the answer categories touch on number of aspects of vote-buying (i.e. percentage of population involved, fraction of parties involved, geographical spread, the extent to which bribes bought turnout and votes), even though the concept is purported to be uni-dimensional. Thus, we use this question as an example because it highlights how reliant vignettes - but also IRT-based methods more generally - are on the qualities of the underlying questions.

Table A.V: Vignettes for V-Dem Question on "Election Vote Buying"

Border Category	Vignette
0-1	In Country A, vote buying was a common strategy for several parties attempting to gather votes in the last election. The largest party is known for bribing people at all levels to vote for them. Some other parties approach only citizens in the poorer areas, offering cash or other handouts to place a vote for them on Election Day. However, one party probably did not engage in vote buying.
1-2	In Country B, a few of the well-resourced parties went around in many areas distributing cash handouts and material gifts such as clothing before the last elections. However, some news media reported that citizens in these areas sometimes accepted gifts from more than one party, so it is unclear how much this was about vote buying or just encouraging the voters in these areas to take the parties and candidates seriously. Yet, other citizens testified that they only took gifts from the party they would vote for.
2-3	In Country C, national laws outlaw the transaction of money for votes. However, in the last election, some parties managed to circumvent these laws by providing other materials, such as food items or cheap personal electronics, and it seems that cash may also have been handed out in certain districts. Due to the small scope, it is unclear if the parties in question sought to persuade potential voters or if it was just about recognizing some citizens who have been loyal supporters. In any case, it is unlikely to have affected election results.
3-4	In Country D, nearly no parties in the national parliament attempt the method of luring voters by material gifts during campaigning. Prior to the latest election, there were some reports of one party providing inducements to attend rallies and that it possibly also tried convincing some of the poorest communities by distributing gifts, but legal action was taken against this party.

Once the vignette texts were ready, we had to decide how to incorporate them in the data collection tool that coders used. Past literature has revealed that it is best to provide respondents with vignettes before they code, as this "anchors" them to a common scale (Hopkins and King, 2010). Accordingly, the coders were asked to code the vignettes for a given question before completing the question. To guard against coders attempting to order the vignettes as they coded them for each question, each question was randomly assigned one of five scrambled sequences for its vignettes.

References

Bakker, R., E. Edwards, S. Jolly, J. Polk, J. Rovny & M. Steenbergen. 2014. "Anchoring Experts: Using Vignettes to Compare Party Ideology Across Countries." *Research and Politics* 1(3):1-8.

Hopkins, Daniel J. & Gary King. 2010. "Improving Anchoring Vignettes Designing Surveys to Correct Interpersonal Incomparability." *Public Opinion Quarterly* 74(2):201-222.

King, Gary, Christopher J L Murray, Joshua a Salomon & Ajay Tandon. 2004. "Enchancing the Validty and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1):191-207.

King, Gary & Jonathan Wand. 2007. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes." *Political Analysis* 15(1):46-66.

Miller, Michael K. "Democratic pieces: Autocratic elections and democratic development since 1815." *British Journal of Political Science* 45.3 (2015): 501-530.

Appendix IV: Robustness tests for analysis on interstate war and democracy

Table A.VI: Regressing interstate war on Polity2 and Polyarchy, adding country-fixed effects to the benchmark.

	(1) Polity2 1817-2006	(2) Polyarchy 1817-2006	(3) Polity2 1817-1918	(4) Polyarchy 1817-1918	(5) Polyarchy 1918-2006
War past 5 yrs.	-0.072	0.004**	0.052	0.002	0.005**
	(-1.21)	(2.52)	(0.69)	(1.48)	(2.03)
Ongoing war	0.106	-0.001	0.148	-0.000	-0.002
	(1.33)	(-0.36)	(1.28)	(-0.04)	(-0.72)
Ln(GDPpc)	0.083**	0.002*	0.203***	0.005***	-0.000
	(2.60)	(1.73)	(3.74)	(4.29)	(-0.23)
Ln(population)	-0.074	-0.001	-0.048	0.002	-0.002
	(-1.49)	(-1.17)	(-0.82)	(1.26)	(-1.02)
Country-FE	Y	Y	Y	Y	Y
Year-FE	Y	Y	Y	Y	Y
N	14116	16202	4786	5094	11030
R^2	0.904	0.956	0.921	0.930	0.933

Notes: ***p<0.01; **p<0.05; *p<0.1. All models are OLS with errors clustered by country. T-values reported in parentheses.