THE QUALITY OF
GOVERNMENT INSTITUTE

UNIVERSITY OF
GOTHENBURG

What Can We Know about the Effects of Democracy Using Cross-National Data?

2023

Working paper series 2023:11

Jonathan Stavnskær Doucette

THE QUALITY OF
GOVERNMENT INSTITUTE

GÖTEBORGS
UNIVERSITET

What Can We Know about the Effects of Democracy Using Cross-National Data?

Jonathan Stavnskær Doucette

WORKING PAPER SERIES 2023:11

QoG THE QUALITY OF GOVERNMENT INSTITUTE
Department of Political Science
University of Gothenburg
Box 711, SE 405 30 GÖTEBORG
October 2023
ISSN 1653-8919

What Can We Know about the Effects of Democracy Using Cross-National Data?

Jonathan Stavnskær Doucette*

August 2023

Abstract: More than 1100 studies have been published that examine the effects of democracy using cross-national data since 2000. This research note examines whether these analyses have sufficient statistical power to detect an effect of democracy. Using Monte Carlo simulation and examining consensus effects previously reported in the literature, it finds that studies lack power to detect anything but strong, non-dynamic, and homogeneous effects of democracy when examining countries over time.

*Associate Professor, Department of Politics, Aalborg University. Email: jostdo@dps.aau.dk

Introduction

More than 1100 studies have been published that examine the effects of democracy using cross-national data since 2000 (see Gerring et al. 2022 for an overview). Despite this, it has not been established whether such analyses have sufficient statistical power to detect an effect of democracy. A lack of power can be problematic, as it implies a high probability of committing a false negative (Type II error). In addition, even statistically significant estimates risk vastly overstating effect size (Type M error) and they may also have the wrong sign (Type S error) when studies are underpowered (Arel-Bundock et al. 2022; Gelman and Carlin 2014). This research note seeks to remedy this lack of knowledge by using simulation to examine variation in the estimates for the effect of democracy. It finds that, with currently available data, analyses are only powered to detect strong, non-dynamic, and homogeneous effects of democracy.

A staggering amount of factors have been theorized to be affected by democracy. However, this note is primarily focused on economic development for several reasons: First, it is the outcome that has been examined most frequently by the literature (Gerring et al. 2022, 367; Colagrossi et al. 2020); second, there are plausible theoretical arguments for finding a substantial and positive impact of democracy (e.g. Knutsen 2012; Gerring et al. 2005; Baum and Lake 2003); and third, data on GDP per capita is availability for more countries and for longer time-spans than most other outcomes. Thus, if there is not sufficient power to detect an effect of democracy on economic development, then it is unlikely that an effect on other outcomes can be analyzed.

I find, using the most extensive data available on democracy and GDP per capita, that analyses are powered to detect a direct effect of democracy (81% power, $\alpha = 0.05$). However, this result is very sensitive to effect size and the number of countries in the analysis. If the effect of democracy is slightly weaker than reported in, for instance, Acemoglu et al. (2019), studies are underpowered. Moreover, if data is missing for a few countries (for datasets of approximately 155 countries or less), power is lacking. This implies that statistical power quickly becomes an issue if one studies outcomes that are less strongly connected to democratization (such as inequality or civil war, e.g. Leipziger 2023; Cederman et al. 2010). In support of this, I show that studies are severely underpowered to detect the effect of democracy on civil war onset even when using the most extensive datasets available.

Note also that the above results assumes non-dynamic and homogeneous effects. If effects change over time after democratization (and thus are dynamic), the use of TWFE may not be appropriate, and the power required to detect an effect using alternative more appropriate estimators increases (see Chiu et al. 2023). This is also the case when studying whether the effect of democracy vary across groups. Thus, the above findings are for a best-case scenario. Assuming that the effect of democratization is dynamic[1], I show that studies are only powered to detect long-run effects when using event-study estimators to capture the dynamic relationship. Assuming that the effect of democratization vary across groups, I find that studies lack the power to detect even very large differences in effect size across groups (and this is when assuming non-dynamic effects).

This has implications for the study of democracy but most likely also for the analysis of other similar cross-national analyses examining factors such as state capacity, civil society, or political party institutionalization (e.g. Andersen and Doucette 2022; Hegre et al. 2020; Bizarro et al. 2018). Scholars should be cautious when interpreting the magnitude of effects found in cross-national analyses of the effects of democracy (and other similar factors) and recognize the uncertainty inherent in such estimates. In addition, one suggestion could be that scholars identify additional implications of the theory that can also be tested. If the pattern is similar across outcomes, it raises confidence in the results. Moreover, one can identify instances where there is sub-national variation in democratization (or at least on the theoretically relevant component of democracy) which can supplement the cross-national analysis (see, for instance, Lankina and Getachew 2012; Grumbach 2023 for data examples).

[1]As suggested by, for example, Figures 1-5 from Acemoglu et al. 2019

Simulation approach

I assess the performance of statistical significance tests based on a panel of countries (i) obersved in different years (t). I adopt the standard power threshold of 80% with a significance level of 0.05 (α = 0.05). An often used approach in the literature is a linear regression of $Ln(GDP/cap)_{it}$ on $Democracy_{it}$ and country and year fixed effects (γ_i,δ_t). This is also termed the Two-Way Fixed Effects (TWFE) estimator. The tests are based on standard errors that cluster on countries. I assume no difference in pre-trends between autocracies and future democracies. I summarize the specification as:

$$Ln(GDP/cap)_{it} = \gamma_i + \delta_t + \beta Democracy_{it} + \epsilon_{it} \quad (1)$$

In the baseline model, β is assumed to be equal to 0.15. This corresponds roughly to the average effect estimate from Acemoglu et al. (2019, 58-59), which represent an updated appraisal of the effect of democracy on economic development.[2] The number of countries in the sample is 185 (C), which corresponds to the observed number of countries with data on both the democracy and the GDP per capita variable in at least one year. ϵ_{it} captures other time-variant factors that affect a country's economic development. I evaluate the variability of $\hat{\beta}$ as follows.

I simulate the steps outlined below 1000 times and save the $\hat{\beta}$ from each repetition:

1. Construct a panel dataset of countries observed from 1789 to 2015

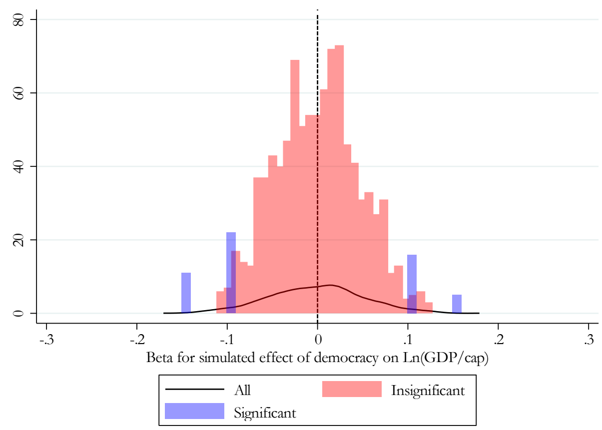
[2]Their estimated effect is dynamic and grows over time. According to their estimates, GDP per capita grows by around 1 percent per year after democratization, and around 20 years after the transition it remains around 20 percent higher. To get an average effect estimate, I first calculate the assumed effect size in each observed country-year observation where a country in the data is democratic ($years - democratic_{it} * 0.01$ if $years - democratic_{it} < 21$ and $democracy_{it} = 1$, and 0.2 if $years - democratic > 20$). Next, I average over the observed effect sizes and get an estimate of 0.15. This corresponds to a Cohen's D of 0.18 (or 0.47 when using the leftover variation in the outcome once country and year fixed effects are partialled out). Note that this is larger than the average effect size reported in Colagrossi et al. 2020 which is 0.12 (see Table 6).

2. Assign logged GDP per capita-year series to each country based on data from Fariss et al. 2022
3. Randomly assign democracy-year series to each country based on data from Boix et al. 2013
4. Extract C number of countries randomly from the dataset and discard the rest
5. Multiply $Ln(GDP/cap)_{it}$ by 0.15 in years where $Democracy_{it}$ is equal to 1
6. Estimate $Ln(GDP/cap)_{it} = \gamma_i + \delta_t + \beta Democracy_{it} + \epsilon_{it}$
7. save $\hat{\beta}$

Note that steps 2 and 3 ensures that $Ln(GDP/cap)_{it}$ and $Democracy_{it}$ are uncorrelated in expectation.[3] Thus, without step 5, estimates of β should center around 0 if the TWFE estimator is unbiased in this case. Figure 1 confirms that this is the case. Step 4 allows me to vary the number of units used in the estimation. As a result, it is possible to assess how the power requirements change as a function of the number of countries that are included in the sample.

[3]Step 2 and 3 are repeated with replacement, thus ensuring that it is possible to simulate a larger dataset of countries than is actually available.

Figure 1: Distribution of estimates when $\beta = 0$



Note: The dashed black line corresponds to the population β. C is set at 185.

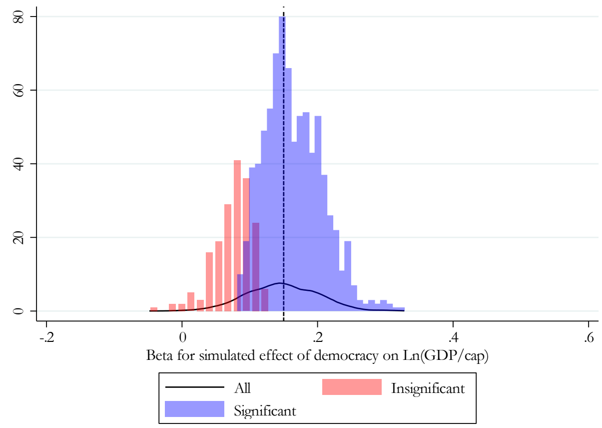
Findings

I now evaluate how the power requirements of this approach vary as a function of i) effect size, ii) the number of countries in the dataset, and iii) the presence of dynamic or heterogenous effects. These represent common differences between studies of the effect of democracy, as i) some outcomes are more loosely connected to democracy (e.g. Leipziger 2023; Paglayan 2021), ii) occasionally outcome data is only available for some countries or periods (e.g. Stasavage 2005), and iii) in some cases the effect changes as democracies age or vary across groups (e.g. Paglayan 2021; Acemoglu et al. 2019).

Figure 2 reports the results for the baseline specification where the assumed effect of democracy is 0.15. A majority of estimates (approximately 81%) are positive and significant. At a first glance this may not appear to pose a problem. However, this is marginally above the standard 80% power threshold. Thus, studies are only just powered to detect this effect, and, as shown below, this is very sensitive to effect size, the availability of data, and the type of effect one expects.

Figure 3 plots how this changes as a function of effect size. Looking at the upper right

Figure 2: Distribution of estimates when $\beta = 0.15$



Note: The dashed black line corresponds to the population β. C is set at 185.

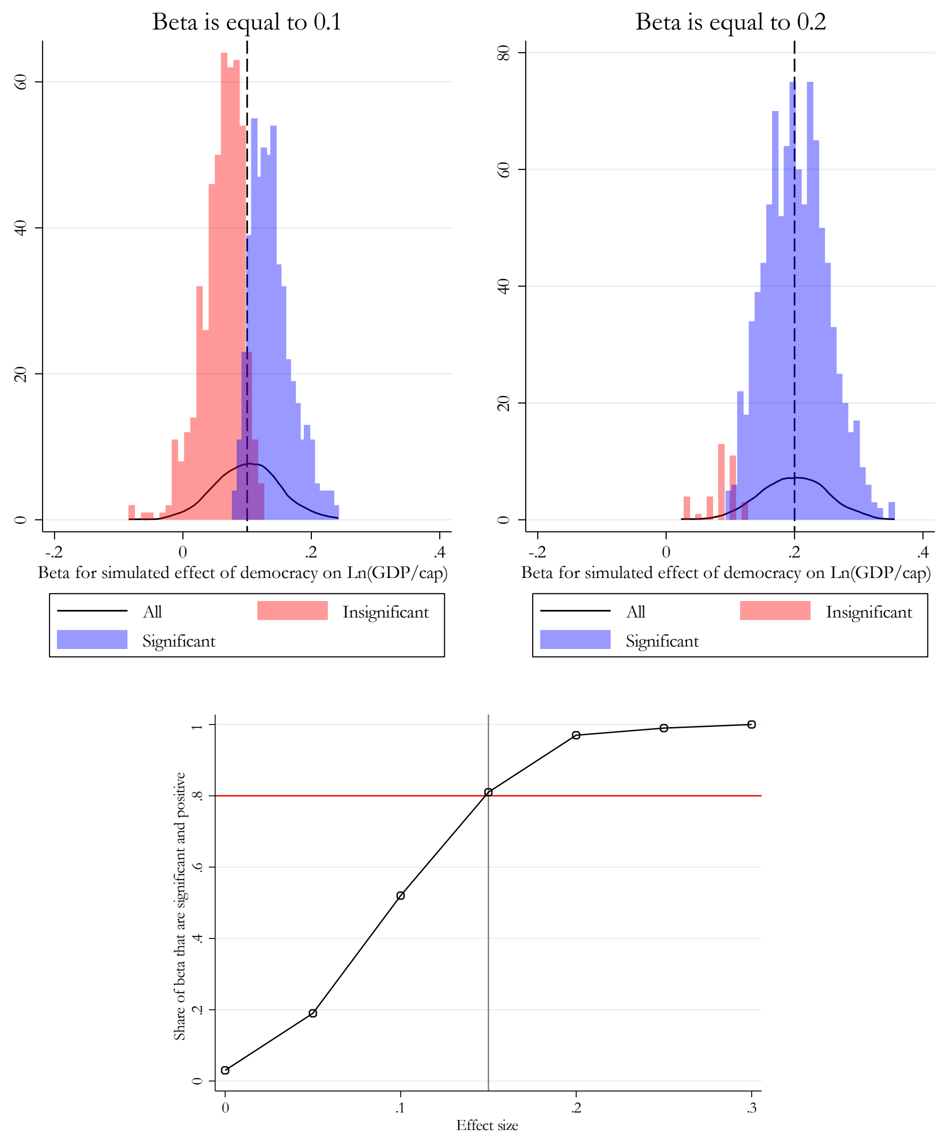
graph, almost all estimates are positive and significant if the population $\beta = 0.2$. Thus, it is unproblematic to detect very strong effects of democracy. However, when looking at relationships where the effect of democracy is slightly smaller as in the upper left graph (β = 0.1), almost half of the estimates become insignificant. Given that economic development is one of the outcomes where we have good theoretical reasons to expect an effect of democracy, it is likely that analyses of many other outcomes risk being underpowered.

Figure 4 shows the relationship between the number of countries in the dataset and the share of coefficients that are significant when the effect in the population is $\beta = 0.15$. A similar pattern appears here as the number of countries in the data only has to drop slightly for the power level to drop below conventional levels.[4] Notice also that the variability of the estimates increases substantially when only 100 countries are included in the data (the standard deviation of the estimate increases from 0.05 in the baseline simulation with 185 countries to 0.072, a 44% increase, with 100 countries).

If an effect of democracy on economic development exist, it is likely to be dynamic and growing over time (see Acemoglu et al. 2019). A common approach to modelling this are

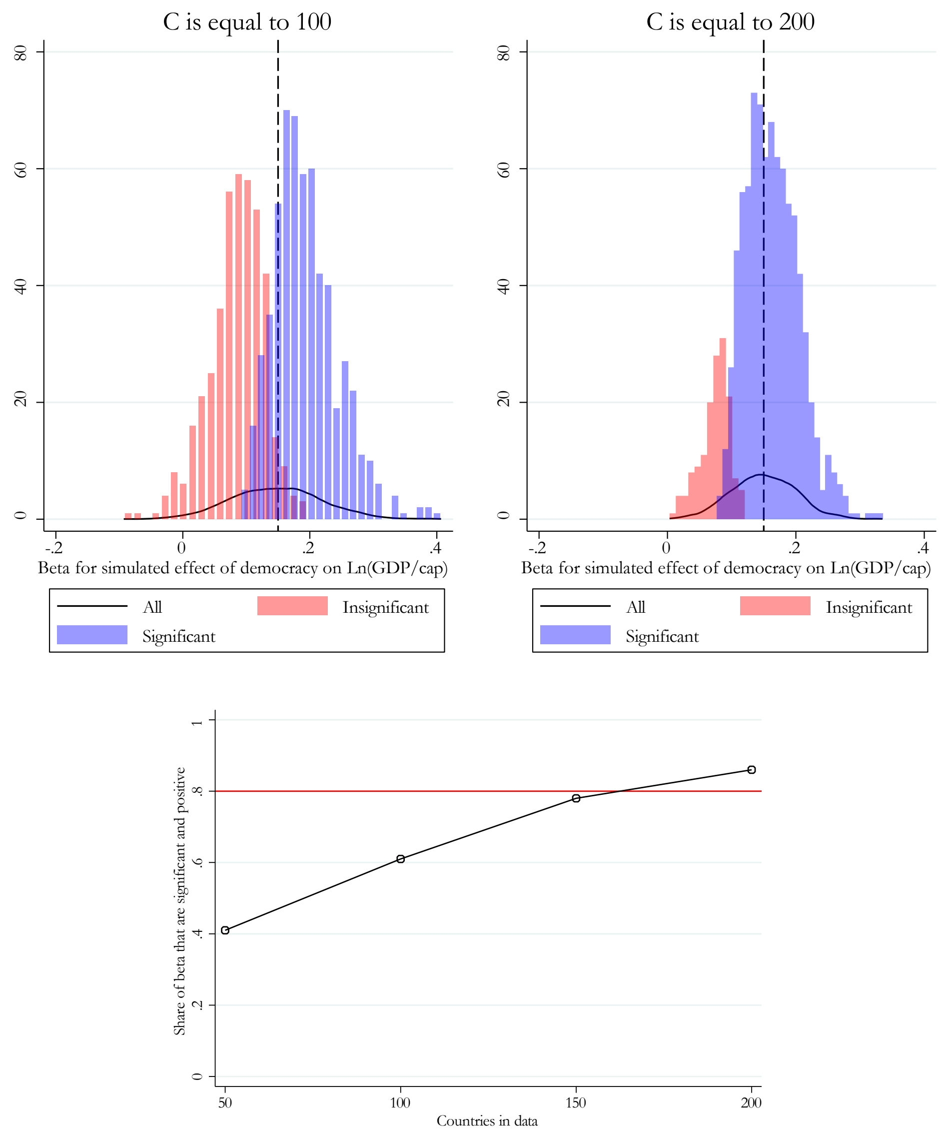
[4]The scenario with additional countries is of less interest here, as the number of countries in current datasets are probably unlikely to increase by much in the future.

Figure 3: Effect size and the share of significant and positive estimates



Note: C is set at 185. The vertical dashed black line in the upper graphs corresponds to the population β. The horizontal red line in the lower graph shows the conventional power threshold.

Figure 4: The number of countries and the share of significant and positive estimates



Note: β is set at 0.15. The vertical dashed black line in the upper graphs corresponds to the population β. The horizontal red line in the lower graph shows the conventional power threshold.

event-studies that include dummies for the relative years prior to and after democratization in addition to country- and year-fixed-effects (often excluding a dummy for the year just before democracy is introduced). How does this alter the power requirements? As before, I use the estimates from Acemoglu et al. (2019), which indicate that GDP per capita grows after democratization in comparison with autocracies until about 20 years after democratization. At this point, democracies remain about 20% richer than autocracies. To simulate this, I assume that β in the population grows by 0.01 each year after democratization and plateaus at 0.2 after 20 years[5]. Figure 5 presents the distribution of estimates for the over time effect of democratization. It indicates that studies are generally underpowered to detect the dynamic effect of democracy from 1 to 15 years after democratization. However, the power level does reach the standard threshold around 15 years after democratization. Thus, it may be possible to recover long-run effects in this case.

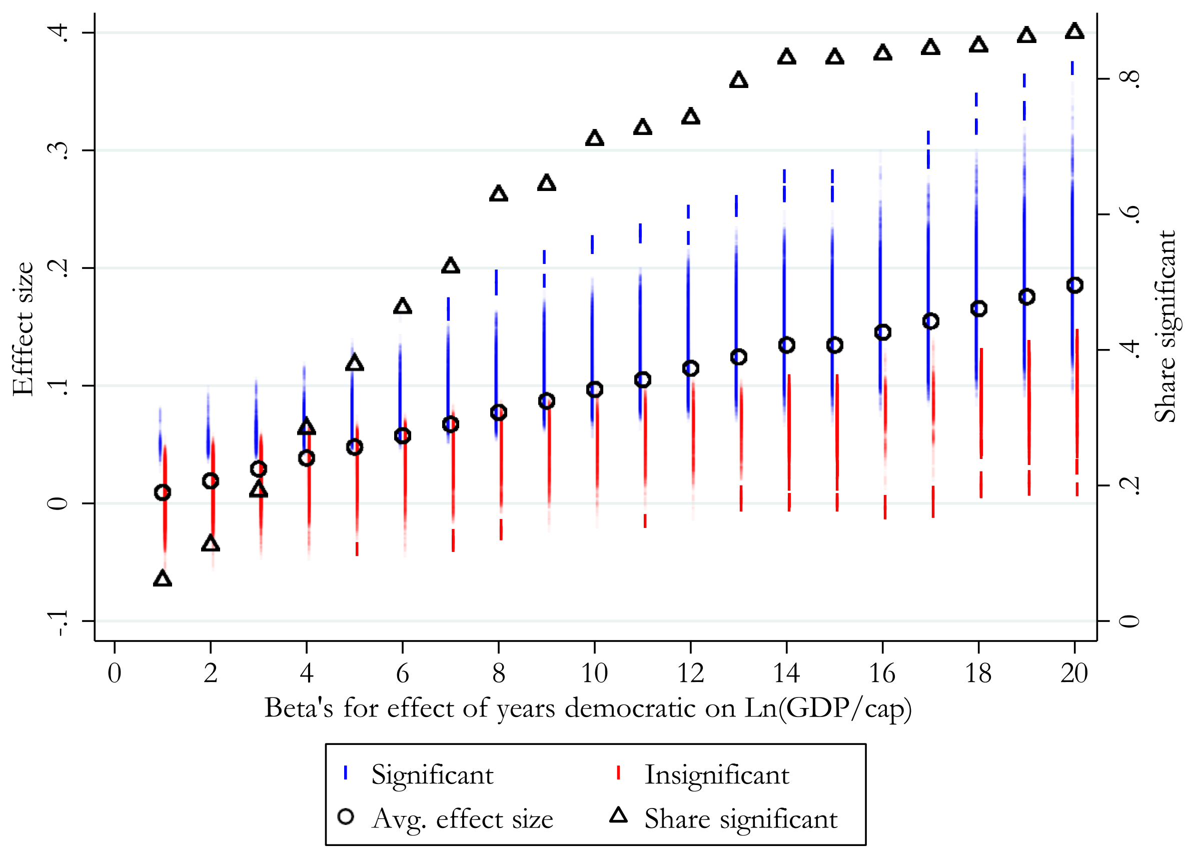
The effect of democracy may differ across groups, and as such, we might be interesting in estimating this. Thus, I randomly assign countries into two groups and vary the size of the effect of democracy within each group across three scenarios based on the size of the difference in effect size between groups[6]. To capture this, I include an interaction term between the democracy indicator and the group indicator in the baseline TWFE model. Figure 6 plots the simulated distribution of estimates for the interaction term. Even when effects are very heterogeneous, we lack the power to detect a difference in the effect of democracy. Therefore, we are unlikely to have sufficient power to detect potential heterogenous effects of democracy.

In the Online Appendix, I show that these results are consistent across different choices available to researchers analyzing the effects of democracy. First, I find a similar pattern (see Figure A1) when using an interval-scaled measure of democracy using the *v2x˙polyarchy* variable from V-Dem (Coppedge et al. 2022). In addition, I analyze an alternative outcome – civil war onset – that has received substantial attention in the democratization literature

[5]This corresponds to the results of Acemoglu et al. (2019, 58)

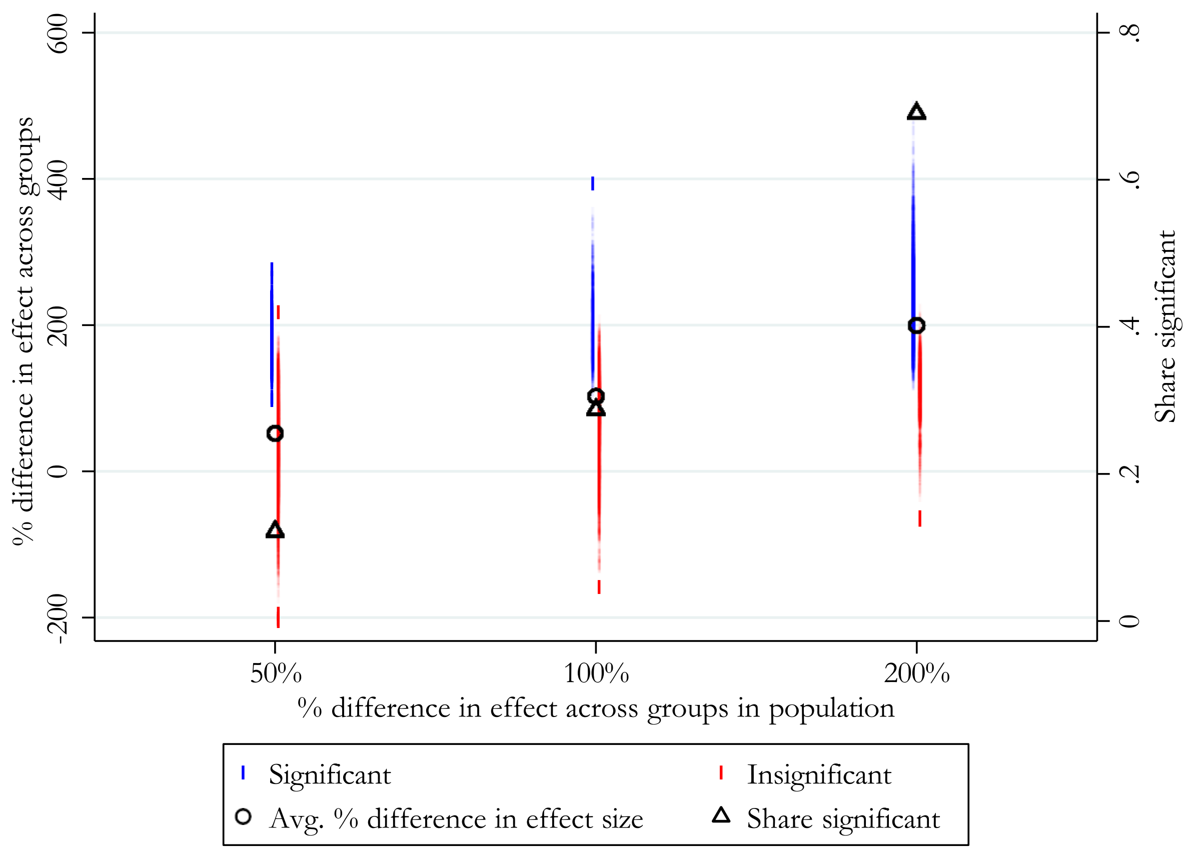
[6]I) The high heterogeneity scenario with a null effect in one group and double the effect in the other group (corresponding to a 0.3 difference in effect size between groups), ii) the medium heterogeneity scenario with a 0.15 difference in effect size between groups, and iii) the low heterogeneity scenario with a 0.075 difference in effect size between groups.

Figure 5: Dynamic effects



Note: β in the population grows by 0.01 each year after democratization and plateaus at 0.2 after 20 years based on Acemoglu et al. 2019. I assume no difference in pre-trends between autocracies and future democracies. The left y-axis shows the distribution of effect sizes, while the right y-axis shows the share of effect estimates that are positive and significant.

Figure 6: Heterogeneous effects



Note: β is assumed to vary by Z_i. The average β is 0.15. The TWFE models thus include an interaction term ($\beta_2(democracy_{it} \times Z_i)$) in addition to the term for democracy ($\beta_1 democracy_{it}$). Z_i is absorbed by country fixed effects. The left y-axis shows the distribution of estimates for β_2, while the right y-axis shows the share of interaction term estimates that are significant.

(see the ”Democracy and civil war” section). The theoretical arguments are less one-sided for civil war and the findings are mixed. As a result, it reflects less ideal conditions for finding an effect of democracy. Based on Bartusevicius and Skaaning (2018), I assume a population β of -0.01 (in a linear probability model). I find that the lack of power is much more severe in this case, as only around 40% of estimates are significant and in the right direction.

Discussion

Taken together these results suggest that analyses are only powered to detect fairly strong direct effects of democracy. Thus, the absence of an effect of democracy for an outcome can in many cases not be considered definitive proof that democracy had no effect on that outcome. Reversely, given the variability of the estimates and their sensitivity to the number of countries included and effect size, it is prudent to be cautious when interpreting effect sizes even when finding a significant effect as this might simply reflect noise. This is further warranted as the effect of democracy is likely to be both heterogeneous and dynamic in many cases (e.g. Leipziger 2023; Acemoglu et al. 2019), which further exacerbating power issues.

What can be done about this? Ideally, it would be advisable to conduct power-analyses at the pre-analysis stage across a set of plausible effect sizes. Yet, in many cases the relationship has been studied extensively by others. First, one should therefore be cautious when interpreting the magnitude of effects found in cross-national analyses of the effects of democracy and recognize the uncertainty inherent in such estimates. In addition, one could supplement the analysis of the effects of democracy on an outcome by identifying additional implications of the theory that can also be tested. If the pattern is similar across outcomes, it raises confidence in the results. Moreover, one can identify instances where there is sub-national variation in democratization (or at least on the theoretically relevant component of democracy) which can supplement the cross-national analysis (see, for instance, Lankina and Getachew 2012; Grumbach 2023 for data examples).

References

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

Andersen, David and Jonathan Doucette (2022) "State First? A Disaggregation and Empirical Interrogation." *British Journal of Political Science* 52(1): 408–415.

Arel-Bundock, Vincent , Ryan Briggs, Hristos Doucouliagos, Marco Avina, and T. Stanley (2022) "Quantitative Political Science Research is Greatly Underpowered." *I4R Dicussion Paper* (6).

Bartusevicius, Henrikas and Svend-Erik Skaaning (2018) "Revisiting Democratic Civil Peace: Electoral Regimes and Civil Conflict." *Journal of Peace Research* 55(5): 625–640.

Baum, Matthew and David Lake (2003) "The Political Economy of Growth: Democracy and Human Capital." *American Journal of Political Science* 47(2): 333–347.

Bizarro, Fernando , John Gerring, Carl Henrik Knutsen, Allen Hicken, Michael Bernhard, Svend-Erik Skaaning, Michael Coppedge, and Staffan Lindberg (2018) "Party Strength and Economic Growth." *World Politics* 70(2): 275–320.

Boix, Carles , Michael Miller, and Sebastian Rosato (2013) "A Complete Dataset of Political Regimes, 1800-2007." *Comparative Political Studies* 46(12): 1523–54.

Cederman, Lars-Erik , Simon Hug, and Lutz Krebs (2010) "Democratization and Civil War: Empirical Evidence." *Journal of Peace Research* 47(4): 377–394.

Chiu, Albert , Xingchen Lan, Ziyi Liu, and Yiqing Xu (2023) "What to Do (and Not to Do) with Causal Panel Analysis under Parallel Trends: Lessons from A Large Reanalysis Study." *SSRN working paper Retrieved from papers.ssrn.com/sol3/papers.cfm?abstract˙id=4490035.*

Colagrossi, Marco , Domenico Rossignoli, and Mario Maggioni (2020) “Does Democracy Cause Growth? A Meta-Analysis (of 2000 regressions).” *European Journal of Political Economy* 61(101824).

Coppedge, Michael , John Gerring, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Teorell, David Altman, Frida Andersson, Michael Bernhard, M. Steven Fish, Adam Glynn, Allen Hicken, Carl Henrik Knutsen, Kelly McMann, Valeriya Mechkova, Farhad Miri, Pamela Paxton, Daniel Pemstein, Rachel Sigman, Jeffrey Staton, and Brigitte Zimmerman (2022) “V-dem Codebook.” *Varieties of Democracy Project v11*.

Fariss, C. J. , T. Anders, J. N. Markowitz, and M. Barnum (2022) “New Estimates of Over 500 Years of Historic GDP and Population Data.” *Journal of Conflict Resolution* 66(3): 553–591.

Gelman, Andrew and John Carlin (2014) “Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors.” *Perspectives on Psychological Science* 9(6): 641–651.

Gerring, John , Philip Bond, William Barndt, and Carola Moreno (2005) “Democracy and Economic Growth: A Historical Perspective.” *World Politics 57*: 323–64.

Gerring, John , Carl Henrik Knutsen, and Jonas Berge (2022) “Does Democracy Matter?.” *Annual Review of Political Science* 25(357).

Grumbach, J. (2023) “Laboratories of Democratic Backsliding.” *American Political Science Review* 117(3): 967–984.

Hegre, Håvard , Michael Bernhard, and Jan Teorell (2020) “Civil Society and the Democratic Peace.” *Journal of Conflict Resolution* 64(1): 32–62.

Knutsen, Carl Henrik (2012) “Democracy and Economic Growth: A Review of Arguments and Results.” *International Area Studies Review* 15(4): 393–415.

Lankina, Tomila and Lullit Getachew (2012) “Misison or Empire, Word or Sword? The Human Capital Legacy in Postcolonial Democratic Development.” *American Journal of Political Science* 56(2): 465–483.

Leipziger, Lasse E. (2023) “Does Democracy Reduce Ethnic Inequality?.” *accpeted for publication in American Journal of Political Science.*

Paglayan, Agustina (2021) “The Non-Democratic Roots of Mass Education: Evidence from 200 Years.” *American Journal of Political Science* 115(1): 179–198.

Stasavage, David (2005) “Democracy and Education Spending in Africa.” *American Journal of Political Science* 49(2): 343.358.

Online Appendix for ”What Can We Know about the Effects of Democracy Using Cross-National Data?”

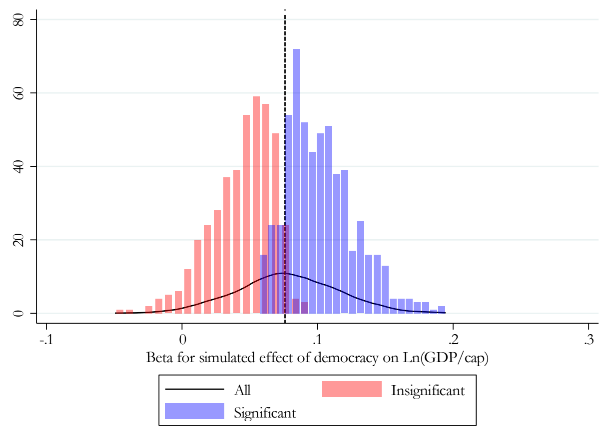
Contents

Results using V-Dem's Polyarchy variable

To ensure consistent effect size across the two measures of democracy, I first convert the baseline effect size for the dichotomous democracy (Cohen's d of 0.18) to a similarly sized effects for a continuous measure (Pearson's R=0.09).[7]. Thus, a standard deviation increase on the polyarchy measure should cause a 0.09 standard deviation increase (0.076) in logged GDP per capita. I therefore standardize the polyarchy variable, and set β = 0.076. Figure A1 shows the results when using the V-Dem measure. Similar power issues persist when using the polyarchy measure of democracy.

[7] Baed on $r = d/\sqrt{d^2 + 4}$, where r is Pearson's R, and d is Cohen's d

Figure A1: Distribution of estimates when β = 0.076



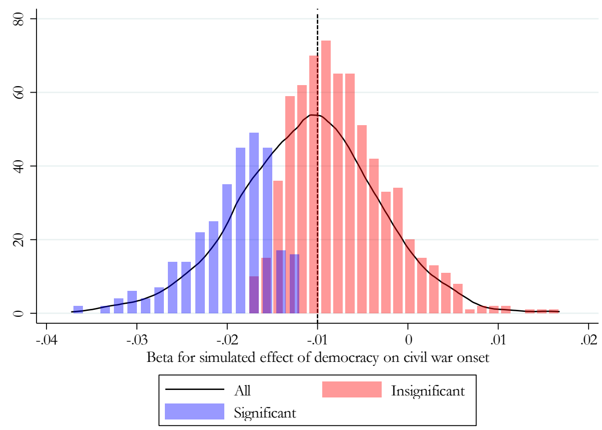
Note: The dashed black line corresponds to the population β. C is set at 185.

Democracy and civil war

According to Bartusevicius and Skaaning (2018), democracy reduces the risk of civil war onset by about 1% percentage point. However, this effect is subject to considerable theoretical dispute, and the expected relationship is not as clear as for economic development. I set the population β to -0.01, as I am estimating a linear probability model. I use a similar simulation approach as earlier changing my outcome to a measure of civil war onset (data on civil conflict from Bartusevicius and Skaaning 2018).[8] Figure A2 show the main findings. Even with an extensive dataset, studies lack power to detect an effect of democracy on civil war onset. Less than half of all estimates are significant even when the population β is equal to -0.01. Figures A3 and A4 further indicate that this is not mitigated even when the effect is substantially larger or the number of countries is increased.

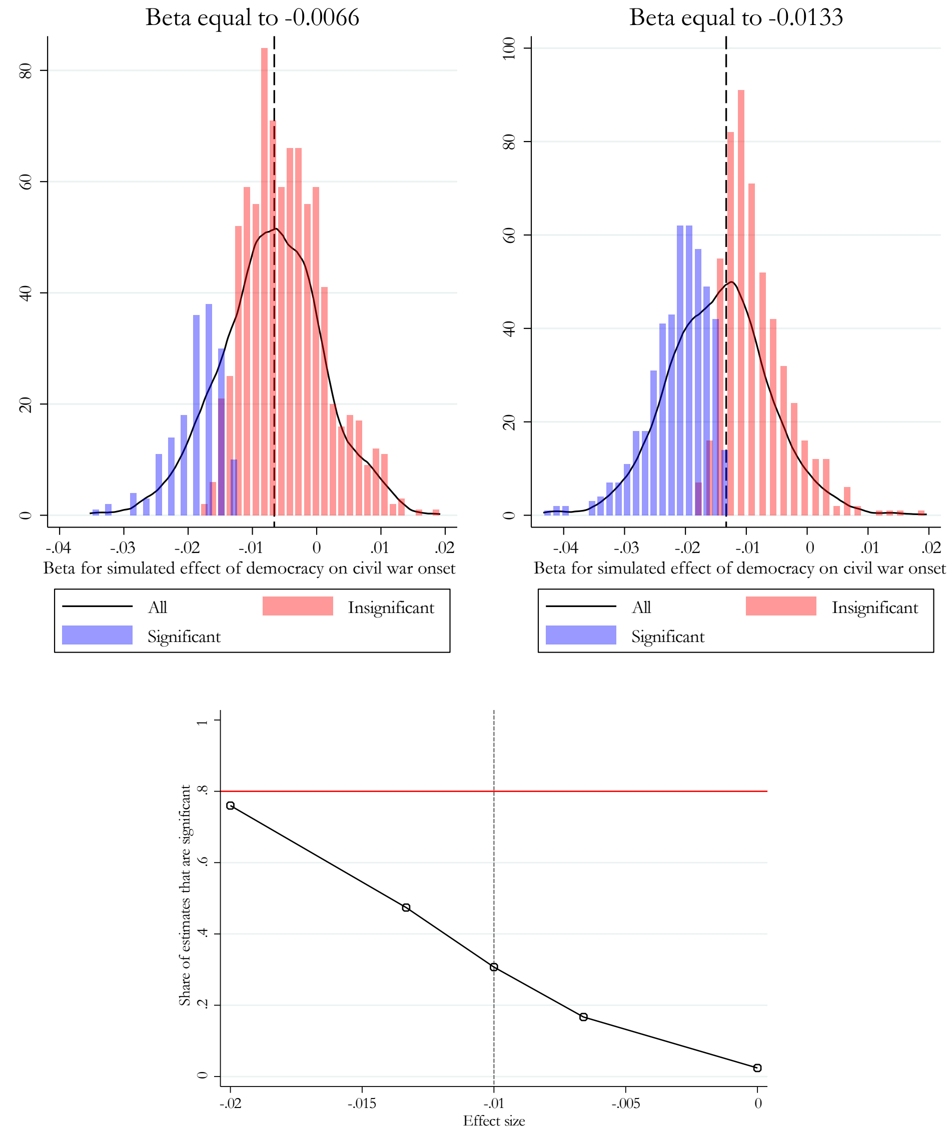
[8]Step 5 has been modified slightly as the outcome is now a dummy. Instead I now create an additional variable that consists of draws from a binomial distribution with a probability of 0.01. Next, I change 0 to 1 in autocracy-years where this additional variable is equal to 1.

Figure A2: Distribution of estimates when $\beta = -0.01$



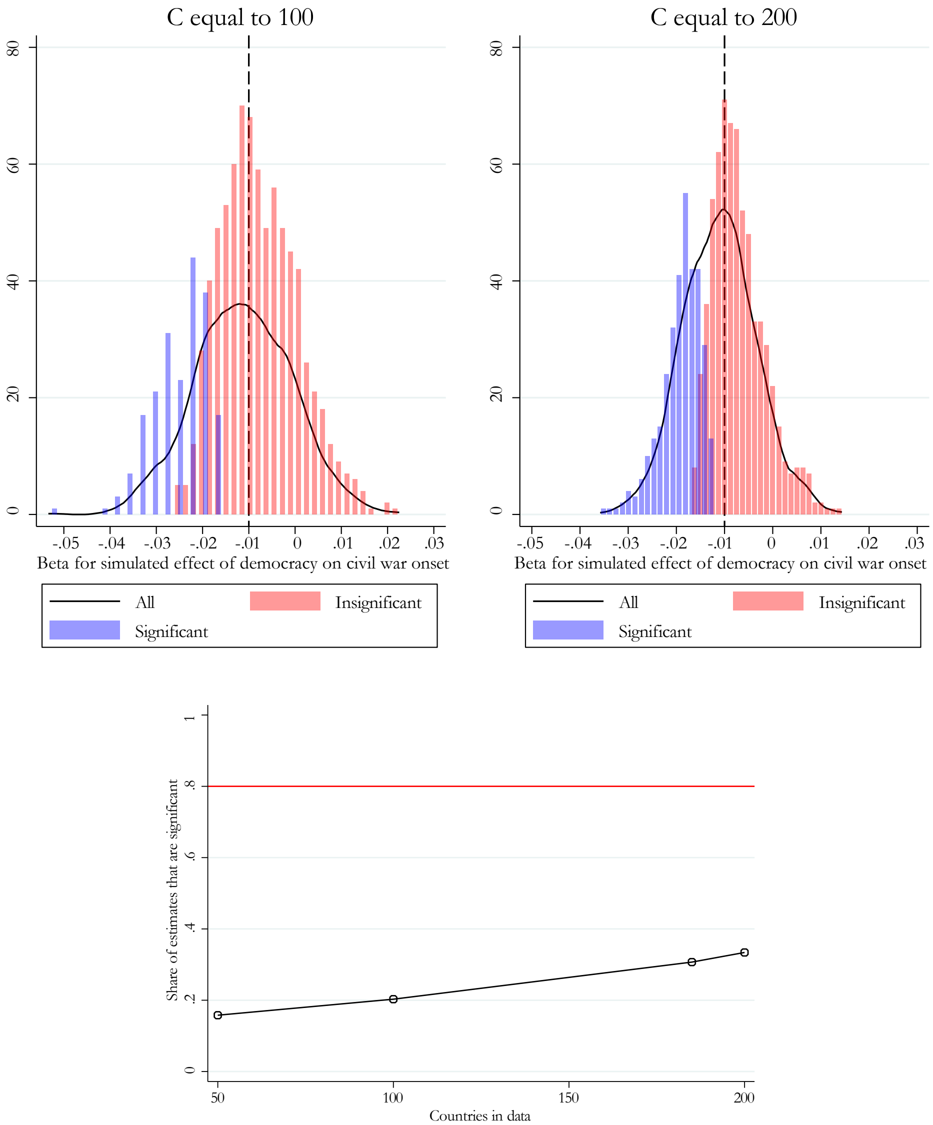
Note: The dashed black line corresponds to the population β. C is set at 185.

Figure A3: Effect size and power



Note: C is set at 185. The vertical dashed black line in the upper graphs corresponds to the population β. The horizontal red line in the lower graph shows the conventional power threshold.

Figure A4: The number of countries and power



Note: β is set at -0.01. The vertical dashed black line in the upper graphs corresponds to the population β. The horizontal red line in the lower graph shows the conventional power threshold.

References

Bartusevicius, Henrikas and Svend-Erik Skaaning (2018) “Revisiting Democratic Civil Peace: Electoral Regimes and Civil Conflict.” *Journal of Peace Research* 55(5): 625–640.