

Gründler, Klaus; Krieger, Tommy

**Working Paper**

## Should We Care (More) About Data Aggregation? Evidence from the Democracy-Growth-Nexus.

CESifo Working Paper, No. 7480

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Gründler, Klaus; Krieger, Tommy (2019) : Should We Care (More) About Data Aggregation? Evidence from the Democracy-Growth-Nexus., CESifo Working Paper, No. 7480, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/198840>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

**Should We Care (More) About  
Data Aggregation? Evidence  
from the Democracy-Growth-  
Nexus.**

*Klaus Gründler, Tommy Krieger*

## **Impressum:**

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email [office@cesifo.de](mailto:office@cesifo.de)

Editor: Clemens Fuest

[www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Should We Care (More) About Data Aggregation? Evidence from the Democracy-Growth-Nexus.

## Abstract

We compile data for 186 countries (1919 - 2016) and apply different aggregation methods to create new democracy indices. We observe that most of the available aggregation techniques produce indices that are often too favorable for autocratic regimes and too unfavorable for democratic regimes. The sole exception is a machine learning technique. Using a stylized model, we show that applying an index with implausibly low (high) scores for democracies (autocracies) in a regression analysis produces upward-biased OLS and 2SLS estimates. The results of an analysis of the effect of democracy on economic growth show that the distortions in the OLS and 2SLS estimates are substantial. Our findings imply that commonly used indices are not well suited for empirical purposes.

JEL-Codes: C260, C430, O100, P160, P480.

Keywords: data aggregation, democracy, economic growth, indices, institutions, machine learning, measurement of democracy, non-random measurement error.

*Klaus Gründler*  
Ifo Institute – Leibniz Institute for Economic  
Research at the University of Munich  
Poschingerstrasse 5  
Germany – 81679 Munich  
gruendler@ifo.de

*Tommy Krieger*  
University of Konstanz  
Department of Economics  
Universitätsstraße 10  
Germany – 78464 Konstanz  
tommy.krieger@uni-konstanz.de

January 22, 2019

We are grateful for the constructive comments from Toke Aidt, Norbert Berthold, Richard Bluhm, Markus Brueckner, Martin Gassebner, Kai Gehring, Jerg Gutmann, Raphael Franck, Stephan Maurer, Niklas Potrafke, Guido Schwerdt, Uwe Sunde, and Heinrich Ursprung. We would also like to thank seminar participants at the Universities of Cambridge and Konstanz, and conference participants at the annual meeting of the Public Choice Society in New Orleans, the annual meeting of the European Public Choice Society in Budapest, the Spring Meeting of Young Economists in Halle, the Silvaplane Workshop in Political Economy, the CESifo Political Economy Workshop in Dresden, and the annual meeting of the Verein für Socialpolitik in Vienna. Finally, we are highly appreciative of the financial support from the Graduate School of Decision Science (University of Konstanz).

# 1 Introduction

The view that authoritarian institutions are more conducive to economic growth than liberal-democratic institutions has recently gained popularity among some voters and politicians (Plattner, 2015, Wike et al., 2017). The Hungarian Prime Minister, Victor Orbán, for example, argued in his speech on July 26, 2014, that:

*“The defining aspect of today’s world can be articulated as a race to figure out a way of organizing communities and to find the state that is most capable of making a nation competitive. This is why, [...] a trending topic in thinking is understanding systems that are not Western, not liberal, not liberal democracies, maybe not even democracies, and yet making nations successful.”*

and attracted global attention by concluding that:

*“What all this exactly means [...] we have to abandon liberal methods and principles of organizing a society, as well as the liberal way to look at the world.”*<sup>1</sup>

Early empirical studies on the effect of democracy on economic growth support the notion that autocratic regimes grow faster than democratic regimes and explain their result with differences in tax policies (Barro, 1996, Tavares and Wacziarg, 2001). Some recent empirical studies challenge this pessimistic view and report results showing that democracy causes long-run growth by increasing human capital and the quality of the economic institutions (Acemoglu et al., 2018, Gründler and Krieger, 2016, Papaioannou and Siourounis, 2008, Persson and Tabellini, 2006, 2008). Knutsen (2012) argues that these recent studies are more sophisticated because of their empirical methods. This paper shows that only improving the identification strategy does not suffice to allay all endogeneity concerns, and thereby raises doubts about the precision of the estimates reported in these studies.

Any empirical study that examines the consequences of a transition from autocracy towards democracy (or vice versa) needs an index that measures the degree to which a regime satisfies democratic principles. Since the literature provides multiple democracy

---

<sup>1</sup>For an English version of the speech, see <https://hungarianspectrum.wordpress.com/2014/07/31/>.

indices, we examined in an earlier study whether the estimated effect of democracy on economic growth depends on the measure of democracy (Gründler and Krieger, 2016). Our results show that replacing a democracy indicator with another indicator creates notable changes in the size of the regression coefficient. However, because of a lack of data, exploring where these differences in the effect size come from was impossible.<sup>2</sup>

To address this question, we compile novel data on political participation, political competition, and the freedom of speech (186 countries, 1919 – 2016) and ask how these regime characteristics can be transformed into an index. Since generally acknowledged guidelines for data aggregation do not exist, we apply different aggregation techniques and compare their performance.<sup>3</sup> We observe that most of the available aggregation methods create indices that are often too favorable for autocracies and too unfavorable for democracies. The sole aggregation method that does not produce these malfunctions is a machine learning technique.

We develop a theoretical model to examine the empirical consequences of using an aggregation method that produces implausibly low (high) indices for pure democracies (autocracies). We find that applying a biased index in a regression analysis causes an upward bias in the ordinary least squares (OLS) estimator. Our model also shows that having an instrument for the degree of democratization does not suffice to correct this upward bias. This result is important because many scholars working on the effects of political transitions claim that their instrumental variable approach allays any concerns about measurement error in the democracy index.

Using our new democracy indices, we estimate the effect of democracy on economic growth to investigate whether the overestimation biases predicted by our model are of practical relevance. We report results from OLS fixed effect and instrumental variable regressions showing that democracy positively affects economic growth. In line with the prediction of our stylized model, we observe that applying an index that is biased for

---

<sup>2</sup>In advance, it is unclear whether the differences in the effect size pose a problem. If, for example, the differences in the effect size are caused by differences in the concept of democracy, we do not have to worry because we can develop an economic theory explaining why using different concepts leads to different results about the effect of democracy on economic growth. If, however, some indices are just more error-prone than other indices, we must be concerned about the different effect sizes because then they may only exist since some indices are unsuited for empirical purposes. Below, we argue that the second case applies.

<sup>3</sup>Our list of aggregation techniques includes: an additive method, a multiplicative method, an additive-multiplicative method, a Bayesian latent variable method, and a machine learning method. We select these five techniques because other scholars used them to create measures of democracy.

clearly autocratic regimes or clearly democratic regimes causes a significant increase in the estimated effect of democracy on economic growth.

This paper addresses economists and political scientists interested in the causes and consequences of political transitions.<sup>4</sup> Our results suggest that widely used democracy indicators—such as the Polity index, the Freedom House indices, the Vanhanen (2000) index, or the indices proposed by Pemstein et al. (2010) and Márquez (2016)—are not well suited for empirical analyses. The reason is that the aggregation methods used for the creation of these indices cause non-random measurement errors. We also present a machine learning index that is not prone to non-random measurement errors and thus more suitable for research purposes.

Our results are also relevant for other applied economists and social scientists since indices belong to the standard tool kit in all humanities. We are convinced that most problems caused by the application of simple data aggregation methods arise in various situations and that scholars should thus be cautious in using simple indices for causal identification.

We organize the paper in the following way. Section 2 presents the basic ingredients of democracy indices and compares the performance of different aggregation methods. Section 3 shows why data aggregation matters for identifying the influence of political regimes. Section 5 concludes.

## 2 Measuring democracy

We use the usual three-stage procedure to measure democracy (Munck and Verkuilen, 2002): first, we define the term “democracy” (*conceptualization*), then collect the data reflecting the components of the definition (*operationalization*), and finally specify the rule that transforms the raw data into an index (*aggregation*).

---

<sup>4</sup>For studies examining how economic development and economic shocks influence democratization, see Acemoglu et al. (2008), Aidt and Franck (2015), Aidt and Leon (2016), Brückner and Ciccone (2011), Cervellati et al. (2014), Gundlach and Paldam (2009), Lipset (1959), Murtin and Wacziarg (2014), and Przeworski (2000).

## 2.1 Conceptualization

The literature distinguishes between narrow (thin, minimalist), realistic (balanced), and broad (thick, maximalist) definitions of democracy (O’Donnell, 2001). Narrow concepts are focused on whether public elections for political mandates are competitive. Realistic concepts also require universal suffrage and basic political rights, while broad concepts also incorporate a wide range of socioeconomic conditions.

Munck and Verkuilen (2002) explain why neither a narrow nor a broad definition of democracy is well suited for analytical purposes. Broad concepts are difficult to operationalize because of insufficient data availability and overlap with other economic concepts (e.g. corruption, economic freedom, inequality), whereas narrow concepts do not sufficiently differentiate between autocratic, democratic, and hybrid regimes.

We share Munck and Verkuilen’s (2002) concerns and assume a realistic concept of democracy. Our concept resembles the concept of Dahl (1971) and includes three core aspects of democracy: political competition, political participation, and freedom of opinion.

## 2.2 Operationalization

We collected data on nine regime characteristics that are available for a comprehensive sample of country-years. Our data meets Munck and Verkuilen’s (2002) guidelines that recommend the use of non-aggregated data from objective and subjective sources.

Political participation is defined as the right of citizens to elect their political leaders and representatives (Dahl, 1971). Suffrage may be limited, either through constitutional restrictions that exclude citizens because of their gender, race, or income, or by non-constitutional restrictions that result from martial law or repression. We collected data on voter turnout and calculated the voter-population-ratio to capture the two types of disenfranchisement.<sup>5</sup>

The political process is competitive if individuals with different party affiliations compete in elections for political support (Przeworski, 1991). We use data on five regime

---

<sup>5</sup>We obtain our data from a number of sources, including African Election Database, the Inter-Parliamentary Union, Adam Carr’s Election Archive, the International Institute for Democracy and Electoral Assistance, the International Foundation of Electoral Systems, Nohlen (2005), Nohlen et al. (2001, 1999), and Nohlen and Stöver (2010). A documentation of the collected data is available upon request.



characteristics to capture this core aspect of democracy: (i) an expert-based measure of party pluralism<sup>6</sup>, (ii) the share of seats in parliament not won by the leading party, (iii) the number of seats won by the second strongest party divided by the number of seats won by the leading party, (iv) the share of votes not won by the leading party (or its presidential candidate), and (v) the share of votes won by the second strongest party (candidate) divided by the share of votes won by the leading party (candidate).<sup>7</sup>

Freedom of opinion means that citizens can freely choose their sources of information and can express their political views even if these views are not compatible with the political views of the government. To evaluate whether the citizens of a country enjoy these human rights, we use gender-specific ratings on the freedom of debate from the Varieties of Democracy (V-Dem) database (Coppedge et al., 2018).

### 2.3 Aggregation

Formally, the degree of democratization ( $d$ ) is a function of a set of observable regime characteristics ( $\mathbf{x}$ ):

$$d = \mathcal{A}(\mathbf{x}) \in [0, 1] \quad \text{with} \quad \mathbf{x} = (x_1, \dots, x_m) \in [0, 1]^m \quad (1)$$

where  $\mathcal{A}: [0, 1]^m \rightarrow [0, 1]$  is the aggregation function and  $m$  the number of regime characteristics. Specifying  $\mathcal{A}$  is difficult because we neither observe the shape of  $\mathcal{A}$  nor can we directly observe the actual level of democracy ( $d$ ) for all regimes.

The literature suggests two ways of meeting this challenge. The standard procedure is to weight the regime characteristics and then to apply a multiplicative or additive aggregation function (Goertz, 2006):

$$d = \prod_{j=1}^m x_j^{\omega_j} \quad \text{or} \quad d = \sum_{j=1}^m \omega_j \cdot x_j \quad \text{with} \quad \sum_{j=1}^m \omega_j = 1 \quad (2)$$

---

<sup>6</sup>This ordinal index of party pluralism has five categories: (i) there are no political parties, (ii) one legal party exists, (iii) there are multiple parties but opposition parties are faced with significant obstacles, (iv) there are multiple parties but opposition parties are faced with small obstacles, and (v) there are multiple parties and virtually no obstacles for opposition parties. The core data come from Coppedge et al. (2018). Supplementary information come from the African Election Database, the Inter-Parliamentary Union, Nohlen (2005), Nohlen et al. (2001, 1999), and Nohlen and Stöver (2010).

<sup>7</sup>We weight parliamentary and presidential elections according to their relevance for the political decision making process. We obtained the data needed for the calculation of the four latter characteristics from multiple sources (for a list, see Footnote 5).

where  $\omega_j \geq 0$  denotes the weight of regime characteristic  $x_j$ .<sup>8</sup>

Pemstein et al. (2010) and Gründler and Krieger (2016) deviate from this standard procedure: Pemstein et al. (2010) propose a Bayesian latent variable (LV) method and Gründler and Krieger (2016) a machine learning technique for pattern recognition. A major advantage of these two approaches is that they relax the assumptions about the functional relationship ( $\mathcal{A}$ ) between the regime characteristics ( $\mathbf{x}$ ) and the degree of democratization ( $d$ ). A disadvantage is, however, that the aggregation process is fully data driven rather than founded in preconceived rules.

## 2.4 Comparing aggregation methods

In this section, we compare the performance of five aggregation methods: an additive method, a multiplicative method, a method that combines additive and multiplicative elements, a Bayesian latent variable approach, and a machine learning technique. We proceed in two steps: first, we apply these aggregation techniques to our nine regime characteristics, and then consider multiple country examples to identify the similarities and the differences between the resulting democracy indices (for computational details, see Appendix A).

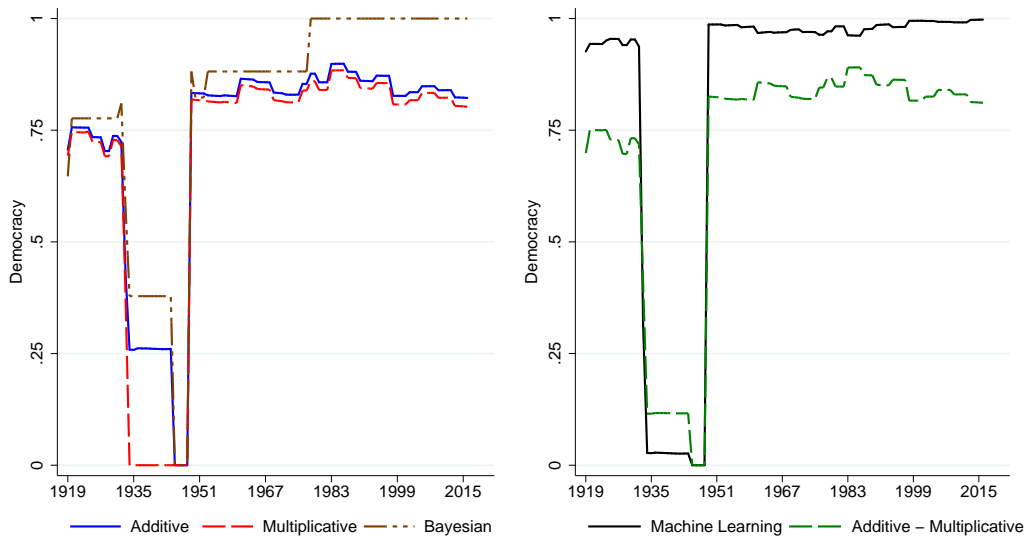
### 2.4.1 Germany

After its defeat in the First World War, Germany adopted a novel constitution and replaced the constitutional monarchy with a semi-presidential system. This constitution introduced universal suffrage and ensured freedom of opinion for all male and female citizens. Competitive multiparty elections were regularly held between 1919 and 1932, bringing politicians from different parties into power (Nohlen and Stöver, 2010). The parliamentary election in March 1933 was won by the Nazi Party and its leader, Adolf Hitler, became chancellor of Germany. In July 1933, Hitler enacted a law banning all political parties apart from the Nazi Party. The elections in December 1933, March 1936, and April 1938 were one-party elections and the Nazi Party punished any form of opposition (Shirer, 1991). After the victory of the Allied Forces in 1945, Germany was divided into four occupation zones. The occupation of the Western Allies ended in May

---

<sup>8</sup>Skaaning et al. (2015) and Teorell et al. (2018) refine the conventional approach by combining additive and multiplicative measures of democracy.

**Figure 1** Democracy in Germany (1919 – 2016).



**Notes:** The figures show the level of democracy of Germany, depending on how we aggregate our nine regime characteristics. For the period from 1949 to 1989, the measures of democracy refer to the Federal Republic of Germany.

1949 when the Federal Republic of Germany was founded. Since then, free multiparty elections were regularly held in the Federal Republic of Germany (Nohlen and Stöver, 2010).<sup>9</sup>

Figure 1 shows the level of democracy in Germany, depending on the aggregation method. All measures of democracy indicate three periods: two are democratic (1919 – 1932, 1949 – 2016) and one is non-democratic (1933 – 1948). This classification seems plausible given the historical facts. Having a closer look at Figure 1, we observe two notable differences. The first difference concerns the Nazi period (1933 – 1945). While the machine learning indicator and the multiplicative indicator suggest the absence of democracy, the other indices indicate the existence of some democratic structures. The latter assessment is incompatible with our concept of democracy since the Nazi party persecuted opponents and restricted the freedom of opinion (Shirer, 1991). The second striking difference exists for the recent past. The additive indicator, the multiplicative indicator, and the additive-multiplicative indicator suggest a democratic deficit, whereas the other indicators do not point to a lack of democracy. Because we are not aware of any policy report or qualitative study indicating a lack of political competition, political

<sup>9</sup>The occupation of the Soviet Union ended in October 1949 as the German Democratic Republic was founded. The reunification of the Federal Republic of Germany and the German Democratic Republic took place in October 1990.

participation, or freedom of opinion, we doubt whether the former indicators correctly reflect what is usually considered as the prevailing level of democracy.

#### 2.4.2 Russia (Soviet Union)

After the abdication of the Russian czar in March 1917, the October Revolution, and the Russian Civil War, the Soviet Union was founded in December 1922. The constitution enacted in December 1936 introduced universal direct suffrage. One year later, the first general election took place. However, no political parties except the ruling Communist Party was permitted to participate. Subsequent elections were also single-party elections. In 1986, Mikhail Gorbachev, the General Secretary of the Communist Party, launched two new policies—*Glasnost* and *Perestroika*<sup>10</sup>—thereby initiating an institutional change (Sakwa, 2005). The democratization process accelerated after the collapse of the Soviet Union. In 1991, Boris Yeltsin became the first freely elected president of the Russian Federation. President Yeltsin was narrowly reelected in July 1996 and stepped down in December 1999. Under his successors, Vladimir Putin and Dmitry Medvedev, party competition decreased and constraints on the freedom of opinion became more frequent (Nohlen and Stöver, 2010, Sakwa, 2014).

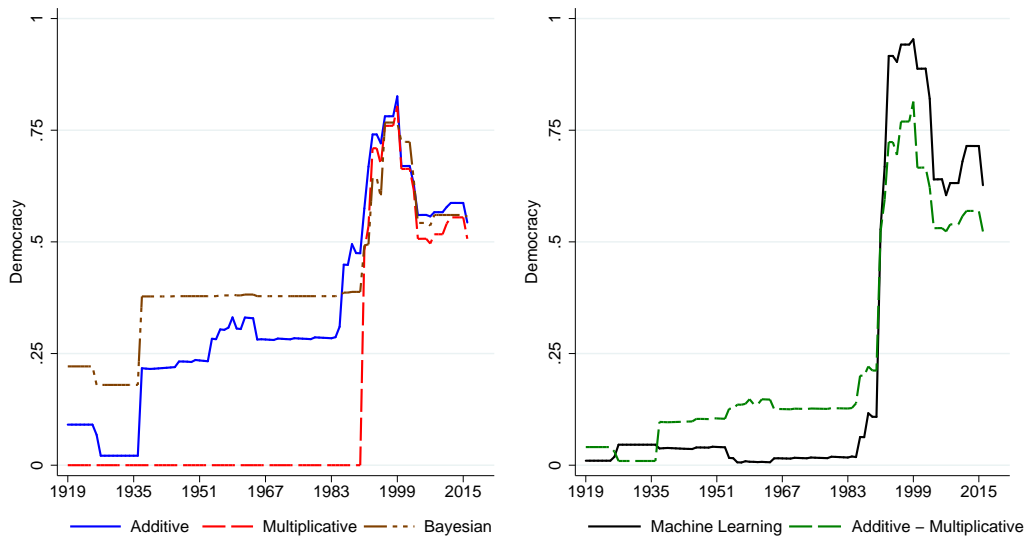
Figure 2 shows the level of democracy of the Russian Federation (Soviet Union) for different aggregation methods. We observe great similarities for the post-Cold War era and significant differences for the Soviet era. After the fall of the Iron Curtain, all five indices indicate a distinct increase in the Russian degree of democratization. We think this increase is plausible since free multi-party elections took place in the early years of the Russian Federation, while single-party elections were held in the Soviet period. We also observe that all five indices decrease after the inauguration of Vladimir Putin and find several expert reports that justify this result (Hale et al., 2004, Sakwa, 2010).

The measurement differences in the Soviet period resemble the differences for the Nazi period in Germany. The additive index, the Bayesian index, and the additive-multiplicative index indicate that the Soviet Union partly respected basic democratic principles, while the multiplicative indicator and the machine learning indicator suggest

---

<sup>10</sup>Perestroika was introduced to modernize the economy of the Soviet Union. Glasnost comprised several political reforms that were implemented to lift restrictions on the freedom of speech and information (Sakwa, 1990).

**Figure 2** Democracy in Russia (1919 – 2016).



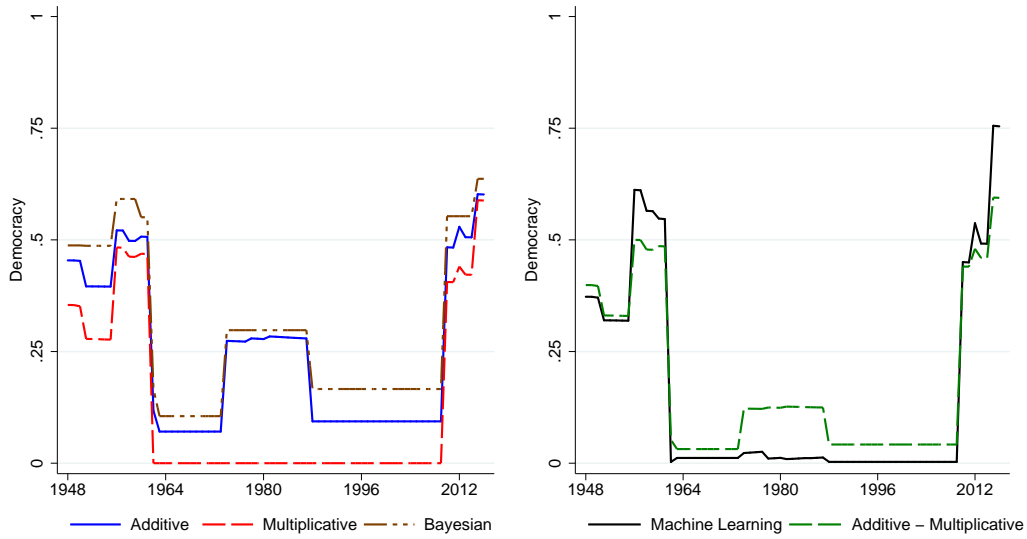
**Notes:** The figures show the level of democracy of Russia, depending on how we aggregate our nine regime characteristics. For the period from 1922 to 1990, the measures of democracy refer to the Soviet Union.

that the Soviet Union had a highly autocratic government. Only the latter assessment is consistent with a realistic concept of democracy because of the absence of political pluralism and freedom of opinion in the Soviet Union.

### 2.4.3 Myanmar (Burma)

After gaining independence from the United Kingdom in 1948, Burmese politics was mainly shaped by the Anti-Fascist People's Freedom League (AFPFL), which won the elections in 1951/52, 1956, and 1960. Dissatisfied with the performance of the AFPFL, General Ne Win overthrew the civilian government in March 1962 and established a military-dominated regime (Nohlen et al., 2001). In 1974, the military government organized elections for the first time since 1960. However, the Burma Socialist Program Party—led by General Ne Win—did not allow for electoral competition (Devi, 2014). In 1988, General Ne Win resigned from office and the interim government announced a multiparty election for 1990. This election was clearly won by the National League for Democracy (NLD), the main opposition force. The military prevented, however, the democratic takeover and military governments stayed in power for another 21 years. In 2010, the military withdrew and thus paved the way for an institutional change (Barany, 2016, Rieffel, 2013).

**Figure 3** Democracy in Myanmar (1948 – 2016).

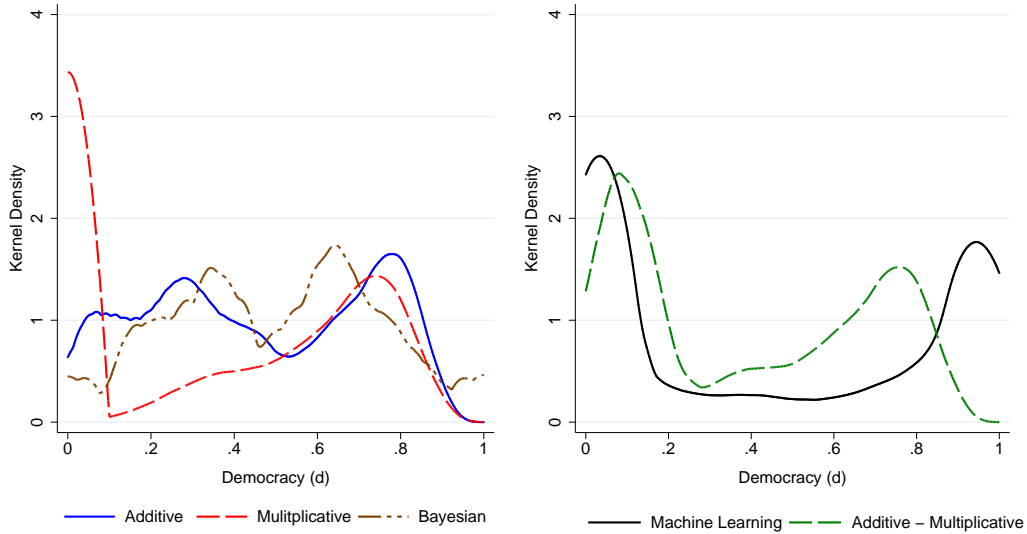


**Notes:** The figures show the level of democracy of Myanmar, depending on how we aggregate our nine regime characteristics. The country was renamed from Burma to Myanmar in 1989.

Figure 3 presents how the five democracy indicators evaluate Myanmar (Burma). We observe that the indicators differ only slightly from each other for the periods in which multiparty elections took place (1948 – 1961, 2010 – 2016).<sup>11</sup> Furthermore, we find two striking differences for the period of military dictatorship. First, the additive index, the Bayesian index, and, to a lesser extent, the additive-multiplicative index indicate a step towards democracy in 1974, whereas the machine learning index and the multiplicative index remain unchanged. We argue that the latter assessment is more consistent with a realistic concept of democracy because political competition did not exist in the single-party regime established in 1974. Second, the additive index and the Bayesian index suggest some democratic structures in the two periods in which the government was not elected (1962 – 1973, 1988 – 2009). We believe that these figures are incompatible with the concept of democracy because electoral legitimacy constitutes a necessary condition for democracy.

<sup>11</sup>Even though multiparty elections were held in these two periods, all five measures indicate democratic deficits. We think this result is compatible with our concept of democracy for two reasons: first, some ethnic groups (e.g. Rohingya) were excluded from the political process and did not enjoy the freedom of opinion, and second, the great influence of the military on the political decision making process. For example, only 168 out of 224 seats in the parliament were up for election in 2010 and 2015. The other seats were reserved for members of the armed forces (Barany, 2016).

**Figure 4** Kernel densities



**Notes:** The figure shows the kernel densities of our five measures of democracy. We use the Epanechnikov kernel to estimate the density functions. Colonies are excluded from the estimation sample.

## 2.5 Discussion

We draw three conclusions from Section 2.4. First, the question of how we transform the regime characteristics into an index matters because the degree of democratization significantly change when we replace an aggregation method with another aggregation method. Second, the additive method and the Bayesian method are likely to produce indicators that are too favorable for highly autocratic regimes to be consistent with the concept of democracy. Finally, the machine learning method is least likely to create an indices that underestimate the level of democracy of a highly democratic regime.<sup>12</sup>

An potential objection to our evaluation is that the three country examples cannot provide sufficient evidence for the second and third conclusion. This objection is not far-fetched, especially because our most favored aggregation method, the machine learning method, is our brainchild (Gründler and Krieger, 2016).<sup>13</sup> Below, we present additional results that support our conclusions and, hopefully, fully alleviate the suspicion that we

<sup>12</sup>Our analysis leaves open the question of which aggregation technique produces the “best” indices for hybrid (partial, semi-democratic) regimes. We think that answering this question is impossible because the true degree of democratization of a hybrid regime cannot be observed (Diamond, 2002, Lindberg et al., 2014). Excluding hybrid regimes from our analysis is not problematic for two reasons: first, the number of hybrid regimes is relatively low, and second, the differences between the five indicators are much smaller for hybrid regimes than for autocracies and democracies.

<sup>13</sup>In Gründler and Krieger (2016), we use the machine learning method to synthesize existing democracy indicators. For data availability reasons, the resulting index only covers the period from 1981 to 2011. Another problem is that combining different measures of democracy leads to conceptual vagueness.

**Table 1** Average changes in the level of democracy — Democratization events

	Machine Learning	Additive	Multiplicative	Add. / Multi.	Bayesian
$\Delta$	0.522	0.276	0.399	0.344	0.284

**Notes:** This table reports results from estimating (3). The figures show how much the democracy indices change, on average, when Papaioannou and Siourounis (2008) indicate a transition from autocracy towards democracy.

strategically selected our three country examples to give an advantage to our preferred aggregation method. Another concern might be that the regime characteristics could be strategically chosen. Appendix B.2 shows that our conclusions remain unchanged if we use the regime characteristics proposed by Teorell et al. (2018).

Appendix B.1 presents additional country examples (Bhutan, Kenya, and Spain) to support the conclusions drawn from Section 2.4. We also examined many other country examples (not reported, but available upon request) and find strong evidence for our conjectures.<sup>14</sup>

Figure 4 presents the results from kernel density estimations. We observe that the density functions of all five indicators have a local maximum in the lower part of the spectrum ( $d < 0.5$ ) and another local maximum in the upper part of the spectrum ( $d > 0.5$ ). However, the exact locations of the local maxima differ between the five density functions. While the density functions of the multiplicative index, the machine learning index, and the additive-multiplicative index have a lower maximum at  $d \approx 0$ , the density functions of the additive index and the Bayesian index have a lower local maximum at  $d \approx 0.3$ . This result is reassuring since it is consistent with the view that the additive method and the Bayesian method are likely to produce implausibly high indices for autocratic regimes. Furthermore, the upper local maximum of the density function of the machine learning index is closer to  $d = 1$  than the other upper local maxima, showing that the machine learning method is least likely to create implausibly low indicators for highly democratic regimes.

Papaioannou and Siourounis (2008) report a list of countries in which an autocratic regime was replaced with a democratic regime in the 1960 – 2005 period. If the conclusions drawn from the case studies generally apply, we should find that the machine

<sup>14</sup>We can explain the performance differences with the assumptions that the aggregation methods make about the functional relationship between the regime characteristics and the level of democracy (for details, see Appendix B.3).



learning index indicates, on average, greater changes in the degree of democratization than the other four indices. We expect such a difference because the machine learning index is least likely to underestimate (overestimate) the degree of democratization of a clearly democratic (autocratic) regime. To test this hypothesis, we calculate a measure reflecting the average change in the level of democracy:

$$\Delta^j = \frac{1}{|\mathcal{D}|} \sum_{(i,t) \in \mathcal{D}} |d_{i,t}^j - d_{i,t-1}^j| \quad (3)$$

where  $d_{i,t}^j$  denotes the level of democracy of country  $i$  in year  $t$  indicated by index  $j$ . The set  $\mathcal{D}$  includes all country-year observations for which Papaioannou and Siourounis (2008) report a transition from autocracy to democracy. The results shown in Table 1 confirm that the machine learning index increases more than the other indices when a country becomes democratic.<sup>15</sup>

### 3 Econometric model

In the previous section, we showed that most aggregation methods are likely to create implausibly low (high) indices for democratic (autocratic) regimes. In this section, we present an econometric model that shows how these malfunctions affect the result of a regression in which the degree of democratization serves as the explanatory variable.

#### 3.1 Ordinary least squares estimator

Consider a model in which the degree of democratization ( $d$ ) influences an outcome variable ( $y$ ) in the following manner:

$$y_i = \alpha + \beta \cdot d_i + \varepsilon_i \quad (4)$$

where  $\alpha > 0$  and  $\beta > 0$  denote unknown parameters and  $\varepsilon$  a randomly distributed error term. For analytical convenience, we also assume that  $m$  of the  $n$  independent

---

<sup>15</sup>Appendix Table E.2 shows that we obtain similar results if we use the Boix et al. (2013) database to identify major regime changes. The Boix et al. (2013) database has two advantages compared to the Papaioannou and Siourounis (2008) list: first, it has a greater coverage (1800 – 2007), and second, it indicates transitions from democracy to autocracy. The disadvantage of the Boix et al. (2013) data is that it does not distinguish between autocratic, hybrid, and democratic regimes.

observations have a degree of democratization of  $d_{low}$  and that the remaining  $n - m$  observations have a degree of democratization of  $d_{high} > d_{low}$ . We consider two indices: the first index ( $d_{1,\cdot}$ ) correctly indicates the true degree of democratization, while the second index ( $d_{2,\cdot}$ ) is biased:

$$d_{1,j} = d_j \quad \text{and} \quad d_{2,j} = d_j + \mathcal{E}(d_j) \quad \text{for} \quad j \in \{low, high\} \quad (5)$$

where  $\mathcal{E}(\cdot)$  indicates the measurement error of the second index. The results presented in Section 2 suggest that most of the available aggregation techniques produce too high (low) indices for autocratic (democratic) regimes. We portray these biases by specifying:

$$\mathcal{E}(d_j) = \begin{cases} -\eta & \text{for } d_j = d_{high} \\ \gamma & \text{for } d_j = d_{low} \end{cases} \quad \text{with } \eta > 0 \quad \text{and} \quad \gamma > 0.^{16} \quad (6)$$

In this simple model, the ordinary least squares (OLS) estimators have the following form:

$$\beta^{k,ols} = \frac{\text{cov}(y, d_{k,\cdot})}{\text{var}(d_{k,\cdot})} = \frac{m \cdot \sum_{i=m+1}^n (y_i - \bar{y}) - (n - m) \cdot \sum_{i=1}^m (y_i - \bar{y})}{(d_{k,high} - d_{k,low}) \cdot m \cdot (n - m)} \quad (7)$$

where  $k$  indicates whether we use the unbiased democracy indicator ( $k = 1$ ) or the biased democracy indicator ( $k = 2$ ).

Equation (7) implies that the OLS estimator of the parameter  $\beta$  increases when we replace the unbiased index ( $d_{1,\cdot}$ ) with the biased index ( $d_{2,\cdot}$ ). The reason is that the biased index underestimates the difference between  $d_{high}$  and  $d_{low}$ :

$$d_{2,high} - d_{2,low} < d_{1,high} - d_{1,low} = d_{high} - d_{low} \quad \Rightarrow \quad \beta^{2,ols} > \beta^{1,ols}. \quad (8)$$

### 3.2 Two-stage least squares estimator

Many economists and political scientists who investigate the consequences of political transitions use an instrumental variable (IV) strategy to alleviate concerns related to measurement errors in the democracy index. This approach works if the measurement

---

<sup>16</sup>Below, we only consider cases in which  $d_{2,high} = d_{high} - \eta > d_{low} + \gamma = d_{2,low}$ .

error is random, i.e. when it does not depend on the level of democracy (Angrist and Pischke, 2009). In Sections 2.4 and 2.5, we showed, however, that the measurement errors caused by most of the available aggregation methods are not random. We thus doubt whether an IV approach suffices to fully compensate for the biases in democracy indicators.

To justify our concerns, we extend the model specified in (4) – (6) by the assumption that we observe a variable  $z \geq 0$  which positively correlates with the true degree of democratization ( $d$ ) and does not affect the outcome variable ( $y$ ). Put differently, the variable  $z$  can serve as an instrumental variable in a two-stage least squares (2SLS) regression. The 2SLS estimator then has the form:

$$\beta^{k,iv} = \frac{\text{cov}(y, z)}{\text{cov}(d_{k,\cdot}, z)} = \frac{\delta^{ols}}{\rho^{k,ols}}, \quad (9)$$

where  $\rho^{k,ols}$  is the OLS estimator of the first-stage model:

$$d_i = \pi + \rho \cdot z_i + \xi_i \quad \text{with} \quad \rho > 0, \quad (10)$$

and  $\delta^{ols}$  denotes the OLS estimator of the reduced-form model:

$$y_i = \zeta + \delta \cdot z_i + \iota_i \quad \text{with} \quad \delta > 0. \quad (11)$$

From

$$\begin{aligned} \rho^{k,ols} &= \frac{\text{cov}(z, d_{k,\cdot})}{\text{var}(z)} \\ &= \frac{\frac{1}{n} \cdot (d_{k,high} - d_{k,low}) \cdot (m \cdot \sum_{i=m+1}^n (z_i - \bar{z}) - (n-m) \cdot \sum_{i=1}^m (z_i - \bar{z}))}{\text{var}(z)} \end{aligned}$$

we can infer that the estimate of the first-stage parameter  $\rho$  decreases if we replace the unbiased index with the biased index, because the biased index underestimates the difference between  $d_{high}$  and  $d_{low}$ . The main consequence of using the biased index is that we overestimate the parameter  $\beta$ :

$$d_{2,high} - d_{2,low} < d_{1,high} - d_{1,low} \quad \Rightarrow \quad \rho^{2,ols} < \rho^{1,ols} \quad \Rightarrow \quad \beta^{2,iv} > \beta^{1,iv}. \quad (12)$$

## 4 Democracy and economic growth

We now turn to the empirical part of the paper, in which we use our new indices to estimate the effect of democracy on economic growth. The motivation for this study is twofold: first, confirming the key predictions of our stylized model, and second, showing that the overestimation biases predicted by our model are of practical relevance.

The question of whether democracy cause economic growth is the subject of many studies. Some theories argue that a democracy grows faster than an autocracy since it has better economic institutions and spends more on education. Other theories suggest, however, that economic growth decreases if a democratic system replaces an autocratic system due to higher taxation and efficiency losses (Acemoglu, 2008, Besley and Coate, 1998, Saint-Paul and Verdier, 1993).

A number of recent empirical studies show that the positive aspects of democracy outweigh the negative aspects (Acemoglu et al., 2018, Flachaire et al., 2014, Gründler and Krieger, 2016, Madsen et al., 2015, Papaioannou and Siourounis, 2008, Persson and Tabellini, 2006, Rodrik and Wacziarg, 2005). Below, we present results confirming the view that democracy positively affects economic growth. The focus of our analysis is, however, on the question of whether the size of the effect depends on the aggregation method used for the creation of the measure of democracy.

### 4.1 Identification strategies

Three endogeneity problems complicate any empirical study that examines the effect of democracy on economic growth. First, autocracies may differ from democracies in non-observable factors that also affect economic growth. Second, causality may run from economic development to democracy. Finally, democratization is often preceded by a temporal decline in GDP per capita (Acemoglu et al., 2018).

Most empirical studies address these endogeneity issues with a dynamic fixed effect model:

$$Y_{i,t} = \sum_{l=1}^L \beta_l \cdot Y_{i,t-l} + \gamma \cdot D_{i,t} + \xi_i + \eta_t + \varepsilon_{i,t} \quad (13)$$

where  $D$  denotes the level of democracy of country  $i$  in year  $t$ ,  $Y$  the log of GDP per

capita,  $\xi$  the country fixed effect,  $\eta$  the year fixed effect, and  $\varepsilon$  the error term.<sup>17</sup>

The effect of democracy on economic growth ( $\gamma$ ) can be identified with the dynamic fixed effect model if the error term is uncorrelated with the past, current, and future realizations of the degree of democratization. Since this condition may not be satisfied because of omitted time-varying factors, several studies use a two-stage least squares approach in which the average level of democracy in neighboring countries serves as the instrument for the domestic degree of democratization (Acemoglu et al., 2018, Persson and Tabellini, 2009):<sup>18</sup>

$$D_{i,t} = \sum_{l=1}^L \delta_l \cdot Y_{i,t-l} + \alpha \cdot Z_{i,t-1} + \zeta_i + \tau_t + \nu_{i,t} \quad (14)$$

with

$$Z_{i,t} = \frac{1}{|\mathcal{R}|} \sum_{j \in \mathcal{R}} D_{j,t} \quad \text{and} \quad \mathcal{R} = \{j : j \neq i, r_j = r_i\} \quad (15)$$

where  $r_i$  denotes the region in which country  $i$  is located.<sup>19</sup>

## 4.2 Results from ordinary least squares regressions

Column 1 of Table 2 shows the results from estimating (13) when using the machine learning indicator to measure the degree of democratization. In line with other recent studies, we find a positive and statistically significant effect of democracy on economic growth. The OLS coefficient implies that a transition from autocracy ( $d = 0$ ) towards democracy ( $d = 1$ ) causes an increase in GDP per capita by about 1.7 percent.<sup>20</sup>

The results presented in Section 2 suggest that the additive index is more likely to underestimate changes in the true degree of democratization than the machine learning index. The reason is that the additive aggregation approach often produces too large (small) indices for autocratic (democratic) regimes. In Section 3.1, we argue that the

<sup>17</sup>The data on GDP per capita comes from the Maddison Project Database 2018 (Bolt et al., 2018).

<sup>18</sup>The motivation for this instrumentation strategy is that transitions from autocracy to democracy (or vice versa) often occurred in regional waves (Huntington, 1993, Teorell, 2010).

<sup>19</sup>In our baseline analysis, we use the classification of the United Nations to divide the world into 19 regions. Results for other classification schemes look similar and are available upon request.

<sup>20</sup>Compared with other studies that present OLS results from a dynamic fixed effect model, the main advantage of our analysis is the comprehensive sample which includes 163 countries and covers the period from 1919 to 2016. Just for comparison, most previous studies use data from 1960 onward (Acemoglu et al., 2018, Gründler and Krieger, 2016, Papaioannou and Siourounis, 2008, Persson and Tabellini, 2006).

**Table 2** Democracy and economic growth — OLS estimates

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Democracy	0.017*** (0.0028)	0.027*** (0.0044)	0.023*** (0.0037)	0.026*** (0.0041)	0.025*** (0.0042)
Observations	10,733	10,733	10,733	10,733	10,733
Countries	163	163	163	163	163
R <sup>2</sup>	0.997	0.997	0.997	0.997	0.997
Equal. (p-val.)	-	0.017	0.090	0.025	0.041

**Notes:** The table presents OLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

estimate obtained from an OLS regression increases and becomes inconsistent when we apply a democracy index that suffers from these malfunctions. Column 2 of Table 2 supports this argument since it shows that the estimate of the effect of democracy on economic growth increase from 1.7 percent to 2.7 percent when we replace the machine learning index with the additive index. The difference between these two estimates is statistically significant at the five percent level.

In Columns 3 – 5, we show the results for the multiplicative index, the additive-multiplicative index, and the Bayesian index. Again, we find that the OLS estimates of the effect of democracy on economic growth significantly increase compared with the estimate reported in Column 1. This result was to be expected because the machine learning method is less likely than the other aggregation methods to produce measures of democracy that underestimate the changes in the true degree of democratization (see Section 2).

### 4.3 Results from two-stage least squares regressions

Table 3 reports results from two-stage least squares regressions in which the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. Column 1 presents the results for the machine learning index. We find a positive and statistically significant 2SLS estimate of the effect of democracy on economic growth (see Panel A).

In Panel B, we report the first-stage result. The estimate suggests that the regional

**Table 3** Democracy and economic growth — 2SLS estimates

	Machine Learning	Additive	Multiplicative	Add./ Multi.	Bayesian
	(1)	(2)	(3)	(4)	(5)
Panel A: Second-stage estimates					
Democracy	0.029*** (0.0047)	0.048*** (0.0073)	0.037*** (0.0060)	0.042*** (0.0065)	0.052*** (0.0080)
Equal. (p-val.)	-	0.011	0.178	0.061	0.005
Panel B: First-stage estimates					
Demo. (reg.)	0.789*** (0.0609)	0.482*** (0.0365)	0.618*** (0.0476)	0.557*** (0.0416)	0.446*** (0.0351)
Equal. (p-val.)	-	0.000	0.001	0.000	0.000
Observations	10,708	10,708	10,708	10,708	10,708
Countries	163	163	163	163	163
SaWi (F.stat.)	167.79	174.43	168.88	179.00	161.44
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000

**Notes:** The table presents 2SLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

level of democracy positively correlates with the domestic level of democracy. We also present results from three weak instrument tests. We report the first-stage F-statistic as suggested by Sanderson and Windmeijer (2016) and find that it exceeds the Stock and Yogo (2005) critical value. Furthermore, we present the p-values of the Anderson and Rubin (1949) test and the Stock and Wright (2000) test. None of these tests points to a weak instrument problem.

Section 3.2 suggests that the first-stage (second-stage) estimate becomes too small (large) when using an indicator that underestimates the changes in the true degree of democratization. Columns 2 – 5 of Table 3 support these predictions. We observe that the first-stage estimate decreases and that the second-stage estimate increases when we replace the machine learning indicator with an alternative measure of democracy. The differences in the first-stage estimates are statistically significant at the one percent level and the differences in the second-stage estimates are—with one exception—statistically significant at conventional levels.

## 4.4 Robustness

We draw two conclusions from the regression results presented in Tables 2 and 3: first, democracy positively affects economic growth, and second, we overestimate the positive effect of democracy on economic growth if we use an additive, multiplicative, additive-multiplicative, or Bayesian index. The latter result is particularly noteworthy because almost all studies that examine the economic consequences of political transitions use these types of indices. Below, we present the results of several robustness checks.

Some economists argue that annual data is inappropriate for studying the causes of economic growth. These scholars prefer data that is averaged over multiple years since data averaging filters out the business cycle fluctuations and mitigates the impact of measurement error in the variables (Durlauf et al., 2005). In Appendix Tables E.3 and E.4, we show that the estimates of the effect of democracy on economic growth remain positive and statistically significant when we use five-year data rather than annual data. These two tables also show that using averaged data does not suffice to eliminate the differences in the effect size caused by a replacement of the machine learning index.

We conduct several subsample analyses to rule out that our results are driven by a particular group of countries or a specific period. Appendix Tables E.5 and E.6 show that our results hold when we drop all countries from either Africa, America, Asia, or Europa. Appendix Tables E.7 and E.8 suggest that our results also remain unchanged when we restrict the analysis to specific periods (1946 – 2016, 1970 – 2016).

In Appendix Tables E.9 and E.10, we extend the regression model by several control variables (civil conflict, population growth, institutional quality, education, investment, trade openness, government consumption). We find that the estimates of the effect of democracy on economic growth slightly decrease and are statistically significant at the five percent level. We also observe that the machine learning indicator still suggests a smaller effect than the other indicators.<sup>21</sup>

Finally, Appendix Tables E.11 and E.12 show results from regressions in which the

---

<sup>21</sup>We use data from Brecke (1999) and the Uppsala Conflict Data Program to create an index of civil conflict. The measure of institutional quality comes from Coppedge et al. (2018) and reflects the extent to which courts operate independently. We use the Barro and Lee (2013) data on years of schooling to measure education. The data on population growth is obtained from different sources, including: Bolt et al. (2018), the Cross-National Time Series Data Archive, the World Bank, and the web page [www.populstat.info](http://www.populstat.info). The data on trade openness, government consumption, and investment comes from Feenstra et al. (2015).



average years of schooling serves as the outcome variable. In line with other empirical studies, we find that education increases in the degree of democratization (Acemoglu et al., 2018, Baum and Lake, 2003, Harding and Stasavage, 2013). We also observe that the positive effect of democracy on education increases when we replace the machine learning index with another index. This result is notable because it suggests that our conclusions about how changes in the aggregation technique affect the OLS and 2SLS estimates have at least some external validity.

## 5 Conclusion

Dissatisfied with the quality of standard measures of democracy, we compiled data on political participation, political competition, and the freedom of opinion (186 countries, 1919 – 2016) to create a new indicator. Since the literature does not provide generally acknowledged guidelines for aggregation, we use different methods and compare their performance. We observe that four out of five well-established aggregation techniques produce indices that underestimate changes in the true degree of democratization. The sole exception is a machine learning method.

We present a model that assumes a linear relationship between democracy and an outcome variable to illustrate the empirical consequence of applying error-prone data aggregation methods. We show that estimating such a model creates excessively large OLS estimates if the democracy indicator underestimates changes in the true degree of democratization. We also explain why having an instrument for the level of democracy does not suffice to correct this upward bias. The results of an analysis of the effect of democracy on economic growth demonstrate that the bias is significant.

Our paper has two main implications: First, data aggregation matters and all social scientists should be careful in drawing causal inferences from regressions in which the variables of interests are measured by an index. Second, researchers interested in the consequences of political transitions should discontinue using simple democracy indices, such as the Polity IV index or the Freedom House indices, because their underlying aggregation methods produce non-random measurement errors.

## References

- Acemoglu, D. (2008). Oligarchic versus democratic societies. *Journal of the European Economic Association*, 6(1):1–44.
- Acemoglu, D., Johnson, S., Robinson, J. A., and Yared, P. (2008). Income and democracy. *American Economic Review*, 98(3):808–42.
- Acemoglu, D., Naidu, S., Restrepo, P., and Robinson, J. A. (2018). Democracy does cause growth. *Journal of Political Economy*, forthcoming.
- Aidt, T. S. and Franck, R. (2015). Democratization under the threat of revolution: Evidence from the great reform act of 1832. *Econometrica*, 83(2):505–547.
- Aidt, T. S. and Leon, G. (2016). The democratic window of opportunity: Evidence from riots in Sub-Saharan Africa. *Journal of Conflict Resolution*, 60(4):694–717.
- Anderson, T. W. and Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of Mathematical Statistics*, 20(1):46–63.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press.
- Barany, Z. (2016). Moving toward democracy: The 2015 parliamentary elections in Myanmar. *Electoral Studies*, 42(1):75–77.
- Barro, R. J. (1996). Democracy and Growth. *Journal of Economic Growth*, 1(1):1–27.
- Barro, R. J. and Lee, J. W. (2013). A new data set of educational attainment in the world, 1950–2010. *Journal of Development Economics*, 104(1):184–198.
- Baum, M. A. and Lake, D. A. (2003). The political economy of growth: Democracy and human capital. *American Journal of Political Science*, 47(2):333–347.
- Besley, T. and Coate, S. (1998). Sources of inefficiency in a representative democracy: a dynamic analysis. *American Economic Review*, 88(1):139–156.
- Boix, C., Miller, M., and Rosato, S. (2013). A complete data set of political regimes, 1800–2007. *Comparative Political Studies*, 46(12):1523–1554.
- Bolt, J., Inklaar, R., de Jong, H., and van Zanden, J. L. (2018). Rebasings ‘maddison’: new income comparisons and the shape of long-run economic development. *GGDC Research Memorandum*.
- Brecke, P. (1999). Violent conflicts 1400 ad to the present in different regions of the world.

- In *1999 Meeting of the Peace Science Society*.
- Brückner, M. and Ciccone, A. (2011). Rain and the democratic window of opportunity. *Econometrica*, 79(3):923–947.
- Cervellati, M., Jung, F., Sunde, U., and Vischer, T. (2014). Income and democracy: Comment. *American Economic Review*, 104(2):707–719.
- Cheibub, J. A., Gandhi, J., and Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, 143(1-2):67–101.
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Skaaning, S.-E., and Teorell, J. (2018). V-dem codebook v8.
- Dahl, R. A. (1971). *Polyarchy: Participation and Opposition*. Yale University Press.
- Devi, K. S. (2014). Myanmar under the military rule 1962-1988. *International Research Journal of Social Sciences*, 3(10):46–50.
- Diamond, L. J. (2002). Thinking about hybrid regimes. *Journal of Democracy*, 13(2):21–35.
- Durlauf, S. N., Johnson, P. A., and Temple, Jonathan R. W. (2005). Growth econometrics. In *Handbook of Economic Growth*. Elsevier.
- Elklit, J. and Svensson, P. (1997). What makes elections free and fair? *Journal of Democracy*, 8(3):32–46.
- Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182.
- Flachaire, E., García-Peñalosa, C., and Konte, M. (2014). Political versus economic institutions in the growth process. *Journal of Comparative Economics*, 42(1):212–229.
- Goertz, G. (2006). *Social science concepts: A user’s guide*. Princeton University Press.
- Gründler, K. and Krieger, T. (2016). Democracy and growth: Evidence from a machine learning indicator. *European Journal of Political Economy*, 45(1):85–107.
- Gundlach, E. and Paldam, M. (2009). A farewell to critical junctures: Sorting out long-run causality of income and democracy. *European Journal of Political Economy*, 25(3):340–354.
- Hale, H. E., McFaul, M., and Colton, T. J. (2004). Putin and the” delegative democracy” trap: Evidence from Russia’s 2003-04 elections. *Post-Soviet Affairs*, 20(4):285–319.
- Harding, R. and Stasavage, D. (2013). What democracy does (and doesn’t do) for ba-

- sic services: School fees, school inputs, and african elections. *Journal of Politics*, 76(1):229–245.
- Huntington, S. P. (1993). *The third wave: Democratization in the late twentieth century*. University of Oklahoma Press.
- Knutsen, C. H. (2012). Democracy and economic growth: A survey of arguments and results. *International Area Studies Review*, 15(4):393–415.
- Lindberg, S. I., Coppedge, M., Gerring, J., and Teorell, J. (2014). V-dem: A new way to measure democracy. *Journal of Democracy*, 25(3):159–169.
- Lipset, S. M. (1959). Some social requisites of democracy. *American Political Science Review*, 53(1):69–105.
- Madsen, J. B., Raschky, P. A., and Skalil, A. (2015). Does democracy drive income in the world, 1500–2000? *European Economic Review*, 78(1):175–195.
- Márquez, X. (2016). A quick method for extending the unified democracy scores. *SSRN Working Paper Series*.
- Munck, G. L. and Verkuilen, J. (2002). Conceptualizing and measuring democracy: Evaluating alternative indices. *Comparative Political Studies*, 35(1):5–34.
- Murtin, F. and Wacziarg, R. (2014). The democratic transition. *Journal of Economic Growth*, 19(2):141–181.
- Nohlen, D. (2005). *Elections in the Americas: A data handbook*. Oxford University Press.
- Nohlen, D., Grotz, F., and Hartmann, C. (2001). *Elections in Asia and the Pacific: A data handbook*. Oxford University Press.
- Nohlen, D. and Stöver, P. (2010). *Elections in Europe: A data handbook*. Nomos.
- Nohlen, D., Thibaut, B., and Krennerich, M. (1999). *Elections in Africa: A data handbook*. Oxford University Press.
- O’Donnell, G. A. (2001). Democracy, law, and comparative politics. *Studies in Comparative International Development*, 36(1):7–36.
- Papaioannou, E. and Siourounis, G. (2008). Democratisation and Growth. *Economic Journal*, 118(532):1520–1551.
- Pemstein, D., Meserve, S. A., and Melton, J. (2010). Democratic compromise: A latent variable analysis of ten measures of regime type. *Political Analysis*, 18(4):426–449.
- Persson, T. and Tabellini, G. (2006). Democracy and development: The devil in the

- details. *American Economic Review*, 96(2):319–324.
- Persson, T. and Tabellini, G. (2008). The growth effect of democracy: Is it heterogenous and how can it be estimated? In *Institutions and Economic Performance*. Harvard University Press.
- Persson, T. and Tabellini, G. (2009). Democratic Capital: The Nexus of Political and Economic Change. *American Economic Journal: Macroeconomics*, 1(2):88–126.
- Plattner, M. F. (2015). Is democracy in decline? *Journal of Democracy*, 26(1):5–10.
- Przeworski, A. (1991). *Democracy and the market: Political and economic reforms in Eastern Europe and Latin America*. Cambridge University Press.
- Przeworski, A. (2000). *Democracy and development: Political institutions and well-being in the world, 1950-1990*. Cambridge University Press.
- Rieffel, L. (2013). Myanmar on the move: An overview of recent developments. *Journal of Current Southeast Asian Affairs*, 31(4):31–49.
- Rodrik, D. and Wacziarg, R. (2005). Do democratic transitions produce bad economic outcomes? *American Economic Review*, 95(2):50–55.
- Saint-Paul, G. and Verdier, T. (1993). Education, democracy and growth. *Journal of Development Economics*, 42(2):399–407.
- Sakwa, R. (1990). *Gorbachev and his Reforms, 1985-1990*. Prentice Hall.
- Sakwa, R. (2005). *The rise and fall of the Soviet Union*. Routledge.
- Sakwa, R. (2010). *The crisis of Russian democracy: The dual state, factionalism and the Medvedev succession*. Cambridge University Press.
- Sakwa, R. (2014). *Putin redux: Power and contradiction in contemporary Russia*. Routledge.
- Sanderson, E. and Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.
- Shirer, W. L. (1991). *The rise and fall of the Third Reich: A history of Nazi Germany*. Random House.
- Skaaning, S.-E., Gerring, J., and Bartusevičius, H. (2015). A lexical index of electoral democracy. *Comparative Political Studies*, 48(12):1491–1525.
- Stock, J. H. and Wright, J. H. (2000). GMM with weak identification. *Econometrica*, 68(5):1055–1096.

- Stock, J. H. and Yogo, M. (2005). Testing for weak instruments in linear iv regression. In *Identification and Inference for Econometric Models*. Cambridge University Press.
- Tavares, J. and Wacziarg, R. (2001). How democracy affects growth. *European Economic Review*, 43(8):1341–1378.
- Teorell, J. (2010). *Determinants of democratization: Explaining regime change in the world, 1972–2006*. Cambridge University Press.
- Teorell, J., Coppedge, M., Lindberg, S., and Skaaning, S.-E. (2018). Measuring polyarchy across the globe, 1900–2017. *Studies in Comparative International Development*, forthcoming.
- Turner, M., Chuki, S., and Tshering, J. (2011). Democratization by decree: the case of Bhutan. *Democratization*, 18(1):184–210.
- Turner, M. and Tshering, J. (2014). Second elections and democratic consolidation: the experience of Bhutan. *Asian Journal of Political Science*, 22(3):315–334.
- Vanhanen, T. (2000). A new dataset for measuring democracy, 1810-1998. *Journal of Peace Research*, 37(2):251–265.
- Wike, R., Simmons, K., Stokes, B., and Fetterolf, J. (2017). *Globally, Broad Support for Representative and Direct Democracy*. PEW Research Center.

# For online publication

## A Supplementary material — Section 2.4

A key objective of our study is to figure out which of the available aggregation methods are suited for the creation of democracy indices. In this section, we provide some methodological details about the five aggregation techniques that we compare in our analysis.

### A.1 Additive approach

The main difficulty in creating an additive index

$$d^{\text{ad}} = \omega_1 \cdot x_1 + \dots + \omega_m \cdot x_m \quad \text{with} \quad \sum_{j=1}^m \omega_j = 1 \quad \text{and} \quad (x_1, \dots, x_m) \in [0, 1]^m \quad (\text{A.1})$$

is to select the weights  $(\omega_1, \dots, \omega_m)$  for the regime characteristics  $(x_1, \dots, x_m)$ . The related literature suggests two weighting methods: (i) assigning the same weight to all regime characteristics, or (ii) extracting the weights from a principle component analysis (PCA). We use the second approach because our regime characteristics are unequally distributed between our three dimensions of democracy (political competition, political participation, and freedom of opinion).

### A.2 Multiplicative approach

Creating a multiplicative index

$$d^{\text{mu}} = x_1^{\omega_1} \cdot \dots \cdot x_m^{\omega_m} \quad \text{with} \quad \sum_{j=1}^m \omega_j = 1 \quad \text{and} \quad (x_1, \dots, x_m) \in [0, 1]^m \quad (\text{A.2})$$

also requires the selection of weights  $(\omega_1, \dots, \omega_m)$ . For the sake of consistency, we use the same weighting scheme as for the additive index.

### A.3 Combining the additive and the multiplicative approach

Teorell et al. (2018) argue that combining an additive approach with a multiplicative approach is the best solution for the aggregation problem:

$$d^{\text{am}} = \lambda \cdot \sum_{j=1}^m \omega_j \cdot x_j + (1 - \lambda) \cdot \prod_{j=1}^m x_j^{\omega_j} \quad (\text{A.3})$$

where  $\lambda \in (0, 1)$  is the weight of the additive component. We use a PCA to calculate this weighting factor.

#### A.4 Bayesian latent variable approach

We use the R package provided by Márquez (2016) to implement the Bayesian latent variable approach. This approach requires ordinal regime characteristics. We therefore transform all continuous variables in five point scale variables.

#### A.5 Machine learning approach

In a previous study, we proposed an aggregation method that is based on a machine learning technique for pattern recognition, known as Support Vector Machines (SVM) (Gründler and Krieger, 2016). Our major objective was to avoid simple assumptions about the functional relationship between the regime characteristics and the degree of democratization. To achieve this objective, we put the aggregation problem into a non-linear optimization problem.

To use SVM for the measurement of democracy, we need some observations—called priming data—with a generally acknowledged level of democracy, based on which the machine learning algorithm can estimate the aggregation function. We argue that the degree of democratization of the most democratic and the most autocratic regimes are uncontroversial<sup>1</sup> and that these regimes can thus be used as priming data.<sup>2</sup> To find suitable regimes, we use the democracy indices developed by Teorell et al. (2018) and Pemstein et al. (2010).<sup>3</sup> We label a country-year as a highly autocratic (democratic) regime if it belongs to the lower (upper) decile of either of the two indices. Appendix

<sup>1</sup>According to Cheibub et al. (2010) and Lindberg et al. (2014), distinguishing the least democratic from the most democratic regimes is a simple exercise: most scholars agree that Sweden is a democracy and Saudi Arabia is not. We use the consensus on the regimes at the ends of the spectrum to justify the assumption that their degree of democratization is directly observable.

<sup>2</sup>A potential concern is whether learning from “extreme” cases leads to reliable democracy indices for “intermediate” cases. The case studies presented in Section 2.4 and Appendix B.1 allay this concern because they show that all aggregation methods produce similar indices for hybrid regimes.

<sup>3</sup>These indicators are most suited because of their continuous scale and their availability. The machine learning indices hardly change when we use alternative measures of democracy for labeling.



Table E.1 lists all labeled observations.<sup>4</sup>

## B Supplementary material — Section 2.5

The three case studies presented in Section 2.4 suggest that replacing an aggregation method with another method can cause significant changes in the indicated degree of democratization. We find in particular that the additive approach and the Bayesian approach often produce too optimistic indices for autocratic regimes, while the additive-multiplicative approach and the multiplicative approach often produce too pessimistic indices for democratic regimes. Furthermore, we find that the machine learning index suffers from none of these malfunctions.

### B.1 Additional country examples

One legitimate concern is that three country examples do not suffice to draw general conclusions about the performance of different aggregation methods. To partly address this concern, we present below three additional country examples (Bhutan, Kenya, and Spain).

#### B.1.1 Bhutan

Bhutan has been an independent country since August 1949. In 1953, Bhutan’s king, Jigme Dorji Wangchuk, established a parliament consisting of 150 nominated members. Even though several kings tried to democratize the regime, it took more than 50 years until the era of authoritarian rule ended (Nohlen et al., 2001, Turner et al., 2011). The first general election was held in 2008 and was won by the Bhutan Peace and Prosperity Party (DPT). In the next election, the DPT lost support and was replaced by the main opposition party (Turner and Tshering, 2014).

Figure C.1 shows the degree of democratization of Bhutan for different aggregation

---

<sup>4</sup>The priming data include 2728 country-years: 1308 are democratic and 1420 are autocratic. We use the election handbooks prepared by Dieter Nohlen (1999, 2001, 2005, 2010) and historical records to check whether the assigned labels are consistent with our concept of democracy. We found only three observations for which the classification may be controversial: The first two are the observation for Switzerland in 1952 and 1970, since women were not entitled to vote in national elections. The last is the observation for Israel in 1999, since Palestinians living in the Gaza Strip and the West Bank were not allowed to participate in the general election for the Israeli parliament. The machine learning indices remain virtually unchanged when we exclude these three country-years from the priming data.

methods. We observe that all five indicators significantly increase in 2008, the year in which the first general election was held. We also observe that the indices differ widely from each other in the 1949 – 2007 period. The additive index, the Bayesian index, and, to a lesser extent, the additive-multiplicative index suggest that Bhutan partly met democratic standards, whereas the multiplicative index and the machine learning index indicate that Bhutan was a non-democratic state. We argue that only the latter indices are compatible with a realistic concept of democracy because no federal elections were held in Bhutan before 2008.

### **B.1.2 Kenya**

Kenya gained independence from the United Kingdom in December 1963. At this time, Kenyan politics was shaped by the rivalry between the Kenyan African National Union (KANU)—the winning party of the parliamentary election in May 1963—and the Kenya African Democratic Union (KADU). After the forced unification of the KANU and the KADU in December 1964 and the ban of the Kenya People’s Union (KPU) in October 1969, the KANU was the only legal political party until December 1991 (Nohlen et al., 1999). Against all expectations, KANU won the multi-party election in December 1992. Elklit and Svensson (1997) and Nohlen et al. (1999) explain this victory with electoral fraud. The dominance of the KANU lasted for another ten years and ended after the general election in December 2002.

Figure C.2 illustrates the Kenyan level of democracy, depending on the aggregation method. All measures of democracy indicate a significant decrease in the degree of democratization in 1969 and a significant increase in 1992. We observe, however, some differences in the size of the changes because the indicators suggest different levels of democracy for the period from 1969 to 1991. The multiplicative index and the machine learning index suggest the absence of democracy, whereas the other indices suggest the existence of weak democratic structures. The latter indices are not consistent with the concept of democracy because Kenya was a single-party regime at this time.

### B.1.3 Spain

After years of restricted political competition, political patronage, and electoral fraud, General Miguel Primo de Rivera seized power in September 1923 and established—with support from the Spanish king—a military regime. Triggered by great social unrest, General Rivera stepped down in January 1930. About one year later, the Spanish king withdrew as well and thus paved the way for the proclamation of the Second Spanish Republic in April 1931 (Nohlen and Stöver, 2010). The Republican era lasted until the end of the Spanish Civil War and the establishment of a military regime led by General Franco in 1939. Under Franco’s rule, which ended with his death in 1975, no general elections were held and all political parties were banned. After a short transition period, Spain established a parliamentary monarchy in 1977. Since then, all governments were legitimated through free and fair elections (Nohlen and Stöver, 2010).

Figure C.3 shows how the Spanish level of democracy depends on the aggregation method. Two differences are notable. First, the additive indicator and the Bayesian indicator suggest that democratic principles were partly met under the rule of General Rivera (1923 – 1930), whereas the multiplicative index and the machine learning index indicate the absence of democratic structures. Because no elections took place under Rivera’s rule, we think that the former indices are incompatible with our concept of democracy. Second, some indices indicate a lack of democracy in the post-Franco era (1977 – 2016). We doubt whether this assessment is reasonable given that multi-party elections were regularly held and no remarkable restrictions on the freedom of opinion existed during this period (Nohlen and Stöver, 2010).

## B.2 Alternative regime characteristics

Another legitimate concern is that we strategically selected our regime characteristics to favor the machine learning technique. To allay this concern, we replace our regimes characteristics with the regime characteristics used by Teorell et al. (2018) and repeat our analysis.<sup>5</sup> We chose the Teorell et al. (2018) data for two reasons: first, the concept of democracy assumed by Teorell et al. (2018) resembles our concept of democracy, and

---

<sup>5</sup>Teorell et al. (2018) use five expert-based ratings, capturing information on: suffrage, the procedure to select government officials, electoral fairness, freedom of expression, and freedom of association.

second, Teorell et al. (2018) provide an additive index, a multiplicative index, and an additive-multiplicative index.

Appendix Figures C.4 – C.6 illustrate that our results about the performance of the aggregation methods are robust to a change in the regime characteristics. We find that the additive index and the Bayesian index overestimate the true level of democracy of autocratic regimes, such as the Nazi regime in Germany, the communist regime in the Soviet Union, and the military regime in Myanmar (Burma). We also observe that the multiplicative index and the additive-multiplicative index underestimate the true degree of democratization of democratic regimes, such as present-day Germany. None of these malfunctions appear when we use the machine learning index.

### B.3 Methodological explanations

How can we explain that most of the available aggregation methods produce indicators that are downward-biased (upward-biased) for democratic (autocratic) regimes? Below, we will show that the answer to this open question lies in the assumptions about the functional relationship between the regime characteristics and the level of democracy. Since the crucial arguments are similar for all concerned aggregation techniques and to keep our explanation as short and simple as possible, we restrict our attention to the additive index in the remainder of this section.

#### B.3.1 Underestimation — Democratic regimes

An additive index is based on the assumption that the level of democracy is equal to 1 only if all regime characteristics reach their maximum:

$$d^{\text{ad}} = 1 \quad \Leftrightarrow \quad x_1 = \dots = x_m = 1.$$

In our case, this assumption implies, for example, that voter turnout must be 100% and that the number of votes won by the leading party must be equal to the number of votes won by the second strongest party. Put differently, the additive index indicates a lack of democracy if eligible voters voluntarily abstain from voting, or if the leading party is simply more popular than the second strongest party. We argue that these conditions

are too demanding to be compatible with a standard definition of democracy and that additive indices thus often underestimate the true level of democracy of democratic regimes.

### **B.3.2 Overestimation — Autocratic regimes**

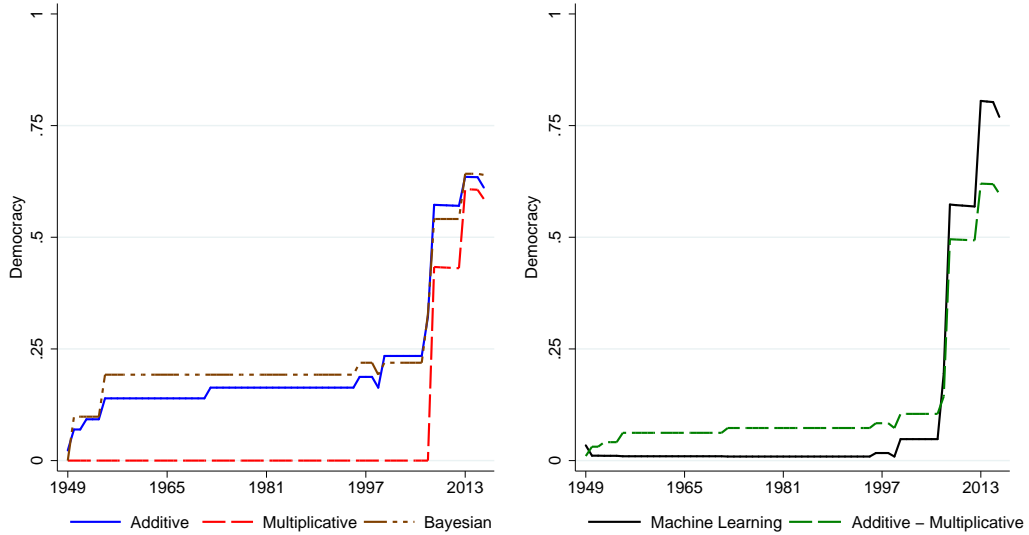
An additive index also assumes that the degree of democratization is only 0 when all regime characteristics reach their minimum:

$$d^{\text{ad}} = 0 \quad \Leftrightarrow \quad x_1 = \dots = x_m = 0.$$

In our case, this assumption implies, for example, that a regime in which single-party elections are regularly held (e.g. Soviet Union, 1937 – 1990) or a regime in which no elections take place but limited freedom of opinion exist (e.g. Bhutan, 1949 – 2007) obtain a democracy score that is greater than 0. Because any definition of democracy includes electoral competition as necessary conditions, we argue that additive indicators often overestimate the true level of democracy of autocratic regimes.

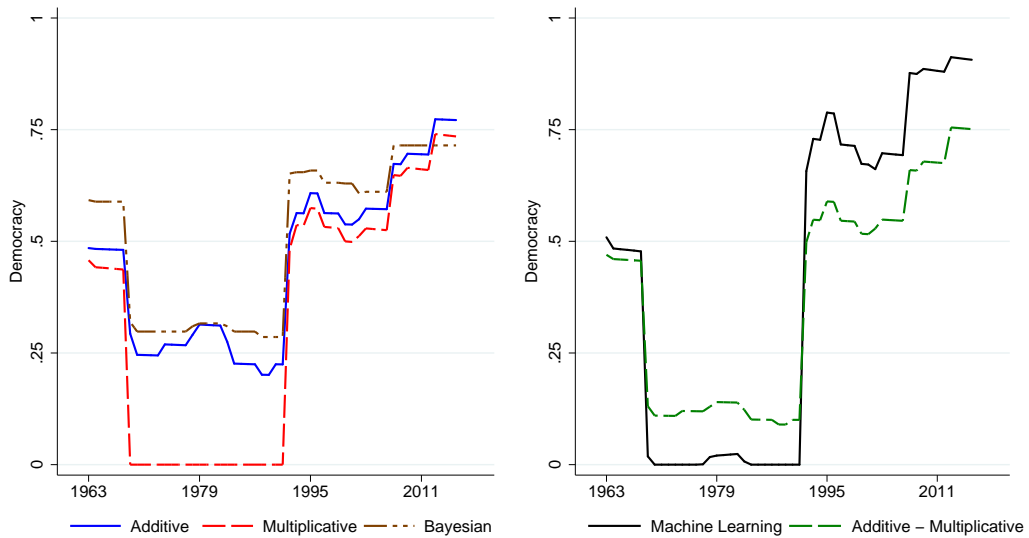
## C Additional figures

Figure C.1 Democracy in Bhutan (1949 – 2016).



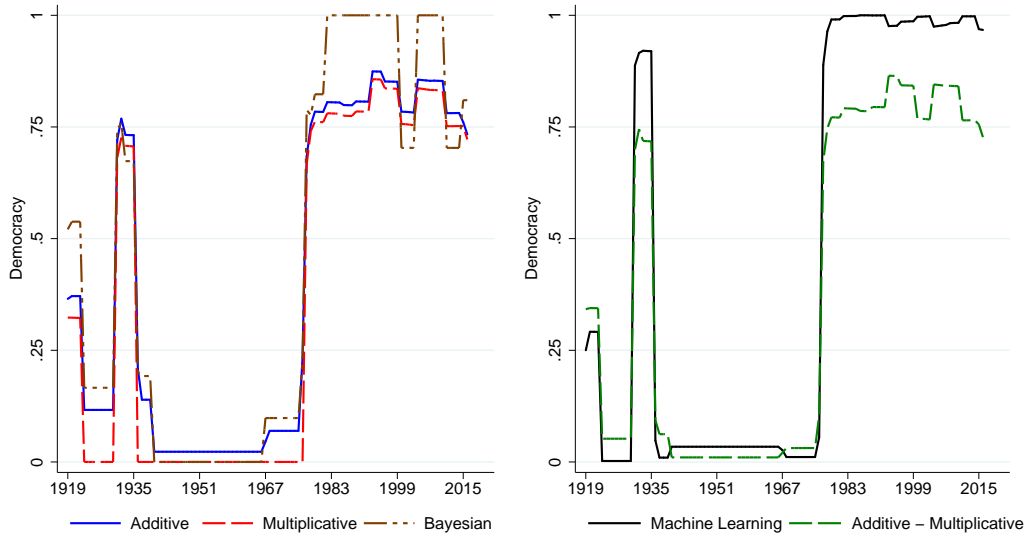
**Notes:** The figures show the level of democracy of Bhutan, depending on how we aggregate our nine regime characteristics.

Figure C.2 Democracy in Kenya (1963 – 2016).



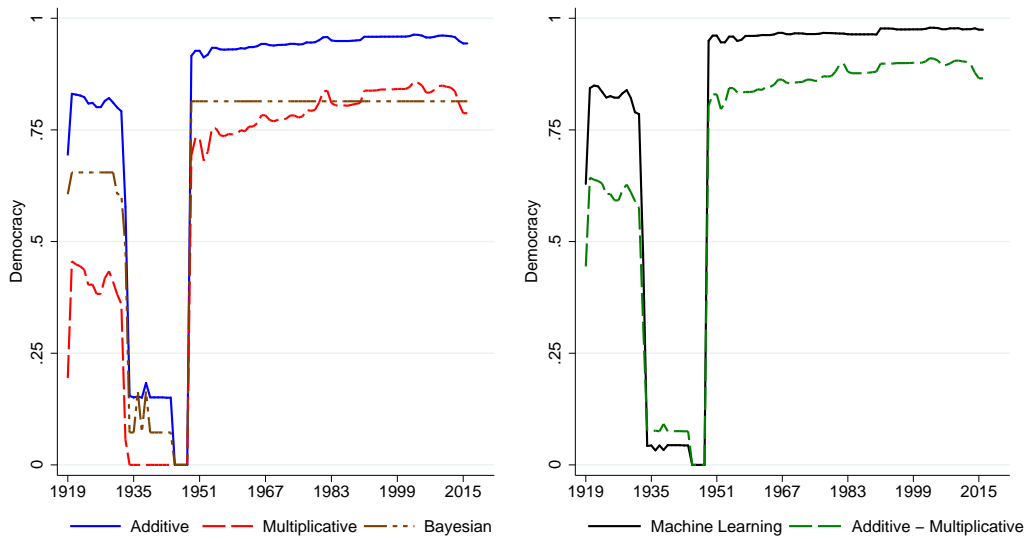
**Notes:** The figures show the level of democracy of Kenya, depending on how we aggregate our nine regime characteristics.

**Figure C.3** Democracy in Spain (1919 – 2016).



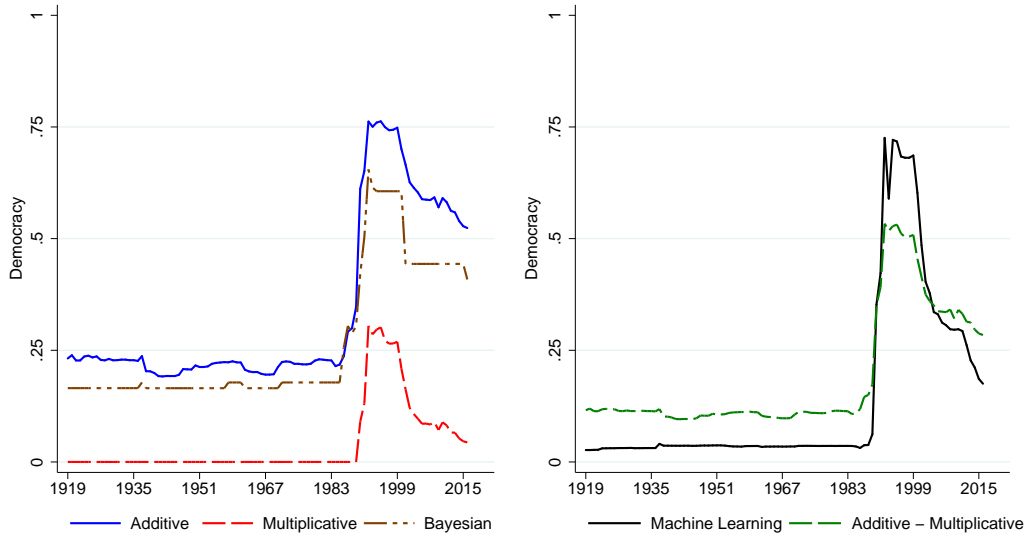
**Notes:** The figures show the level of democracy of Spain, depending on how we aggregate our nine regime characteristics.

**Figure C.4** Democracy in Germany (1919 – 2016) — Alternative regime characteristics.



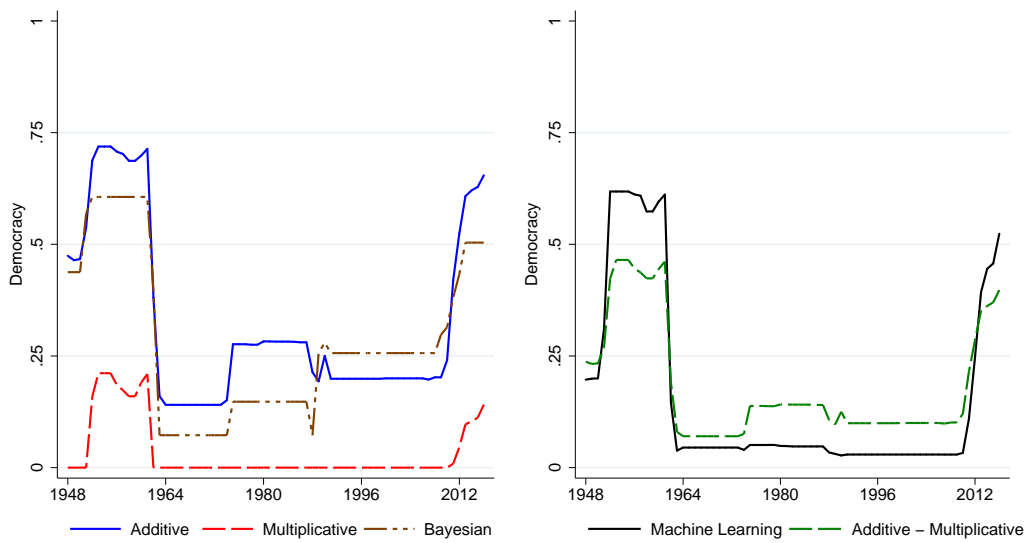
**Notes:** The figures show the level of democracy of Germany, depending on how we aggregate the five regime characteristics used by Teorell et al. (2018). For the period from 1949 to 1989, the measures of democracy refer to the Federal Republic of Germany.

**Figure C.5** Democracy in Russia (1919 – 2016) — Alternative regime characteristics.



**Notes:** The figures show the level of democracy of Russia, depending on how we aggregate the five regime characteristics used by Teorell et al. (2018). For the period from 1922 to 1990, the measures of democracy refer to the Soviet Union.

**Figure C.6** Democracy in Myanmar (1948 – 2016) — Alternative regime characteristics.



**Notes:** The figures show the level of democracy of Myanmar, depending on how we aggregate the five regime characteristics used by Teorell et al. (2018). The country was renamed from Burma to Myanmar in 1989.



## D Additional tables

Table E.1 Priming data — Selected country-years

Country	Observations	Years
<i>Democratic regimes (1308 observations)</i>		
Australia	51	1964, 1966–67, 1969–2016
Austria	63	1950–59, 1961, 1963–65, 1967–2015
Barbados	2	1981–82
Belgium	55	1955, 1963–2016
Brazil	19	1995–98, 2000–14
Canada	42	1972–81, 1983–2013, 2016
Chile	15	2002–16
Costa Rica	41	1975–80, 1982–2016
Cyprus	22	1989–2010
Czech Republic	23	1991–2013
Denmark	67	1950–2016
Estonia	24	1993–2016
Finland	35	1967–1971, 1987–2016
France	47	1970–2016
Germany	48	1950–52, 1972–2016
Greece	28	1984–2011
Hungary	8	1994–1997, 2004–05, 2007, 2009
Iceland	47	1967–2013
Ireland	40	1968, 1976–81, 1983–2015
Israel	1	1999
Italy	48	1960–61, 1963–1976, 1983–1992, 1994–2012, 2014–16
Japan	14	1980–90, 2010–14
Latvia	2	2013–14
Luxembourg	60	1951–53, 1960–2016
Malta	27	1973–75, 1989–2012
Mauritius	1	2004
Netherlands	65	1952–2016
New Zealand	54	1963–2016
Norway	49	1968–2016
Poland	24	1991–2014
Portugal	33	1984–2016
Slovakia	12	1999–2012
Slovenia	12	1997, 2002–11, 2013–15
Spain	29	1980–82, 1987–2012
St. Kitts and Nevis	1	2010
Sweden	50	1967–2016
Switzerland	48	1952, 1970–2016
Taiwan	1	2005
United Kingdom	35	1961, 1964, 1975–89, 1999–2016
United States	37	1968–69, 1976–77, 1983–86, 1988–2016
Uruguay	28	1989–2016
<i>Autocratic regimes (1420 observations)</i>		
Afghanistan	34	1950–63, 1978–91, 1996–2001
Albania	29	1951–53, 1955, 1957–61, 1965, 1970, 1981–89
Algeria	14	1965–76, 1985, 1994
Angola	18	1975–92
Argentina	12	1966–71, 1977–82
Bahrain	28	1971–73, 1976–2000
Benin	10	1974–79, 1984, 1986–87, 1989
Bhutan	57	1950–2006
Bolivia	7	1972–77, 1980
Brazil	2	1968–69
Bulgaria	8	1980, 1983–89
Burkina Faso	1	1965
Burma (Myanmar)	39	1963–1974, 1983–2008, 2010
Burundi	16	1967–81, 1988
Cambodia	16	1953, 1966–67, 1979–91
Cameroon	1	1988
Central African Rep.	15	1967–1980, 1988
Chad	17	1962, 1970, 1972–74, 1976–77, 1980–89
Chile	14	1975–88
China	44	1950–78, 1985, 1989–97, 2000, 2013–16
Democratic Rep. of Congo	16	1965–76, 1983–87, 1989
Rep. of Congo	7	1968–1972, 1978–79
Cuba	48	1958, 1960–2005, 2007
Czech Republic	5	1950–1954
Dominican Rep.	10	1950–1953, 1955–1960
Egypt	3	1954–56
Equatorial Guinea	17	1973–82, 1985–91
Eritrea	15	2002–16
Ethiopia	32	1950–59, 1961–62, 1965, 1970–72, 1974–86, 1988–90
⋮	⋮	⋮
⋮	⋮	⋮

**Table E.1** Priming data — Selected country-years (continued)

Country	Observations	Years
⋮	⋮	⋮
Gabon	1	1968
Germany (East)	25	1953, 1960–66, 1969, 1972–88
Ghana	1	1965
Greece	6	1968–73
Guatemala	3	1956, 1964–54
Guinea	18	1958–60, 1967–68, 1972–84
Guinea-Bissau	2	1980, 1985
Haiti	20	1950, 1963–64, 1967, 1971–84, 1992–93
Indonesia	1	1965
Iran	19	1953–64, 1966–1974, 1976
Iraq	40	1963–2002
Ivory Coast	1	1966
Jordan	22	1950, 1957–59, 1961–64, 1967–68, 1970, 1974–83
Korea (North)	59	1958–2016
Kuwait	10	1965–66, 1976, 1979–80, 1986–1990
Laos	40	1976–89, 1991–2016
Lesotho	2	1970–71
Liberia	2	1980–81
Libya	45	1952–55, 1970–2010
Malawi	22	1964, 1966, 1972–87, 1989–92
Maldives	4	1965–68
Mali	7	1975–78, 1980–81, 1985
Mauritania	4	1979–82
Mongolia	9	1950–51, 1980, 1983–88
Morocco	12	1956–1962, 1965–69
Mozambique	15	1976–1990
Nepal	10	1950–51, 1960–65, 1967–68
Niger	4	1979–82
Nigeria	1	1966
Oman	31	1970–2000
Pakistan	1	1980
Panama	3	1969–71
Paraguay	5	1954, 1956–57, 1960–61
Peru	1	1969
Philippines	5	1973–77
Portugal	16	1951–54, 1956–59, 1961–64, 1966–68, 1970
Qatar	47	1970–2016
Romania	6	1983–88
Russia (Soviet Union)	9	1950–52, 1967–69, 1984–86
Rwanda	5	1974–1977, 1980
Saudi Arabia	67	1950–2016
Serbia (Yugoslavia)	10	1950, 1954–57, 1975–79
Somalia	19	1970–80, 1983–90
Spain	18	1950–67
Sudan	16	1959–63, 1989–97, 1999, 2001
Swaziland	11	1974–77, 1984, 1986–87, 1989–92
Syria	28	1961, 1965, 1970–73, 1983–2002, 2011–12
Taiwan	20	1950–69
Togo	14	1967–79, 1985
Tunisia	5	1956–1959, 1962
Turkmenistan	20	1992–2012
Uganda	9	1972–1980
United Arab Emirates	41	1971–2011
Uruguay	3	1976–78
Uzbekistan	16	1995–06, 2008, 2010–12
Venezuela	3	1954–56
Vietnam	20	1954–59, 1985–86, 1989–2000
Yemen (Yemen North)	31	1950–70, 1978–1987

**Notes:** This table reports the country-years that are part of the priming data. The selection is based on the indices developed by Pemstein et al. (2010) and Teorell et al. (2018).

**Table E.2** Average changes in the level of democracy — Regime changes

	Machine Learning	Additive	Multiplicative	Add. / Multi.	Bayesian
$\Delta^j$	0.437	0.287	0.350	0.322	0.305

**Notes:** The table reports results from estimating

$$\Delta^j = \frac{1}{|\mathcal{D}|} \sum_{(i,t) \in \mathcal{D}} |d_{i,t}^j - d_{i,t-1}^j|$$

where  $d_{i,t}^j$  denotes the level of democracy of country  $i$  in year  $t$  indicated by index  $j$ . The set  $\mathcal{D}$  includes all country-year observations for which Boix et al. (2013) report a political transition.

**Table E.3** Democracy and economic growth — OLS estimates, five-year data

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Democracy	0.055*** (0.0144)	0.102*** (0.0221)	0.078*** (0.0195)	0.091*** (0.0208)	0.090*** (0.0225)
Observations	2,250	2,250	2,250	2,250	2,250
Countries	163	163	163	163	163
R <sup>2</sup>	0.982	0.982	0.982	0.982	0.982
Equal. (p-val.)	-	0.037	0.258	0.093	0.121

**Notes:** The table presents OLS estimates. The dependent variable is the log of GDP per capita. All regressions include the first lag of the dependent variable, country fixed effects, and period fixed effects. All democracy indicators are continuous and range from 0 to 1. The data is averaged over five-year periods (1920–24, 1925–29, ...). Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.4** Democracy and economic growth — 2SLS estimates

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A: Second-stage estimates					
Democracy	0.135*** (0.0256)	0.223*** (0.0407)	0.173*** (0.0328)	0.192*** (0.0358)	0.241*** (0.0449)
Equal. (p-val.)	-	0.030	0.242	0.107	0.018
Panel B: First-stage estimates					
Demo. (reg.)	0.757*** (0.0615)	0.457*** (0.0373)	0.589*** (0.0475)	0.530*** (0.0421)	0.422*** (0.0356)
Equal. (p-val.)	-	0.000	0.001	0.000	0.000
Observations	2,244	2,244	2,244	2,244	2,244
Countries	163	163	163	163	163
SaWi (F.stat.)	151.44	149.54	153.56	158.64	140.55
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000

**Notes:** The table presents 2SLS estimates. The dependent variable is the log of GDP per capita. All regressions include the first lags of the dependent variable, country fixed effects, and period fixed effects. All democracy indicators are continuous and range from 0 to 1. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. The data is averaged over five-year periods (1920–24, 1925–29, ...). We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.5** Democracy and economic growth — OLS estimates, subsample analysis (continent)

	Machine Learning	Additive	Multiplicative	Add./ Multi.	Bayesian
	(1)	(2)	(3)	(4)	(5)
Panel A: Exclude Africa					
Democracy	0.015*** (0.0035)	0.027*** (0.0054)	0.021*** (0.0043)	0.024*** (0.0048)	0.025*** (0.0048)
Observations	7,539	7,539	7,539	7,539	7,539
Countries	112	112	112	112	112
Equal. (p-val.)	-	0.041	0.252	0.094	0.041
Panel B: Exclude America					
Democracy	0.020*** (0.0035)	0.034*** (0.0056)	0.028*** (0.0046)	0.032*** (0.0050)	0.030*** (0.0051)
Observations	8,807	8,807	8,807	8,807	8,807
Countries	136	136	136	136	136
Equal. (p-val.)	-	0.014	0.099	0.023	0.055
Panel C: Exclude Asia					
Democracy	0.013*** (0.0029)	0.021*** (0.0047)	0.019*** (0.0039)	0.021*** (0.0043)	0.021*** (0.0046)
Observations	8,249	8,249	8,249	8,249	8,249
Countries	120	120	120	120	120
Equal. (p-val.)	-	0.067	0.114	0.062	0.079
Panel D: Exclude Europe					
Democracy	0.019*** (0.0031)	0.028*** (0.0049)	0.025*** (0.0041)	0.028*** (0.0045)	0.026*** (0.0047)
Observations	8,164	8,164	8,164	8,164	8,164
Countries	123	123	123	123	123
Equal. (p-val.)	-	0.056	0.127	0.053	0.143

**Notes:** The table presents OLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.6** Democracy and economic growth — Second-stage estimates, subsample analysis (continent)

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A: Exclude Africa					
Democracy	0.029*** (0.0064)	0.050*** (0.0105)	0.038*** (0.0085)	0.043*** (0.0092)	0.058*** (0.0122)
Observations	7,514	7,514	7,514	7,514	7,514
Countries	112	112	112	112	112
SaWi (F.stat.)	77.13	80.77	71.79	77.89	73.21
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.043	0.307	0.147	0.018
Panel B: Exclude America					
Democracy	0.030*** (0.0055)	0.050*** (0.0084)	0.038*** (0.0067)	0.043*** (0.0074)	0.051*** (0.0085)
Observations	8,411	8,411	8,411	8,411	8,411
Countries	136	136	136	136	136
SaWi (F.stat.)	215.05	203.01	253.14	252.75	204.21
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.018	0.264	0.097	0.017
Panel C: Exclude Asia					
Democracy	0.023*** (0.0045)	0.038*** (0.0070)	0.029*** (0.0056)	0.033*** (0.0062)	0.041*** (0.0077)
Observations	8,235	8,235	8,235	8,235	8,235
Countries	120	120	120	120	120
SaWi (F.stat.)	153.96	158.60	153.14	162.62	144.16
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.034	0.264	0.119	0.018
Panel D: Exclude Europe					
Democracy	0.038*** (0.0063)	0.059*** (0.0093)	0.049*** (0.0080)	0.053*** (0.0085)	0.063*** (0.0101)
Observations	8,152	8,152	8,152	8,152	8,152
Countries	123	123	123	123	123
SaWi (F.stat.)	99.05	108.62	99.82	105.48	103.87
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.027	0.167	0.080	0.012

**Notes:** The table presents 2SLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.7** Democracy and economic growth — OLS estimates, subsample analysis (periods)

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A: 1970 – 2016					
Democracy	0.021*** (0.0030)	0.037*** (0.0055)	0.028*** (0.0041)	0.033*** (0.0047)	0.033*** (0.0054)
Observations	7,041	7,041	7,041	7,041	7,041
Countries	163	163	163	163	163
Equal. (p-val.)	-	0.005	0.092	0.014	0.024
Panel B: 1946 – 2016					
Democracy	0.015*** (0.0026)	0.026*** (0.0044)	0.021*** (0.0034)	0.024*** (0.0038)	0.024*** (0.0042)
Observations	9,679	9,679	9,679	9,679	9,679
Countries	163	163	163	163	163
Equal. (p-val.)		0.019	0.094	0.026	0.039

**Notes:** The table presents OLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.8** Democracy and economic growth — Second-stage estimates, subsample analysis (periods)

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A: 1970 – 2016					
Democracy	0.042*** (0.0056)	0.072*** (0.0090)	0.052*** (0.0069)	0.060*** (0.0076)	0.076*** (0.0096)
Observations	7,020	7,020	7,020	7,020	7,020
Countries	163	163	163	163	163
SaWi (F.stat.)	254.10	205.49	276.11	264.99	208.24
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.001	0.129	0.020	0.000
Panel B: 1946 – 2016					
Democracy	0.027*** (0.0047)	0.045*** (0.0075)	0.035*** (0.0060)	0.039*** (0.0066)	0.048*** (0.0082)
Observations	9,656	9,656	9,656	9,656	9,656
Countries	163	163	163	163	163
SaWi (F.stat.)	202.64	212.95	204.72	221.09	191.88
AR (p-val.)	0.000	0.000	0.000	0.000	0.000
StWr (p-val.)	0.000	0.000	0.000	0.000	0.000
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.021	0.210	0.085	0.010

**Notes:** The table presents 2SLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, country fixed effects, and year fixed effects. All democracy indicators are continuous and range from 0 to 1. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .



**Table E.9** Democracy and economic growth — OLS estimates, additional controls

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A — Controls: conflict, pop. growth, education, institutional quality					
Democracy	0.008** (0.0037)	0.022*** (0.0062)	0.013*** (0.0047)	0.0173*** (0.0054)	0.019*** (0.0054)
Observations	8,908	8,908	8,908	8,908	8,908
Countries	140	140	140	140	140
Equal. (p-val.)	-	0.029	0.360	0.116	0.065
Panel B — Controls: Panel A + government consumption, investment, trade openness					
Democracy	0.006* (0.0031)	0.015** (0.0060)	0.009** (0.004)	0.012** (0.0049)	0.014*** (0.0052)
Observations	7,094	7,094	7,094	7,094	7,094
Countries	137	137	137	137	137
Equal. (p-val.)	-	0.123	0.399	0.210	0.105

**Notes:** The table presents OLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, the first lag of various control variables (civil conflict, education, population growth, court independence). In Panel B, we additionally control for government consumption, investment, and trade openness. All democracy indicators are continuous and range from 0 to 1. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.10** Democracy and economic growth — Second-stage estimates, additional controls

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A — Controls: conflict, pop. growth, education, institutional quality					
Democracy	0.021** (0.0089)	0.038** (0.0165)	0.026** (0.0112)	0.031** (0.0130)	0.044** (0.0188)
Observations	8,908	8,908	8,908	8,908	8,908
Countries	140	140	140	140	140
SaWi (F.stat.)	54.33	44.61	59.09	55.89	41.29
AR (p-val.)	0.026	0.026	0.026	0.026	0.026
StWr (p-val.)	0.013	0.013	0.013	0.013	0.013
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.285	0.640	0.461	0.226
Panel B — Controls: Panel A + government consumption, investment, trade openness					
Democracy	0.023** (0.0098)	0.042** (0.0182)	0.029** (0.0123)	0.034** (0.0143)	0.047** (0.0202)
Observations	7,094	7,094	7,094	7,094	7,094
Countries	137	137	137	137	137
SaWi (F.stat.)	53.44	44.72	56.31	54.59	43.17
AR (p-val.)	0.018	0.018	0.018	0.018	0.018
StWr (p-val.)	0.007	0.007	0.007	0.007	0.007
KP (p-val.)	0.000	0.000	0.000	0.000	0.000
Equal. (p-val.)	-	0.296	0.638	0.464	0.248

**Notes:** The table presents 2SLS estimates. The dependent variable is the log of GDP per capita. All regressions include four lags of the dependent variable, the first lag of various control variables (civil conflict, education, population growth, court independence). In Panel B, we additionally control for government consumption, investment, and trade openness. All democracy indicators are continuous and range from 0 to 1. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero:  $*p < .10$ ,  $**p < .05$ ,  $***p < .01$ .

**Table E.11** Democracy and education — OLS estimates, five-year data

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Democracy	0.212*** (0.0441)	0.348*** (0.0606)	0.273*** (0.0558)	0.315*** (0.0595)	0.282*** (0.0555)
Observations	1,979	1,979	1,979	1,979	1,979
Countries	141	141	141	141	141
R <sup>2</sup>	0.989	0.989	0.989	0.989	0.989
Equal. (p-val.)	-	0.027	0.280	0.084	0.207

**Notes:** The table presents OLS estimates. The dependent variable is the average years of schooling. All regressions include the first lag of the dependent variable, the first lag of the log of GDP per capita, country fixed effects, and period fixed effects. All democracy indicators are continuous and range from 0 to 1. The democracy indices are averaged over five-year periods (1920–24, 1925–29, ...) since annual data on the average years of schooling does not exist. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .

**Table E.12** Democracy and education — 2SLS estimates, five year data

	<b>Machine Learning</b>	<b>Additive</b>	<b>Multiplicative</b>	<b>Add./ Multi.</b>	<b>Bayesian</b>
	(1)	(2)	(3)	(4)	(5)
Panel A: Second-stage estimates					
Democracy	0.337*** (0.1282)	0.569*** (0.2093)	0.436*** (0.1679)	0.487*** (0.1839)	0.650*** (0.2495)
Equal. (p-val.)	-	0.267	0.556	0.415	0.210
Panel B: First-stage estimates					
Demo. (reg.)	0.599*** (0.0666)	0.355*** (0.0398)	0.463*** (0.0513)	0.415*** (0.0450)	0.311*** (0.0380)
Equal. (p-val.)	-	0.000	0.009	0.000	0.000
Observations	1,976	1,976	1,976	1,976	1,976
Countries	141	141	141	141	141
SaWi (F.stat.)	81.15	79.52	81.60	85.14	66.90
AR (p-val.)	0.011	0.011	0.011	0.011	0.011
StWr (p-val.)	0.008	0.008	0.008	0.008	0.008
KP (p-val.)	0.000	0.000	0.000	0.000	0.000

**Notes:** The table presents 2SLS estimates. The dependent variable is the average years of schooling. All regressions include the first lag of the dependent variable, the first lag of the log of GDP per capita, country fixed effects, and period fixed effects. All democracy indicators are continuous and range from 0 to 1. The democracy indices are averaged over five-year periods (1920–24, 1925–29, ...) since annual data on the average years of schooling does not exist. The first lag of the regional (jack-knifed) degree of democratization serves as the instrument for the domestic degree of democratization. We report different first-stage diagnostics to indicate the strength and validity of our instrumental variable. Standard errors clustered by country are reported in parentheses. We report results from a Wald test to show whether the estimates reported in Columns 2 – 5 are significantly different from the estimates reported in Column 1. The following notation is used to highlight coefficients that are significantly different from zero: \* $p < .10$ , \*\* $p < .05$ , \*\*\* $p < .01$ .