The determinants of democracy: a sensitivity analysis

Espen Geelmuyden Rød¹ · Carl Henrik Knutsen^{2,3} · Håvard Hegre^{1,3}

Received: 30 October 2018 / Accepted: 15 October 2019 / Published online: 25 October 2019
© The Author(s) 2019

Abstract

Numerous studies—operating with diverse model specifications, samples and empirical measures—suggest different economic, social, cultural, demographic, institutional and international determinants of democracy. We distinguish between democratization and democratic survival and test the sensitivities of 67 proposed determinants by varying the control variable set, democracy measure, and sample time period. Furthermore, we go beyond existing sensitivity analyses and unpack the aggregate results by analyzing how theoretically motivated control variables affect sensitivity for two prominent determinants in the democracy literature: income and Islam. Overall, our results reveal a far larger number of robust determinants of democratization than of democratic survival. For democratic survival, the only robust factors are income and a law-abiding bureaucracy. In addition, our results highlight uncertainty surrounding the relationship between income and democratization, but show that broader development processes enhance the chances of democratization. Moreover, chances of democratization are lower in countries with large Muslim populations, but that relationship is sensitive to controlling for natural resources, education and neighborhood characteristics. Other results of the sensitivity analysis show that political protests, a democratic neighborhood, and the global proportion of democracies positively influence democratization, while natural resources, majoritarian systems, and long-tenured leaders make countries less likely to democratize.

Keywords Democratization · Democracy · Development · Religion · Sensitivity analysis

JEL Classification E02 · D74 · Z12

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11127-019-00742-z>) contains supplementary material, which is available to authorized users.

✉ Espen Geelmuyden Rød
espen.g.roed@pcr.uu.se

¹ Department of Peace and Conflict Research, Uppsala University, Uppsala, Sweden

² Department of Political Science, University of Oslo, Oslo, Norway

³ Peace Research Institute Oslo, Oslo, Norway

1 Introduction

The Arab Spring uprisings in 2011 were important political events, re-igniting old academic debates. The demands for democracy in those countries made some observers declare Huntington's (1996) argument on the incompatibility of Islam and democracy to be dead. Others dusted off Lipset's (1959) classic modernization argument, pointing to high education levels and a large middle class in Tunisia, but voicing scepticism regarding prospects of democracy in poor Yemen and oil-producing Libya. But how much do we actually know about the determinants of democratization and democratic survival? Two decades before the Arab Spring, numerous long-lived autocracies were challenged in Sub-Saharan Africa and electoral democracies emerged (Bratton and van de Walle 1997). However, African countries were, according to prevalent political science theories, unlikely to democratize, with low incomes, resource-based economies, deep ethnic cleavages, and problematic colonial histories.

Discrepancies between theoretical expectations and empirical events are not too surprising when considering the empirical literature on democracy. While consensus exists on some empirical relationships, studies often find weak or conflicting results despite plausible theoretical claims. For instance, several studies—using different model specifications, samples, and measures of democracy—yield divergent conclusions on whether income enhances democratization (Przeworski et al. 2000; Boix and Stokes 2003; Boix 2011). A key issue in the empirical literature on democracy is that theories often are tested without controlling for major alternative explanations (Gassebner et al. 2013). That problem is amplified by disagreements on how to measure democracy and that some determinants may have stable effects across time, whereas others operate only in specific time periods.

With so many open questions concerning appropriate control variables, democracy measurement, and time period, it is critical to examine systematically how robust the results are across different specifications. Assessing robustness is not only important for literatures with mixed results, but also when consensus emerges. For the latter, verifications of findings and investigations of their boundaries and assumptions are crucial for scholars and policy-makers alike. In this paper, we build on the designs of Sala-i-Martin (1997) and Hegre and Sambanis (2006) to study the sensitivities of 67 proposed determinants of democracy. We estimate 2.7 million dynamic logit regressions on 171 countries with time series extending from 1960 to 2015. With this approach, we mainly make an *empirical contribution* to the theoretically rich literature on democratization. Yet, we also highlight how our analysis can have important theoretical implications for key debates.

The study most closely related to ours is the extreme bounds analysis (EBA) by Gassebner et al. (2013), exploring 59 determinants of democratization and democratic survival using data from 1976 to 2002 for 165 countries.¹ Yet, our analysis is different from theirs in several important respects. What is most important, in addition to varying the control variable set, we gauge sensitivity using four prominent operationalizations of democracy

¹ Relying on the Democracy–Dictatorship measure from Cheibub et al. (2010), Gassebner et al. find the number of past democratic transitions (+), OECD membership (+), Muslim share of population (–), fuel exports (–), and economic growth (–) to be robust determinants of democratization. For democratic stability, only income level (+), democratic neighborhood (+), former military leader (–), and number of past transitions (–) are robust. A similar EBA approach also has been applied in studying terrorism (Gassebner and Luechinger 2011) and coups (Gassebner et al. 2016).

and four different time periods. Our study also differs in its assessment and interpretation of results. In addition to investigating aggregate robustness across specifications, we scrutinize the problem of omitted variable bias—and the converse issue of post-treatment bias—by considering “facilitating” and “fatal” controls. When entered into the specifications, *facilitating controls* increase coefficient size and turn otherwise non-significant predictors statistically significant. In contrast, *fatal controls* attenuate coefficients and make otherwise significant predictors turn insignificant. By using this approach, we systematically can investigate how the inclusion or exclusion of theoretically motivated controls affect our empirical conclusions (Plümper and Trautmüller 2016). Furthermore, in order to reduce concerns about unobserved country-specific confounding, in an extension we use the approach from Mundlak (1978) to estimate between- and within-country effects.

In addition to discussing the results of the sensitivity analysis as a whole, we demonstrate the usefulness of our approach by focusing on two high-profile debates, namely on (1) income and democratization/democratic survival (Przeworski et al. 2000; Boix and Stokes 2003) and (2) Islam and democratization (Huntington 1996; Hariri 2015). Our investigation provides important empirical insights with implications for theory development. First, many specifications yield no clear relationship between income and democratization. However, that non-finding is sensitive to the choice of democracy measure, the time period under study, and controlling for natural resources and industrialization, which we consider to be plausible facilitating controls. Moreover, the association between income and democratization is stronger when we consider within-country rather than between-country variation, using the Mundlak approach. Second, our results indicate that the strong relationship between income and democratic stability is to a large extent driven by between-country variation. In other words, while it is not clear that democratic countries become more stable as they grow richer, more well-to-do democracies are more stable than poor ones. Third, the aggregate results indicate that Islam impedes democratization, but we show that the relationship is sensitive to controlling for natural resources, education and level of democracy in the neighborhood (fatal controls).

More generally, our results reveal far more robust determinants of democratization than of democratic survival. To mention a few robust results, democratic neighborhood and the global proportion of democracies positively influence democratization, while resource-rich regimes and countries with majoritarian systems are less likely to democratize. As Gassebner et al. (2013) do, we find that short-term economic growth is negatively related to democratization, but only for two of four democracy measures. Moreover, that negative relationship is stronger during the Cold War period and vanishes afterwards. Furthermore, political protests are associated with democratization. For democratic survival, the only truly robust factor—aside from income level—is a state bureaucracy that is law-abiding. In addition to those key results, we provide a template for democracy researchers to assess sensitivity in a comprehensive manner. The setup of our supplementary material reduces the threshold for researchers wanting systematically to assess sensitivity pertaining to control variable selection, sample time series, and democracy measurement.

Yet, we caution against expectations that ours is a fully automated approach to drawing robust inferences—we are the first to acknowledge that our approach is a complement to, and not a substitute for, theory development and carefully constructed designs. We assess the relevance of control variable strategies in order to account for biases stemming from observable confounders, in addition to democracy measurement and sample time series. We also expand parts of the analysis to address country-specific characteristics and post-treatment bias. However, particular sources of bias remain, and researchers should keep them in mind when interpreting our results. Even with the extensive and carefully selected

sets of control variables and our efforts to address unobserved country-specific unobservables, countries may still differ on other unobservable factors that affect regime type *and* several of our covariates. Reverse causality, namely that our covariates are affected by regime type or recent regime change is another potential source of bias. In other words, even the sets of robust correlations established in this paper need to be scrutinized further in terms of unobserved confounding and reverse causality before they can be accepted as estimates of causal effects. Unfortunately, addressing reverse causality by employing alternative designs such as instrumental variable (IV) regressions or (natural) experiments is beyond the scope of this study. With those caveats in mind, we remind that we are, in fact, dealing with several key sources of bias and uncertainty that turn out to substantially affect results across a wide variety of proposed determinants of democracy.

In the following, we summarize the determinants of democracy briefly. Thereafter, we introduce the methodology, data and discuss limitations, before presenting the results.

2 Determinants of democracy and two prominent debates

The democratization literature is vast. The proposed determinants cover socio-economic factors such as income inequality and communications infrastructure, political-institutional factors such as colonial heritage and state capacity, and cultural factors related to the dominant religion. We cannot cover that literature in its entirety, but highlight 18 key concept categories, including six social and economic concepts, seven political and institutional concepts, and five concepts pertaining to demographic, cultural and international-contextual factors.²

Modernization theory is probably the most widely debated theory pertaining to democracy; most empirical studies employ GDP per capita to proxy for economic development (Przeworski et al. 2000; Boix and Stokes 2003; Acemoglu et al. 2008). Still, as Lipset (1959) highlighted in his classic exposition, other factors also related to development could affect democracy. Our first four concepts—(1) communications technology proliferation, (2) industrialization and urbanization (with growth of the urban middle- and industrial working classes), and (3) education—all have played prominent roles in debates on whether economic development leads to democratization or stabilizes democracy (Moore 1966; Rueschemeyer et al. 1992; Inglehart and Welzel 2005; Boix 2003; Acemoglu et al. 2005; Teorell 2010).

Much attention likewise has been devoted to how (4) natural resources impede transitions to democracy (Ross 2001). Furthermore, the study of (5) inequality and democratization has spawned different elaborate theoretical models and arguments (Boix 2003; Acemoglu and Robinson 2006), but the empirical evidence for any clear net effect remains elusive (Houle 2009). The final economic concept relates to (6) economic policy and performance—short-term economic growth and inflation, for example, may have deleterious effects on both autocracies and democracies (Przeworski and Limongi 1997; Kennedy 2010; Teorell 2010).

Concerning political and institutional concepts, we include (7) political system features. Many studies focus on how parties and elections in autocracies (Geddes 1999) or the forms of democratic governments (presidential or parliamentary systems, see Linz 1990; Cheibub

² See Table A-2 in the online appendix for an overview.

2007) affect democratization and democratic survival. The next concepts cover (8) ethnic exclusion, closely related to notions of power-sharing arrangements, and (9) colonial heritage, for which the most prominent hypothesis is that British colonial heritage improves prospects for democracy (Bernhard et al. 2004). Another widely held notion (Huntington 1968; Fukuyama 2014) is that high levels of state capacity are required for well-functioning and enduring democracies, pertaining to (10) institutional capacity. Finally, (11) political (in)stability, (12) mass activism, and (13) armed conflict cover events often considered detrimental to the survival of democracies and autocracies alike (Huntington 1991; Chenoweth and Stephan 2011).

Demographic factors, such as (14) population size, growth and density (Dahl and Tufte 1973), or ethnic, linguistic and religious (15) heterogeneity in the population (Alesina et al. 2003; Merkel and Weiffen 2012; Gerring et al. 2018), also could influence democratization and democratic survival. Cultural explanations of democracy are prevalent (Inglehart and Welzel 2005); cultural factors relevant for democracy often are linked to particular (16) religions, notably Protestantism, Catholicism, and Islam (Huntington 1996; Rowley and Smith 2009; Potrafke 2011). Values and attitudes with potential implications for regime type also may be clustered broadly by (17) geographic region. Moreover, regional patterns of regime change can occur through the mechanisms of (18) diffusion in neighboring or similar countries (Huntington 1991; Gleditsch and Ward 2006; Brinks and Coppedge 2006).

When presenting our results, we discuss the most robust findings on the determinants of democracy in light of the foregoing concept categories. Before that, however, we focus our discussion on two particular debates that have drawn much scholarly attention, namely the impact of income and Islam on democracy. That discussion allows us to scrutinize and compare various empirical estimates and discuss and illustrate the multiple issues that can be addressed by using our sensitivity approach carefully, including methodical assessments of fatal and facilitating variables. The income and Islam debates have drawn considerable attention from scholars (as our literature reviews below attest to) and policy-makers. Moreover, the scholarly debates suggest that various sources of sensitivity investigated in our analysis, e.g., related to control variables entered and lengths of the time series, matter for results on those relationships. Finally, both literatures contain extensive discussions of plausible mechanisms, which is especially interesting for us when studying the different, potential fatal and facilitating controls. Thus, results pertaining to the two relationships discussed here should both be inherently interesting to a broad audience, and, what is even more important, allow us to display the various benefits of our methodological approach.

2.1 Does higher income lead to democratization or democratic stability?

Income level is the most intensely studied determinant of democracy. Lipset (1959) famously argued that higher income levels increase the chances of countries being democratic. While Lipset's conclusions were qualified (Moore 1966) and challenged (O'Donnell 1973) early on, several studies replicated the positive relationship using more data, alternative controls, and more complex estimation techniques (Diamond 1992; Londregan and Poole 1996). By the mid-1990s, that relationship was considered to be a rare stylized fact of political science.

Despite these results—and various plausible explanations for why higher income enhances democratization, ranging from changing popular values (Inglehart and Welzel 2005) to autocrats' lower opportunity costs of exiting office (Boix and Stokes 2003)—no

consensus any longer exists that income leads to democracy. The relationship was challenged forcefully by Przeworski and Limongi (1997) and Przeworski et al. (2000), reporting evidence that income relates to the survival of democracies, but not to democratic transitions.

The relationship between income and democratic survival, more specifically, remains widely accepted today (but see Acemoglu et al. 2009). However, the null finding with respect to democratization has been contested by various studies. Yet, the income-democratization link has been re-established in other studies adjusting the statistical model and using alternative democracy measures (Epstein et al. 2006; Hadenius and Teorell 2005), and in studies extending the time series (Boix and Stokes 2003; Boix 2011). Other studies find the relationship to be conditional on the incumbent regime breaking down or the leader dying (Kennedy 2010; Treisman 2015). Boix and Stokes (2003) elaborate on how the null result of Przeworski and colleagues may stem from omitted variable bias, for instance driven by natural resource production, which hinders democratization. Moreover, time-invariant political-historical or geographical features may bias the income-democratization relationship upwards; Acemoglu et al. (2008) report that the income-democracy correlation disappears once accounting for country fixed effects. However, that finding also has been contested by studies using different estimators (Heid et al. 2012) or longer time series (Boix 2011).

Below, we contribute to this literature by scrutinizing how the relationship depends on choice of controls, but also the democracy measure and the sampled time series.

2.2 Is democratization less likely in countries that are predominantly Muslim?

Cultural explanations of democracy often highlight how “democratic” values, attitudes and norms are linked to religion. One argument is that conformity, obedience to authority, and other authoritarian values that predispose individuals to more easily accept autocratic regimes (Inglehart and Welzel 2005), are linked to religiosity (Saroglou et al. 2004). Other scholars highlight how *particular* religions make citizens and societies more or less predisposed to democratic rule. Some decades ago, Catholicism was considered to hinder democracy, especially when compared to Protestantism (Lipset and Lakin 2004).

More recently, both popular and scholarly discourses have focused on possible anti-democratic effects of Islam. Huntington (1996) famously argued that countries belonging to the “Islamic Civilization” inherently are less susceptible to democracy because of the religion’s doctrinal characteristics and its unique value system. Several other studies have proposed that Muslim-dominated countries are less likely to democratize, suggesting mechanisms ranging from treatment of women (Fish 2002)—discrimination against women is more pronounced in Muslim-majority countries where Islam is the source of legislation (Gouda and Potrafke 2016)—how religious services have been financed historically (Rothstein and Broms 2013), how autocratic incumbents exploit fears of political Islam among secular opponents to fend off liberalization (Lust 2011), to how Islamic legal institutions may impede political development by slowing down economic development (Kuran 2012). The hypothesis that Islam hinders democratization has received support in several cross-country studies (Rowley and Smith 2009; Potrafke 2011, 2013).

Other studies have contested that hypothesis. The debate mainly has revolved around whether “something else” about Muslim-majority countries makes them less susceptible to democratization, and the related issue of including the appropriate controls in regression analyses. One contention is that many majority-Muslim countries are Arab and that it

is those countries—which have other anti-democratic features—is driving the relationship (Stepan and Robertson 2003). Many Muslim countries also are oil-rich, leading scholars to suspect that resource-curse mechanisms drive the negative correlation between Islam and democratization (Ross 2001). More recently, Hariri (2015) has presented evidence that the strengths of pre-colonial Middle Eastern state institutions account for the negative association between majority-Muslim status and democratization.

3 Methodology of sensitivity analysis

The mixed evidence on many proposed determinants of democracy and the absence of consensus on how to specify empirical “democracy models” correctly, constitute strong reasons for conducting sensitivity analysis. Different sensitivity designs have been used by economists to study, for example, the determinants of economic growth (Leamer 1985; Levine and Renelt 1992). We follow the design of Sala-i-Martin (1997), who estimated a large number of models of the following form:

$$\gamma_j = \alpha_j + \beta_{yj}y + \beta_{zj}z_j + \beta_{xj}x_j + \epsilon \quad (1)$$

Here, γ is the dependent variable and j refers to the model, y is a vector of three variables appearing in all regressions, z is the variable of interest (“focus variable”), and x is a vector of up to three variables (“core variables”) drawn from a larger pool of candidate variables.

We adapt that model in several ways. First, while Sala-i-Martin applied the method to a continuous dependent variable, we model the relationship between predictors and dichotomous dependent variables using the logit-link function and standard errors clustered by country. Second, in order to investigate transitions to and from democracy simultaneously, we employ a “dynamic logit” model (Przeworski et al. 2000). That model has regime type (D_t) as dependent variable and includes the lagged dependent variable (D_{t-1}), along with interaction terms between D_{t-1} and all explanatory variables as regressors. Thus, the x variables in Eq. 1 become *pairs* of variables. We include two additional core variables (C_{t-1}). First, to take into account that regimes consolidate over time, we enter a logged indicator of time since transition to autocracy/democracy. Second, GDP per capita also is included as a core variable:

$$\begin{aligned} D_{j,t} = & \alpha_j + \beta_{Cj}C_{t-1} + \beta_{Dj}D_{t-1} + \beta_{DCj}C_{t-1}D_{t-1} + \beta_{zj}z_{j,t-1} \\ & + \beta_{Dzj}z_{j,t-1}D_{t-1} + \beta_{xj}x_{j,t-1} + \beta_{Dxj}x_{j,t-1}D_{t-1} + \epsilon \end{aligned} \quad (2)$$

The main transition term coefficients represent changes in log odds of democracy at t for observations that were non-democratic at $t - 1$ (democratization) and the sum of the main and interaction term coefficients represents the log odds of democracy *if* the country was democratic at $t - 1$ (democratic stability).

To mitigate multicollinearity and post-treatment bias we restrict the *total* number of explanatory variables to six pairs (main and interaction terms of core variables (C), focus variable (z), and controls (x)) plus the lagged dependent variable. Our data contain different measures of the same concept. For instance, we measure education using both illiteracy rates and secondary school-enrollment ratios. Therefore, we restrict combinations of x -variables to those measuring *different* concepts. If we were to control for alternative

operationalizations of the same concept, the estimated effect of the focus variable likely would be attenuated.

Building on Plümper and Trauttmüller (2016), we take the sensitivity approach one step further by exploring which of the included controls x_j affect a given focus variable z_j the most. To structure that discussion we identify variables that are ‘facilitating’ or ‘fatal’. A variable x_j is *facilitating* a variable z_j if z_j tends to be significantly different from 0 if x_j is included in the model, but not when x_j is excluded. A variable x_j is *fatal* for z_j if z_j tends to be insignificant if x_j is entered, but significantly different from 0 when x_j is excluded.

Our extensive specification tests reflect that we do not *know* which control variables should be included when testing various determinants of democracy. We assume only that lagged regime type, time since transition, GDP per capita, and interactions should be entered. While most scholars would agree that such a model is under-specified, they would disagree over which other variables to add. However, the choice of controls is likely to affect results. Our analysis explicitly incorporates a full range of theoretically relevant variables and reports how the inclusion of each (potentially fatal/facilitating) control affects the results. That approach does not replace the need to develop theory, but it allows us to gauge empirically the robustness of evidence in a field where numerous theoretical models and contradicting findings exist.

3.1 Limitations

Before presenting the data and results, we address limitations to our main setup, alongside strategies for mitigating those issues. First, the estimates could be affected by post-treatment bias if the controls include variables that (I) are consequences of our focus variable and (II) affect the dependent variable. That problem is most apparent when entering indicators for mass activism and leadership change, but also persists for structural factors. For example, we underestimate the *total effect* of bureaucratic quality on democratic survival if bureaucratic quality enhances income levels (always controlled for), and income enhances survival. The aggregate coefficients therefore represent *direct effects*. Yet, our investigation of facilitating and fatal variables facilitates theoretically informed discussions of which of the controls represent likely confounders and which represent likely mediators. In an extension, we also drop GDP per capita as a core variable and discuss implications for post-treatment and omitted variable bias from differences between those results and our main analysis.

Second, unmeasured country-specific confounders could affect the results and we therefore implement the approach in Mundlak (1978) to distinguish between- and within-country variation in an extension. Since many theoretical arguments in the democracy literature consider differences between countries rather than changes within countries, we prefer that approach over country-fixed effects (which wipe out differences between countries) in this analysis. Whether countries with predominantly Muslim populations are less likely to democratize than other countries is one prominent example of a question that must draw on cross-country variation.

A third issue is reverse causality. Unfortunately, we cannot deal with that problem for all variables in our setup. Thus, we have to assume exogeneity for all (lagged) independent variables as most large-n studies on democracy do. However, IV estimation is compatible with the approach in this paper; researchers who focus on one particular X–Y relationship and have identified valid instruments can implement it easily.

4 Data

Our dataset covers 171 countries (1960–2015) and 67 proposed indicators of democracy's determinants grouped into the 18 concepts discussed in Sect. 2 (see Table A-1 in the online appendix for details on all indicators). We standardized all continuous indicators to make estimates more easily comparable.

Numerous measures of democracy exist, including several new ones that draw on different conceptual schemes and methodological approaches such as IRT-modelling or machine learning (Pemstein et al. 2010; Gründler and Krieger 2016). We employ four widely adopted democracy measures. First, Democracy and Dictatorship (DD/ACLP Cheibub et al. 2010) has long been the most prominent minimalist democracy measure, categorizing regimes as democratic only if they have multi-party elections *and* incumbents have *demonstrated* willingness to concede electoral defeat and transfer power constitutionally to the opposition. We use updated data from Bjørnskov and Rode (2019). Second, the Boix-Miller-Rosato (BMR) measure (Boix et al. 2012) differs from DD/ACLP by evaluating factors other than governmental change to decide whether a regime is competitive or not, and by also requiring the granting of voting rights to the majority of male citizens so as to code a regime as democratic.

DD/ACLP and BMR are dichotomous by construction. In contrast, our third measure, Polity2 from Marshall et al. (2018) ranges from -10 to $+10$. Polity measures a more expansive concept of democracy than DD/ACLP/BMR do by incorporating executive-branch constraints, which, according to liberal understandings of democracy, is critical for avoiding abuses of power. We use the most common threshold and construct a dummy scored 1 if *Polity2* ≥ 6 . Although the cutoff point admittedly is arbitrary (Cheibub et al. 2010), and other thresholds have been adopted (Bogaards 2012), our specification classifies the more ambiguous, and often controversial, cases as non-democracies. Finally, we use a categorical measure based on V-Dem data (Coppedge et al. 2019), namely the “Regimes of the World” measure in Lührmann et al. (2018), which we dichotomize (electoral democracies and liberal democracies are counted as democracies). To count as an electoral democracy, a country must pass three thresholds on indicators of multiparty elections (*v2elmulpar_osp* > 2), free and fair elections (*v2elfrefair_osp* > 2), and an Electoral Democracy index (*v2x_polyarchy* > 0.5). The resulting measure overlaps strongly (about 90% agreement) with the three other democracy measures, but typically leads to fewer countries being classified as democracies (Lührmann et al. 2018, 68).

We also estimate our models on four different time periods: (1) 1960–2015 (entire period), (2) 1974–2015 (Third wave), (3) 1960–1990 (Cold War), and (4) 1991–2015 (post-Cold War). We introduce different sample time periods because the determinants of democracy may change over time, for instance with changes in the power structure of the international system.³

Four democracy measures combined with four time periods leave us with 16 subsets of results. Thus, the total number of possible variable combinations is in the tens of millions, executing which is very intensive computationally. We therefore hardwire our setup to 2500 regressions per focus variable/democracy measure/time period combination, drawing

³ Given the lack of pre-1960 data for numerous indicators, we unfortunately cannot study earlier periods in our setup, which could affect our conclusions (see Boix 2011).

controls randomly for each regression. That approach results in approximately 2.7 million regressions. This random sampling is highly unlikely to affect our conclusions qualitatively.

Many variables have numerous missing observations (Table A-2 in online appendix). Faced with that problem, we could have added all variables independent of missing values (Gassebner et al. 2013). However, we want to isolate the sensitivity of parameter estimates to model specification changes, and varying the sample can influence the estimates. Second, we could have excluded indicators with many missing observations, thereby ignoring substantively interesting explanatory variables. But, excluded variables may correlate systematically with the included variables; leaving some variables out may affect results for those that are entered. Hence, we opted for multiple imputation using AMELIA II software (Honaker and King 2010). We provide details on the imputation model in online Appendix B. We created ten different imputed datasets and drew randomly one of the ten datasets for each regression to avoid the problem of particular imputations influencing the results.

5 Results

In order to focus our discussion, we first present and evaluate results pertaining to the debates on income and Islam. Next, we identify indicators that are robust across all democracy measures in our main setup on the full time series (1960–2015), before we present various extensions. The online appendix presents summary tables and figures for readers interested in evaluating the robustness of additional specific independent variables. The supplementary syntax can be used to reproduce graphs similar to those presented below for any variable of interest. We report robustness in terms of (unweighted median) $\tilde{\beta}$ - and $\tilde{\sigma}$ -values, or the equivalent $\tilde{t}$ -values, where $\tilde{t} = \frac{\tilde{\beta}}{\tilde{\sigma}}$ across all relevant estimations for which a variable is a focus variable. In order to avoid dichotomizing evidence according to conventional significance thresholds, we also discuss the magnitudes of $\tilde{\beta}$ and how they change when particular controls are entered (see Wasserstein and Lazar 2016).

We visualise findings with coefficient plots of $\tilde{\beta}$ for each of the 16 possible democracy measure and time period combinations.⁴ The 95% confidence intervals are calculated using $\tilde{\sigma}$. Coefficient colors (black, dark grey, and grey) indicate level of significance (.05, .10, and insignificant). Similarly, we display the results for facilitating and fatal variables as coefficient plots, wherein colors indicate how findings change when controlling for a specific variable. A major change (black) occurs when an aggregate insignificant variable becomes significant at .05 (facilitating) or when an aggregate significant variable at .05 becomes insignificant (fatal). A minor change (dark grey) occurs when a insignificant variable becomes significant at .10 (facilitating), or when a significant variable at .05 changes to significant at .10 (fatal).

5.1 Do higher incomes lead to democratization and democratic stability?

Figure 1 summarizes results for the relationships between GDP per capita and democratization (left panel) and democratic survival (right panel). The left panel illustrates that our conclusions regarding the income–democratization relationship depend on the choice of

⁴ The plots are based on subsets of 9000–10000 specifications for GDP per capita which enters all regressions, and 80–120 specifications for Muslim population shares.

Fig. 1 Coefficient plots of $\bar{\beta}$ with 95% CI across all specifications for GDP pc and democratization (left) and GDP pc on democratic stability (right). Black = significant at .05 level. Dark grey = significant at .10 level. Grey = insignificant

democracy measure and sample time period. In the full time series, income does not relate to democratization systematically when using the Polity, DD/ACLP and BMR democracy measures. However, we note that $\bar{\beta}$ is positive across democracy measures and, when the V-Dem measure is used, also statistically significant. Compared to V-Dem, the odds of obtaining a positive and significant (at .05) coefficient of income on democratization (across 1960–2015) is reduced by 74% to 85% when instead using Polity, DD/ACLP, or BMR.⁵ Sample differences cannot account for this pattern as they are similar across the different measures.

What time period we draw data from also affects results: $\bar{\beta}$ is consistently positive and statistically significant across democracy measures in the Cold War period. The coefficient also is quite large; one standard deviation increase in GDP per capita, on average, increases the odds of democratization (in the Cold War period) by 1.39. That result could, for instance, be related to several developed Eastern European countries democratizing at the end of the Cold War (Boix and Stokes 2003). In the post-Cold War period, by contrast, the effect of GDP per capita on democratization is negative for three of the four democracy measures. These findings highlight that the effect may have changed over time, potentially reflecting changes in the international system (Boix 2011) and new technology (Rød and Weidmann 2015). One plausible hypothesis is the following: Until the last few decades, increasing income levels have corresponded strongly with changes in social structures (e.g., the rise of organized industrial workers and the middle class), which led to stronger pressures for democracy. With the growing automatization and digitization of

⁵ To calculate the numbers, we ran a logistic regression on our estimated results. The dependent variable takes the value 1 if a given estimation shows that income is significant at .05, and 0 otherwise. We follow a similar approach for other results reported below; see online “Appendix C” for results.

One potential problem with the aggregate results presented above is that they can “hide” subsets of regressions in which income relates to democratization. Do any controls *facilitate* a positive and significant relationship? Fig. 2 plots $\bar{\beta}$ and $\bar{\sigma}$ for each of the four democracy measures, for 67 subsets of specifications defined by having a specific

control included. Several interesting patterns emerge, with numerous controls facilitating noticeable changes in $\bar{\beta}$ and their significance levels. The most important pattern in Fig. 2 is that the largest positive and significant $\bar{\beta}$ values appear for regressions that include controls identifying whether income is a function of natural resource extraction or not (natural resource, industrialization and urbanization variables). We are about 2.5 times more likely to find that income is significant at .05 when comparing models controlling for any resource curse indicator to other models. Furthermore, across all four democracy measures, the coefficient enlarges more than fourfold, from a median of 0.05 in specifications excluding resource-related controls to 0.22 in specifications including them. Since natural resource income can impede broader economic development *and* make autocracies less likely to democratize (Ross 2001; Boix and Stokes 2003), one may argue plausibly that a proper test should include such controls. The results indicate a clear relationship between GDP per capita and democratization, and the odds of democratization increases by 1.25 when GDP per capita increases by one standard deviation. A second notable pattern is that negative $\bar{\beta}$ values appear when controlling for concepts identifying other features of economic development (communication technology proliferation, education). The latter result indicates that the relationship between income and democratization is likely to be related to broader development processes, and that any *direct effect* observed when accounting for other aspects of development is zero or negative.

Regarding income and democratic survival, the right panel in Fig. 1 shows that the aggregate $\bar{\beta}$ is positive and statistically significant for all four democracy measures for the 1960–2015 period. $\bar{\beta}$ is quite large; a standard deviation change in GDP per capita increases the odds that a democracy survives the next year by a median of 1.7. Although the 95% confidence intervals overlap zero for the Cold War sample when V-Dem is used, the finding is robust. The remarkably consistent results corroborate the widely held notion that democracy is less prone to collapse in rich than in poor countries.

However, particular controls could be “fatal” to the income–democratic survival result. Figure 3 shows, *first*, that when democracy is measured by PolityIV and DD/ACLP, income has a large, positive coefficient on democratic survival *no matter which control* enters the model. In contrast, $\bar{\beta}$ declines and significance levels drop when controlling for two and six variables if BMR and V-Dem, respectively, are used. *Second*, $\bar{\beta}$ is smaller and more often insignificant when controlling for concepts that tap into other features of development (communications technology proliferation, industrialization and urbanization, education, health, administrative capacity). Thus, as for the income–democratization link, the evidence is (even) clearer for the notion that broader development processes, rather than only income more narrowly, stabilize democracy.

5.2 Is democratization less likely in countries that are predominantly Muslim?

The aggregate results, shown in Fig. 4, support the proposition that democratization is less likely in predominantly Muslim countries: $\bar{\beta}$ is negative across all democracy measures and time periods; it is significant at .10 in ten of 16 samples. The estimated effect also is quite large: A standard deviation change in the Muslim population percentage reduces the odds of democratization by 27%. In the Third Wave sample, the negative relationship is remarkably consistent across democracy measures, while more uncertainty exists surrounding the effect in the Cold War sample.

To inspect the negative aggregate relationship more closely, Fig. 5 plots fatal variables for the Muslim population share effect on democratization. As anticipated from the

Fig. 3 Fatal variables for GDP pc on democratic stability. Black = insignificant. Dark grey = significant at .10 level. Grey = significant at .05 level

literature review, the aggregate result is sensitive to the inclusion of specific controls. *First*, $\bar{\beta}$ shifts toward zero and loses significance—consistently across democracy measures—in models controlling for resource curse variables, suggesting that the aggregate relationship is inflated by abundant natural resources in Muslim countries. The odds of finding that the Muslim share has a significant negative relationship with democratization drops by 82% when controlling for resource curse variables, and $\bar{\beta}$ drops from a median of -0.33 to -0.21 . *Second*, controlling for neighborhood democracy levels reduces the odds of finding a

Fig. 4 Coefficient plots of $\bar{\beta}$ with 95% CI across all specifications for Muslim share of population on democratization. Black = significant at .05 level. Dark grey = significant at .10 level. Grey = insignificant

significant result by 96%. Here, $\bar{\beta}$ also drops substantially, from -0.32 to -0.17 . *Third*, the significance of the Muslim variable drops across democracy measures when controlling for education. However, if having a Muslim population affects educational outcomes, for instance by depressing female school enrollment (Fish 2011), the estimates are afflicted by post-treatment bias. Other, more complex stories may underlie the finding, as dependence on natural resources reduces the need for an educated, specialized workforce. In regressions from which natural resource measures are excluded, education might pick up such variation and contribute to reducing omitted variable bias.

In sum, closer scrutiny reveals that the robust aggregate negative effect of Muslim population on democratization is weakened when theoretically relevant controls are entered. When controlling for resource abundance, education, and neighboring regime type, the relationship is attenuated and most specifications fail to yield a significant result. At the same time, we refrain from concluding too forcefully since $\bar{\beta}$ for the Muslim-share variable consistently is negative and relatively large even when the above-discussed “fatal” controls are entered.

5.3 “Robust” determinants

We will now discuss “robust” determinants of democratization and democratic survival. Our criterion for applying that label is that the variable is significant at .05 for *all four democracy measures* in regressions run on the full time series (1960–2015). All robust

Fig. 5 Fatal variables for Muslim share of population on democratization. Black = insignificant. Dark grey = significant at .10 level. Grey = significant at .05 level

variables are listed in Table 1 (democratization *or* democratic stability), and ranked according to t -values for regressions using PolityIV. The tables also contain information on the number of fatal variables for each robust determinant ('Nr. fat'). Overall, we identify many more robust determinants of democratization than of democratic survival (20 versus two). That conclusion reflects, in part, less variation and, thus, greater uncertainty in estimates associated with the fewer transitions to dictatorship from democracy than vice versa. Yet, the large difference in robust determinants suggests additional reasons for the pattern. The large difference could reflect that previous theoretical efforts and empirical studies, which

Table 1 List of highly robust determinants for all democracy measures, results for democratization and democratic stability. Highly robust = significant at .05 for all four democracy measures in regressions run on the full time-series (1960–2015)

Focus variable	PolityIV coef	PolityIV t	ACLP coef	ACLP t	BMR coef	BMR t	VDem coef	VDem t	Nr. fat. Polity	Nr. fat. ACLP	Nr. fat. BMR	Nr. fat. VDem
Democratization												
Nonviolent campaign	2.244	8.467	2.134	8.009	2.271	8.447	1.528	5.652	0	0	0	0
Demonstrations	0.481	7.194	0.433	5.875	0.457	7.106	0.240	3.078	0	0	0	0
Neighborhood Polity score	0.741	5.274	0.526	4.083	0.621	4.430	0.903	6.258	0	0	0	0
Energy consumption pc	− 0.541	− 4.676	− 0.568	− 5.867	− 0.330	− 2.503	− 0.491	− 3.744	0	0	6	0
Plurality/majority system	− 1.071	− 4.654	− 0.928	− 4.361	− 0.811	− 3.657	− 0.952	− 4.277	0	0	0	0
TV's pc	0.822	4.598	0.560	3.353	0.587	3.500	0.597	3.021	0	0	0	1
Radios pc	0.673	3.904	0.462	3.020	0.536	3.172	0.633	4.402	0	0	0	0
Average years of education, female	0.632	3.826	0.342	2.316	0.364	2.479	0.494	3.302	0	13	9	3
Resource rents, % GDP	− 0.742	− 3.820	− 0.487	− 3.376	− 0.570	− 3.374	− 0.700	− 4.548	0	0	0	0
Phones pc	0.680	3.805	0.364	1.977	0.405	2.388	0.771	4.520	1	20	7	2
Industry's share of economy, %	− 0.453	− 3.703	− 0.426	− 3.586	− 0.537	− 4.015	− 0.654	− 5.412	3	2	1	1

Table 1 (continued)

Focus variable	PolityIV coef	PolityIV t	ACLP coef	ACLP t	BMR coef	BMR t	VDem coef	VDem t	Nr. fat. Polity	Nr. fat. ACLP	Nr. fat. BMR	Nr. fat. VDem
Western Europe/North America	1.028	3.657	0.854	3.410	0.693	2.855	0.459	2.470	0	1	4	15
Oil/gas value pc	-0.542	-3.613	-0.375	-2.647	-0.284	-1.985	-0.625	-4.455	0	4	26	0
Length of executive tenure	-0.409	-3.446	-0.327	-3.204	-0.254	-2.557	-0.506	-4.853	0	0	0	0
Average years of education, male	0.516	3.091	0.306	2.101	0.315	2.143	0.457	2.756	4	18	17	7
Oil rents, % GDP	-1.291	-2.997	-1.095	-3.861	-1.159	-4.193	-1.283	-3.081	0	0	0	0
Global proportion of democracies	0.390	2.988	0.375	2.863	0.290	2.237	0.619	4.905	3	0	12	0
Central/South America	0.729	2.720	0.901	3.276	0.808	2.867	0.490	2.128	6	1	3	14
Time since ILC	0.680	2.144	0.879	3.194	0.979	3.387	0.871	3.932	9	0	0	0
Middle East/North Africa	-1.456	-2.114	-1.946	-2.782	-1.687	-2.807	-1.878	-3.295	11	1	1	1

Table 1 (continued)

Focus vari- able	PolityIV coef	PolityIV t	ACLPL coef	ACLPL t	BMR coef	BMR t	VDem coef	VDem t	Nr. fat. Polity	Nr. fat. ACLPL	Nr. fat. BMR	Nr. fat. VDem
Democratic stability												
Impartial public adm.	1.346	4.320	0.774	3.841	1.042	3.956	1.515	5.877	0	0	0	0
Political cor- ruption	-0.669	-4.188	-0.423	-2.203	-0.464	-2.594	-0.740	-4.193	0	5	1	0

have informed our variable selection, are more attuned to explaining democratization (see Knutsen and Nygård 2015). But, democratic breakdowns also could be processes that inherently are harder to explain than democratization episodes. Nonetheless, our analysis highlights that, in addition to GDP per capita, only the indicators for political corruption and impartial public administration are robust determinants of democratic survival. The results thus suggest that the most robust determinants of democratic survival relate to features of the political and institutional history of a country and, most notably, the extent to which a law-abiding bureaucracy has developed.

For democratization, Table 1 reveals that the length of executive tenure is robust and, moreover, that variable is not associated with *any* fatal controls. The measure of time since irregular regime change reveals a similar result. Our analysis thus leaves little doubt that autocratic incumbents who have entrenched their positions over longer periods of time are less likely to experience democratization. Furthermore, being located in a democratic neighborhood increases the likelihood of democratization (Gleditsch and Ward 2006); also that result is remarkably robust.

Other robust determinants of democratization include measures of domestic unrest, especially non-violent mass campaigns (corroborating Chenoweth and Stephan 2011). The domestic unrest variables reveal little sensitivity to including particular controls. In addition, we find that having a majoritarian electoral system is related negatively to democratization. Indeed, the literature on autocratic elections highlights how majoritarian systems may be easier to manipulate for autocrats, yield large regime-supporting parties seat premiums in most contexts, and mitigate the formation and growth of new opposition parties that later may challenge the regime (Higashijima 2019).

Regarding economic indicators, various measures of communication technologies, namely televisions, radios, or phones per capita, also are positive and robust. Similarly, variables capturing social development, in particular average years of education for women/men, are robust. Thus, evidence indeed exists favoring a more nuanced interpretation of the Lipset (1959) thesis; at least some forms of economic development are associated robustly with democratization. Moreover, we note that both the robust education measures, as well as less robust measures of communications technology (extension of radios, telephones, or televisions) display a stronger positive correlation with democratization in all other sample specifications than the post-Cold War sample. In other words, those aspects of development are more strongly related to democratization in the earlier than in later parts of our sample.

Regarding other economic determinants, our results lend support to the resource curse argument. All three of our variables related to resource wealth display negative and robust relationships with democratization (Ross 2001). One noteworthy *omission* from our list of most robust indicators is economic growth. Gassebner et al. (2013) find strong evidence that economic success stabilizes autocracies (but not democracies). Also in our analysis, growth is related negatively to democratization and significant at conventional levels for Polity and DD/ACLP, but *not* for BMR and V-Dem. Moreover, that coefficient is by far the largest in the Cold War period (− 0.28) and much smaller during the Third Wave (− 0.10) and the post-Cold War (0.00) periods. That reduction in the size of the growth coefficient is consistent across all democracy measures.

5.4 Additional tests

In this section, we discuss variations on our main analysis (see online Appendix E, F, G for results). First, we estimated our models excluding GDP per capita as a core variable in order to reduce concerns of post-treatment bias. Yet, for the democratization results, omitting income has only a minor impact on which of the indicators are robust. Only three variables drop off the robust list, namely energy consumption per capita, oil/gas per capita and global proportion of democracies. For democratic survival, however, several indicators of concepts associated with broader economic modernization become robust (industrialization and urbanization, education) once income is omitted.

The latter result provides additional evidence that broader economic development enhances democratic survival. However, when it comes to disentangling which features of development have a clear relationship with democratic stability, we run into issues of complex causality. For instance, it is plausible that richer countries can afford to build better educational systems, suggesting that income should be controlled for. Simultaneously, improved education may enhance economic efficiency and, thus, income (Bils and Klenow 2000). Controlling for GDP per capita therefore also could produce downward post-treatment bias. What we *can* say, however, is that development is related to democratic durability—be it through increased incomes or other channels. And, the sensitivity analysis with and without GDP per capita may be considered to represent, respectively, lower and upper bounds for the direct relationship between the different features of development and democratic survival.

Second, we re-ran the analysis using GDP data from the Maddison project (Bolt et al. 2018) rather than World Bank (2019). Changing the data source for GDP has some effects on the estimates, for instance on the coefficients on GDP per capital itself for democratization. These results are reported in online Appendix G. While the democratic survival results are robust, the democratization results are weakened. That evidence adds to the cautionary note that the relationship between income and democratization is not robust across specifications.

Third, we assess the extent to which between- or within-country differences drive the results by applying the Mundlak (1978) estimator, which calculates the mean value for each country (between-variation) and subtracts country-year values from this mean (within-variation). Both the between- and within-variables subsequently are entered into our original setup. We focus our discussion on the effects of income on democracy since the effect of Muslim population on democratization—unsurprisingly owing to its stability over time—exists only between countries. The Mundlak estimation reveals that the income–democratic survival relationship is driven mainly by variation between countries. A relatively large average estimated within-country effect also is found (with considerable uncertainty) when using Polity, DD/ACLP and BMR, but not when using V-Dem. We note that the uncertainty for the within-country findings can stem from relatively few transitions from democracy to autocracy in the post-1960 data. Yet, in combination, the results add nuance the findings above; it is not clear that a democratic country becomes less prone to breaking down as it becomes wealthier. But, when comparing across countries, richer democracies are more stable than poorer ones.

For income and democratization, between-country estimates do not suggest a relationship; $\hat{\beta}$ hovers around 0. But we find evidence of a positive within-country relationship; the income coefficient is positive and statistically significant across all democracy measures. If we consider only models controlling for natural resources, a further

increase in $\bar{\beta}$ is evident across all democracy measures. Thus, despite the overall sensitivity of the relationship, the latter results point in the direction that higher incomes are related to higher probabilities of democratization.

6 Conclusion

The determinants of democracy are among the most studied topics in political science, with several well-known theoretical arguments relating to both democratization and democratic survival. Yet, the picture from the empirical literature is unclear, and the mixed evidence has spawned numerous debates between scholars reporting diverging findings. Many results are sensitive to model specification choices, democracy measure used, and time period under study—this goes even for literatures where a consensus on empirical relationships has emerged, but where this consensus is based on results from a narrow range of specifications. The overview that we have presented here on how such choices matter for particular results is an important contribution to the large, and growing, enterprise of democracy studies.

We have conducted a comprehensive sensitivity analysis, investigating the robustness of 67 indicators tapping into 18 proposed social, economic, political, institutional, demographic, cultural and international-contextual determinants of democracy over the 1960–2015 period. To illustrate our approach we focused on two prominent debates, namely (1) the impact of income on democratization and democratic stability and (2) the impact of Islam on democratization. We also discussed the overall results, identifying far greater numbers of determinants of democratization than of democratic survival. Finally, we make available supplementary materials to facilitate investigation of parameter instability for scholars interested in particular determinants of democratization or democratic survival.

While our analysis has covered several important issues and highlighted different sources of variation that influence the empirical analysis of democracy, other sources of variation should be explored further in similar systematic studies. Notably, future sensitivity analysis could use a set-up similar to ours to explore whether the use of different continuous variables—capturing smaller and larger “upturns” and “downturns” (Teorell 2010), also within the binary regime categories studied herein—matter for the empirical results. When employing continuous democracy measures, one pertinent test to consider is the sensitivity of results to controlling for country fixed effects.

Nonetheless, our approach improves upon existing sensitivity analyses—on democracy and on other outcomes studied by social scientists—in at least two ways. First, we address, simultaneously, robustness to varying the control specification, sample and operationalization of the outcomes. Second, going beyond the aggregate results, we address how the inclusion of specific control variables influences the sensitivity of the aggregate results. By doing so, we, for instance, highlight how findings on income levels or Muslim population shares are influenced by time period under study, choice of democracy measure, or the inclusion of theoretically relevant control variables. As a consequence, our setup is more likely to help uncover sources of disagreement in existing scholarly debates; such more nuanced results are likely to help spur further theory development. Thus, we urge democracy scholars to familiarize themselves with our approach, with how the different specification choices tend to affect results, and explicitly to acknowledge the related uncertainty. We hope that our contribution serves

to move contentious debates on various proposed determinants of democracy one step forward.

Acknowledgements Open access funding provided by Uppsala University. The research was funded by the Research Council of Norway (Projects 204454/V10 and 217995/V10), and by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. 694640). Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). We thank three anonymous reviewers for insightful comments and suggestions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2005). From education to democracy. *American Economic Review*, 95(2), 44–49.
- Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3), 808–842.
- Acemoglu, D., Johnson, S., Robinson, J. A., & Yared, P. (2009). Reevaluating the modernization hypothesis. *Journal of Monetary Economics*, 56(8), 1043–1058.
- Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. New York: Cambridge University Press.
- Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., & Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(June), 155–194.
- Bernhard, M., Reenock, C., & Nordstrom, T. (2004). The legacy of western overseas colonialism on democratic survival. *International Studies Quarterly*, 48(1), 225–250.
- Bils, M., & Klenow, P. J. (2000). Does schooling cause growth? *The American Economic Review*, 90(5), 1160–1183.
- Bjørnskov, C., & Rode, M. (2019). Regime types and regime change: A new dataset. *Review of International Organizations*. Forthcoming.
- Bogaards, M. (2012). Where to draw the line? From degree to dichotomy in measures of democracy. *Democratization*, 19(4), 690–712.
- Boix, C. (2003). *Democracy and redistribution*. Cambridge: Cambridge University Press.
- Boix, C. (2011). Democracy, development, and the international system. *American Political Science Review*, 105(4), 809–828.
- Boix, C., Miller, M., & Rosato, S. (2012). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, 46(12), 1523–1554.
- Boix, C., & Stokes, S. C. (2003). Endogenous democratization. *World Politics*, 55(4), 517–549.
- Bolt, J., Inklaar, R., de Jong, H., & van Zanden, J. L. (2018). Rebasing maddison: new income comparisons and the shape of long-run economic development. *GGDC Research Memorandum*, p. 174.
- Bratton, M., & van de Walle, N. (1997). *Democratic experiments in Africa: Regime transitions in comparative perspective*. Cambridge: Cambridge University Press.
- Brinks, D., & Coppedge, M. (2006). Diffusion is no illusion: Neighbor emulation in the third wave of democracy. *Comparative Political Studies*, 39(4), 463–489.
- Cheibub, J., Gandhi, J., & Vreeland, J. (2010). Democracy and dictatorship revisited. *Public Choice*, 143(1–2), 67–101.
- Cheibub, J. A. (2007). *Presidentialism, parliamentarism and democracy*. Cambridge: Cambridge University Press.
- Chenoweth, E., & Stephan, M. J. (2011). *Why civil resistance works: The strategic logic of nonviolent conflict*. New York: Columbia University Press.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Fish, S. M., Glynn, A., Hicken, A., Lührmann, A., Marquardt, K. L., McMann, K., Paxton, P., Pemstein, D., Seim, B., Sigman, R., Skaaning, S.-E., Staton, J., Wilson, S., Cornell, A., Gastaldi, L., Gjerløw, H., Ilchenko, N., Krusell, J., Maxwell, L., Mechkova, V., Medzihorsky, J., Pernes, J., Römer, J. v.,

- Stepanova, N., Sundström, A., Tzelgov, E., Wang, Y.-t., Wig, T., & Ziblatt, D. (2019). V-dem [country-year/country-date] dataset v9.
- Dahl, R. A., & Tufte, E. R. (1973). *Size and democracy*. Stanford: Stanford University Press.
- Diamond, L. (1992). Economic development and democracy reconsidered. *American Behavioral Scientist*, 35(4), 450–499.
- Epstein, D. L., Bates, R., Goldstone, J., Kristensen, I., & O'Halloran, S. (2006). Democratic transitions. *American Journal of Political Science*, 50(3), 551–569.
- Fish, M. S. (2002). Islam and authoritarianism. *World Politics*, 55(1), 4–37.
- Fish, M. S. (2011). *Are muslims distinctive? A look at the evidence*. Oxford: Oxford University Press.
- Fukuyama, F. (2014). *Political order and political decay*. London: Profile Books.
- Gassebner, M., Gutmann, J., & Voigt, S. (2016). When to expect a coup d'état? An extreme bounds analysis of coup determinants. *Public Choice*, 409, 1–21.
- Gassebner, M., Lamla, M. J., & Vreeland, J. R. (2013). Extreme bounds of democracy. *Journal of Conflict Resolution*, 57(2), 171–197.
- Gassebner, M., & Luechinger, S. (2011). Lock, stock, and barrel: A comprehensive assessment of the determinants of terror. *Public Choice*, 149(3–4), 235.
- Geddes, B. (1999). What do we know about democratization after twenty years? *Annual Review of Political Science*, 2, 115–144.
- Gerring, J., Hoffman, M., & Zarecki, D. (2018). The diverse effects of diversity on democracy. *British Journal of Political Science*, 48, 283–314.
- Gleditsch, K. S., & Ward, M. D. (2006). Diffusion and the international context of democratization. *International Organization*, 60(4), 911–933.
- Gouda, M., & Potrafke, N. (2016). Gender equality in muslim-majority countries. *Economic Systems*, 40(4), 683–698.
- Gründler, K., & Krieger, T. (2016). Democracy and growth: Evidence from a machine learning indicator. *European Journal of Political Economy*, 45(S), 85–107.
- Hadenius, A., & Teorell, J. (2005). *Assessing alternative indices of democracy*. *Quality of Government Institute*. Gothenburg: University of Gothenburg.
- Hariri, J. G. (2015). A contribution to the understanding of middle eastern and muslim exceptionalism. *The Journal of Politics*, 77(2), 477–490.
- Hegre, H., & Sambanis, N. (2006). Sensitivity analysis of empirical results on civil war onset. *Journal of Conflict Resolution*, 50(4), 508–535.
- Heid, B., Langer, J., & Larch, M. (2012). Income and democracy: Evidence from system gmm estimates. *Economics Letters*, 116(2), 166–169.
- Higashijima, M. (2019). *The Dictator's Dilemma at the ballot box: Electoral manipulation, economic maneuvering, and political order in autocracy*. Book manuscript.
- Honaker, J., & King, G. (2010). What to do about missing values in time-series cross-section data. *American Journal of Political Science*, 54(3), 561–581.
- Houle, C. (2009). Inequality and democracy: Why inequality harms consolidation but does not affect democratization. *World Politics*, 61(4), 589–622.
- Huntington, S. P. (1968). *Political order in changing societies*. New Haven: Yale University Press.
- Huntington, S. P. (1991). *The third wave: Democratization in the late twentieth century*. Norman: University of Oklahoma Press.
- Huntington, S. P. (1996). *The clash of civilizations and the remaking of world order*. New York: Simon and Schuster.
- Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change and democracy—The human development sequence*. Cambridge: Cambridge University Press.
- Kennedy, R. (2010). The contradiction of modernization: A conditional model of endogenous democratization. *The Journal of Politics*, 72(3), 785–798.
- Knutsen, C. H., & Nygård, H. M. (2015). Institutional characteristics and regime survival: Why are semi-democracies less durable than autocracies and democracies? *American Journal of Political Science*, 59(3), 656–670.
- Kuran, T. (2012). *The long divergence: How Islamic law held back the Middle East*. Princeton: Princeton University Press.
- Leamer, E. (1985). Sensitivity analyses would help. *American Economic Review*, 75(3), 308–313.
- Levine, R., & Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, 82(4), 942–963.
- Linz, J. J. (1990). The perils of presidentialism. *Journal of Democracy*, 1(1), 51–69.
- Lipset, S. M. (1959). Some social requisites of democracy: Economic development and political legitimacy. *American Political Science Review*, 53(1), 69–105.

- Lipset, S. M., & Lakin, J. M. (2004). *The democratic century*. Oklahoma: The University of Oklahoma Press.
- Londregan, J. B., & Poole, K. T. (1996). Does high income promote democracy? *World Politics*, 49(3), 1–30.
- Lührmann, A., Tannenberg, M., & Lindberg, S. I. (2018). Regimes of the world (row): Opening new avenues for the comparative study of political regimes. *Politics and Governance*, 6, 60–77.
- Lust, E. (2011). Missing the third wave: Islam, institutions, and democracy in the middle east. *Studies in Comparative International Development*, 46(2), 163–190.
- Marshall, M. G., Gurr, T. R., & Jaggers, K. (2018). Polity IV project: Political regime characteristics and transitions, pp. 1800–2017. Center for Systemic Peace.
- Merkel, W., & Weiffen, B. (2012). Does heterogeneity hinder democracy? *Comparative Sociology*, 11(3), 387–421.
- Moore, B. (1966). *Social origins of democracy and dictatorship: Lord and peasant in the making of the modern world*. Boston: Beacon Press.
- Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society*, 46, 69–85.
- O'Donnell, G. A. (1973). *Modernization and bureaucratic-authoritarianism: Studies in South American politics*. Berkeley: UC Berkeley Press.
- Pemstein, D., Meserve, S. A., & Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, 18(4), 426–449.
- Plümper, T. and Traunmüller, R. (2016). The sensitivity of sensitivity analysis. *Working Paper*.
- Potrafke, N. (2011). Islam and democracy. *Public Choice*, 151(1), 185–192.
- Potrafke, N. (2013). Democracy and countries with muslim majorities: A reply and update. *Public Choice*, 154(3–4), 323–332.
- Przeworski, A., Alvarez, M. E., Cheibub, J. A., & Limongi, F. (2000). *Democracy and development political institutions and well-being in the world, 1950–1990*. Cambridge: Cambridge University Press.
- Przeworski, A., & Limongi, F. (1997). Modernization: Theory and facts. *World Politics*, 49(2), 155–183.
- Rød, E. G., & Weidmann, N. B. (2015). Empowering activists or autocrats? The internet in authoritarian regimes. *Journal of Peace Research*, 52(3), 338–351.
- Ross, M. (2001). Does oil hinder democracy. *World Politics*, 53(April), 325–61.
- Rothstein, B., & Broms, R. (2013). Governing religion: The long-term effects of sacred financing. *Journal of Institutional Economics*, 9(04), 469–490.
- Rowley, C., & Smith, N. (2009). Islam's democracy paradox: Muslims claim to like democracy, so why do they have so little. *Public Choice*, 139(3), 273–299.
- Rueschemeyer, D., Stephens, E. H., & Stephens, J. D. (1992). *Capitalist development and democracy*. Chicago: University of Chicago Press.
- Sala-i-Martin, X. X. (1997). I just ran two million regressions. *American Economic Review*, 87(2), 178–183.
- Saroglou, V., Delpierre, V., & Dernelle, R. (2004). Values and religiosity: A meta-analysis of studies using schartz's model. *Personality and Individual Differences*, 37(4), 721–734.
- Stepan, A. C., & Robertson, G. B. (2003). An “arab” more than a “muslim” democracy gap. *Journal of Democracy*, 14(3), 30–44.
- Teorell, J. (2010). *Determinants of democracy: Explaining regime change in the world, 1972–2006*. Cambridge: Cambridge University Press.
- Treisman, D. (2015). Income, democracy, and leader turnover. *American Journal of Political Science*, 59(4), 927–942.
- Wasserstein, R. L., & Lazar, N. A. (2016). The asa's statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.
- World Bank (2019). World development indicators 2019.