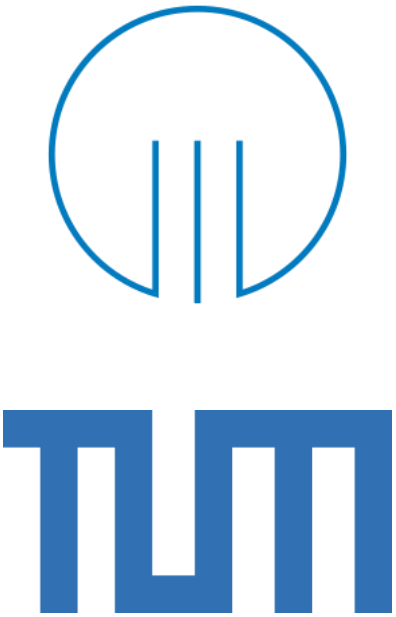
Technische Universität München
Hochschule für Politik
Richard-Wagner-Straße 1, 80333 München



Master's Thesis

A Machine Learning Approach to Measuring Democracy

Author
Name: Joao Lucas Dziekan Rodrigues Hilgert
[redacted] [redacted]

Supervisor
Name: [redacted]
Chair: [redacted]

Date of Submission: 02.07.2022

Table of Contents

1. Introduction

There are major challenges that revolve around measuring democracy. First and foremost, and arguably most obviously, how broad the concept should be. And it shouldn't be overlooked, since it is crucial for related research (Gründler & Krieger, 2021; Merkel, 2004; Pérez-Liñán, 2017, 2020). Often it requires "an operational definition of democracy precise enough to classify specific countries as we observe during particular historical periods" (Pérez-Liñán, 2017, p. 86). The minimalist, i.e. narrow, approaches focus mainly on competitive and public elections for political office. On the other end there are the maximalist, i.e. broad, approaches that might include socioeconomic aspects, e.g. civil liberties, inequality, rule of law, and even economic metrics. In between that are a wide gamma of variations. (Gründler & Krieger, 2021; Merkel, 2004; Munck & Verkuilen, 2002; O'Donnell, 2011). Conceptually speaking, all aforementioned approaches are equally valid, since there is little in the way of an objective method to evaluate them (Gründler & Krieger, 2021; Guttman, 1994; Munck & Verkuilen, 2002). Another issue relates to the object being researched. If, for instance, the research revolves around the relationship between democracy and economy, a definition that includes economic metrics might not be appropriate as it could skew results. This means that the answer as to what type of conceptualization is more appropriate can often be contextual (Bjørnskov & Rode, 2020; Gründler & Krieger, 2021; Gutmann & Voigt, 2018).

Secondly, there remains the question on how the different dimension relate to each other. In this instance there are basically two approaches, namely *formative* and *reflective* (Gründler & Krieger, 2021; Teorell et al., 2019). The former assumes that every condition of democracy is necessary. For instance, if election results are completely fabricated or the right to candidacy is severely restricted, the degree of suffrage amongst the citizenry is not relevant (Gründler & Krieger, 2021; Teorell et al., 2019). This in turn means that a "low score on any of the component indices thus suppresses the value of the overall index" (Teorell et al., 2019, p. 81). The latter assumes that all conditions are the result of a common factor and hence partially interchangeable. It follows that the aggregation rule should be additive and not multiplicative. Different than with the formative approach, one low value in the reflective approach would have less of an impact on the resulting index. Much like the breadth of the concept, there isn't an objective answer to the question on which approach is best. Teorell et al. (2019) argue that a combination of the two alleviates issues with both, though recent research indicates that that might not be the case (Gründler & Krieger, 2020). This was also further

explored in this contribution in Section 6. They all have their own merits, however the distinction should not be overlooked. From the perspective of providers of democracy indices, the choice in approach directly affects the aggregation of data down the line. For users, however, it is important to be cognizant of the approach used, because it affects how data is interpreted (Gründler & Krieger, 2021; Teorell et al., 2019).

Lastly, it is important to be cognizant of the resulting scale type. It can be dichotomous, ordinal, graded or continuous. Each of them has strengths and weaknesses. However, dichotomous scales lose a lot of discriminating power (Elkins, 2000; Gründler & Krieger, 2021). For instance, if one would take the continuous scales of the *Varieties of Democracy*'s (V-Dem) *Electoral Democracy Index* (EDI) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021) and the *Freedom House* (FH) indicator (Freedom House, 2021) in the year 2020, from the 173 countries included in both datasets a fourth of them would fall in the middle third of both scales.

All aforementioned challenges play a role when deciding how to measure democracy. Results vary depending on the scope of the conceptualization, the aggregation, and the scaling (Elkins, 2000; Gründler & Krieger, 2016, 2020, 2021; Munck & Verkuilen, 2002). However, some of these concerns can be alleviated. For instance, as just alluded to, with growing apprehension over hybrid regimes, i.e. "political regimes that fall 'somewhere in between' full democracy and overt dictatorship" (Pérez-Liñán, 2020, p. 90), dichotomous measures often lack the discriminatory power to make the necessary distinction. Obviously, there is little debate on how to categorize countries in the extreme, e.g. Switzerland or the Democratic People's Republic of Korea in 2020. But a dichotomous scale lacks the nuance to classify borderline cases (Pérez-Liñán, 2017, 2020). Furthermore, they are not inherently more reliable than graded measures, for instance (Elkins, 2000; Gründler & Krieger, 2016, 2021). Lastly, dichotomous scales overlook the process of democratization over time due to their lack of gradation (Doucouliagos & Ulubasoglu, 2007; Gründler & Krieger, 2016). Hence, there are objective reasons that can be rationalized in the pursuit of choosing a scale.

The same cannot necessarily be said for conceptualization and the aggregation. As mentioned previously the former is based on non-objective criteria, and the latter to an extent depends on the former. It is important to understand the relationship between the conditions for democracy, however that is dependent on which conditions are chosen. Furthermore, many of the current, notorious indices choose their aggregation mechanisms rather arbitrarily (Gründler & Krieger, 2016, 2021; Munck & Verkuilen, 2002). Building on the work by Gründler and Krieger (2016, 2021), this contribution attempted to partially overcome this issue by employing

novel approaches to measure democracy based on *machine learning* (ML). There are multiple advantages. For one, these sorts of tools have a more pronounced ability to analyze datasets with a broad degree of dimensions and features. Further, they are better equipped to recognize patterns in the data and learn from them, without being explicitly programmed beforehand (Gründler & Krieger, 2016; Samuel, 1959).

For the purposes of this contribution, and building on aforesaid research, this contribution employed *Support Vector Machines* (SVM). The practical applications thereof can hardly be overstated. As pointed out by Gründler and Krieger (2016, p. 89), it has been employed with promising results in a plethora of other applications in a variety of scientific fields (Cortes & Vapnik, 1995; Gualtieri, n.d.; Guyon et al., 2002; Joachims, 2002; Orrù et al., 2012). However, in social sciences researchers have only relatively recently begun to understand and benefit from the use of machine learning (Grimmer, 2015; Grimmer et al., 2021; Molina & Garip, 2019). In this case, it provided with a suitable strategy to find the usually subjective and inherently unknowable aggregation function, hence avoiding arbitrary assumptions, and improving the general gradation quality of the democracy indicator. Machine learning algorithms in general and SVM in particular are designed to solve problems where the functional form is not known and where the functional relationship between input and output is not privy to researchers (Gründler & Krieger, 2016; Steinwart & Christmann, 2008).

The most appropriate aggregation function was found by transferring the problem of aggregation to the context of optimization. In order to do it, the necessary conditions were twofold. It was necessary to have input data for all observations in the dataset and further a limited set of observations with a known output, which was used to train the algorithm. To that end a set of countries that were indisputably held as highly democratic and highly autocratic were selected as priming data. These were distinguished from other regimes by using a set of two commonly accepted indices, namely the *Unified Democracy Score* (UDS) (Marquez, 2022; Pemstein et al., 2010) and EDI (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). Countries that belonged to the upper or lower decile of either index's scale was used as priming data. This data was validated in two ways. First, the priming data was compared to other established indices, to check for any noticeable incongruencies (Gründler & Krieger, 2016, 2021). Second, previous research and reports were used to confirm, in cases of incongruencies, that there weren't any regimes in the priming data that were outwardly unreasonable. The countries in the priming data, both democracies and autocracies, were checked for representativeness to make sure that it is not excluding regimes with specific

characteristics, such as regime type or durability. An SVM regression algorithm was then used to obtain the aggregation function (Gründler & Krieger, 2016, 2021).

The regression was applied to four datasets. One took the same variables as an existing index, namely the EDI. This was done to isolate the aggregation functions and assess their distinct effects on the overall indices. The other three encompassed variables that related to different breadths of conceptualizations of democracy as proposed by Wolfgang Merkel (Merkel, 2004) and his concept of *Embedded Democracy* offers key advantages. Due to its arguably modular characteristic, it is possible to test for different changes in conceptualization, without altering the core features, i.e. the *Electoral Regime*. It also offers the opportunity to pinpoint whether specific partial regimes play a larger role in the changes of the gradation of democracies. Further, it allows for the determination of changes in specific partial regimes, that might be overlooked by more minimalistic approaches. Hence, the extent to which an SVM based democracy index could effectively measure democracy and identify its different gradations was assessed by comparing the distinct SVM indices, not only to each other, but also to the already existing *Support Vector Machine Democracy Index* (SVMDI) by Gründler & Krieger (2021), which was described further in Section 2, and upon which this contribution attempted to improve.

Some of the data used in this contribution has already been mentioned, however the main data used to apply the SVM, and measure democracy's partial regimes came from the V-Dem Dataset (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). For starters, the dataset encompasses 202 countries in the period of 1789 to 2020. Though not every country-year combination was used, the wide range of data was a beneficial asset. Furthermore, it encompasses 470 unique indicators, 363 of which are coded from 1900 on (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Gründler & Krieger, 2021). This contribution only included a small fraction of them, i.e. only those that related to the internal partial regimes and external evironment as defined by Merkel (2004), as well as no aggregated metrics to avoid the aggregation issue mentioned previously. Further, both objective data as well as expert-based subjective data were included, which in turn satisfied Munck and Verkuilen's (2002) guidelines for the use of regime characteristics (Gründler & Krieger, 2021). This contribution benefited from V-Dem's transparency regarding their data and their coding (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Gründler & Krieger, 2021).

This contribution, hence, used the aforementioned tools and assets in the pursuit of answering the question: To what extent can a machine learning based approach to measuring democracy be applied to broader conceptualizations of democratic regimes? In that effort this contribution isolated the aggregation function to assess its potential benefits and pitfalls in relation to traditional aggregation procedures. It also built upon Gründler and Krieger's (2016, 2021) concept of SVMDI by expanding the defining characteristic of democracy and modularizing different areas of democratic conceptualization. This was done to isolate characteristics of democracy that can be conceptually related to other fields of study, without changing the core of the minimalistic concept. Finally, this contribution also attempted to determine, whether different breadths of democratic conceptualization, when aggregated via SVM regression, can reasonably detect variations in democratic quality.

2. Literature Review

2.1. Conceptualization

The conceptual breadth of democracy is a very controversial topic, where "it is rather unlikely to expect agreement on the 'one' concept" (Geissel et al., 2016, p. 572). However, this is still important to discern. Both narrow and broad (and everything in between) conceptions have their merit and pitfalls (Gründler & Krieger, 2020, 2021; Munck & Verkuilen, 2002; O'Donnell, 2001). Depending on the kind of research, a narrower approach might be more appropriate as to avoid conceptual overlap and avoid insufficient data availability. A narrow concept, or a too restrictive criteria selection, might lead an indicator do lose discriminatory power and overlook developments in democratization or democratic backsliding (Gründler & Krieger, 2020, 2021; Munck & Verkuilen, 2002). "If the definition of democracy involves only a low standard, then differences among established democracies can hardly be identified. In other words, gradations of democratic quality cannot be detected" (Geissel et al., 2016, p. 572). This contribution, hence, did not intend to outright answer this controversial question, but to shed a light on the perils of a low-standard selection criteria for democratic characteristics (Geissel et al., 2016; Gründler & Krieger, 2016; Merkel, 2004).

For the purposes of this contribution Wolfgang Merkel's (Merkel, 2004) concept of Embedded Democracy was employed. It presupposes, that democratic systems are embedded in two ways. For one, it is embedded internally, where "the specific interdependence/independence of the different partial regimes of a democracy secures its normative and functional existence" (Merkel, 2004, p. 36). Furthermore, it is also embedded

externally, where "these partial regimes are embedded in spheres of enabling conditions for democracy that protect it from outer as well as inner shocks and destabilizing tendencies" (Merkel, 2004, p.36).

2.1.1. Internal Embeddedness

There are 5 partial regimes of a liberal democracy: *political rights, civil rights, horizontal accountability, the effective power to govern,* and *a democratic electoral regime* (Merkel, 2004). These regimes are embedded internally, meaning that one regime is dependent on other regimes. They support each other's functioning, e.g. civil rights and rights of participation supporting the conduction of democratic elections. However, they also guarantee that the political actors of one does not infringe on the other, e.g. by guaranteeing horizontal accountability between powers, the likelihood of a breach of civil rights or effective power to govern can be dramatically reduced. Meaning that these regimes are simultaneously independent and interdependent (Merkel, 2004).

For starters, the political rights regime is a precondition for elections (Merkel, 2004). At its core it relates to the freedom of expression and freedom of association. Societies are complex and this regime enables diverse opinions and positions to be discussed publicly and to coordinate in order to influence the political landscape. Public and private media should not only exist, but also be independent from political pressure (Merkel, 2004). Furthermore, citizens should have the freedom to form organizations and coordinate in order to influence political outcomes, without interference. This includes political parties, which should have the freedom to coordinate and express themselves, without politically motivated restrictions (Merkel, 2004). This guarantees that the recurring check on political power, embodied by elections, is supplanted by a soft and steady check in between them (Merkel, 2004).

The civil rights regime protects individual rights from the reach of the state or even a 'tyranny of the majority' (Merkel, 2004). Rule of law has to be guaranteed, i.e. the state's power is limited by upholding the laws effectively according to clearly delineated prerogatives, as well as protecting certain laws from the reach of the legislature, the executive and even a majority of citizens. Basic constitutional rights lie at the core of the rule of law (Merkel, 2004). "Individual rights of protection grant legal protection of life, freedom and property – the threefold meaning of Locke's term property – as well as protection against illegitimate arrest, exile, terror, torture or unjustifiable intervention into personal life, both on behalf of the state and on behalf of private or anti-establishment forces and individual actors. Equal access to the law and equal treatment by the law are basic civil rights" (Merkel, 2004, p. 40). However, this

alone does not suffice. To uphold the three powers, i.e. legislature, executive and judiciary, need to be distinct and accountable to each other (Merkel, 2004).

The regime of division of powers and horizontal accountability makes sure that no power encroaches on the political and civil liberties of the citizens (Merkel, 2004). The exercise of executive power is particularly limited in this case. To that end an independent and functional judiciary is necessary, which can review executive and legislative action. Though, in the concept of Embedded Democracy, both civil rights and horizontal accountability require an independent judiciary (Merkel, 2004), for the purposes of this contribution the independence of the judiciary was included exclusively in horizontal accountability, as it relates more closely to how the three powers relate to each other. Vertical accountability, i.e. elections, control the government periodically. Civil rights create a barrier against state abuse towards the individual. Horizontal accountability implies that the three bodies check on each other reciprocally, however none of them should dominate or interfere with the core-sphere of responsibility of others (Merkel, 2004). These constitutional powers maintain a constant check on the responsiveness and accountability of the government (Merkel, 2004).

The partial regime of the effective power to govern, as the name already suggests, underscores the requirement that those who are elected to high office are also those that govern in the end (Merkel, 2004). While in a not-too-distant past this might have seemed self-evident, in the present more complex power structure make it so that this partial regime cannot be taken for granted. Extra-constitutional actors, e.g. the military or other powerful actors not subject to democratic accountability, should not have final decision-making power (Merkel, 2004).

Lastly, but arguably most importantly, the electoral regime grants access to positions of state power incumbent on the results of open and competitive electoral contests (Merkel, 2004). This regime is positioned at the center of all other regimes. It is the *sine qua non* that at its simplest separates potential democracies from autocracies (Merkel, 2004). This regime borrows from Dahl's (1989) concept of *polyarchy* and hence presupposes that democratic elections are supported by 4 elements: *elected representatives*, *free and fair elections*, *universal and active suffrage*, as well as *universal and passive right to vote* (Merkel, 2004). Elections are periodical sanctioning mechanism and when coupled with the partial regime of political rights, it embodies the essence of vertical accountability. By virtue of being periodic, though, it does not alone suffice to guarantee a democratic regime, as mentioned previously. While politicians might align themselves politically with the priorities of citizens, in the interest of getting reelected, this on its own is not reliable safeguard for democratic government (Merkel, 2004).

2.1.2. External Embeddedness

Furthermore, democracies are embedded externally (Merkel, 2004). They exist in the socioeconomic and civic-societal environment. These environments are not composed of defining characteristics of a democracy, however they encompass, enable and stabilize a democratic regime (Merkel, 2004).

The Socio-Economic Context

For one, a democracy benefits substantially from a stable and fruitful socioeconomic environment, or as Wolfgang Merkel calls it the "Socio-Economic Context" (Merkel, 2004, p. 44). In short, the "more well-to-do a nation, the greater is the chances that it will sustain democracy" (Lipset, 1959). This relationship has been observed often in the last four decades and has proven remarkably stable (Cheibub et al., 1996; Cutright, 1963; Dahl, 1971; Lipset, 1959, 1993; Merkel, 2004; Vanhanen, 1984, 1989; Welzel, 2000). Beyond the usual metrics of economic development, e.g. *gross domestic product* (GDP), there is also the influence played by socioeconomic inequality. Unequal distribution of economic resources can lead to a deterioration in the quality of a democracy if a consistent part of the populace is pushed below the poverty line (Merkel, 2004). The rationale is that extreme socioeconomic inequality can serve as an effective hurdle for political equality. Citizens will more likely be able to form their own opinions as active members of the citizenry, if they do not fear for their physical and economic security and if they are sufficiently educated through a sufficiently developed social and economic status (Merkel, 2004). The economic underprivileged part of the populace, in the exercise of civil and political rights, are at a disadvantage (Merkel, 2004).

There are caveats to be made, though. First, the influence of inequality has notable exceptions, namely developing countries and countries in the so-called third world, tend to have an economically disadvantaged population that shows "low-intensity citizenship" (Merkel, 2004; O'Donnell, 1998). Second, a developed economy is not the *sine qua non* for democratic systems. Alone, it cannot predict whether a country is democratic or not (Merkel, 2004). Neither is there a direct correlation with the level of democratization. Furthermore, economic development cannot on its own predict thresholds or transition zones for democratization (Merkel, 2004).

However, these factors can create a socioeconomic environment that is conducive to a healthy democratic regime (Merkel, 2004). "To sum up, a developed economy, the prevention of extreme poverty, the pluralization of the social structure, and the fair distribution of the material and cognitive resources of society create a shield for democracy and, in most cases,

enhances the quality of democracy with regards to the rule of law and participation. Inversely, both the lack of a well-developed economy and abrupt downward economic change endanger the stability and the quality of a liberal democracy" (Merkel, 2004, p. 45).

Civil Society

The notion that a well-developed civil society strengthens democratic systems is not novel. It is based on arguments brought forth by philosophers of the past, as well as more modern theorists (Merkel, 2004). The environment of civil society has to fulfill certain functions to support a democratic regime (Merkel, 2004). For starters, it has to protect individuals from a state wielding its power arbitrarily (Merkel, 2004). Second, it should support the balance of powers and rule of law (Merkel, 2004). Third, civil society should adequately educate the citizenry, as well as recruit political elites (Merkel, 2004). Lastly, it should establish a 'pre-institutional' public sphere of critical discourse, where self-determined forms of participation are able to influence political agendas beyond the traditional political and business interests (Merkel, 2004). "In a strict sense, civil society does not belong to the defining core of a constitutional democracy. It is rather outside of this core and therefore may be regarded as externally embedding it. The functions civil society carries out, however, have considerable implications for the sustainability of democratic constitutional institutions" (Merkel, 2004, p. 47).

International and Regional Integration

The stability and quality of a democracy can be affected by the integration of nations into international as well as regional organizations, that stand on an economic or politically democratic basis. There are caveats, nonetheless. Military or security focused alliances do not impart a similar effect on democratic systems, even if largely occupied by a democratic member base or seemingly committed to democratic principles (Merkel, 2004). Arguably the most prominent case in that regard is the *North Atlantic Treaty Organization* (NATO) (Merkel, 2004). Historically, the organization has been dominated by democratic states (Merkel, 2004; North Atlantic Treaty Organization, 2020). Furthermore, the alliance was founded as bulwark against an authoritarian state, namely the Union of Soviet Socialist Republics (USSR) (North Atlantic Treaty Organization, 2022). This is also officially embodied by the North Atlantic Treaty, which states in its preamble that its parties are "determined to safeguard the freedom, … founded on the principles of democracy, individual liberty and the rule of law" (North Atlantic Treaty Organization, 2022), as well as the Membership Action Plan published in 1999 stipulated that as a requirement for prospective members (North Atlantic Treaty Organization,

1999). These are both, however, non-binding and can be ignored at will, which the alliance has not refrained from doing in the past. While dominated by democracies, the pool of founding members included Portugal, which in 1949 was notably under the leadership of dictator António Salazar (Wagner, 2010). More recently and perhaps notably, Turkey's democracy has deteriorated substantially in the last decade (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Elçi, 2019; Merkel, 2004). Both of these cases have largely been ignored, likely due to their strategic importance, as foreign policy security is the focus of the alliance (Merkel, 2004).

On the other hand, organizations that are founded or promote on democratic principles have a higher likelihood to positively influence the stability of democracy. That is no guarantee though. For an instance the promotion of democracy, human rights and the rule of law are mentioned prominently in the constitutive act of the *African Union* (AF) and the institution has also implemented many instruments to that end (African Union, 2000; Mangu, 2014). However, since its founding, the continent has seen multiple, recurring breaches of constitutional rights, election integrity and human rights (Mangu, 2014; Sarkin, 2018). A similar dynamic applies to *Association of Southeast Asian Nations* (ASEAN), for instance. Despite multiple references to democracy and human rights as core principles of the organization, most members are not considered democratic and a measly half of them have commissions focused on human rights (Poole, 2019).

MERCOSUR[1] at first might seem like an exception to that trend. Though not part of its founding document (MERCOSUL, 1991), in 1998 the MERCOSUR adopted the Ushuaia Protocol, which established democratic principles as a cornerstone of the organization as it envisions healthy democratic institutions as "essential condition for the development of integration processes between States party to the … Protocol" (MERCOSUL, 1998). Furthermore, it establishes sanctioning mechanisms against states that rupture their own democratic order, as other member states can vote to suspend rights accordingly (MERCOSUL, 1998; van der Vleuten & Hoffmann, 2010). However, in 2012 the MERCOSUR underwent its first expansion when Venezuela joined the organization. This was controversial on its own, since Venezuela not only lacked a lot of the characteristics of a healthy democracy, it had also been on a steady downward trend for the preceding decade, in stark contrast to the other member states (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; García-Guadilla & Mallen, 2019). Four years after its accession, though, the other member states made use of

[1] Also referred to as MERCOSUL in the Portuguese acronym.

the suspension power against the left-wing government in Venezuela indefinitely. However, this action came after right-wing governments took office in Brazil and Argentina, hence "the politics of enforcement shouldn't be conflated with the politics of adoption of democracy clauses" (Closa & Palestini, 2018, p. 26). Furthermore, many of its members have experienced downward trends in many democracy indices in the recent past, including its arguably most important members, i.e. Brazil and Argentina (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Freedom House, 2021).

Merkel (2004, p. 48) argues that neither "the Association of South-east Asian Nations (ASEAN) nor South America's MERCOSUR and NATO have comparable effects [to the EU's], because they are not committed to the principle of democratic values in the same way". The de facto European Union was established with the signing of the *Treaty on European Union*[2] in 1992 (European Union, 1992), however its predecessor, the *European Economic Community* (EEC), already embodied some of its tenets, specifically on matters pertinent to the economy and trade (European Economic Community, 1958). However, it does not include commitments towards a democratic political order (Kubicek, 2003). This changed with the *Treaty of Maastricht* which established the promotion and maintenance of democracy, the rule of law, and human rights a core objective of the EU, as well as the 1993 Copenhagen Criteria, which established democratic criteria for membership (European Union, 1992; Kubicek, 2003).

This democratic tradition, despite being only formalized in the early 1990s (Kubicek, 2003), has been rooted in the roster of members of the EU. Since its founding literally all member states scored above 0.5 in the EDI and were ranked as "free" in FH indicator. This only changed in 2018 when Hungary got an EDI score of 0.489 and FH ranked Hungary as "partly free", which means it stands alone amongst all 28 current and former EU members to do so (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Freedom House, 2021). A decline has also been seen in Poland. Once in office the ruling parties of both countries, i.e. the *Fidesz* party in Hungary and *Law and Justice* (PiS[3]), began bearing its influence on key independent public institutions to impede their use of veto power (Kovács & Scheppele, 2018). First and foremost, their sights were set on the constitutional judiciaries. Once neutralized, they were unable to serve as a check on the power of those in the executive and legislative. Furthermore, this allowed ordinary judiciaries to be molded as to make it more difficult for individuals and opposition group to mount a challenge against the government (Kovács &

[2] Also referred to as the *Treaty of Maastricht.*
[3] in its Polish acronym

Scheppele, 2018). In short, "[j]udicial independence, once quite strong in both Poland and Hungary, is now a thing of the past" (Kovács & Scheppele, 2018). This encroachment on the separation of power has taken place in the form of constitutional reform that amended the role of higher courts, changing selection procedures for constitutional judges and election rules of presidents of the court, hence putting more power in the hands of the legislature (usually dominated by the ruling party), increasing the number of judges (so-called "court packing"), restricting the competencies of the court by restricting the power of constitutional review, nullifying case law established by the Constitutional Court prior to the constitutional reform, removing abstract review (once the most common type of constitutional review) , by which any citizen could bring a constitutional matter to the court, reducing the age of retirement for judges, hence packing lower courts, amongst other changes (Kovács & Scheppele, 2018).

These developments were not left without a response, albeit belated and generally ineffective. As Hungary's "attacks on the independence of the judiciary continued over eight long years, some of the European Union institutions took note, made repeated criticisms, but ultimately did not succeed in altering the course of events substantially" (Kovács & Scheppele, 2018). When Hungary replaced judges by lowering the retirement ages and firing sitting judges *en masse*, the European Commission charged Hungary with violating the prohibition on age discrimination. Though the commission won the case, this sort of violation only incurs a compensation fine, which Hungary paid to the aggrieved party, but it did nothing to reverse the decision or penalize Hungary in any other way (Kovács & Scheppele, 2018). However, they couldn't go further legally, since, at least the Commission perceived it then, the treaties that EU members are bound to presume that member states have independent judiciaries, so there was no explicit provision in EU law that would give the Commission the power to charge Hungary with interference of the judiciary (Kovács & Scheppele, 2018). As the Hungarian government consolidated its power over the judiciary, the European Parliament did pass several resolution denouncing its actions, but they were largely ignored by the Commission and the Council (Kovács & Scheppele, 2018). In 2018 the Parliament decided with a supermajority to trigger Article 7, which could strip Hungary of its voting rights though the procedure is still ongoing (Makszimov, 2022; Staudenmaier, 2018), and in 2022 the Commission triggered the Rule of Law Framework against Hungary, which in turn could suspend its EU funding (Deutsche Welle, 2022).

When Polish President Andrzej Duda, however, kicked off a constitutional fight with the President of the Constitutional Tribunal over the decision to force PiS to swear in three judges nominated by the previous ruling party, the European Commission, unlike with

Hungary, acted rather quickly (Kovács & Scheppele, 2018). Soon thereafter, and after the Polish government ignored recommendations by the Commission, it invoked its then new tool namely Rule of Law Framework. Despite it being designed to be implemented against Hungary, which it ultimately was, its first target was the newly inaugurated PiS government (Kovács & Scheppele, 2018). These actions, rather, did little to deter the Polish Government, which quickly took over the Constitutional Tribunal and in turn allowed the PiS Government to make changes to the judiciary in general, e.g. baseless firing of judges and reduction of retirement age (Kovács & Scheppele, 2018). All throughout this process, the Commission escalated its implementation of Rule of Law Framework, the (rather quiet) Council approved the decision of the Commission and the Parliament passed resolutions against the aforementioned actions (Kovács & Scheppele, 2018). However, the Polish government "barreled ahead, unchecked" (Kovács & Scheppele, 2018). This reached its zenith when, in 2017, the Commission triggered Article 7 proceedings against Poland, which are, much like in Hungary's case, still on-going (Bobinski, 2022; Kovács & Scheppele, 2018).

These facts present a problem for international organizations as indicators of democracy. While both MERCOSUR and the EU have overwhelmingly members that could arguably be considered democracies, both organizations have proven ineffective at preventing democratic backsliding within their ranks and, in the case of the former, not accepting member with considerable breaches of democratic principles (Closa & Palestini, 2018; García-Guadilla & Mallen, 2019; Kovács & Scheppele, 2018). The latter has had its ranks filled with democratic states. However, it has proven itself ineffective at preventing backsliding, as the cases of Poland and Hungary show (Kovács & Scheppele, 2018). The reasons for it are many. For starters, though some agreements of the EU do put an emphasis on democracy and its inherent principles, the organization does not have the necessary tools to enforce them (Kovács & Scheppele, 2018). The EU "was designed to protect Member States from an overreaching Union rather than to protect the Union from failing Member States" (Kovács & Scheppele, 2018). Hence, the European Commission has few tools at its disposal to intervene. Furthermore, member states are very reluctant to judge each other since that opens space for the judgement to be turned on themselves. While action has come from the Parliament and the Commission, the Council has remained nearly absolutely silent (Kovács & Scheppele, 2018). It has also largely ignored calls from the aforementioned institutions to do so. Lastly, but perhaps more importantly, the EU is not a federation. EU institutions were given competencies by member states, however national constitutional structure does not fall under that purview (Kovács &

Scheppele, 2018). The opposite is in fact the case, i.e. national constitutions are protected from EU interference by the EU's own legal framework (Kovács & Scheppele, 2018).

Hence while Merkel (2004, p. 48) argued that "the European Union (EU) and its precursor organizations have proven the most successful in the international embedding of democracies. … This economic and institutional embedding into community-based, strong common interests and values will also considerably stabilize the ten central and eastern European new member countries after 2004", this contribution concludes that recent developments have shown that said stabilization effect is not guaranteed, that there is no international organization today that fulfills the expectation of this externally embedded context, and hence it was not considered in further sections.

2.2. Aggregation

A significant issue in measuring democracy, as alluded to previously, is how to aggregate those characteristics. In their Working Paper, Gründler and Krieger indicated that a machine learning based index can avoid underestimating changes in the degree of democratization, when compared to standard methods of aggregation, i.e. additive, multiplicative, additive and multiplicative, and a Bayesian latent variable approach (Gründler & Krieger, 2020). Their published research indicated that a machine learning index can produce much more consistent and plausible results, with a high degree of detail (Gründler & Krieger, 2016, 2021).

However, their previous approaches have points that can be improved upon. First of all, while the SVMDI has been compared to other approaches, the index used entirely different conceptions of democracy and variables. Hence, one could not isolate the aggregation function in order to assess its potential benefits or issues in relation to an actual existing index. Furthermore, both SVMDIs, i.e. the one created in 2016 (Gründler & Krieger, 2016) and the one created in 2021 (Gründler & Krieger, 2021), had their own specific set of issues.

While the first attempt used a relatively broader concept of democracy, the theoretical underpinning of the data selection was not well-grounded. First of all, since the variables come from a variety of different sources (Gründler & Krieger, 2016), there is a potential for conceptual overlap (Gründler & Krieger, 2021). Second, some of these variables are aggregated themselves (Gründler & Krieger, 2016), which defeats the purpose of using SVM to aggregate the data. Third, it's timespan is rather restricted as it only encompasses 30 year, from 1981 to 2011 (Gründler & Krieger, 2016), an issue the authors themselves bring up in their later publication (Gründler & Krieger, 2021). Lastly, for a relatively broader concept of democracy,

the number of variables is relatively low, i.e. 11 (Gründler & Krieger, 2016). For context, the SVM index with the narrowest conception in this contribution, based on Merkel's (2004) Electoral Regime, also had 11 variables. This does not mean it is necessary better, but it does indicate that various characteristics of democracy might have been overlooked or simplified.

The updated SVDMI solves some of these issues, such as the restricted timespan and previously aggregated variables (Gründler & Krieger, 2021). However, in their effort to avoid conceptual overlap with other fields of research and the variable selection related to Dahl's (1989) concept of Polyarchy, they ran into the issue raised by Geissel et al. (2016), i.e. successfully assessing gradations of democratic quality, a weakness that was further explored in this contribution in Section 6.

2.2.1. Indicator Review

One question arises, namely what is the purpose of another democracy index? Beyond the ones hitherto mentioned, there are already several ones available (Gründler & Krieger, 2021). Some of which have already been used in this contribution to exemplify variations in the level of democracy between countries. Furthermore, they are by no means monolithic. The indices of democracy use conceptions of democracy range from Dahl's (1989) minimalistic polyarchy to broader concepts. They also vary in aggregation methods and timespans (Gründler & Krieger, 2021). This section, inspired by and updating the work done by Gründler and Krieger (2016, 2021), attempted to answer this question by going over prominent democracy indices, reviewing them and, at last, explain what gaps a new machine-learning focused index with a wider as well as modular conceptualization of democracy might fill. This section does not intend to address all existing indices or all debates around them, but this is supposed to give a short overview of the state of affairs and pinpoint possible weaknesses.

V-Dem's Electoral Democracy Index

One of the two indices already mentioned and used in this contribution, V-Dem's EDI[4] is arguably one of the most renowned. The research project responsible for it, i.e. V-Dem, aims at providing institutional information on all countries worldwide (Gründler & Krieger, 2021). In its entirety, the V-Dem database contains more than 450 regime characteristics, which are well grounded in theory, and some of which are used to create five distinct democracy indices (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021). Arguably the most prominent under these indices is the EDI (Gründler & Krieger, 2021;

[4] Often referred to as Polyarchy Index (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021)

Teorell et al., 2019). It also serves as a basis for all other democracy indices in the V-Dem roster (Gründler & Krieger, 2016). This index is continuous, offers a confidence interval, is available for a total of 202 countries between 1789 and 2021, and uses an arguably minimalistic approach. It is built out of five characteristics, all of which are expert-based, with the exception of one, namely *suffrage*. The others are *extent to which executive and legislative are elected*, *associational autonomy*, *free and fair elections*, and *freedom of expression and alternative sources of information* (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021; Teorell et al., 2019). They themselves, apart from suffrage, are composed of several sub-indicators, the aggregation of which depends on whether the sub-indicator in question is additive or multiplicative. The index itself is then calculated by taking the mean of the indices (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Gründler & Krieger, 2021; Teorell et al., 2019). This is explained in further detail in section 6.

And herein lies a few weaknesses. Combining additive and multiplicative indicators does not suffice to outright remove their individual methodological weakness (Gründler & Krieger, 2020, 2021). Furthermore, Teorell et al. (2019) do not provide a clear-cut reasoning for why these indicators should be equally weighed and additively combined (Gründler & Krieger, 2021). Lastly, this index is arguably minimalistic. The aforementioned characteristics do not include factors such as an independent judiciary, economic stability, socioeconomic inequality, or the role of civil society. Though, as mentioned previously, this does not necessarily mean it's worse, it might lead the index to overlook certain developments, e.g. loss of judicial independence.

Freedom House (FH)

The second index mentioned in this contribution is Freedom House's Freedom in the World Index. In actuality it is composed of two indices, i.e. a graded index as well as an ordinal index, though the latter is just an extension of the former. The grading ranges from 100 (*most democratic*) to 0 (*most autocratic*)[5] and the ordinal scale contains three categories, namely *free*, *partly free* and *not free* (Freedom House, 2022; Gründler & Krieger, 2021). To come up with these indices it separates between two distinct aspects of democracy, i.e. *Political Rights* (PR) and *Civil Liberties* (CL). They are then added to form the final index. Each of these aspects are

[5] It is possible for a country or territory's total political rights score to be less than zero (between –1 and –4) if it receives mostly or all zeros for each of the 10 political rights questions and it receives a sufficiently large negative score for the political rights discretionary question (Freedom House, 2022).

formed by 10 and 15 characteristics, respectively, each being given a score between zero and four. In the end PR can have a maximum of 40 and CL a maximum of 60 points[6] (Freedom House, 2022; Gründler & Krieger, 2021).

The index includes nearly all independent countries and self-governing territories. The timespan however is much shorter than the EDI, since it starts in 1972 (Gründler & Krieger, 2021). Conceptually it is broader, encompassing characteristics such as minority rights, the existence of safeguards against corruption, and judicial independence (Freedom House, 2022; Gründler & Krieger, 2021). However, this indicator suffers from methodological weaknesses. Chief amongst these is its aggregation rule as it is a simple additive function and hence can create dubious classifications at the lower end of the spectrum (Gründler & Krieger, 2020; Munck & Verkuilen, 2002; Teorell et al., 2019). A country can theoretically score high on certain aspects but score the lowest possible on what one could argue are necessary conditions for democracy, such as free and fair elections. Furthermore, 25 expert-based regime characteristics are only publicly available from 2006 onwards (Gründler & Krieger, 2021). Finally, Freedom House tend to follow American research on democracy, which neglects "the connection between poverty and low intensity citizenship. Thus, while Germany has been scolded or even downgraded for surveying a dubious quasi-religious sect like Scientology with its intelligence service, the fact that almost 20 per cent of US citizens (predominantly African Americans) are disadvantaged through poverty in the exercise of their civil and political rights is completely ignored" (Merkel, 2004, p. 58). So while the concept might be relatively broad, when compared to the EDI, even that breadth misses components that might affect the quality of a democratic system.

Boix–Miller–Rosato indicator (BMR)

The BMR is a dichotomous indicator that categorizes a regime as a democratic, if it meets three conditions, i.e. the majority of male and female citizens are eligible to vote, the executive was directly or indirectly elected in popular elections and is responsible to voters or to a legislature, and the legislature (as well as the executive) was chosen in free and fair elections (Boix et al., 2013, 2018; Gründler & Krieger, 2021). This indicator has a few major strengths. For instance, the database is relatively comprehensive, including indices for 222 countries spanning the timespan between 1800 to 2015, and thus one of the best measures in

[6] "Prior to the 2020 edition, Freedom in the World assigned a country or territory two ratings—one for political rights and one for civil liberties—based on its total scores for the political rights and civil liberties questions. Each rating of 1 to 7, with 1 representing the greatest degree of freedom and 7 the smallest degree of freedom, corresponded to a specific range of total scores. … The ratings are still included in the raw data available for download" (Freedom House, 2022).

terms of data coverage. It also is methodologically consistent, as they argue why the aforementioned institutional aspects are considered, then operationalized, and they aggregate them in accordance with their conceptual assumptions (Gründler & Krieger, 2021).

However, this indicator also possesses major weaknesses. First of all, it is dichotomous, which, as mentioned previously, lacks a lot of discriminating power. This becomes problematic, when assessing the cause or effect of democracy, as well as its development over time. It is hard (if not impossible) to distinguish gradual processes of democratization or democratic backsliding (Gründler & Krieger, 2021). Furthermore, since the conditions that form the BMR are also theoretically dichotomous, the use of subjective conditions becomes problematic, e.g. free and fair elections. For some observations that might be clear, but for many it isn't. In such cases, there should be at least an uncertainty dummy variable, which would indicate that coders were not certain about their conclusion. One could go further and provide a confidence interval that reflects the measurement of uncertainty. However, the BMR provides neither, therefore making the identification of borderline cases impossible (Boix et al., 2013, 2018; Gründler & Krieger, 2021). The suffrage condition also only stipulates that half of the combined male and female populations have the right to vote. This leaves space for the exclusion of parts of the citizenry, e.g. minorities, from the electoral process, without technically breaking this condition. Finally, Boix et al. (Boix et al., 2013, 2018) do not provide their raw data, as recommended by Munck and Verkuilen (2002), hence impeding users of the indicator from checking whether "their results depend on the choice of the aggregation method nor whether they are caused by a specific aspect of democracy" (Gründler & Krieger, 2021, p. 4).

Democracy-Dictatorship Indicator (DD)

Much like the BMR the DD is a binary indicator, its consistency between the definition of democracy and the aggregation procedure is a positive characteristic, and it uses four conditions to classify a regime as democratic, i.e. *the head of government was elected directly or indirectly*, *the legislature was elected through a popular elections*, *the parliamentary elections must be disputed by more than one party*, and *an alternation in power should be under the regiment of the same rules that brought the incumbent government into office* (Bjørnskov & Rode, 2020; Cheibub et al., 2010; Gründler & Krieger, 2021; Merkel, 2004). The latest edition of the index encompasses 192 countries, 16 self-governing territories, 96 colonies and a timespan between 1950 to 2020. Despite being less comprehensive than the BMR in that regard, it does provide which conditions a specific country violates, if it's not classified as a democracy (Bjørnskov & Rode, 2020; Gründler & Krieger, 2021). The obvious transparency

benefit notwithstanding, this allows users to pinpoint whether their results are due to a specific area of democracy (Gründler & Krieger, 2021).

This indicator, by virtue of being dichotomous, is plagued by the issues therewith attached and shared by the BMR (Gründler & Krieger, 2021). Furthermore, the indicator has a couple of conceptual issues. For one, it does not have any requirement on political participation, e.g. suffrage, which can lead to misclassifications (Gründler & Krieger, 2021; Munck & Verkuilen, 2002). Furthermore, the requirement relating to an alternation in power can also lead to misclassification, as young democracies in a transitional phase might be disregarded as there might not have been enough time for one to take place (Gründler & Krieger, 2021; Knutsen & Wig, 2015).

Polity IV and Polity V

Much like FH, this indicator is graded, ranging from -10 (*most autocratic*) to +10 (*most democratic*) and it is composed of five institutional aspects, i.e. *constraints on chief executive*, *competitiveness of the chief executive recruitment*, *openness of the chief executive recruitment*, *regulation of participation*, and *competitiveness of participation*. Every country with more than 500 inhabitants is evaluated in relation to these characteristics and that is available from 1800 to 2018 (Gründler & Krieger, 2021; Marshall et al., 2019; Marshall & Gurr, 2020). The characteristics are also publicly available, meaning users can check whether their results are affected by one of these particular aspects of democracy (Gründler & Krieger, 2021).

There are a few conceptual and methodological issues, though. For starters, many observations are given so called "standardized authority scores" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021; Marshall & Gurr, 2020). These include countries with a systemic interruption due to foreign influence, cases of interregnum or anarchy, and countries in systemic transition. They are given categorical values outside of the normal range of the index. An alternative variable is proposed by Marshall and Gurr (2020), which would treat cases of foreign interruption as "system missing", cases of interregnum or anarchy are given a neutral value of 0, and cases of transition are given equidistant scores in between the initial and final values (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Marshall & Gurr, 2020). However, these solutions might still lead to spurious changes in the level of democracy (Gründler & Krieger, 2021). The database also does not offer confidence intervals or any form of measurement uncertainty, which makes it difficult to classify anocracies (Gründler & Krieger, 2021). Finally, a major issue with this index is its aggregation rule. All regime characteristics are weighted equally, the

aggregation rule is simply additive, and the aggregation procedure in general lacks theoretical foundation, which results in doubtful classification, particularly at the lower end of the spectrum (Cheibub et al., 2010; Gründler & Krieger, 2020, 2021; Munck & Verkuilen, 2002; Teorell et al., 2019; Treier & Jackman, 2008).

Acemoglu–Naidu–Restrepo–Robinson indicator (ANRR)

This dichotomous indicator varies substantially from the previously mentioned indices, as it combines multiple indicators, i.e. Polity IV, FH[7], DD, and BMR (Gründler & Krieger, 2021). A regime is democratic when the Polity index is greater than zero and the FH index is below 10. Should one of these indices not be available for the country in question, then the country also has to be to categorized by the remaining index, as well as be categorized as democratic by either DD or BMR (Gründler & Krieger, 2021). Should both FH and Polity not be available, then the country has to be categorized as democratic by both DD and BMR. Should none of these conditions apply, then the country is categorized as autocratic (Acemoglu et al., 2019; Gründler & Krieger, 2021).

This database contains indices for 184 countries between 1960 and 2010 (Gründler & Krieger, 2021). There is caveat to this though. Since all the underlying measures are available to the public, users can extend the timespan with relative ease, should it be needed (Gründler & Krieger, 2021). However, there are shortcomings. For one, the concepts of democracy underlying these indicators are by no means the same. FH, for instance, has a much broader definition of democracy than any of the other three (Gründler & Krieger, 2021). Furthermore, since FH only publishes data as far back as 1972, there are many countries that are not on FH. This in turn means there is the possibility that an observation is categorized as a democracy by the other indices, which would otherwise not be categorized as democracy by FH but is not on its database (Gründler & Krieger, 2021). Lastly, the classification creates spurious regime changes, a factor that Acemoglu et al. (2019) themselves pointed out (Gründler & Krieger, 2021). To overcome this issue, they manually recoded transitions based on whether they were plausible or not from their point of view, however there is no selection criteria and hence it is rather arbitrary (Acemoglu et al., 2019; Gründler & Krieger, 2021). Finally, there are no uncertainty measure for this indicator (Gründler & Krieger, 2021).

[7] Acemoglu et al. (2019) uses the FH ratings as they were prior to 2020, i.e. from 2 (most democratic) to 14 (most autocratic) (Freedom House, 2022; Gründler & Krieger, 2021).

Unified Democracy Score (UDS)

The UDS is a composite index. It uses a variety of indices[8] and its original version offers data for a timespan from 1946 to 2000, though it has been relatively recently updated to expand it all the way back to 1815 (Gründler & Krieger, 2021; Marquez, 2022; Pemstein et al., 2010). A major strength is that the UDS produces a measure for each country-year combination and also provides a measure of uncertainty. That being said, the confidence intervals increase towards the extreme ends of the spectrum and narrow for hybrid regimes, which, as pointed out by Teorell et al. (2019), is counterintuitive as the level of uncertainty increases towards the middle of the spectrum. Hence, these confidence levels should be interpreted with caution (Gründler & Krieger, 2021).

Lexical Index of Electoral Democracy (LIED)

Developed by Skaaning et al. (2015), distinguishes between seven political regime types, i.e. regimes without elections, with no-party or one-party elections, with multiparty parliamentary elections, with multiparty parliamentary and presidential elections, minimally competitive multiparty election for legislature and executive, with minimally competitive multiparty election and full male or female suffrage for legislature and executive, and lastly with minimally competitive multiparty election and full male and female suffrage for legislature and executive (Gründler & Krieger, 2021).

This indicator possesses one of the widest temporal coverages of all by starting in 1789 and also provides data on all sovereign countries in the world (Gründler & Krieger, 2021). Despite being complimented by clear coding rules and a well-grounded aggregation procedure (Gründler & Krieger, 2021). However, it can hardly be used in applied studies effectively. The main problem lies it its categorical scale. If used as a single variable, it presents an issue for a regression model, since the "democratic distance" between categories is not equidistant (Gründler & Krieger, 2021). There is a negligible jump, if any, from a country without elections to a country with no-party/one-party elections, in terms of its degree of democratization (Gründler & Krieger, 2021). However, that changes dramatically when a country has multiparty elections. This presents a challenge for users of that index (Gründler & Krieger, 2021).

[8] It employs the DD, and FH, Polity, as well as further indicators proposed by Vanhanen (2000) Arat (1991), Bollen (1999), Bowman et al. (2005), Coppedge & Reinicke (1990), Hadenius (1992), and Gasiorowski (1996) though these have fallen into disuse (Gründler & Krieger, 2021; Pemstein et al., 2010).

2.3. Summary

Many of these indicators are frequently used in the scientific literature. Building upon the work done by Gründler and Krieger (2021), this contribution has reviewed and updated the overview of eight prominent democracy indicators. Further, it has pointed out weaknesses in many of them. These usually revolve around the aggregation method, or the conceptual choices therewith attached. This contribution attempted to alleviate the aggregation issue, by using SVM to aggregate characteristics, and to an extent the conceptual issue by applying a conception of democracy, i.e. Embedded Democracy (Merkel, 2004), that can increases in scope but distinguishes between different important aspects of democracy. Hence, the SVM index in this contribution avoids overlap in characteristics within itself, as well as offer researchers the opportunity to avoid conceptual overlap with their own fields, while still being able to identify gradations of democratic quality.

3. Methodology

3.1. Machine Learning Indicator

As mentioned in Section 1, SVM is a supervised machine learning technique that reveals unknown relationships between input characteristic and an outcome (Gründler & Krieger, 2021; Steinwart & Christmann, 2008), and one of its key features is that the function that maps input onto output does not have to be specified beforehand, as opposed to conventional statistical methods (Gründler & Krieger, 2021). These characteristics assist in the objective of alleviating the issue of simplistic and arbitrary aggregation methods (Gründler & Krieger, 2021).

The idea behind SVM regression is as follows. Given a certain dataset $F = \{(x_1, y_1); \dots ; (x_n, y_n)\}$, the objective is to find a functional relationship $\hat{f}(x_i) = y_i$ that approximates the "true" function $f(x_i) = y_i$ for all $i = 1, \dots , n$. In that effort, said functional relationship should have at most ε deviation from y_i in the training data and at the same time the shape of that functional relationship should be as flat as possible (Gründler & Krieger, 2016; Smola & Schölkopf, 2004; Vapnik, 1995). What this means is that the basic rationale of SVM is, given a plane onto which the data is plotted, the algorithm creates a boundary around a line, i.e. a hyperplane, that tries to encompass the highest amount of data without violating the margins of said boundary. The distance of the boundaries of the hyperplane from its center line are defined by ε (Gründler & Krieger, 2016; Smola & Schölkopf, 2004; Vapnik, 1995). This ensures, in a hypothetical scenario, that one would "want to be sure not to lose more than ε money when dealing with exchange rates, for instance"(Smola & Schölkopf, 2004, p. 200). The support vectors are the

points on the plane that fall outside of the boundary of the hyperplane. The objective of SVM regression is to reduce the error to a minimum, by keeping as many data points inside the hyperplane as possible while keeping the distance to the support vectors to a minimum. That is where the cost parameter *C* comes into play. It determines to what extent are violations of the boundaries, namely the points outside the hyperplane, penalized. If $C = 0$ then there is no cost to these violations. However, a large *C* can have a substantial effect on the functional relationship (Gründler & Krieger, 2016; Smola & Schölkopf, 2004).

Though sometimes the functional relationship might be linear, sometimes this relationship can also be non-linear. To deal with this issue Boser et al. (1992) suggest using a higher dimensional space instead of the plane used previously (Gründler & Krieger, 2016; Smola & Schölkopf, 2004). However, how to map the data onto a higher dimensional space is usually unknown as well as computationally infeasible. Boser et al. (1992) propose a method to overcome these issues by using a so-called kernel trick. The kernel, at its simplest, is a function that takes that non-linear problem and transforms into a linear one within a space of higher dimensionality and that is then handled by the algorithm (Gründler & Krieger, 2016; Smola & Schölkopf, 2004). Though there are many kernels, this contribution used the Gaussian Radial Basis Function (RBF) kernel, as it is broadly accepted, recommended by Guenther & Schonlau (2016) and it has, in previous research, shown promising results in robustness checks (Gründler & Krieger, 2016, 2021).

This is a very simplified explanation of the process by which SVM regression is conducted, as the mathematical background does not fall under the scope of this contribution. By applying said methods and parameters the function $\hat{f}(x_i)$ that approximates the true function is calculated (Gründler & Krieger, 2016).

The SVM regression, for the purposes of this contribution, produced a measure of democracy by taking a sample Z that included $n \gg 0$ country-year combinations of observations *(i,t),* for which the regime characteristics $z^1, \ldots, z^r$ were available (Gründler & Krieger, 2021). What this means is that the SVM takes a sample that contains a substantial amount of country-year observations, all of which need to have available data for the selected characteristics. The SVM regression was developed and the data wasworked on using the programming language *R* and the platform *RStudio*, as well as several packages available for it, e.g. *dplyr*, *ggplot2*, *e1071* and *caTools*. Furthermore, a specific R package with

specialized functions was created for the purposes of this contribution[9]. The functions created and the code used for this contribution can be found in the Appendix.

3.1.1. Priming Data

To train the model, there was a necessity for a set of cases whose level of democracy was directly observable. This was needed so that unknown functional relationship between the regime characteristics and the level of democracy could be learned (Gründler & Krieger, 2021). Regimes that laid on either side of the autocracy-democracy-spectrum were an appropriate choice to this end. This is because it is well-established that their classification is rather uncontroversial (Cheibub et al., 2010; Gründler & Krieger, 2021; Lindberg et al., 2014). For an instance, there is little debate on whether Nazi Germany in 1941 was an authoritarian state, much alike there isn't much controversy on whether Sweden in 2020 should be considered a democratic country. As Gründler and Krieger (2021) argue, the low level of disagreement about these cases implies that their degree of democracy is directly observable.

The selection of regimes that belong was based on two indices, i.e. the EDI and UDS. These indicators were chosen mainly for two reasons. First, they have a relatively wide time coverage. Second, they both are continuous indices, which provided a finer level of detail. Regimes that were in either the upper or lower decile of either indicator were labeled as democratic or autocratic, respectively. The reason that two indices were used is that previous research has shown that this increases the quality and heterogeneity of the priming data and hence of the final index (Gründler & Krieger, 2021). The EDI was also used as target variable for the priming data.

Consensus on the priming data

It was paramount to correctly label the priming data as that determined how characteristics influence the overall index. The first step was to check whether there is a sizeable discrepancy between the priming data and other indicators. For this purpose, five other indicators were used: the DD, the BMR, the Polity V Index, the FH Index and the LIED (Gründler & Krieger, 2021). First, the priming data was compared to other established indices, to check for any noticeable incongruencies (Gründler & Krieger, 2016, 2021). Second, previous research was used to confirm, in cases of incongruencies, that there weren't any regimes in the priming data that were outwardly unreasonable.

[9] It is available at https://github.com/jlhilgert/demMA and can be installed onto R using the *devtools* package and the *install_github* function that comes with it using the command *install_github("jlhilgert/demMA")*.

Representativeness of regime types

The data was also checked for representativeness in both ends of the spectrum. Autocratic regimes types can differ substantially from one another (Cheibub et al., 2010; Gründler & Krieger, 2021). The same goes for democratic types (Gründler & Krieger, 2021; Pérez-Liñán, 2020). To produce plausible democracy indices, the priming data had to account for these different types of regimes. Following the work of Gründler and Krieger (2021), the autocratic regimes in the priming data were tested thrice.

First, following the definitions of Cheibub et al. (2010), autocratic regimes were distinguished by different types of dictatorships, i.e. *civil*, *military*, *communist* and *royal*. This was done using the data provided by Bjørnskov & Rode (2020) and Geddes et al. (2014). Second, the regimes were checked for the presence of an electoral regime. Lastly, they were checked for heterogeneity in durability (Gründler & Krieger, 2021).

Democracies were controlled for institutional heterogeneity (Gründler & Krieger, 2021). Much like the autocratic regimes the steps were threefold. First, they were checked for form of government, i.e. presidential, semi-presidential and parliamentarian (Gründler & Krieger, 2021). Second, they were controlled for unicameral and bicameral parliamentary chambers (Gründler & Krieger, 2021). Finally, they were controlled for proportional and majoritarian voting systems (Gründler & Krieger, 2021).

3.1.2. Aggregation Procedure

Having selected the priming data, the next step was to transform the characteristics of individual regimes into a measure of democracy. The scale of the final index was continuous as that provides a higher degree of discrimination between regimes and it is a better scale to observe changes over time. The next step was to randomly extract a training set T out of the priming data P (Gründler & Krieger, 2021). A uniformly distributed random number generator determined the total number of observations in the training set, which was no smaller than 20% and no greater than 50% of the total number of observations in P (Gründler & Krieger, 2016, 2021).

This training set was in turn used to estimate the aggregation function, for which purpose the SVM regression tool was applied. However, there are a few methodological choices made beforehand. The penalization parameter was set to one as proposed by Mattera & Haykin (1999) (Gründler & Krieger, 2021) and the kernel to be employed was the RBF, for the reasons mentioned in Section 2 (Gründler & Krieger, 2021). The margin parameter ε was set to 0.025

(Gründler & Krieger, 2021). The data was also centered and scaled, as the variables had varying scales.

This returned a model containing the aggregation function f_s, which in turn could be used to calculate the democracy index for all other country-year observations in sample *Z*. Lastly, these steps were repeated for 2000 iterations (Gründler & Krieger, 2021). This ensured that the aggregation procedure converged and hence reduced the error to a minimum (Gründler & Krieger, 2021).

The aggregation output consisted of three categories (Gründler & Krieger, 2021). The first was the continuous index of democracy of each country-year combination in the sample *Z*. This was done by taking the median of the 2000 iterations that were produced by the aggregation procedure (Gründler & Krieger, 2021). The second output was the standard deviation of said iterations for each observation (Gründler & Krieger, 2021). The third was all percentiles of all iterations for each observation (Gründler & Krieger, 2021). This in the end was a dataset with the identification information for each iteration, e.g. country and year, the resulting SVM index, the standard deviation of all iterations, 100 columns representing each percentile of the aggregation results for all 2000 iterations. The second and third categories are useful to reflect on the extent of measurement uncertainty and can be used to calculate a confidence interval for each observation (Gründler & Krieger, 2021).

3.2. Data Selection

The primary source of data was V-Dem's database (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021), which includes a plethora of information for a variety of countries and a relatively long timeframe, as alluded to previously. Following Munck's and Verkuilen's (2002) guidelines and Gründler's and Krieger's (2021) previous work, the set of variables included both objective and subjective/expert-based data. Naturally, not all of the over 4000 variables in V-Dem's dataset were used. The variables were chosen as they related to the constitutive aspects of Embedded Democracy (Merkel, 2004) as outlined in Section 2. Most of them related to mid-level indices in V-Dem's dataset. In such cases, all corresponding low-level indices were used, so as not use mid-level indices aggregated using BFAM, and hence isolate the aggregation function. The only two mid-level indices employed, were either manually coded or were objective values. Furthermore, this contribution focused on the timespan from 1919 to 2020. The reason for this was twofold. The simplest one was data availability, particularly when checking for consensus with other indices and representativeness of the data (Bjørnskov & Rode, 2020; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada,

et al., 2021; Geddes et al., 2014). Furthermore, it facilitates comparison with the SVDMI index in Section 5.

3.3. Analysis of Results

The data was used to create four datasets. One took the same variables as an existing index, namely the EDI. This was done to isolate the aggregation functions and assess their distinct effects on the individual indices. In that effort the SVM index was compared to the EDI, the reasons for which are described further in Section 5. The observations with deltas larger than 0.01 were separately analyzed using *Ordinary Least Squares* (OLS) regression to assess which variables had a higher impact on each individual index. The results are further exemplified with select cases from this subset. This was described in further detail in Section 5.

The other three encompassed variables that relate to individual democratic regimes and contexts as established by Merkel (2004) and discussed in Section 2. There was no overlap between regimes/contexts in terms of variables, meaning that a specific characteristic of democracy was only related to one regime/context of an Embedded Democracy. The result was three datasets. One including only the electoral regime, a second including all internally embedded regimes and a third including the externally embedded regimes as well. This was later applied to assess whether a broader definition of democracy helps identify gradations in democratic quality. Three select cases exemplify how conceptualization, regardless of aggregation function, might lead to a decreased ability of an index to identify variations in the quality of a democracy. This was described in further detail in Section 5.

3.4. Summary

In this section the concept of SVM was shortly explained, to contextualize in simple terms how the regression calculated the score of democratic quality for each observation. It was explained how that took place in technical terms and what parameters are important in that effort, based on recommendations established by the current literature (Gründler & Krieger, 2021; Guenther & Schonlau, 2016; Mattera & Haykin, 1999). Paramount in this case was the selection of the training data. In this section it was outlined how this data, i.e. the priming data containing reasonable cases of democracy or autocracy, was selected. This priming data was controlled for consensus among democracy indices and its representativeness in terms of specific regime characteristics. Furthermore, the overall aggregation procedure was described as well as the output of the SVM regression. Lastly it was described how the data that was used for the analysis, how it was selected, as well as how the analysis was to be conducted.

4. Priming data

The priming data, used to train the models, originated from the full V-Dem dataset (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021), as it not only was necessary in the selection process but also it contained many of the other indicators used to check for consensus and representativeness, including the UDS (Marquez, 2022; Pemstein et al., 2010), which was used to select the observations (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). Furthermore, for the consensus check as well as for controlling institutional characteristics, information was extracted from the DD index and further data made available by Bjørnskov and Rode (2020). For the calculation of regime length, data was collected from the datasets provided by Geddes, Wright and Frantz (2014).

In total, the priming dataset contained information on 4643 different country-year observations. It spanned from 1919 to 2020 and included 135 countries, though not every country had an observation for each year. Not all observations could be included in the datasets used to check for representativeness, as the data on institutional characteristics and regime duration were less comprehensive (Bjørnskov & Rode, 2020; Geddes et al., 2014). The Bjørnskov and Rode (2020) dataset starts in 1950 and the Geddes, Wright and Frantz (2014) only spans to 2014. However, since both still encompassed a substantial amount of the original dataset, they were still useful in assessing the representativeness of the priming data. The validation naturally also did not include the entirety of the timespan for all comparisons, since, as shown in Section 2, different indicators have encompassed different timeframes.

4.1. Consensus

To check the priming data for consensus, a binary classification was created. Should a country-year observation meet the requirements for a democratic regime it was given a classification of one. Otherwise, it would receive a classification of zero. This was possible since the observations were in either extreme end of the spectrum and it made checking for overlaps or disparities with other indicators simpler and easier to identify. These were then compared with the BMR Index (Boix et al., 2013, 2018), the DD Index (Bjørnskov & Rode, 2020), the LIED (Skaaning et al., 2015), the Polity V Index (Marshall & Gurr, 2020), and the Freedom House Index (Freedom House, 2021, 2022).

Boix–Miller–Rosato indicator (BMR)

Being a binary indicator, it was very simple to compare the priming data with it. However, there was a caveat. For the purposes of this contribution the regular BMR democracy

variable was not used, but rather a variation of it with omitted data. This was due to how BMR codes data. If there is a military occupation in a particular country, the regular BMR democracy indicator maintains the same regime type prior to the occupation (Boix et al., 2013, 2018). Czechoslovakia, for instance, is coded as a democracy during its occupation by Nazi Germany between 1940 and 1944 (Boix et al., 2013, 2018; Horáková, 2003). The variation of it codes occupied countries with NA values (Boix et al., 2013, 2018; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). The reason that this is not done for the main indicator by Boix et al. (2013, 2018) is to avoid coding democratic transitions at the end of military occupations.

The comparison subset of the data had a total of 1991 observations, 1989 of which were congruent. Only 0.1% of the data differed. All of the differing data came from the autocratic group, meaning that the priming data and the BMR indicator concurred on all democratic regimes. The differing observations related to Uganda in 1980 and Greece in 1944. Both discrepancies weren't unreasonable, but also not entirely uncontroversial. In 1980 Uganda suffered a military coup in May and also held a multiparty election in December (Gründler & Krieger, 2021). Greece in 1944 was under Nazi occupation until October and a civil war broke out in December (Encyclopaedia Britannica, n.d.). Hence, it is not a stretch to conclude that different indicators, with different criteria, coded these cases differently. This was borne out by inconsistent data relating to these two cases across indicators (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021).

Democracy-Dictatorship Indicator (DD)

Much like BMR, the DD Index was also fairly easy to compare. There were a total of 2173 observations. And much like BMR there was full consensus on all regimes labelled as democratic. For autocratic regimes there were 4 differing observations, or just shy of 0.2% of data. All of these referred to Benin between 1952 and 1955. The reason for the discrepancy was actually fairly straightforward. It was due to the DD indicator's concept of democracy as formulated in Section 2. Benin in that time period satisfied the indicator's four conditions for democracy, however Benin, for instance, only provided male suffrage, which would severely affect its democracy score under broader concepts of democracy. Hence, the discrepancy was not unreasonable.

Lexical Index of Electoral Democracy (LIED)

By virtue of being ordinal, the comparison was slightly less straightforward. The comparison dataset contained 2019 observations. Democracies in the priming data all fell under

the highest level on the LIED scale, i.e. “regimes with minimally competitive multiparty election and full male and female suffrage for legislature and executive” (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021; Skaaning et al., 2015). This means that there was again consensus on the democratic end of the spectrum.

On the authoritarian end it got more complex. The overwhelming majority of instances of autocracy, i.e. 96.4% of them, fell at or below the second lowest level on the scale, namely “regimes with no-party or one-party elections” (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021; Skaaning et al., 2015). Of the 56 remaining cases 55 didn’t reach the fourth lowest level, i.e. “regimes with minimally competitive multiparty elections” (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Gründler & Krieger, 2021; Skaaning et al., 2015). The remaining observation, i.e. Nicaragua in 1933, much like Benin in the DD indicator, fell short of the highest category since it had yet to grant women the right to vote.

If only the two lowest categories are considered autocracies and only the highest democracy, the consensus rate between the priming data and the LIED indicator laid at 97.2% of all observations. Even then the categories in between would fall short of many conceptualizations of democracy, as described in Section 2. In short there was a high degree of reasonable agreement between the priming data and the aforementioned indicator.

Freedom House (FH)

The priming data was compared to FH’s categorical scale, which distinguishes countries as free, partly free, and not free. The combined dataset included 1207 observations. Following the trend established by previous indicators, all democratically labeled observations in the priming data were coded as free in the FH index.

The autocratically labeled regimes see less of a consensus. Though none were coded as free, just shy of 10% of autocratic regimes were coded as partly free, meaning there was a consensus among that subset of over 90%. Of the overall data, around 8.7% were given a combined average of the Political Rights and Civil Liberties variables between 5.5 and 5, meaning they are on the lower end of the full partly free range, i.e. from 5.5 to 3.0. Only 1.3% of the observations ranked above 5 and none above 4. These cases were categorized as autocracies by all of the other indicators used in this section, with the sole exception of Uganda in 1980, which has already been discussed previously. Hence, it is reasonable to conclude that there was considerable consensus between the priming data and this indicator as well, albeit to a lesser extent than the indicators discussed so far.

Polity Index

The Polity V index is also ordinal and hence also nuanced. In its totality, it encompassed 2032 observations. Notably, though, 112 of these observations fell under one of the "standardized authority scores" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Marshall & Gurr, 2020). These include countries with a systemic interruption due to foreign influence, cases of interregnum or anarchy, and countries in transition. Much like all other indicators, there was overwhelming consensus on the top end of the scale. All democracies in the priming data had a score of +9 or above, with the sole exception of France in 1946 which was coded as being in transition.

Again, much like previous indices, the situation was far more complex in the bottom end of the spectrum. Nearly 7% of all observations identified as autocracies in the priming data were coded as being part of the aforementioned standardized group. A vast majority of the remaining observations, i.e. around 91%, were given a Polity rating of -6 or lower. Around another 7.5% scored between -5 and -3, which would put them in the low end of the anocracy category as established by the Polity Project. However, 1.4% scored between 0 and +4 and one observation scored a +8. This last case relates to Greece in 1944 and has been discussed before. As mentioned in Section 2, the aggregation rule of the Polity IV and V indices is methodologically problematic. By weighing all regimes equally and using an additive rule, there are doubtful classifications at the lower end of the scale (Gründler & Krieger, 2020; Munck & Verkuilen, 2002; Teorell et al., 2019), a conclusion that is supported by other indicators (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). Apart from the sole cases of Greece om 1944 and Uganda in 1980, which have already been discussed, all other regimes selected as autocracies in the priming data and with a Polity score above 0 were not coded as democratic by any other indicator discussed in this section.

Summary

Table 1 - Consensus between priming data and other indicators

	Democratic regimes				
	Polity (≥ 6)	LIED (=6)	FH (Free)	BMR (=1)	DD (=1)
Overlap	1.0	1.0	1.0	1.0	1.0
	Autocratic regimes				
	Polity (≤ -6)	LIED (≤ 1)	FH (Not Free)	BMR (=0)	DD (=0)
Overlap	0.93	0.9721	0.9	0.9989	0.9981

Notes: This table summarizes the rate of overlap between the share of country-year observations that were labeled "autocratic" or "democratic" and classified as such by alternative indices. The Polity score does not include countries that were labelled with the "standardized authority scores" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Marshall & Gurr, 2020).

Table 1 summarizes the results of this section. It shows that there was a high degree of overlap, i.e. consensus, between the priming data and other indicators. There was full consensus for all

observations in the top end of the spectrum. The same could not be said for the bottom end. However, despite there not being full consensus, the rate of overlap was high for all other indicators. The little disagreement there was, was not completely unreasonable, since, as pointed out in this section, the vast majority of these discrepancies could be traced either to uncertainties on how to code transitional regimes or on methodological differences between the indicators.

4.2. Representativeness

The priming dataset was controlled for representativeness in terms of distinct institutional characteristics, as well as timespan. The timespan was calculated in two ways. Once as the total ultimate duration of the regime, regardless as to what year the observation pertains. And once as to the duration of the regime up to the observation year. If a regime had not ended by the time the original dataset (Geddes et al., 2014) was publicized, the end year is considered to be 2014, when that took place. The objective of this section was not to check whether all regime characteristics are equally represented in the data, but rather to simply confirm whether the characteristics were heterogenous or not.

Autocracies

For the countries that scored below the first decile of the scale of either the EDI (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Teorell et al., 2019) or the UDS (Marquez, 2022; Pemstein et al., 2010), they were controlled for type, electoral regime, and durability.

In terms of regime types, all of them were included in the dataset, though to different extents. Royal and military dictatorships were the best represented in the priming data, together encompassing nearly 80%. Communist dictatorships comprised 15.35% and civilian autocracies came last with just over 8%. That being said, civilian autocracies comprised 101 country-year observations.

In terms of electoral systems, 65.9% of them had no elections, whereas 34.1% did. In the latter group 21.15% had non-democratic multi-party elections and 12.69% had single-party elections. 0.25%, i.e. 4 observations, had democratic elections. It is important to point out that, in all of these 4 observations suffrage was only available for the male part of the population. Since Bjørnskov and Rode (2020) distinguish between free elections and suffrage, this was possible. However, this wasn't problematic as many conceptions of free and fair elections consider universal suffrage a characteristic of paramount importance.

In terms of duration, summary statistics were calculated for both overall duration of regime, and duration up until the year of observation. For both categories the minimum duration was less than a year. For the former, though, the first quartile was under 23 years, the median was 44, the third quartile under 65, and the mean was just shy of 60. For the latter, the first quartile laid under 5 years, the median 14, the third quartile under 35 and the mean nearly 34. For both categories the max was over 250 years, i.e. 272 and 258 respectively.

Democracies

For the countries that scored above the ninth decile of the scale of either the EDI (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Teorell et al., 2019) or the UDS (Marquez, 2022; Pemstein et al., 2010), they were controlled for form of government, number of parliamentary chambers, voting system and durability.

For the form of government, the distribution was not uniform but there were comprehensive number of regimes for all three options. Nearly 67% of observations were parliamentary democracies, and both mixed, i.e. semi-presidential, and presidential democracies encompassed 16.9 and 16.25 percent of observations, respectively. The distribution of the total number of chambers was nearly cut down the middle. Nearly 49% of legislatures were unicameral, whereas just over 51% were bicameral. On the other hand, the variations of voting systems were more unequally distributed. While just shy of 91% of systems were proportional, a little over 9% of them were majoritarian.

In terms of duration democratic regimes also varied substantially, however not as extremely as autocracies. The total length varied from 21 to 143 years, where the first quartile was below 68, the median was 106, the third quartile was below 128, and the mean was 101. The duration of the regime at the time of the observation spanned from 7 to, as well, 143 years. The first quartile was below 56, the median was 80, the third quartile was below 106, and the mean was 81.

Summary

All regime characteristics were to varying extents included in the priming dataset. Though said characteristics were not always equally distributed, neither democracy nor autocracy in the dataset were a homogenous group, at least in the terms herein established. In the one shared characteristic, i.e. duration, they differed both internally and in relation to each other. The spread of regime duration varied substantially in both, thought the overall variation was larger for autocracies. That being said while democracies were fairly well distributed, since

mean and median have similar values in both of the aforementioned types of length, autocracies overall tended to last less. The mean for both types of duration came close to the values for their individual third quartiles. This indicated that while some outliers pulled the mean up, the vast majority of autocracies did not last relatively long, with a duration at time of observation for $^{3}/_{4}$ of the data under 35 years and an overall duration under 65. While these caveats are important to point out, and indicate that democratic regimes were generally more stable, this also shows that both democracies and autocracies were also not homogenic in term of their duration.

5. Analysis

The analysis of the results was conducted in two main ways. First it was important to analyze the aggregation function. One concern with the aggregation procedure herein used is that it is not possible to provide a formula which describes explicitly how individual characteristics affect the final level of democracy (Gründler & Krieger, 2021). Hence, the aggregation cannot be justified by theoretical assumptions (Gründler & Krieger, 2021). This is gap that this contribution attempted to fill. To do that it compared the aggregation method applied here to an already existing and established aggregation method.

In that effort the index that served as baseline of comparison was the EDI (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). There were many reasons for this. For one, as mentioned previously, the breadth of data made available in the V-Dem dataset can hardly be overstated. It contains more than 4000 variables for more than 200 countries from 1789 to 2020 (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). Furthermore, the selection and construction of regime characteristics is theoretically well grounded, each rating is given a confidence interval, the raw data is publicly available, and its codebook clearly describe coding and aggregation methods (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Gründler & Krieger, 2021). Lastly, the EDI employs a combination of most traditional forms of aggregation (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Gründler & Krieger, 2021).

The second step was to compare the results herein developed, not only with the already existing SVMDI index, but with their different breadths of Merkel's (2004) concept of Embedded Democracy. This was an effort to conclude to what extent can an SVM based democratic index with broader concepts of democracy effectively identify gradations of

democracy. Since there is no simple objective way to measure it, this contribution analyzed specific cases, which could indicate strengths and weaknesses of the different variations.

5.1. Aggregation

The EDI is composed of five mid-level composite indicators, three of them are measures taken from various sub-indicators and aggregated by a Bayesian Factor Analysis Model (BFAM), one manually coded measure, and one objective measure (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021). These were described further as they constitute the variables used for both the EDI and the SVM index. Since the main issue the SVM index was attempting to address was the unknowable aggregation function, no composite index was used, but rather their subindices. The reason for this was simple. These composite indices are themselves aggregated. If the objective was for the SVM to decide how to aggregate variables and to then compare the aggregation results with the EDI, then it was important to use the lowest level variables available.

The first of the aggregated indices measures freedom of association (v2x_frassoc_thick) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). It is itself composed of six low-level indicators, i.e. party ban (v2psparban), barriers to parties (v2psbars), opposition parties autonomy (v2psoppaut), elections multiparty (v2elmulpar), Civil society organization (CSO) entry and exit (v2cseeorgs) and CSO repression (v2csreprss) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). CSOs are considered "interest groups, labor unions, religiously inspired organizations (if they are engaged in civic or political activities), social movements, professional associations, and classic non-governmental organizations (NGOs), but not businesses, political parties, government agencies, or religious organizations that are primarily focused on spiritual practices" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021, p. 192).

The second aggregated index measures clean, free, and fair elections (v2xel_frefair) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). It is composed of eight low-level indices, i.e. *Election Management Body* (EMB) autonomy (v2elembaut), EMB capacity (v2elembcap), election voter registry (v2elrgstry), election vote buying (v2elvotbuy), election other voting irregularities (v2elirreg), election government intimidation (v2elintim), non-state electoral violence (v2elpeace), and election free and fair (v2elfrfair) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). It is important to mention at this point that many of the characteristic that relate to elections in these two composite indices were

only available during election years and if there was an electoral regime to speak of (v2x_elecreg). Those characteristics were modified so that the data repeated itself in the years in between elections, as recommended by V-Dem's codebook (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). Furthermore, should an election data be missing in the first year of analysis, the data was backfilled with the most recent value, as V-Dem does with the US and the UK (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

The third aggregated index measures freedom of expression and alternative sources of information (v2x_freexp_altinf) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). It is composed of 9 low-level indicators, i.e. media censorship effort (v2mecenefm), harassment of journalists (v2meharjrn), media bias (v2mebias), media self-censorship (v2meslfcen), print/broadcast media critical (v2mecrit), and print/broadcast media perspectives (v2merange), freedom of discussion for men (v2cldiscm), freedom of discussion for women (v2cldiscw) and freedom of academic and cultural expression (v2clacfree) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

The last non-objective indicator measures whether the chief executive and the legislature are elected, directly or indirectly, through popular elections (v2x_elecoff). This indicator is a manually coded, expert-based indicator. It is very minimally defined and only explicates whether there are elections for the aforementioned high offices or not. Some elections that one could call controversial are still counted in this indicator (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

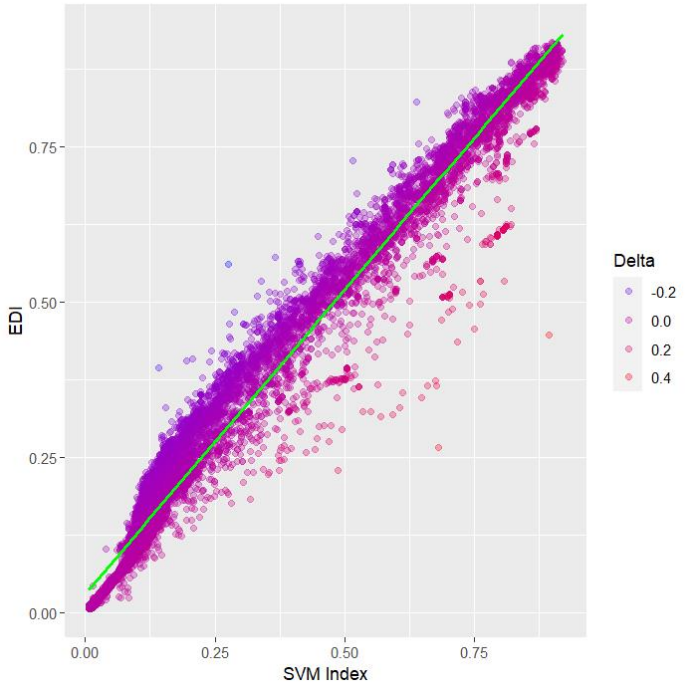
The last indicator measures the degree of suffrage within the populace (v2x_suffr) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). It is objective, composed of factual indicators and provides the share of the adult population that is granted suffrage. This only encompasses *de jure* suffrage, not *de facto*. This means that it does not take into consideration restrictions based on age, residence, criminal record, or legal incompetence (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

These indicators are aggregated in three steps. First, they are aggregated into a multiplicative indicator (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021). All values are multiplied equally. Second, they are aggregated into a weighted additive index. The BFAM aggregated indices are multiplied by 0.25, while the other two are multiplied by 0.125, which is supposed to penalize the latter (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021). "The additive part of the formula lets the two components that can achieve high scores based on the

fulfillment of formal institutional criteria (elected officials and suffrage) together weigh half as much as the other components that enjoy a stronger independent standing in terms of respect for democratic rights (clean elections, freedom of organization and expression)" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021, p. 5). An average is then taken of the two indices to form the final EDI (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021).

The SVM index, on the other hand, used the low-level indices without aggregating them into mid-level indices. They were, hence, considered as individual variables. However, due to the coding rules for some of them, a fair amount of the data had to be adapted. For starters, to be able to track different country-year observations reliably, a form of code had to be used to identify each country. Country names are not ideal in that effort, since names can change over time or a country might dissolve into multiple entities, e.g. Germany, Czechoslovakia, Vietnam (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). To address this issue, this contribution made use of the ISO 3166-1 Alpha-3 standard (International Organization for Standardization [ISO], 2021). For countries that no longer exist as formal entities, e.g. the German Democratic Republic, their historic codes were used (ISO, 2021). These were then combined with the variable related to the year in question to create individual identifiers. Furthermore, one could assume to make the coding of variable less cumbersome, some expert-based variables are coded in the V-Dem dataset as missing based on other expert-based variables. For instance, the autonomy of opposition parties (v2psoppaut) is generally only coded if the barriers to forming parties (v2psbars) are not so restrictive as to not allow non-government affiliated parties to form. The same is true for many electoral variables. Unless there are elections, they are not at times coded. If all of those observations were removed, many countries in the lower end of the spectrum would be excluded from the dataset. To compensate for this,

Figure 1 - Distribution between EDI and the SVM Index



such variables were given the lowest possible score, as that relates to the intent of not coding them (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

A further, yet minor problem, is that about 0.03% of the datapoints in the observations were missing, hence there were minor gaps in the data. To deal with this issue and avoid the removal of entire observations due to occasional minor gaps, cubic spline interpolation was used to fill them. This technique is often used for such purposes (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Gründler & Krieger, 2021). The objective here was to calculate a function that best described the points in a dataset by finding the polynomial of the lowest possible degree and utilize that to fill gaps in the data (Korstanje, 2021). While in a large scale that would skew the data, a minimal amount of the dataset was interpolated and hence shouldn't have a substantial effect.

Taken that into account, the final dataset had 16390 observations, and the priming data had 4616. The priming data, as described in Section 3, was used to train the dataset and then it calculated the index for each observation in the total dataset 2000 times. The final index was then extracted by taking the median of all iterations for each observation.

These SVM results were then compared with the original EDI. An additional variable was created by calculating the delta between both

Figure 2 - Index development over time for Germany

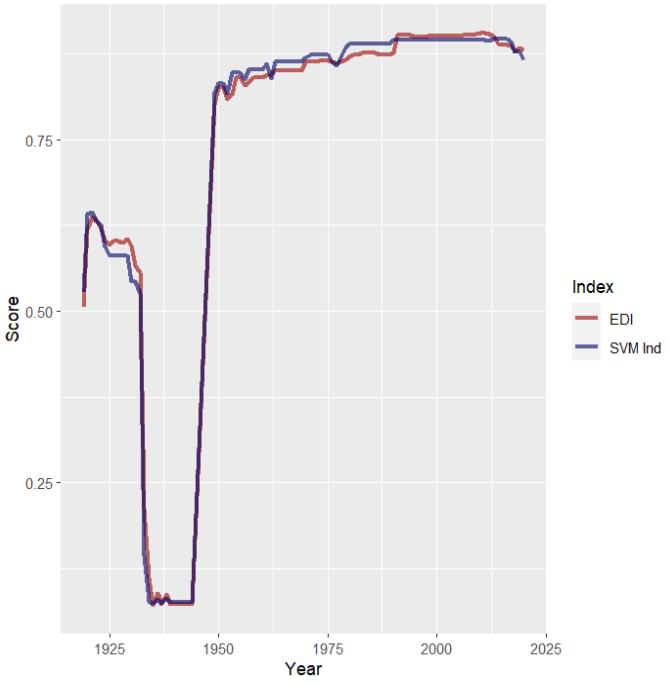


Table 2 - OLS Regression for EDI and SVM Ind.

	Dependent variable:	
	EDI (1)	SVM Ind. (2)
v2psparban_ord	-0.008***	-0.018***
	(0.001)	(0.001)
v2psbars_ord	0.004***	0.004***
	(0.001)	(0.001)
v2psoppaut_ord	0.021***	0.023***
	(0.001)	(0.001)
v2elmulpar_ord	-0.009***	-0.010***
	(0.001)	(0.001)
v2cseeorgs	0.006***	0.008***
	(0.001)	(0.001)
v2csreprss	0.017***	0.018***
	(0.001)	(0.001)
v2x_suffr	0.111***	0.073***
	(0.003)	(0.003)
v2elembaut	0.023***	0.028***
	(0.001)	(0.001)
v2elembcap	0.008***	0.008***
	(0.001)	(0.001)
v2elrgstry_ord	0.008***	0.009***
	(0.001)	(0.001)
v2elvotbuy_ord	0.010***	0.010***
	(0.001)	(0.001)
v2elirreg_ord	0.021***	0.025***
	(0.001)	(0.001)
v2elintim_ord	-0.003**	-0.007***
	(0.001)	(0.001)
v2elpeace_ord	0.006***	0.011***
	(0.001)	(0.001)
v2elfrfair_ord	0.014***	0.015***
	(0.001)	(0.001)
v2x_elecoff	0.119***	0.086***
	(0.002)	(0.002)
v2mecenefm	0.005***	0.010***
	(0.001)	(0.001)
v2meharjrn	0.008***	0.017***
	(0.001)	(0.001)
v2mebias	-0.001	-0.003**
	(0.001)	(0.001)
v2meslfcen	-0.002*	-0.003**
	(0.001)	(0.001)
v2mecrit	0.013***	0.015***
	(0.001)	(0.001)
v2merange	0.020***	0.019***
	(0.001)	(0.001)
v2cldiscm	0.006***	0.011***
	(0.001)	(0.002)
v2cldiscw	0.014***	0.008***
	(0.001)	(0.001)
v2clacfree	0.007***	0.007***
	(0.001)	(0.001)
Constant	0.030***	0.062***
	(0.004)	(0.005)
Observations	10,203	10,203
R^2	0.941	0.938
Adjusted R^2	0.941	0.937
Residual Std. Error (df = 10177)	0.060	0.066
F Statistic (df = 25; 10177)	6,498.867***	6,114.241***
Signif. codes:	*p<0.1; **p<0.05; ***p<0.01	

Notes: Table created using the stargazer function (Hlavac, 2022)

indices. As shown by Figure 1 and as expected, the two indices correlated substantially, particularly in the extremes of the spectrum. This was expected as there was little disagreement on the top and bottom ends of the scale, as mentioned previously (Cheibub et al., 2010; Gründler & Krieger, 2021; Lindberg et al., 2014). An example of this dynamic is, as Figure 2 shows, Germany, as it relates to the ISO 3611-1 Alpha 3 code "DEU". This does not include, for instance, the German Democratic Republic, i.e. "DDR" (ISO, 2021). Throughout its modern history, Germany has inhabited both ends of the democracy-autocracy spectrum. Hence, it exemplified the similarity between indices in the extremes of the scale. However, there were minor discrepancies in its middle section. Overall, no country averaged a difference in score of more than a decile. Out of those, the majority had a higher EDI than SVM index.

In the interest of pinpointing which variables might have played a role in the disparity between the two indices, the observations that diverged by more than 0.01 were analyzed separately. An OLS regression, evidenced on Table 2, showed both expected and insightful results. The correlation between the variables and the indices was noticeable and statistically relevant. This was obviously to be expected as both indices were made up of said variables. The sole exception was the influence of media bias and media self-censorship on the EDI, which were not statistically relevant. The models for both indices were themselves statistically relevant and explained a very high degree of the indicators' variance, again also expectedly.

The model for the EDI explained an even higher degree of its variance than the one for the SVM Index, which indicated that the adaptations introduced to the

dataset did not have a sizeable effect in the relationship between the variables and the index. One would expect that, should the adaptations have had a big impact on said relationship, that the index composed directly of those variables, i.e. the SVM index, would have a higher degree of variance explained by them. That was, however, not the case, albeit by a very slim margin[10].

More interestingly for the purposes of this contribution was the effect that individual variables had on the individual indices. In the SVM index no variable had an effect on the dependent variable larger than nine hundredths of the scale. The two variables with the largest effect were, in decreasing order, presence of elected officials (v2x_elecoff) and share of population with suffrage (v2x_suffr). This makes sense from a theoretical standpoint. These two characteristics are generally accepted as necessary conditions for a democratic regime (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021; Dahl, 1989; Gründler & Krieger, 2021; Merkel, 2004). However, they were but two of 15 variables that also had an effect in the hundredth range.

A similar dynamic played out with the EDI, but with key differences. The effect of nearly all variables ranged from the thousandth to the hundredth range. Much like the SVM index, the same variables in the same descending order, i.e. v2x_elecoff and v2x_suffr, were the source of the highest degrees of variance. However, unlike the SVM index, these were the only variables that had an effect in the tenth range. Furthermore, apart from two other variables, i.e. v2merange and v2cldiscw, all other variables had a similar or lower influence on the variance of the EDI in comparison to the SVM index.

Figure 3 - Index development over time for Russia

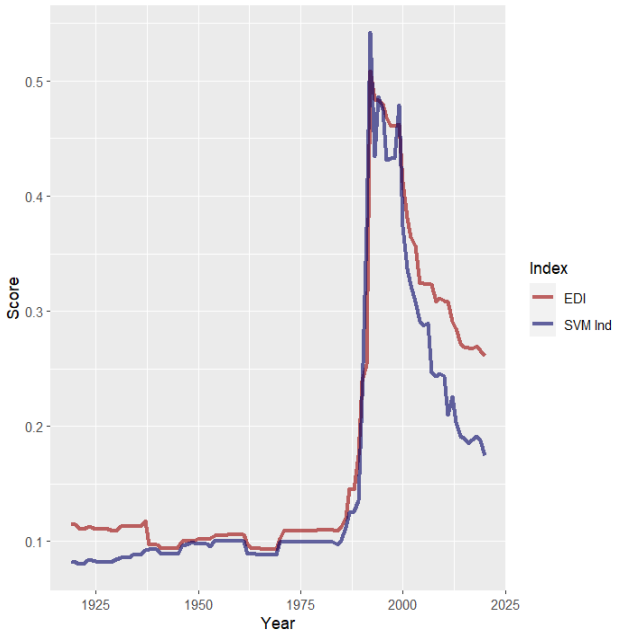
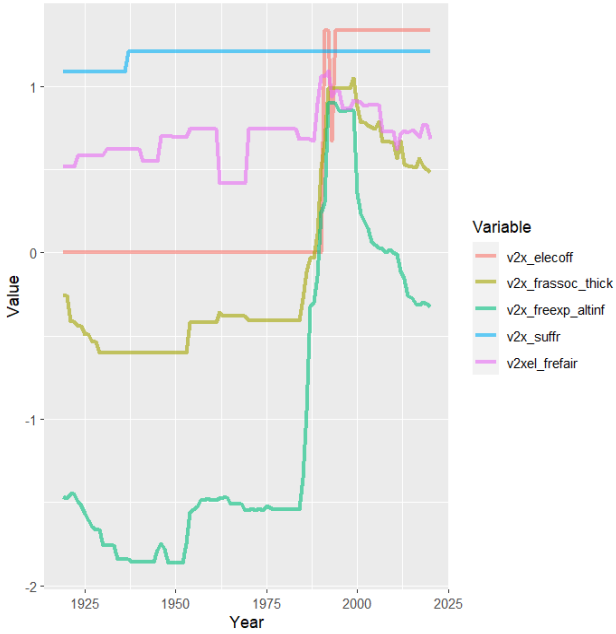


Figure 4 - Development of mid-level variables for Russia



[10] A similar dynamic appeared in the OLS model for all observations.

This indicated that, despite the efforts to penalize v2x_elecoff and v2x_suffr in the additive index, these two variables still exert a big influence on the overall EDI due to the multiplicative index. While it's important that these variables play a big role in assessing the level of democracy, they often undermine the remaining characteristics. This was further exemplified by cases such as Russia and Colombia. These cases were chosen because they undergone considerable movement in the middle part of the spectrum. Furthermore, they were good examples of different ways the disparate aggregation between the indices might have affected the results.

Russia, for instance, as shown in Figure 3, differed most noticeably between 1919 and 1937 and after 1999. During the first period the value of the SVM index laid below 0.1 whereas the EDI laid above that value. In fact, the former stayed below the latter all the way up to the early 1990s. After 1999 the SVM index dropped more pronouncedly than the EDI. The discrepancy in the early 20th century, considering the information of the OLS regression, it could be argued, could be attributed to the high levels of suffrage in this period. Figure 4 shows the development of the individual categories that compose the EDI over time. These categories were composed of the average centralized value of the low-level indices for each corresponding mid-level index. While said figure did not allow for the comparison of their specific values with each other, they did allow one to attempt to identify which of the categories might have had an effect on the overall index. Indeed, while the aggregated indices had a larger effect on the SVM index, the EDI was propped up by the high level of suffrage. Throughout the 20th century the SVM index seems to have been particularly sensitive to the v2xel_frefair category, which relates to free and fair elections. The variation of the SVM index in both the 1930s and 1990s

Figure 5 - Index development over time for Colombia

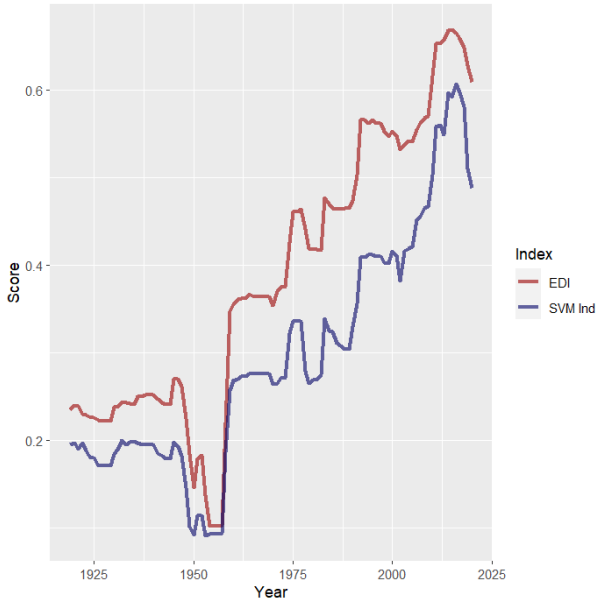
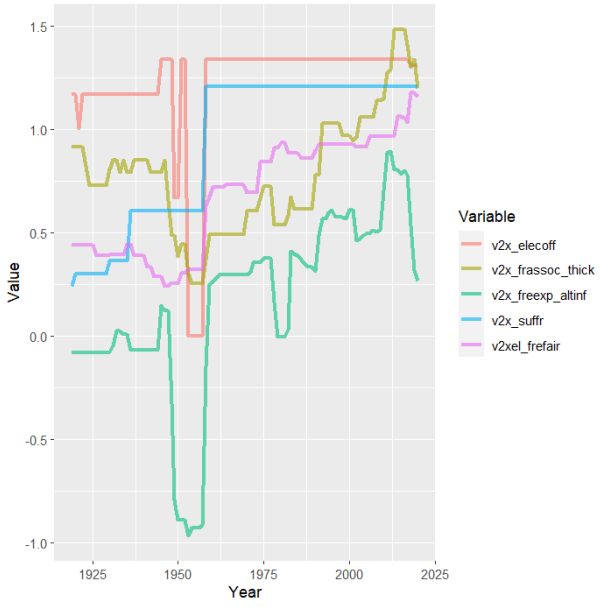


Figure 6 - Development of mid-level variables for Colombia



was an indication of that. However, this dynamic was further evidenced by the development of the categories post 1999. While v2x_elecoff and v2x_suffr, which are not aggregated using a BFAM on low-level indices, remain relatively high, the other three aggregated categories all decrease in that period, which led to a much steeper curve in the SVM index.

Colombia was a further example, though the dynamics of its development were slightly different than Russia's. As Figure 5 indicates, there was no period in which the SVM index went above the EDI. In fact, the SVM index was consistently and considerably below the EDI. However, both followed relatively similar patterns. This showed that both are similarly susceptible to the variation in the individual categories, as indicated by Figure 6. The difference laid in the relatively high levels of the v2x_elecoff variable throughout this period. The closest the two indices come to each other was during the mid-1950s where there was a steep drop-off in the v2x_elecoff and a middling v2x_suffr variable. However, when both of them rose simultaneously, there was a much steeper climb in the EDI, than in the SVM index. Thereafter, due to high levels of v2x_elecoff and v2x_suffr, the EDI remained consistently higher than the SVM index.

Both these cases indicated that, in line with the findings in the OLS regression, while all categories affected the variation in the individual indices, their difference in aggregation produced disparate results. Though the weighted average of the additive index of the EDI attempts to compensate for the BFAM aggregation of certain variables by punishing the v2x_suffr and v2x_elecoff, this is not enough to overcome the effect of the multiplicative part of the aggregation. While this is less of a problem for countries in the extremes of the spectrum, it does present an issue for hybrid regimes. As long as a country has suffrage and elected officials, the effects of freedom of expression, freedom of association, and free and fair elections are much less pronounced. By taking all low-level indices into account without the need of a somewhat arbitrary preceding aggregation step, as well as aggregating them without resorting to additive or multiplicative rules, the SVM index was more sensitive to the level as well as changes of the latter three categories.

5.2. SVM Indices

Having established that the aggregation by SVM regression did not under weigh certain variables, it was important to check if it was also useful at measuring differing conceptual breadths of democracy. At this stage, the same parameters of the SVM regression were applied to three different datasets, each encompassing different breadths of the theory of democracy proposed by Merkel (2004). The narrowest only encompassed the electoral regime, which

related to the concept of Polyarchy as proposed by Dahl (1989) and delineated by Merkel (2004). The second included all factors that encompass the internal embeddedness of democracy. Lastly the third dataset went beyond the other two and included the external embeddedness as well.

All datasets were given the same treatment in terms of identifying observations as the dataset in the previous section. Country names were converted into their ISO 3166-1 Alpha 3 codes (ISO, 2021). Countries that do not exist as formal entities anymore were given their historic codes (ISO, 2021). To identify each observation the ISO codes were merged with the year to create a unique identifying value. Furthermore, much like the in the previous sub-section, certain variables were only coded if other variables had a certain value, e.g. certain electoral variables (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). This was deal with the same way as in Section 5.1.

To avoid conceptual overlap and to take advantage of V-Dem's regime characteristics, which are well grounded in theory (Gründler & Krieger, 2021), this contribution attributed said characteristics to the different regimes established by Merkel (2004). Many of the variables that constituted the first dataset, rather expectedly, were also part of the EDI. Characteristics such as v2x_suffr, v2x_elecoff, and the low-level indices that composed v2xel_frefair were all used for this dataset. However, since Merkel (2004) categorized freedom of association in the Political Rights regime and left the right to candidacy in the Electoral Regime, the low-level indices of v2x_frassoc_thick were added to the former and v2elrstrct, which measures candidate restriction by ethnicity, race, religion, or language, was added to the latter (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

The second dataset was, expectedly, much more comprehensive as it was based on all internally embedded regimes. Obviously, it included the characteristics of the Electoral Regime. Since it included all internally embedded regimes, it, as alluded previously, included all low-level indices of v2x_frassoc_thick (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). Furthermore, that regime also included freedom of expression, which is embodied by the mid-level characteristic v2x_freexp, hence all its low-level indices were also included. These are print/broadcast censorship effort (v2mecenefm), harassment of journalists (v2meharjrn), media self-censorship (v2meslfcen), freedom of discussion for men/women (v2cldiscm, v2cldiscw) and freedom of academic and cultural expression (v2clacfree) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). The next regime, namely the regime of Civil Rights was just as comprehensive, however it was composed of all low-level indices of one single mid-level indicator, namely v2xcl_rol, which measures equality

before the law and individual liberty, i.e. the two categories of this regime. These indicators were rigorous and impartial public administration (v2clrspct), transparent laws with predictable enforcement (v2cltrnslw), access to justice for men/women (v2clacjstm, v2clacjstw), property rights for men/women (v2clprptym, v2clprptyw), freedom from torture (v2cltort), freedom from political killings (v2clkill), from forced labor for men/women (v2clslavem v2clslavef), freedom of religion (v2clrelig), freedom of foreign movement (v2clfmove), and freedom of domestic movement for men/women (v2cldmovem, v2cldmovew) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). The second to last regime was the regime of Horizontal Accountability, which was measured by two mid-level indicators, namely the judicial constraints on the executive index (v2x_jucon) and the legislative constraints on the executive index (v2xlg_legcon) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). Hence, the full list of variables was executive respects constitution (v2exrescon), compliance with judiciary (v2jucomp), compliance with high court (v2juhccomp), high court independence (v2juhcind), lower court independence (v2juncind), legislature questions officials in practice (v2lgqstexp), and legislature investigates in practice (v2lginvstp) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). The only two exceptions were the indicator measuring legislature opposition parties (v2lgoppart), "as this aspect is part of vertical accountability" (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021, p. 288), and executive oversight (v2lgotovst) as it does not relate to the legislature (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). Finally, the last regime was the least comprehensive one, namely the regime of Effective Power to Govern. It included just two low-level variables, namely whether the head of government and whether the head of state sought approval from non-democratic actors in the conduction of domestic policy (v2exctlhg_0, v2exctlhs_0). If the head of state is also the head of government (v2exhoshog), then both variables were given the same value.

The third dataset, i.e. the fully embedded one, included all the previously discussed variables in this section, but it also included the two of the three externally embedded regimes as discussed in Section 2. The civil society component was embodied by the mid-level index civil society participation (v2x_cspart). It includes the characteristics candidate selection — national/local (v2pscnslnl), CSO consultation (v2cscnsult), CSO participatory environment (v2csprtcpt), and CSO women participation (v2csgender) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021).

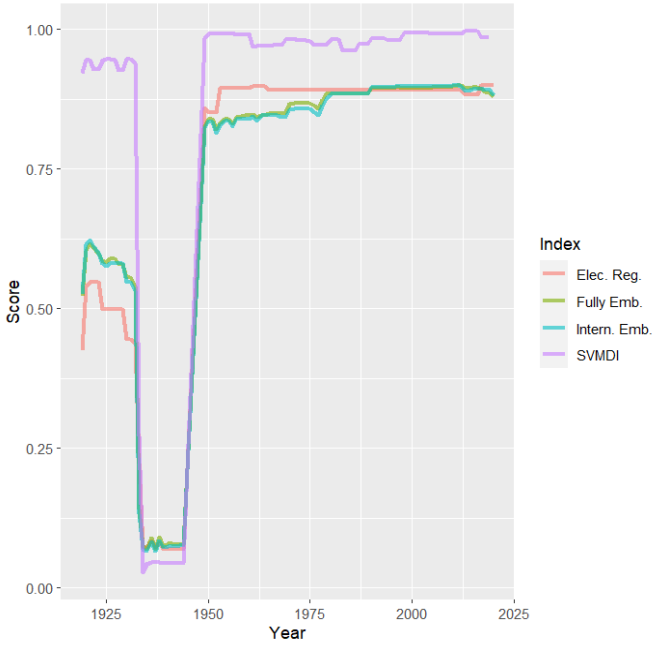
It becomes slightly more complicated with the socioeconomic context. There is no V-Dem mid-level indicator that measures concepts related to that regime. However, the V-Dem

dataset made external variables available that cut across the issues mentioned by Merkel (2004). These are educational inequality - Gini (e_peedgini), GDP per capita (e_migdppc), and life expectancy (e_pelifeex) (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021). They were chosen for different reasons. GDP per capita, for instance, was named by Merkel (2004) as a measure of the socioeconomic context. Life expectancy is a good measure of general poverty as there is a strong association between it and wealth distribution (Canudas-Romo, 2018). Lastly, as a measure of education that could lead to low-intensity citizenship (Merkel, 2004) as well as a measure that correlates with general income inequality (Gregorio & Lee, 2002; Lin, 2007), educational inequality was selected.

All datasets had gaps. For the first two datasets the gap was no larger than 0.21% of datapoints. These gaps were filled either using polynomial or liner interpolation, as it is done for some variables in the V-Dem dataset (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021) and accepted by Gründler & Krieger (2021). The former was applied to continuous variables, and the latter was applied to ordinal variables with missing data. For the last dataset, though it got slightly more complicated. However, it got more complicated with the last dataset. The wide spectrum of data meant that there are more gaps in the data than in previous examples. Where it was possible, these gaps were filled much like with the previous datasets. In the end, the first two dataset retained more than 99% of its original number of observations. The third dataset, on the other hand, had to lose over 4000 observations, since it missed a lot of datapoints particularly related to the economic measures and hence couldn't be interpolated.

The three datasets were compared not only with each other but also with the already existing SVMDI, which had a narrow conception of democracy (Gründler & Krieger, 2021). The objective was two-fold. First, to conclude whether or not the SVM index in this contribution could come to reasonable conclusion with a large amount of data. Second, if it could reasonably identify changes in democratic quality compared to the SVMDI. In that effort three cases were analyzed more closely. Germany was analyzed as it was a good example of a country that has been in both ends of the democracy-dictatorship spectrum and hence it served the purpose of analyzing how the indices perform on the extremes. Two cases of recent democratic backsliding that have already been discussed at length in this contribution, Hungary and Poland in the 21st century, presented a great opportunity to assess how these indices react in the middle of the spectrum.

Figure 3- Index development over time for Germany



Germany, for starters, was an example of potential pitfalls of narrow conceptions of democracy. As Figure 7 shows, during the 1920s and early 1930s the reigning government was the Weimar Republic. Though having a very modern constitution by the standards of its day, in practice its democracy was far from perfect (Sheehan, 2022). This was reflected in the results of the SVM indices herein created. The index to penalize this regime the most was the narrowest SVM index, based solely on the Electoral Regime. The one to penalize it the least was the SVMDI. This underscores the importance of a solid theoretical basis for variable selection. Though both indices are minimalistic, the SVM index based on the Electoral Regime considered the fact that during many years of the Weimar Republic there were comparatively serious issues of voting irregularities and, particularly in the later years, with non-state electoral violence (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021; Schumann, 2001). In fact, in the variables related to correctly organized, free and fair elections, the Weimar Republic failed to achieve the highest score in any of its metrics (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). However, by focusing mostly on ratio of voters in relation to population and shares of seats in parliament, the SVMDI gave it a near perfect score (Gründler & Krieger, 2021). The other, broader indices gave the Weimar Republic a slightly better score, since it did score relatively higher in other metrics related to Civil, and Political Rights, and Horizontal Accountability (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021). Thereafter, all indices followed similar trends, though not exactly the same. For the reasons mentioned above the SVMDI gave post-war Germany at times a perfect score, whereas the other three indices did not. The Electoral Regime Index rose fairly quickly after 1949 as well, probably due to its aforementioned narrow conceptual focus. The broader indices rose more slowly, likely due to Germany's comparatively lackluster score in some metrics related to Press Freedom and Civil Rights (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021).

Hungary and Poland in the 21st century, as shown in Figure 8 and Figure 9 respectively, were examples of another issue related to narrow conceptual breadth. Considering what was discussed in Section 2, it is not impressive, for instance, that Hungary declined in all indices. However, the decline was steeper in the two conceptually broader indices. This was probably due to the steep drop in variables related to Horizontal Accountability. However, recent issues with the electoral process (Bayer, 2022) have brought the Electoral Regime SVM index even further down than the broader conceptions. The SVDMI, hence, has also decreased, though not nearly to same extent, likely due to the same conceptual issues described previously. While the other indices dropped by almost four tenths, the SVDMI, reached its zenith at a near perfect democracy in around 2008 and dropped less than two tenths by 2019.

Figure 5 - Index development over time for Hungary

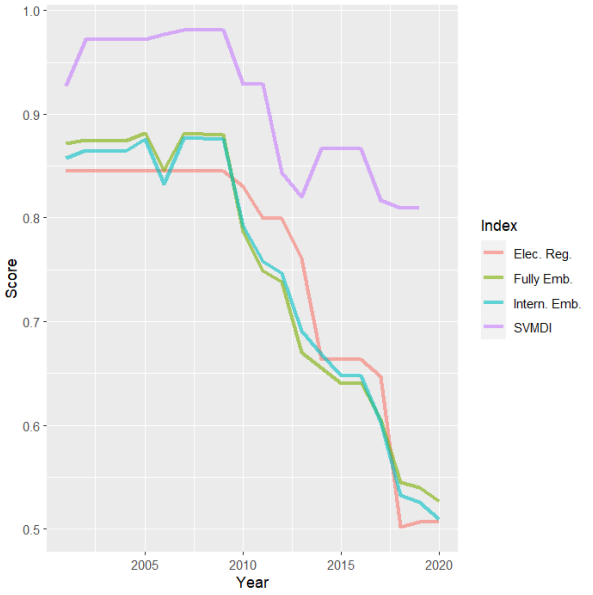
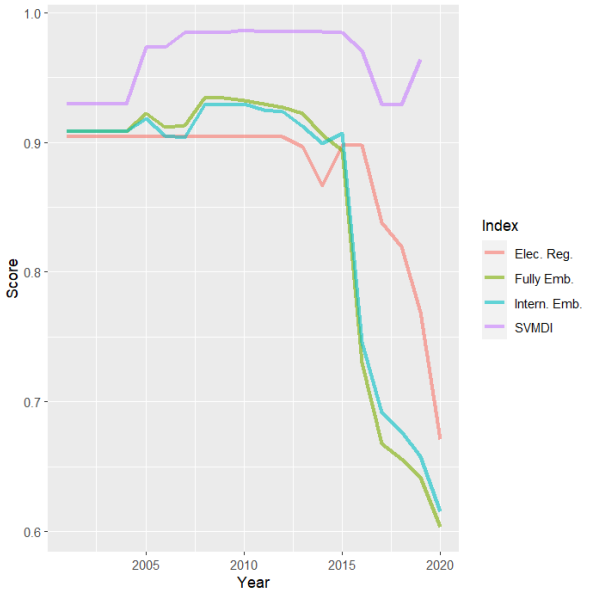


Figure 5 - Index development over time for Poland



Poland, in the same time period was an even better example of this dynamic. For most of the 21st century all indices categorized Poland in the top tenth of the scale. However, for all of them with the sole exception of the SVMDI, the early 2010s saw a slow decline in the indices. By the middle of the decade the two broader indices went into a steep decline, likely due to the issues discussed in Section 2, related to its judiciary. This is supported by the fast decline of variables related to Horizontal Accountability, which the two narrower indices do not consider. The SVM index based on the Electoral Regime only drops when the capacity and, more notably, the autonomy of the EMB go down. The SVDMI, however, stayed comfortably in the top tenth, and mostly ignored these developments. This shows a big conceptual gap in the original SVDMI, that was not missed by the other indices.

6. Conclusion

Measuring democracy is not a straightforward task, but it is of paramount importance. To measure it, one requires "an operational definition of democracy precise enough to classify specific countries as we observe during particular historical periods" (Pérez-Liñán, 2017, p. 86). Furthermore, if this definition is not well thought out, this conceptual low standard might lead to unreasonable results. "In other words, gradations of democratic quality cannot be detected" (Geissel et al., 2016, p. 572). A similar dynamic plays out when it comes to the aggregation. One needs a way to aggregate the characteristics of democracy so as to come up with reasonable results (Gründler & Krieger, 2016, 2021). This is a major issue and point of contention in the current landscape, as described in Section 2 and evidenced by the current literature (Gründler & Krieger, 2021). One can employ a concept of democracy of a high standard and still come up with unreasonable results due to a problematic aggregation. Conversely one can employ a well theorized aggregation and still be faced with unreasonable results, due to a poor selection of characteristics.

Conceptualization and aggregation, hence, are central issues in this debate (Gründler & Krieger, 2021). This contribution by no means attempted to outright solve these issues. However, it did attempt to employ and improve upon novel methods of aggregation, specifically machine learning in general and SVM regression in particular. In recent years such tools have increasingly been employed in the attempt to measure democracy (Gründler & Krieger, 2016, 2020, 2021). However, one could argue, they are still in their infancy in this field. The contribution herein provided was to show potential benefits of such methods in relation to an established aggregation method, as well as expand on its potential by employing broader definitions of democracy than had hitherto been used.

Hence, in that effort, the aggregation function was isolated as a factor of influence in the indicator, by selecting a prominent index, i.e. the EDI, using the same variables as it does, calculating the SVM index, and comparing the two. This comparison indicated that two variables, the only ones that were either manually coded or objective, had a very high degree of influence over the resulting EDI, as opposed to other variables. This happened, despite overt efforts in the aggregation process to reduce their influence (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Altman, et al., 2021; Coppedge, Gerring, Knutsen, Lindberg, Teorell, Marquardt, et al., 2021). Observations that had low scores on the other variables were given a relatively high overall EDI, because the aforementioned two characteristics had high scores.

The SVM index, on the other hand, could use all manually coded, objective, and low-level indices to calculate the result, without needing prior aggregation into mid-level indices. This does not mean that every variable had the same effect, but that no arguably arbitrary decision in that regard had to be made *a priori*. Hence, this further confirms that the aggregation method of the SVM provided a benefit, as it removes the necessity for arbitrary decisions as to the unknown relationship between characteristics and shifts that responsibility into the context of nonlinear optimization (Gründler & Krieger, 2016, 2021).

The conceptualization issue, on the other hand, is harder to improve as it is often contextual (Bjørnskov & Rode, 2020; Gründler & Krieger, 2021; Gutmann & Voigt, 2018). There are usually two counteracting interests. If a concept is too broad, it might conceptually overlap with the object of study. Furthermore, the broader a concept is the more likely it is that there are gaps in the data and that, hence, the coverage of the resulting index is restricted in terms of the number of countries and timespan, which, as shown in Section 2, is also a detriment. However, if the conceptualization is too narrow, then the index might be unable to identify variations in the gradation of democracy. Section 5 exemplified both of these issues by comparing SVM indices composed of different breadths of democracy to each other, as well as to the existing SVMDI (Gründler & Krieger, 2021) which has a narrow conception. This comparison took place in the form of three cases. They exemplified why, despite an improvement in aggregation, it is paramount to apply it to high-standard input variables. Without that, no matter how effective the aggregation might be, the index might be unable to identify variations in the gradation of democracy. The SVMDI overlooks serious issues that even the SVM index with the narrowest conception of democracy was able to capture. While the objective of the SVMDI was to avoid conceptual overlap, it also often gave completely unreasonable scores. Though covering all countries in the dataset does not fall under the purview of this contribution, many are notable examples of this dynamic. Beyond the cases mentioned in Section 5, Russia in 1999, for instance was given by SVMDI a score only 0.05 shy of a perfect democracy (Gründler & Krieger, 2021). This not only contradicted established democracy indicators (Coppedge, Gerring, Knutsen, Lindberg, Teorell, Alizada, et al., 2021), it also contradicted real political issues at the time (Riasanovsky, 2022), as well as the indices herein developed.

Despite the comparative large number of variables, the SVM index based on all internally embedded regimes was capable of identifying variations in democratic quality, when, for instance, the independence of the judiciary was restricted. This was shown by the cases of Hungary and, particularly, Poland. Despite the high number of characteristics, the SVM

regression was able to analyze them effectively and assess the quality of democracy and was sensitive to changes in individual regimes. On the other hand, the fully embedded SVM index suffered. While it did also come up with reasonable results that were often in line with the internally embedded SVM index, the externally embedded variables played a very small role in the result and the index had to be restricted substantially in terms of coverage.

This contribution then showed that a machine learning based approach to measuring democracy can be applied to broader conceptualizations of democratic regimes and still effectively identify variations of democratic quality. Furthermore, its aggregation method has proven itself capable of aggregating a large number of characteristics without *a priori* arbitrary decision on aggregation. While it improved the aggregation concern, this contribution also showed that one cannot escape the issue of conceptualization. It is paramount to base the decision on which characteristics to include on a solid theoretical basis. A too narrow approach might lose its ability to identify variations and a too wide approach might restrict the number of observations that can be analyzed. This contribution attempted to alleviate the issue of conceptual overlap, by providing future researchers with the option to focus their conceptualization on specific regimes or contexts inherent to the concept of Embedded Democracy (Merkel, 2004). This somewhat modular approach allows researchers to isolate variables that do not overlap with their object of study, while still keeping as many characteristics as possible and without have to be overly concerned with how to aggregate them.

The field of machine learning is improving rapidly, and hence future research could contribute by employing new or alternative techniques in the effort of measuring democracy. The role of individual regimes of democracy as established by Merkel (2004) could be further investigated individually, a limitation that this contribution did not address. This contribution was also limited by the fact that observations were treated individually by the SVM regression, which did not consider that all of them belong to a specific country and timeseries. There are a series of robustness checks that have not been conducted, e.g. using alternative priming data or using incorrect labels. Furthermore, while there is consensus on the level of democracy of countries on the ends of the scale, there is disagreement on exactly how democratic/autocratic they are.

This is by no means whatsoever a definitive index to measure democracy. However, this contribution attempted to improve upon existing measures and methods and hopefully inspire future research to better identify variations in the quality of democracy.

7. Abbreviations

Abbreviation	Definition
AF	African Union
ANRR	Acemoglu-Naidu-Restrepo-Robinson indicator
ASEAN	Association of Southeast Asian Nations
BFAM	Bayesian Factor Analysis Model
BMR	Boix-Miller-Rosato indicator
CL	Civil Liberties
CSO	Civil society organization
DD	Democracy-Dictatorship Indicator
ECSC	European Coal and Steel Community
EDI	Electoral Democracy Index
EEC	European Economic Community
EMB	Election Management Body
EU	European Union
FH	Freedom House
GDP	Gross Domestic Product
LIED	Lexical Index of Electoral Democracy
MERCOSUR	Southern Common Market
ML	Machine Learning
NATO	North Atlantic Treaty Organization
OLS	Ordinary Least Squares
PR	Political Rights
RBF	Gaussian Radial Basis Function
SVMDI	Support Vector Machine Democracy Index
UDS	Unified Democracy Score
USSR	Union of Soviet Socialist Republics
V-Dem	Varieties of Democracy
SVM	Support Vector Machines

8. Bibliography

Acemoglu, D., Naidu, S., Restrepo, P., & Robinson, J. A. (2019). Democracy does cause growth. *Journal of Political Economy*, *127*(1). https://doi.org/10.1086/700936

African Union. (2000). *Constitutive Act of the African Union*.

Arat, Z. F. (1991). *Democracy and human rights in developing countries*. Boulder and London. https://xmarquez.github.io/democracyData/articles/Replicating_and_extending_the_UD_scores.html

Bayer, L. (2022, April 4). Hungarian election wasn't a fair fight, observers say. *POLITICO*. https://www.politico.eu/article/hungary-election-level-playing-field-fair-observer/

Bjørnskov, C., & Rode, M. (2020). Regime types and regime change: A new dataset on democracy, coups, and political institutions. *The Review of International Organizations*, *15*(2), 531–551. https://doi.org/10.1007/s11558-019-09345-1

Bobinski, A. (2022, April 22). Poland's Frozen Conflict over Rule of Law. *German Marshall Fund*. https://www.gmfus.org/news/polands-frozen-conflict-over-rule-law

Boix, C., Miller, M., & Rosato, S. (2013). A Complete Data Set of Political Regimes, 1800–2007. *Comparative Political Studies*, *46*(12), 1523–1554. https://doi.org/10.1177/0010414012463905

Boix, C., Miller, M., & Rosato, S. (2018). *Boix-Miller-Rosato Dichotomous Coding of Democracy, 1800-2015*. Harvard Dataverse. https://doi.org/10.7910/DVN/FJLMKT

Bollen, K. A. (1999). *Cross-national indicators of liberal democracy, 1950-1990*. Inter-University Consortium for Political and Social Research.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory - COLT '92*, 144–152. https://doi.org/10.1145/130385.130401

Bowman, K., Lehoucq, F., & Mahoney, J. (2005). Measuring Political Democracy. *Comparative Political Studies*, *38*(8), 939–970. https://doi.org/10.1177/0010414005277083

Canudas-Romo, V. (2018). Life expectancy and poverty. *The Lancet Global Health*, *6*(8), e812–e813. https://doi.org/10.1016/S2214-109X(18)30327-9

Cheibub, J. A., Gandhi, J., & Vreeland, J. R. (2010). Democracy and dictatorship revisited. *Public Choice*, *143*(1–2), 67–101. https://doi.org/10.1007/s11127-009-9491-2

Cheibub, J. A., Przeworski, A., Limongi Neto, F. P., & Alvarez, M. M. (1996). What Makes Democracies Endure? *Journal of Democracy*, *7*(1), 39–55. https://doi.org/10.1353/jod.1996.0016

Closa, C., & Palestini, S. (2018). Tutelage and Regime Survival in Regional Organizations' Democracy Protection. *World Politics*, *70*(3), 443–476. https://doi.org/10.1017/S0043887118000035

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Alizada, N., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Hindle, G., Ilchenko, N., Krusell, J., Lührmann, A., Maerz Seraphine F. and Marquardt, K. L., … Ziblatt, D. (2021). *V-Dem Country-Year Dataset v11.1*. Varieties of Democracy (V-Dem) Project. https://doi.org/https://doi.org/10.23696/vdemds21

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Bernhard, M., Cornell, A., Fish, M. S., Gastaldi, L., Gjerløw, H., Glynn, A., Hicken, A., Lührmann, A., Maerz, S. F., Marquardt, K. L., McMann, K., Mechkova, V., Paxton, P., … Ziblatt, D. (2021). *V-Dem Codebook v11.1*.

Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Marquardt, K. L., Medzihorsky, J., Pemstein, D., Alizada, N., Gastaldi, L., Hindle, G., Pernes, J., von Römer, J., Tzelgov, E., Wang, Y., & Wilson, S. (2021). *V-Dem Methodology v11.1*.

Coppedge, M., & Reinicke, W. H. (1990). Measuring Polyarchy. *Studies In Comparative International Development*, *25*(1), 51–72. https://doi.org/10.1007/BF02716905

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Cutright, P. (1963). National Political Development: Measurement and Analysis. *American Sociological Review*, *28*(2), 253. https://doi.org/10.2307/2090612

Dahl, R. (1971). *Polyarchy. Participation and Opposition*. Yale University Press.

Dahl, R. (1989). *Democracy and its critics*. Yale University Press.

Deutsche Welle. (2022, April 27). EU triggers rule of law procedure against Hungary. *Deutsche Welle*. https://www.dw.com/en/eu-triggers-rule-of-law-procedure-against-hungary/a-61607618

Doucouliagos, C. (Hristos), & Ulubasoglu, M. A. (2007). Democracy and Economic Growth: A Meta-Analysis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1014333

Elçi, E. (2019). The Rise of Populism in Turkey: A Content Analysis. *Southeast European and Black Sea Studies*, *19*(3), 387–408. https://doi.org/10.1080/14683857.2019.1656875

Elkins, Z. (2000). Gradations of Democracy? Empirical Tests of Alternative Conceptualizations. *American Journal of Political Science*, *44*(2), 293–300.

Encyclopaedia Britannica. (n.d.). *Greek Civil War*. Encyclopaedia Britannica. Retrieved June 10, 2022, from https://www.britannica.com/event/Greek-Civil-War

European Economic Community. (1958, January 1). *Treaty establishing the European Economic Community*.

European Union. (1992, February 7). *Treaty on European Union*.

Freedom House. (2021). *Freedom in the World*.

Freedom House. (2022). *Freedom in the World 2022 Methodology*.

García-Guadilla, M. P., & Mallen, A. (2019). Polarization, Participatory Democracy, and Democratic Erosion in Venezuela's Twenty-First Century Socialism. *The ANNALS of the American Academy of Political and Social Science*, *681*(1), 62–77. https://doi.org/10.1177/0002716218817733

Gasiorowski, M. J. (1996). An Overview of the Political Regime Change Dataset. *Comparative Political Studies*, *29*(4), 469–483. https://doi.org/10.1177/0010414096029004004

Geddes, B., Wright, J., & Frantz, E. (2014). Autocratic Breakdown and Regime Transitions: A New Data Set. *Perspectives on Politics*, *12*(2), 313–331. https://doi.org/10.1017/S1537592714000851

Geissel, B., Kneuer, M., & Lauth, H.-J. (2016). Measuring the quality of democracy: Introduction. *International Political Science Review*, *37*(5), 571–579. https://doi.org/10.1177/0192512116669141

Gregorio, J. de, & Lee, J.-W. (2002). Education and Income Inequality: New Evidence From Cross-Country Data. *Review of Income and Wealth*, *48*(3), 395–416. https://doi.org/10.1111/1475-4991.00060

Grimmer, J. (2015). We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together. *PS: Political Science & Politics*, *48*(1), 80–83. https://doi.org/10.1017/S1049096514001784

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine Learning for Social Science: An Agnostic Approach. *Annual Review of Political Science*, *24*(1), 395–419. https://doi.org/10.1146/annurev-polisci-053119-015921

Gründler, K., & Krieger, T. (2016). Democracy and growth: Evidence from a machine learning indicator. *European Journal of Political Economy*, *45*, 85–107. https://doi.org/10.1016/j.ejpoleco.2016.05.005

Gründler, K., & Krieger, T. (2020). *Should We Care (More) About Data Aggregation? Evidence from the Democracy Indices* (No. 7480; CESifo Working Papers).

Gründler, K., & Krieger, T. (2021). Using Machine Learning for measuring democracy: A practitioners guide and a new updated dataset for 186 countries from 1919 to 2019. *European Journal of Political Economy*, *70*. https://doi.org/10.1016/j.ejpoleco.2021.102047

Gualtieri, J. A. (n.d.). The Support Vector Machine (SVM) Algorithm for Supervised Classification of Hyperspectral Remote Sensing Data. In *Kernel Methods for Remote Sensing Data Analysis* (pp. 49–83). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470748992.ch3

Guenther, N., & Schonlau, M. (2016). Support Vector Machines. *The Stata Journal: Promoting Communications on Statistics and Stata*, *16*(4), 917–937. https://doi.org/10.1177/1536867X1601600407

Gutmann, J., & Voigt, S. (2018). The rule of law: Measurement and deep roots. *European Journal of Political Economy*, *54*, 68–82. https://doi.org/10.1016/j.ejpoleco.2018.04.001

Guttman, L. (1994). *Louis Guttman on theory and methodology: Selected writings*. Dartmouth Publishing Company.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, *46*, 389–422. https://doi.org/10.1023/A:1012487302797

Hadenius, A. (1992). *Democracy and development*. Cambridge University Press.

Hlavac, M. (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables. In *R package version 5.2.3*. https://CRAN.R-project.org/package=stargazer

Horáková, P. (2003, October 13). *Forced displacement of Czech population under Nazis in 1938 and 1943*. Czech Radio. https://english.radio.cz/forced-displacement-czech-population-under-nazis-1938-and-1943-8079114

ISO. (2021). *Country Codes (ISO standard no. 3166-1)*. https://www.iso.org/iso-3166-country-codes.html

Joachims, T. (2002). *Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms*. Kluwer Academic Publishers.

Knutsen, C. H., & Wig, T. (2015). Government Turnover and the Effects of Regime Type. *Comparative Political Studies*, *48*(7), 882–914. https://doi.org/10.1177/0010414014565891

Korstanje, J. (2021, June 22). *Polynomial Interpolation*. Towards Data Science. https://towardsdatascience.com/polynomial-interpolation-3463ea4b63dd

Kovács, K., & Scheppele, K. L. (2018). The fragility of an independent judiciary: Lessons from Hungary and Poland—and the European Union. *Communist and Post-Communist Studies*, *51*(3), 189–200. https://doi.org/10.1016/j.postcomstud.2018.07.005

Kubicek, P. (2003). The European Union & Democratization. In *The European Union & Democratization*. Routledge.

Lin, C.-H. A. (2007). Education Expansion, Educational Inequality, and Income Inequality: Evidence from Taiwan, 1976–2003. *Social Indicators Research*, *80*(3), 601–615. https://doi.org/10.1007/s11205-006-0009-8

Lindberg, S. I., Coppedge, M., Gerring, J., & Teorell, J. (2014). V-Dem: A New Way to Measure Democracy. *Journal of Democracy*, *25*(3), 159–169. https://doi.org/10.1353/jod.2014.0040

Lipset, S. M. (1959). Some Social Requisites of Democracy: Economic Development and Political Legitimacy. *American Political Science Review*, *53*(1), 69–105. https://doi.org/10.2307/1951731

Lipset, S. M. (1993). Reflections on Capitalism, Socialism & Democracy. *Journal of Democracy*, *4*(2), 43–55. https://doi.org/10.1353/jod.1993.0019

Makszimov, V. (2022, May 24). No progress in democracy and EU values in Hungary, Commission says. *EURACTIV*. https://www.euractiv.com/section/politics/short_news/no-progress-in-democracy-and-eu-values-in-hungary-commission-says/

Mangu, A. M. (2014). The African Union and the promotion of democracy and good political governance under the African Peer-Review Mechanism: 10 years on. *Africa Review*, *6*(1), 59–72. https://doi.org/10.1080/09744053.2014.883757

Marquez, X. (2022, May 26). *Replicating and Extending the UD scores of Pemstein, Meserve, and Melton*.

Marshall, M. G., & Gurr, T. R. (2020). Polity V: Political Regime Characteristics and Transitions, 1800-2018. In *Center for Systemic Peace*.

Marshall, M. G., Gurr, T. R., & Jaggers, K. (2019). Polity IV project: Political regime characteristics and transitions, 1800–2015. *Center for Systemic Peace*.

Mattera, D., & Haykin, S. (1999). *Support vector machines for dynamic reconstruction of a chaotic system*.

MERCOSUL. (1991, March 26). *Tratado para a Constituição de um Mercado Comum*.

MERCOSUL. (1998, July 24). *Protocolo de Ushuaia sobre Compromisso Democrático*.

Merkel, W. (2004). Embedded and defective democracies. *Democratization*, *11*(5), 33–58. https://doi.org/10.1080/13510340412331304598

Molina, M., & Garip, F. (2019). Machine Learning for Sociology. *Annual Review of Sociology*, *45*(1), 27–45. https://doi.org/10.1146/annurev-soc-073117-041106

Munck, G., & Verkuilen, J. (2002). Conceptualizing and Measuring Democracy: Evaluating Alternative Indices. *Comparative Political Studies*, *35*(1), 5–34. https://doi.org/https://doi.org/10.1177/001041400203500101

North Atlantic Treaty Organization. (1999). *Membership Action Plan (MAP)*. https://www.nato.int/cps/en/natohq/official_texts_27444.htm

North Atlantic Treaty Organization. (2020, September 24). *Member countries*. https://www.nato.int/cps/en/natohq/topics_52044.htm?

North Atlantic Treaty Organization. (2022, May 3). *Founding treaty*. https://www.nato.int/cps/en/natohq/topics_67656.htm?

O'Donnell, G. A. (1998). Horizontal Accountability in New Democracies. *Journal of Democracy*, *9*(3), 112–126. https://doi.org/10.1353/jod.1998.0051

O'Donnell, G. A. (2001). Democracy, Law, and Comparative Politics. *Studies in Comparative International Development*, *36*(1), 7–36. https://doi.org/10.1007/BF02687583

O'Donnell, G. A. (2011). Democracy, Law, and Comparative Politics. *Studies in Comparative International Development*, *36*(1), 7–36.

Orrù, G., Pettersson-Yeo, W., Marquand, A. F., Sartori, G., & Mechelli, A. (2012). Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. *Neuroscience & Biobehavioral Reviews*, *36*(4), 1140–1152. https://doi.org/10.1016/j.neubiorev.2012.01.004

Pemstein, D., Meserve, S. A., & Melton, J. (2010). Democratic Compromise: A Latent Variable Analysis of Ten Measures of Regime Type. *Political Analysis*, *18*(4), 426–449. https://doi.org/10.1093/pan/mpq020

Pérez-Liñán, A. (2017). Democracies. In D. Caramani (Ed.), *Comparative Politics* (4th ed., pp. 83–98). Oxford University Press.

Pérez-Liñán, A. (2020). Democracies. In D. Caramani (Ed.), *Comparative Politics* (5th ed., pp. 86–102). Oxford University Press.

Poole, A. (2019). Conceptions of Democracy and Human Rights in Southeast Asia. In *Democracy, Rights and Rhetoric in Southeast Asia* (pp. 1–18). Springer International Publishing. https://doi.org/10.1007/978-3-030-15522-3_1

Riasanovsky, N. v. (2022, June 25). *Russia*. Riasanovsky. https://www.britannica.com/place/Russia

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, *3*(3), 210–229. https://doi.org/10.1147/rd.33.0210

Sarkin, J. (2018). The Need to Reform the Political Role of the African Union in Promoting Democracy and Human Rights in Domestic States: Making States More Accountable and Less Able to Avoid Scrutiny at the United Nations and at the African Union, Using Swaziland to Spotlight the Issues. *African Journal of International and Comparative Law*, *26*(1), 84–107. https://doi.org/10.3366/ajicl.2018.0221

Schumann, D. (2001). *Politische Gewalt in der Weimarer Republik 1918-1933: Kampf um die Straße und Furcht vor dem Bürgerkrieg*. Klartext.

Sheehan, J. J. (2022, June 22). *Germany*. Encyclopaedia Britannica. https://www.britannica.com/place/Germany

Skaaning, S.-E., Gerring, J., & Bartusevičius, H. (2015). A Lexical Index of Electoral Democracy. *Comparative Political Studies*, *48*(12), 1491–1525. https://doi.org/10.1177/0010414015581050

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Staudenmaier, R. (2018, September 12). EU Parliament votes to trigger Article 7 sanctions procedure against Hungary. *Deutsche Welle*. https://www.dw.com/en/eu-parliament-votes-to-trigger-article-7-sanctions-procedure-against-hungary/a-45459720

Steinwart, I., & Christmann, A. (2008). *Support Vector Machines*. Springer New York. https://doi.org/10.1007/978-0-387-77242-4

Teorell, J., Coppedge, M., Lindberg, S., & Skaaning, S.-E. (2019). Measuring Polyarchy Across the Globe, 1900–2017. *Studies in Comparative International Development*, *54*(1), 71–95. https://doi.org/10.1007/s12116-018-9268-z

Treier, S., & Jackman, S. (2008). Democracy as a Latent Variable. *American Journal of Political Science*, *52*(1), 201–217. https://doi.org/10.1111/j.1540-5907.2007.00308.x

van der Vleuten, A., & Hoffmann, A. R. (2010). Explaining the Enforcement of Democracy by Regional Organizations: Comparing EU, Mercosur and SADC. *Journal of Common Market Studies*, *48*(3). https://doi.org/10.1111/j.1468-5965.2010.02071.x

Vanhanen, T. (1984). *The Emergence of Democracy. A Comparative Study of 119 States, 1850–1979*. Finnish Society of Sciences and Letters.

Vanhanen, T. (1989). The Level of Democratization Related to Socioeconomic Variables in 147 States in 1980-85. *Scandinavian Political Studies*, *12*(2), 95–127. https://doi.org/10.1111/j.1467-9477.1989.tb00084.x

Vanhanen, T. (2000). A New Dataset for Measuring Democracy, 1810-1998. *Journal of Peace Research*, *37*(2), 251–265. https://doi.org/10.1177/0022343300037002008

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*. Springer New York. https://doi.org/10.1007/978-1-4757-2440-0

Wagner, T. (2010, November 19). *Portugal und die NATO*. Deutschlandfunk. https://www.deutschlandfunk.de/portugal-und-die-nato-100.html

Welzel, C. (2000). Humanentwicklung und Demokratie: Welcher Index erfasst die „humane“ Dimension der Demokratie am besten? In *Demokratiemessung* (pp. 132–162). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-322-89590-5_7

9. Appendix

9.1. R Functions

```
# This function was created to more easily download different indeces
downloadDem <- function(data_source, fh_data = "FIH_hist", dd_data =
"col_sheet",
                        remove_file = TRUE) {

  # Freedom House
    # There are different FH datasets and the fh_data parameter specifies
which one to download
  if(tolower(data_source) == "fh") {
    if(fh_data == "FIH_hist") {
      download.file(url =
"https://freedomhouse.org/sites/default/files/2021-02/
Country_and_Territory_Ratings_and_Statuses_FIW1973-2021.xlsx",
                    destfile = "fh.xlsx", method = "curl")

      temp <- read.xlsx(file = "fh.xlsx", sheetIndex = 2, startRow = 3)

      colnames(temp) <- c("country", apply(expand.grid(c("PR","CL",
"Status"),
                                                       c(1972:1981,1983:2020)),
                                           1, paste, collapse="."))
    }

    if(fh_data == "FIH_recent") {
      download.file(url =
"https://freedomhouse.org/sites/default/files/2021-02/
All_data_FIW_2013-2021.xlsx",
destfile = "fh.xlsx", method = "curl")

      temp <- read.xlsx(file = "fh.xlsx", sheetIndex = 2, startRow = 2)
    }

    if(fh_data == "FIH_aggregate") {
      download.file(url =
"https://freedomhouse.org/sites/default/files/2021-02/
Aggregate_Category_and_Subcategory_Scores_FIW_2003-2021.xlsx",
destfile = "fh.xlsx", method = "curl")

      temp <- read.xlsx(file = "fh.xlsx", sheetIndex = 2)
    }

    if(remove_file == TRUE) {
      file.remove("fh.xlsx")
    }
  }

  # V-Dem
  if(tolower(data_source) == "vdem") {
    download.file(url = "https://v-dem.net/media/datasets/Country_Year_V-
Dem_Full_others_R_v11.1.zip",
                  destfile = "vdem.zip", method = "curl")
    unzip("vdem.zip")

    temp <- readRDS("Country_Year_V-Dem_Full+others_R_v11.1/V-Dem-CY-
Full+Others-v11.1.rds")

    if(remove_file == TRUE) {
      unlink("Country_Year_V-Dem_Full+others_R_v11.1/", recursive = TRUE)
```

```
      file.remove("vdem.zip")
    }
  }

  # PolityV
  if(tolower(data_source) == "polityv") {
    download.file(url = "http://www.systemicpeace.org/inscr/p5v2018.sav",
                  destfile = "polityV.sav", method = "curl")

    temp <- read.spss("polityV.sav", to.data.frame = TRUE)

    if(remove_file == TRUE) {
      file.remove("polityV.sav")
    }
  }

  # UDS
  if(tolower(data_source) == "uds") {
    download.file(url = "http://www.unified-democracy-scores.net/files/
20140312/z/uds_summary.csv.gz",
                  destfile = "uds_summary.csv.gz")

    temp <- read.table("uds_summary.csv.gz", header = T, sep = ",")

    if(remove_file == TRUE) {
      file.remove("uds_summary.csv.gz")
    }
  }

  # DD
    # There are different DD datasets and the dd_data parameter specifies
which one to download
  if(tolower(data_source) == "dd") {

    download.file(url =
"http://www.christianbjoernskov.com/wp-content/uploads/2020/09/
Bj%C3%B8rnskov-Rode-integrated-dataset-v3.2.xlsx",
destfile = "dd.xlsx", method = "curl")

    if(dd_data == "reg_char") {
      temp <- read_excel(path = "dd.xlsx", sheet = 1)
    } else {
      temp <- read_excel(path = "dd.xlsx", sheet = 3)
    }

    if(remove_file == TRUE) {
      file.remove("dd.xlsx")
    }
  }

  return(temp)

}

# This function just facilitates the prinint of how far along a loop is
print_progress <- function(c_state, max) {
  print(paste(round((c_state/max)*100, digits = 2), "%"))
}
```

```
# This function runs the SVM regression
  # It needs identifying and support variables, as well as varaibles for
the SVM
  # regression, a dataset with the input data and a dataset with the
priming
  # observations
SVM_index <- function(svm_dataframe, id_vars, support_vars, svm_data_vars,
                      priming_selec, save_path = "") {
  # Combining all variables
  var_SVM <- c(id_vars,support_vars, svm_data_vars)

  # Extranting the priming observation from the dataset with input
variables
  priming_df <- svm_dataframe %>%
    filter(iso_year %in% priming_selec$iso_year) %>%
    select(-all_of(c(id_vars,support_vars)))

  # Creating dataset to be filled with the data from the 2000 iterative
runs of
  # the SVM regression
  result_full_df <- svm_dataframe %>% select(all_of(id_vars))

  # Loop that takes a sample of the priming data, trains the dataset,
applies it
  # to the overall input data, iteratively 2000 times
  for(i in 1:2000) {
    sample <- sample.split(1:nrow(priming_df), SplitRatio = runif(1,
min=0.2, max=0.5))

    train <- subset(priming_df, sample==TRUE)

    train_x <- train %>%
      select(-v2x_polyarchy)

    train_y <- train[,"v2x_polyarchy"]

    modell <- svm(x = train_x, y = train_y, scale = TRUE, kernel =
"radial",
                  cost = 1, na.action = na.omit, epsilon = 0.025)

    resl <- predict(modell, svm_df %>% select(-
all_of(c(id_vars,support_vars,
                                                          "v2x_polyarchy"))))

    result_full_df <- cbind(result_full_df, resl)

    colnames(result_full_df)[ncol(result_full_df)] <- paste("iter",i,sep =
"_")

    print_progress(i, 2000)
  }

  # Creating a reult dataset that will include the output data
  result_output_df <- svm_dataframe %>% select(all_of(c(id_vars,
"v2x_polyarchy")))

  temp_output_df <- data.frame()
```

```
  # Loop that takes the 2000 prediction for each observation, takes the
median
  # as value of the SVM index, calcualtes the standard deviation and all
percentiles
  for(j in 1:nrow(result_full_df)) {
    res_vec <- as.numeric(result_full_df[j,7:ncol(result_full_df)])

    temp_obs_df <- data.frame(med = median(res_vec), sd = sd(res_vec),
                              t(quantile(res_vec, probs = seq(from = 0, to
= 1, by = 0.01)
                                                 )
                                )
                              )

    colnames(temp_obs_df)[3:103] <- paste(seq(from = 0, to = 100, by = 1),
"p",
                                          sep = "")

    temp_output_df <- rbind(temp_output_df, temp_obs_df)

    print_progress(j, nrow(result_full_df))
  }

  # Binding the columns of the output with its identifying variables
  result_output_df <- cbind(result_output_df, temp_output_df)

  # Creating a list that includes the predicted values for all 2000
iterations
  # for each observation in the Full_Result and the output data in the
Output
  result_SVM <- list(
    Full_Result = result_full_df,
    Output = result_output_df
  )

  # If statement that saves the result to an RDS file, should a path or
filename
  # be specified
  if (save_path != "") {
    saveRDS(result_SVM, save_path)
  }

  # Returns the result
  return(result_SVM)
}
```

9.2. R Script for priming data, SVM regression, and analysis

```
#### Libraries ####
library(dplyr)
library(tidyr)
library(readxl)
library(e1071)
library(caTools)
library(countrycode)
library(haven)
library(descr)
library(ggplot2)
library(zoo)
library(ggplot2)
library(ggpubr)

#### Standard variables/datasets ####

vDem <- downloadDem(data_source = "vdem") # Custom function created to
easily download datasets

start_year <- 1919

set.seed(44)

#### Section 4: Priming data ####

# This variable contains the variable names for all of the indicators used
to
# create the priming data with the extremes of the UDS and the EDI as well
as
# the variables used to check for consensus
var_vDem <-
c("country_name","country_text_id","country_id","year","COWcode",
              "v2x_polyarchy","e_uds_mean","e_boix_regime",
"e_democracy_omitteddata",
              "e_lexical_index","e_p_polity","e_fh_status", "e_fh_cl",
"e_fh_pr")


priming_selection <- vDem %>%
  select(all_of(var_vDem)) %>%
  filter(year >= 1919)

# Finding the decile of the UDS scale
ten_p_uds <- -(min(priming_selection$e_uds_mean, na.rm = TRUE) -
                 max(priming_selection$e_uds_mean, na.rm = TRUE)) * 0.1

decile_UDS <- c(
  min(priming_selection$e_uds_mean, na.rm = TRUE)+ten_p_uds,
  max(priming_selection$e_uds_mean, na.rm = TRUE)-ten_p_uds
)


priming_selection <- priming_selection %>%
  # Filtering by the deciles of both scales
  filter(
    (e_uds_mean < decile_UDS[1] | e_uds_mean > decile_UDS[2]) |
      (v2x_polyarchy < 0.1 | v2x_polyarchy > 0.9)
  ) %>%
```

```
  # Correcting missing histoprical iso3c codes
  mutate(iso3c = countryname(country_name, destination = "iso3c")) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Zanzibar", "EAZ", iso3c)) %>%
  # Creating a dichotomous variable to idedntify democracies and
autocracies
  # based on their EDI and UDS scores
  mutate(prim_dem = 0) %>%
  mutate(prim_dem = ifelse(v2x_polyarchy > 0.9 &
                             !is.na(v2x_polyarchy), 1, prim_dem)) %>%
  mutate(prim_dem = ifelse(e_uds_mean > decile_UDS[2] &
                             !is.na(e_uds_mean), 1, prim_dem)) %>%
  # Creating identifying variable and moving certain variables closer to
the
  # first columns
  mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
  relocate(c("iso3c","iso_year","prim_dem"), .after = COWcode) %>%
  # Creating the FH score by calculating it using the CL and PR variables
  mutate(fh_avg = (e_fh_cl + e_fh_pr)/2)

# Downloading the DD dataset
dd_index <- downloadDem("dd", dd_data = "reg_char") %>%
  # Correcting disparities in nomenclature
  mutate(country = ifelse(country == "Korea, People's Republic",
                          "North Korea", country)) %>%
  # Creating the iso3c codes and using them to create unique observation
identifiers
  mutate(iso3c = countryname(country, destination = "iso3c")) %>%
  mutate(iso_year = paste(iso3c, year, sep = "_"))
# Removing columns with NA (due to the format of the file some columns are
just blank)
dd_index <- dd_index[colSums(!is.na(dd_index)) > 0]

# Creating the final priming_selection dataset by combining the
priming_selection
# with the DD dataset
priming_selection <- priming_selection %>%
  left_join(dd_index %>% select(iso_year, Democracy), by = "iso_year") %>%
  rename(dd_index = Democracy)

  # Section 4.1: Consensus #+
  # The following lines create tables used to check for consensus between
  # indicators

    ## Boix ##
crosstab(priming_selection$prim_dem,
priming_selection$e_democracy_omitteddata)

    ## DD ##
crosstab(priming_selection$prim_dem, priming_selection$dd_index)

    ## Lexical ##
crosstab(priming_selection$prim_dem, priming_selection$e_lexical_index)

    ## FH ##
crosstab(priming_selection$prim_dem, priming_selection$e_fh_status)
```

```
    ## Polity ##
crosstab(priming_selection$prim_dem, priming_selection$e_p_polity)


  # Section 4.2: Representativeness #

# Downloading the DD dataset that contains regime information
dd_reg_char <- downloadDem("dd", dd_data = "reg_char")
# Removing empty columns (again due to the formatting of the original file)
dd_reg_char <- dd_reg_char[colSums(!is.na(dd_reg_char)) > 0] %>%
  # Correcting disparities in nomenclature
  mutate(country = ifelse(country == "Korea, People's Republic",
                          "North Korea", country)) %>%
  mutate(iso3c = countryname(country, destination = "iso3c")) %>%
  mutate(iso_year = paste(iso3c, year, sep = "_"))

# Creating a new file based on the priming_selection that just contains
identifying
# variables and the dichotomous variable created in the previous section
priming_repr <- priming_selection %>%
  select(country_name, iso3c, year, iso_year, prim_dem) %>%
  # Joining the regime info dataset with the priming_selection
  left_join(dd_reg_char, by = "iso_year") %>%
  rename(dd_index = Democracy) %>%
  # Removing observations that are not in the DD dataset
  filter(!is.na(`DD category`)) %>%
  # Creating a character variable for regime type from the original numeric
  mutate(regime_type = ifelse(
    `DD regime` == 0, "Parliamentary democracy",
    ifelse(`DD regime` == 1, "Mixed democracy",
           ifelse(`DD regime` == 2, "Presidential democracy",
                  ifelse(`DD regime` == 3, "Civilian autocracy",
                         ifelse(`DD regime` == 4, "Military dictatorship",
                                ifelse(`DD regime` == 5, "Royal
dictatorship", NA
                                )
                         )
                  )
           )
    )
  )
  ) %>%
  mutate(regime_type = ifelse(`DD regime` <= 3 & Communist == 1,
                              "Communist dictatorship", regime_type)) %>%
  # Creating a character variable for election type from the original
numeric
  mutate(
    election_type = ifelse(electoral == 0, "No elections",
                           ifelse(electoral == 1, "Single-party elections",
                                  ifelse(electoral == 2,
                                         "Non-democratic multi-party
elections",
                                         ifelse(electoral == 3,
                                                "Democratic elections", NA)
                                  )
                           )
    )
```

```
  ) %>%
  # Creating a character variable for whether there are elections from the
original numeric
  mutate(election_type2 = ifelse(election_type == "No elections",
                                 "No elections", "Elections")) %>%
  # Creating a character variable for voting system from the original
numeric
  mutate(voting_system = ifelse(`proportional voting` == 1,
                                "Proportional", "Majoritarian")
         )

# Creating dataset to check for regime length
reg_length <- read_excel(
  "data/Autocracy Dataset/GWF Autocratic Regimes 1.2/GWF All Political
Regimes Case List.xls",
  sheet = 1)
# Removing uneeded columns
reg_length <- reg_length[,1:5] %>%
  # Correcting disparities in nomenclature
  mutate(Country = ifelse(Country == "Lietchtenstein", "Liechtenstein",
Country)) %>%
  mutate(iso3c = countryname(Country, destination = "iso3c", warn = F)) %>%
  mutate(iso3c = ifelse(iso3c == "Yugoslavia", "SRB", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Czechoslovakia", "CZE", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "German Democratic Republic", "DDR",
iso3c)) %>%
  mutate(iso3c = ifelse(Country == "Cen African Rep", "CAF", iso3c)) %>%
  # Removing empty rows
  filter(!is.na(Country)) %>%
  # Establishing 2014 as ending year
  mutate(`end yr` = ifelse(is.na(`end yr`), 2014, `end yr`)) %>%
  # Calculating total length
  mutate(tot_length = `end yr` - `beg yr`)

# Creating the final regime length dataset by running a loop that creates a
# country and year observation for each regime
reg_length_final <- data.frame()
for(i in 1:nrow(reg_length)) {
  year_seq <- seq(from=reg_length$`beg yr`[i], to=reg_length$`end yr`[i])

  temp_df <- data.frame(
    Country = reg_length$Country[i],
    iso3c = reg_length$iso3c[i],
    year = year_seq,
    tot_length = reg_length$tot_length[i]
  ) %>%
    # creating regime length up to the date of observation
    mutate(cur_length = year - reg_length$`beg yr`[i])

  reg_length_final <- rbind(reg_length_final, temp_df)
}

# Creating identifying variable for the regime length dataset
reg_length_final <- reg_length_final %>%
  mutate(iso_year = paste(iso3c, year, sep = "_"))
```

```
# Joining the regime length dataset with the priming_selection
priming_selection <- priming_selection %>%
  left_join(
    reg_length_final %>%
      mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
      select(iso_year, tot_length, cur_length),
    by = "iso_year"
  )

# In the following section the countries are checked for their regime
characteristics
# Two tables are calculated: one with absolute numbers and one with
percentages
# for both democracies and autocracies

    ## Autocracies ##
      ### Dictatorship type: ###
table(priming_repr$regime_type[priming_repr$prim_dem==0])
round(prop.table(
  table(priming_repr$regime_type[priming_repr$prim_dem==0])
)*100, digits = 2)

      ### Electoral Systems: ###
# Whether there are election or not
table(priming_repr$election_type2[priming_repr$prim_dem==0])
round(prop.table(
  table(priming_repr$election_type2[priming_repr$prim_dem==0])
)*100, digits = 2)

# Type of election system
table(priming_repr$election_type[priming_repr$prim_dem==0])
round(prop.table(
  table(priming_repr$election_type[priming_repr$prim_dem==0])
)*100, digits = 2)

    ## Democracies:
      ###Democracy type:
table(priming_repr$regime_type[priming_repr$prim_dem==1])
round(prop.table(
  table(priming_repr$regime_type[priming_repr$prim_dem==1])
)*100, digits = 2)

      ###Amount of chambers:
table(priming_repr$`No. of chambers in
parliament`[priming_repr$prim_dem==1])
round(prop.table(
  table(priming_repr$`No. of chambers in
parliament`[priming_repr$prim_dem==1])
)*100, digits = 2)

      ###Voting System:
table(priming_repr$voting_system[priming_repr$prim_dem==1])
round(prop.table(
  table(priming_repr$voting_system[priming_repr$prim_dem==1])
)*100, digits = 2)

    ##Length for Autocracies and Democracies:
```

```
# The length is checked by running a summary function for democracies and
autocracies
# separately
      ###Total length:
tapply(priming_selection$tot_length,priming_selection$prim_dem, summary)

      ###Current length:
tapply(priming_selection$cur_length,priming_selection$prim_dem, summary)

####  Section 5.1: Aggregation comparison with EDI ####

# List of mid-level indices and their constituting low-level indices
polyarchy_list <- list(
  v2x_frassoc_thick = c("v2psparban_ord", "v2psbars_ord", "v2psoppaut_ord",
                        "v2elmulpar_ord", "v2cseeorgs","v2csreprss"),

  v2xel_frefair = c("v2elembaut", "v2elembcap", "v2elrgstry_ord",
"v2elvotbuy_ord",
                    "v2elirreg_ord", "v2elintim_ord", "v2elpeace_ord",
                    "v2elfrfair_ord"),

  v2x_freexp_altinf = c("v2mecenefm", "v2meharjrn", "v2mebias",
"v2meslfcen",
                        "v2mecrit", "v2merange", "v2cldiscm", "v2cldiscw",
                        "v2clacfree")
)

# Identifying variables
id_vars <- c("country_name","country_text_id","country_id","iso3c","year",
             "iso_year")

# Variables used for the SVM regression
svm_data_vars <- c("v2x_polyarchy",

                    "v2psparban_ord", "v2psbars_ord", "v2psoppaut_ord",
                   "v2elmulpar_ord", "v2cseeorgs","v2csreprss",

                    "v2x_suffr",

                    "v2elembaut", "v2elembcap", "v2elrgstry_ord",
"v2elvotbuy_ord",
                   "v2elirreg_ord", "v2elintim_ord", "v2elpeace_ord",
                   "v2elfrfair_ord",

                    "v2x_elecoff",

                    "v2mecenefm", "v2meharjrn", "v2mebias", "v2meslfcen",
                   "v2mecrit", "v2merange", "v2cldiscm", "v2cldiscw",
                   "v2clacfree"
)

# Support variables needed for adaptations to V-Dem's coding rules
support_vars <- c("v2x_elecreg")

# Complete variable list
var_SVM <- c(id_vars,support_vars, svm_data_vars)
```

```
# Creating the dataset to be used in the SVM regression from the original
V-Dem
# dataset
svm_df <- vDem %>%
  # Creating iso3c identifying variable
  mutate(iso3c = countryname(country_name, destination = "iso3c")) %>%
  # Correcting disparities in nomenclature
  mutate(iso3c = ifelse(iso3c == "Yugoslavia", "SRB", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Czechoslovakia", "CZE", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "German Democratic Republic", "DDR",
iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Zanzibar", "EAZ", iso3c)) %>%
  # Filtering for the start year and removing observations without an EDI
score
  filter(year >= start_year, !is.na(v2x_polyarchy)) %>%
  # Creating identifying variable for each observation
  mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
  # Selecting the needed variables
  select(all_of(var_SVM)) %>%
  # Relocating certain variables for better visualization
  relocate(c("iso3c", "iso_year"), .after = country_id) %>%
  # Grouping by country name so that operations aren't conducted cross
country
  group_by(country_name) %>%
  # Filling missing electoral variables as mentioned by V-Dem's coding
rules
  mutate(v2elmulpar_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elmulpar_ord, na.rm = F), 0))
%>%
  mutate(v2elrgstry_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elrgstry_ord, na.rm = F), 0))
%>%
  mutate(v2elvotbuy_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elvotbuy_ord, na.rm = F), 0))
%>%
  mutate(v2elirreg_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elirreg_ord, na.rm = F), 0)) %>%
  mutate(v2elintim_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elintim_ord, na.rm = F), 0)) %>%
  mutate(v2elpeace_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elpeace_ord, na.rm = F), 0)) %>%
  mutate(v2elfrfair_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elfrfair_ord, na.rm = F), 0))
%>%
  # Filling up electoral variables that did not have an election in the
first year
  # of the dataset
  fill(c("v2elmulpar_ord","v2elrgstry_ord", "v2elvotbuy_ord",
"v2elirreg_ord",
         "v2elintim_ord", "v2elpeace_ord", "v2elfrfair_ord"), .direction =
"up") %>%
  # Filling missing party variables as mentioned by V-Dem's coding rules
  mutate(v2psoppaut_ord = ifelse(v2psbars_ord <= 1 & is.na(v2psoppaut_ord),
0,
                                 v2psoppaut_ord)) %>%
```

```
  mutate(v2elembcap = ifelse(v2x_elecreg == 0 & is.na(v2elembcap), 0,
                             v2elembcap)) %>%
  mutate(v2elembaut = ifelse(v2x_elecreg == 0 & is.na(v2elembaut), 0,
                             v2elembaut)) %>%
  # Interpolating missing data
  mutate(v2mebias = na.spline(v2mebias)) %>%
  mutate(v2merange = na.spline(v2merange)) %>%
  mutate(v2meharjrn = na.spline(v2meharjrn)) %>%
  mutate(v2mecenefm = na.spline(v2mecenefm)) %>%
  mutate(v2elembcap = na.spline(v2elembcap)) %>%
  mutate(v2mecrit = na.spline(v2mecrit)) %>%
  mutate(v2x_suffr = na.spline(v2x_suffr)) %>%
  # Ungrouping before SVM regression
  ungroup() %>%
  # Removing still missing data
  drop_na()

# Running SVM function created for the purpose of this contribution
SVM_index(svm_dataframe = svm_df, id_vars = id_vars, support_vars =
support_vars,
          svm_data_vars = svm_data_vars, priming_selec = priming_selection,
          save_path = "result/result_SVM_polyarchy3.RDS")

# Reading saved result
result_EDI <- readRDS("result/result_SVM_polyarchy3.RDS")$Output

# Creating a dataset with the results of the SVM regression, its delta from
# the EDI, and the input variables
graph_EDI <- result_EDI[,1:8] %>%
  mutate(diff = med - v2x_polyarchy) %>%
  left_join(svm_df)

# Scaling the data for graphical representation later on
graph_EDI_scaled <- graph_EDI
graph_EDI_scaled[,11:35] <- scale(graph_EDI[,11:35], center = F, scale = T)

# Keeping variables that will not be averaged into mid-level indices
graph_EDI_mid_level <- graph_EDI_scaled[,c("country_name","iso3c","year",
                                           "iso_year","v2x_polyarchy","med",
                                           "diff","v2x_suffr","v2x_elecoff")]
# This loop creates mid-level indices out of the scaled low-level indices
# (only for graphical representation)
for(i in 1:length(polyarchy_list)){
  temp <- data.frame(rowMeans(graph_EDI_scaled[,polyarchy_list[[i]]]))
  colnames(temp) <- names(polyarchy_list)[i]

  graph_EDI_mid_level <- cbind(graph_EDI_mid_level,temp)
}

  # Figures #

    ## Figure 1 ##
agg_plot_1 <- ggplot(graph_EDI, aes(x = med, y = v2x_polyarchy, color =
diff)) +
  geom_point(alpha = 0.3) +
  geom_smooth(method = "lm", color = "green") +
  scale_color_gradient(low="blue", high="red") +
  labs(y = "EDI", x = "SVM Index") +
```

```
  guides(color = guide_legend(title = "Delta"))

agg_plot_1

# Saving plot as a png
png("img/figure1.png", res = 120, width = 700, height = 700)
agg_plot_1
dev.off()

    ## Figure 2 ##
agg_plot_2 <- graph_EDI %>%
  filter(iso3c == "DEU") %>%
  select(iso3c, year, v2x_polyarchy, med) %>%
  rename(EDI = v2x_polyarchy, `SVM Ind` = med) %>%
  pivot_longer(cols = c("EDI","SVM Ind")) %>%
  as.data.frame(.) %>%
  ggplot(aes(x = year,y = value, color = name)) +
  #ggplot(aes(x = year,y = value, linetype = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index")) +
  scale_color_manual(values=c("#990000", "#000066"))

agg_plot_2

png("img/figure2.png", res = 120, width = 700, height = 700)
agg_plot_2
dev.off()

    ## Figure 3 ##
agg_plot_3 <- graph_EDI %>%
  filter(iso3c == "RUS") %>%
  select(iso3c, year, v2x_polyarchy, med) %>%
  rename(EDI = v2x_polyarchy, `SVM Ind` = med) %>%
  pivot_longer(cols = c("EDI","SVM Ind")) %>%
  as.data.frame(.) %>%
  ggplot(aes(x = year,y = value, color = name)) +
  #ggplot(aes(x = year,y = value, linetype = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index")) +
  scale_color_manual(values=c("#990000", "#000066"))

agg_plot_3

png("img/figure3.png", res = 120, width = 700, height = 700)
agg_plot_3
dev.off()

  ## Figure 4 ##

agg_plot_4 <- graph_EDI_mid_level %>%
  filter(iso3c == "RUS") %>%
  select(c("year","v2x_suffr","v2x_elecoff","v2x_frassoc_thick","v2xel_frefair",
           "v2x_freexp_altinf")) %>%
  pivot_longer(cols = c("v2x_suffr","v2x_elecoff","v2x_frassoc_thick",
                        "v2xel_frefair","v2x_freexp_altinf")) %>%
  as.data.frame(.) %>%
```

```
  ggplot(aes(x = year,y = value, color = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Value", x = "Year") +
  guides(color = guide_legend(title = "Variable"))

agg_plot_4

png("img/figure4.png", res = 120, width = 700, height = 700)
agg_plot_4
dev.off()

    ## Figure 5 ##
agg_plot_5 <- graph_EDI %>%
  filter(iso3c == "COL") %>%
  select(iso3c, year, v2x_polyarchy, med) %>%
  rename(EDI = v2x_polyarchy, `SVM Ind` = med) %>%
  pivot_longer(cols = c("EDI","SVM Ind")) %>%
  as.data.frame(.) %>%
  ggplot(aes(x = year,y = value, color = name)) +
  #ggplot(aes(x = year,y = value, linetype = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index")) +
  scale_color_manual(values=c("#990000", "#000066"))

agg_plot_5

png("img/figure5.png", res = 120, width = 700, height = 700)
agg_plot_5
dev.off()

    ## Figure 6 ##

agg_plot_6 <- graph_EDI_mid_level %>%
  filter(iso3c == "COL") %>%
  select(c("year","v2x_suffr","v2x_elecoff","v2x_frassoc_thick","v2xel_frefair",
           "v2x_freexp_altinf")) %>%
  pivot_longer(cols =
c("v2x_suffr","v2x_elecoff","v2x_frassoc_thick","v2xel_frefair",
                          "v2x_freexp_altinf")) %>%
  as.data.frame(.) %>%
  ggplot(aes(x = year,y = value, color = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Value", x = "Year") +
  guides(color = guide_legend(title = "Variable"))

agg_plot_6

png("img/figure6.png", res = 120, width = 700, height = 700)
agg_plot_6
dev.off()

#### Section 5.2 SVM Indices ####

# Loading the variable list as they relate to Embedded Democracy
svm_data_table <- read_excel("data/SVM_variables.xlsx")

  # Internally embedded #
```

```
# Variables used for the SVM regression
svm_data_vars <- c(svm_data_table$Code[svm_data_table$Embeddedness ==
"Internal"],
                   "v2x_polyarchy")

# Identifying variables
id_vars <- c("country_name","country_text_id","country_id","iso3c","year",
             "iso_year")

# Variables used for the SVM regression
support_vars <- c("v2x_elecreg", "v2exhoshog", "v2lgbicam")

# Complete variable list
var_SVM <- c(id_vars,support_vars, svm_data_vars)

# Creating the dataset to be used in the SVM regression from the original
V-Dem
# dataset
svm_df <- vDem %>%
  # Creating iso3c identifying variable
  mutate(iso3c = countryname(country_name, destination = "iso3c")) %>%
  # Correcting disparities in nomenclature
  mutate(iso3c = ifelse(iso3c == "Yugoslavia", "SRB", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Czechoslovakia", "CZE", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "German Democratic Republic", "DDR",
iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Zanzibar", "EAZ", iso3c)) %>%
  # Removing empty variables related to the DDR before its formal founding
in 1949
  filter(!(iso3c == "DDR") | year > 1948) %>%
  # Filtering for the start year
  filter(year >= start_year) %>%
  # Creating identifying variable for each observation
  mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
  # Selecting the needed variables
  select(all_of(var_SVM)) %>%
  # Relocating certain variables for better visualization
  relocate(c("iso3c", "iso_year"), .after = country_id) %>%
  # Grouping by country name so that operations aren't conducted cross
country
  group_by(country_name) %>%
  # Filling missing electoral variables as mentioned by V-Dem's coding
rules
  mutate(v2elmulpar_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elmulpar_ord, na.rm = F), 0))
%>%
  mutate(v2elrgstry_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elrgstry_ord, na.rm = F), 0))
%>%
  mutate(v2elvotbuy_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elvotbuy_ord, na.rm = F), 0))
%>%
  mutate(v2elirreg_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elirreg_ord, na.rm = F), 0)) %>%
```

```
  mutate(v2elintim_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elintim_ord, na.rm = F), 0)) %>%
  mutate(v2elpeace_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elpeace_ord, na.rm = F), 0)) %>%
  mutate(v2elfrfair_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elfrfair_ord, na.rm = F), 0))
%>%
  # If HoS and HoG are the same, both have the same value
  mutate(v2exctlhg_0 = ifelse(v2exhoshog == 1, v2exctlhs_0, v2exctlhg_0))
%>%
  # Filling up electoral variables that did not have an election in the
first year
  # of the dataset
  fill(c("v2elmulpar_ord","v2elrgstry_ord", "v2elvotbuy_ord",
"v2elirreg_ord",
         "v2elintim_ord", "v2elpeace_ord", "v2elfrfair_ord"), .direction =
"up") %>%
  # Filling missing party and legislature variables as mentioned by V-Dem's
coding rules
  mutate(v2psoppaut_ord = ifelse(v2psbars_ord <= 1 & is.na(v2psoppaut_ord),
0,
                                 v2psoppaut_ord)) %>%
  mutate(v2lgqstexp_ord = ifelse(v2lgbicam == 0,0,v2lgqstexp_ord)) %>%
  mutate(v2lginvstp_ord = ifelse(v2lgbicam == 0,0,v2lginvstp_ord)) %>%
  # Interpolating missing data
  mutate(v2lgbicam = na.approx(v2lgbicam, na.rm = F)) %>%
  mutate(v2x_elecoff = na.approx(v2x_elecoff, na.rm = F)) %>%
  mutate(v2x_suffr = na.spline(v2x_suffr, na.rm = F)) %>%
  mutate(v2elrstrct = na.approx(v2elrstrct, na.rm = F)) %>%
  mutate(v2elembcap = na.spline(v2elembcap, na.rm = F)) %>%
  mutate(v2elembaut = na.spline(v2elembaut, na.rm = F)) %>%
  mutate(v2meharjrn = na.spline(v2meharjrn)) %>%
  mutate(v2mecenefm = na.spline(v2mecenefm)) %>%
  mutate(v2meslfcen = na.spline(v2meslfcen, na.rm = F)) %>%
  mutate(v2exrescon = na.spline(v2exrescon, na.rm = F)) %>%
  mutate(v2jucomp = na.spline(v2jucomp, na.rm = F)) %>%
  mutate(v2juhccomp = na.spline(v2juhccomp, na.rm = F)) %>%
  mutate(v2juhcind = na.spline(v2juhcind, na.rm = F)) %>%
  mutate(v2juncind = na.spline(v2juncind, na.rm = F)) %>%
  mutate(v2lgqstexp_ord = na.approx(v2lgqstexp_ord, na.rm = F)) %>%
  mutate(v2lginvstp_ord = na.approx(v2lginvstp_ord, na.rm = F)) %>%
  # Ungrouping before SVM regression
  ungroup() %>%
  # Removing still missing data
  drop_na()

# Running SVM function created for the purpose of this contribution
SVM_index(svm_dataframe = svm_df, id_vars = id_vars, support_vars =
support_vars,
          svm_data_vars = svm_data_vars, priming_selec = priming_selection,
          save_path = "result/result_SVM_intern_emb4.RDS")



  # Electoral Regime #

# Variables used for the SVM regression
```

```
svm_data_vars <- c(svm_data_table$Code[svm_data_table$Regime ==
"Electoral"],
                   "v2x_polyarchy")

# Identifying variables
id_vars <-
c("country_name","country_text_id","country_id","iso3c","year","iso_year")

# Variables used for the SVM regression
support_vars <- c("v2x_elecreg", "v2exhoshog", "v2lgbicam")

# Complete variable list
var_SVM <- c(id_vars,support_vars, svm_data_vars)

# Creating the dataset to be used in the SVM regression from the original
V-Dem
# dataset
svm_df <- vDem %>%
  #  Creating iso3c identifying variable
  mutate(iso3c = countryname(country_name, destination = "iso3c")) %>%
  # Correcting disparities in nomenclature
  mutate(iso3c = ifelse(iso3c == "Yugoslavia", "SRB", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Czechoslovakia", "CZE", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "German Democratic Republic", "DDR",
iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Zanzibar", "EAZ", iso3c)) %>%
  # Filtering for the start year
  filter(year >= start_year) %>%
  # Creating identifying variable for each observation
  mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
  # Selecting the needed variables
  select(all_of(var_SVM)) %>%
  # Relocating certain variables for better visualization
  relocate(c("iso3c", "iso_year"), .after = country_id) %>%
  # Grouping by country name so that operations aren't conducted cross
country
  group_by(country_name) %>%
  # Filling missing electoral variables as mentioned by V-Dem's coding
rules
  mutate(v2elrgstry_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elrgstry_ord, na.rm = F), 0))
%>%
  mutate(v2elvotbuy_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elvotbuy_ord, na.rm = F), 0))
%>%
  mutate(v2elirreg_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elirreg_ord, na.rm = F), 0)) %>%
  mutate(v2elintim_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elintim_ord, na.rm = F), 0)) %>%
  mutate(v2elpeace_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elpeace_ord, na.rm = F), 0)) %>%
  mutate(v2elfrfair_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elfrfair_ord, na.rm = F), 0))
%>%
```

```
  # Filling up electoral variables that did not have an election in the
first year
  # of the dataset
  fill(c("v2elrgstry_ord", "v2elvotbuy_ord", "v2elirreg_ord",
"v2elintim_ord",
         "v2elpeace_ord", "v2elfrfair_ord"), .direction = "up") %>%
  # Interpolating missing data
  mutate(v2x_elecoff = na.approx(v2x_elecoff, na.rm = F)) %>%
  mutate(v2x_suffr = na.spline(v2x_suffr, na.rm = F)) %>%
  mutate(v2elrstrct = na.approx(v2elrstrct, na.rm = F)) %>%
  mutate(v2elembcap = na.spline(v2elembcap, na.rm = F)) %>%
  mutate(v2elembaut = na.spline(v2elembaut, na.rm = F)) %>%
  # Ungrouping before SVM regression
  ungroup() %>%
  # Removing still missing data
  drop_na()

# Running SVM function created for the purpose of this contribution
SVM_index(svm_dataframe = svm_df, id_vars = id_vars, support_vars =
support_vars,
          svm_data_vars = svm_data_vars, priming_selec = priming_selection,
          save_path = "result/result_SVM_elec_reg4.RDS"
)
```

```
  # Fully Embedded #

# Variables used for the SVM regression
svm_data_vars <- c(svm_data_table$Code,"v2x_polyarchy")

# Identifying variables
id_vars <- c("country_name","country_text_id","country_id","iso3c","year",
             "iso_year")

# Variables used for the SVM regression
support_vars <- c("v2x_elecreg", "v2exhoshog", "v2lgbicam")

# Complete variable list
var_SVM <- c(id_vars,support_vars, svm_data_vars)

# Creating the dataset to be used in the SVM regression from the original
V-Dem
# dataset
svm_df <- vDem %>%
  # Creating iso3c identifying variable
  mutate(iso3c = countryname(country_name, destination = "iso3c")) %>%
  # Correcting disparities in nomenclature
  mutate(iso3c = ifelse(iso3c == "Yugoslavia", "SRB", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Czechoslovakia", "CZE", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Republic of Vietnam", "VDR", iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Yemen People's Republic", "YMD", iso3c))
%>%
  mutate(iso3c = ifelse(iso3c == "German Democratic Republic", "DDR",
iso3c)) %>%
  mutate(iso3c = ifelse(iso3c == "Zanzibar", "EAZ", iso3c)) %>%
```

```
  # Removing empty variables related to the DDR before its formal founding
in 1949
  filter(!(iso3c == "DDR") | year > 1948) %>%
  # Filtering for the start year
  filter(year >= start_year) %>%
  # Creating identifying variable for each observation
  mutate(iso_year = paste(iso3c, year, sep = "_")) %>%
  # Selecting the needed variables
  select(all_of(var_SVM)) %>%
  # Relocating certain variables for better visualization
  relocate(c("iso3c", "iso_year"), .after = country_id) %>%
  # Grouping by country name so that operations aren't conducted cross
country
  group_by(country_name) %>%
  # Filling missing electoral variables as mentioned by V-Dem's coding
rules
  mutate(v2elmulpar_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elmulpar_ord, na.rm = F), 0))
%>%
  mutate(v2elrgstry_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elrgstry_ord, na.rm = F), 0))
%>%
  mutate(v2elvotbuy_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elvotbuy_ord, na.rm = F), 0))
%>%
  mutate(v2elirreg_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elirreg_ord, na.rm = F), 0)) %>%
  mutate(v2elintim_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elintim_ord, na.rm = F), 0)) %>%
  mutate(v2elpeace_ord = ifelse(v2x_elecreg == 1,
                                na.locf(v2elpeace_ord, na.rm = F), 0)) %>%
  mutate(v2elfrfair_ord = ifelse(v2x_elecreg == 1,
                                 na.locf(v2elfrfair_ord, na.rm = F), 0))
%>%
  # If HoS and HoG are the same, both have the same value
  mutate(v2exctlhg_0 = ifelse(v2exhoshog == 1, v2exctlhs_0, v2exctlhg_0))
%>%
  # Filling up electoral variables that did not have an election in the
first year
  # of the dataset
  fill(c("v2elmulpar_ord","v2elrgstry_ord", "v2elvotbuy_ord",
"v2elirreg_ord",
         "v2elintim_ord", "v2elpeace_ord", "v2elfrfair_ord"), .direction =
"up") %>%
  # Filling missing party and legislature variables as mentioned by V-Dem's
coding rules
  mutate(v2psoppaut_ord = ifelse(v2psbars_ord <= 1 & is.na(v2psoppaut_ord),
0,
                                 v2psoppaut_ord)) %>%
  mutate(v2lgqstexp_ord = ifelse(v2lgbicam == 0,0,v2lgqstexp_ord)) %>%
  mutate(v2lginvstp_ord = ifelse(v2lgbicam == 0,0,v2lginvstp_ord)) %>%
  # Interpolating missing data
  mutate(v2meharjrn = na.spline(v2meharjrn)) %>%
  mutate(v2mecenefm = na.spline(v2mecenefm)) %>%
  mutate(e_peedgini = na.spline(e_peedgini, na.rm = F)) %>%
  mutate(e_migdppc = na.spline(e_migdppc, na.rm = F)) %>%
  mutate(e_pelifeex = na.spline(e_pelifeex, na.rm = F)) %>%
  mutate(v2csprtcpt = na.spline(v2csprtcpt, na.rm = F)) %>%
```

```
  mutate(v2elembcap = na.spline(v2elembcap, na.rm = F)) %>%
  mutate(v2elembaut = na.spline(v2elembaut, na.rm = F)) %>%
  mutate(v2elrstrct = na.approx(v2elrstrct, na.rm = F)) %>%
  mutate(v2x_elecoff = na.approx(v2x_elecoff, na.rm = F)) %>%
  mutate(v2lgbicam = na.approx(v2lgbicam, na.rm = F)) %>%
  mutate(v2x_suffr = na.spline(v2x_suffr, na.rm = F)) %>%
  mutate(v2meslfcen = na.spline(v2meslfcen, na.rm = F)) %>%
  mutate(v2exrescon = na.spline(v2exrescon, na.rm = F)) %>%
  mutate(v2jucomp = na.spline(v2jucomp, na.rm = F)) %>%
  mutate(v2juhccomp = na.spline(v2juhccomp, na.rm = F)) %>%
  mutate(v2juhcind = na.spline(v2juhcind, na.rm = F)) %>%
  mutate(v2juncind = na.spline(v2juncind, na.rm = F)) %>%
  mutate(v2lgqstexp_ord = na.approx(v2lgqstexp_ord, na.rm = F)) %>%
  mutate(v2lginvstp_ord = na.approx(v2lginvstp_ord, na.rm = F)) %>%
  mutate(v2pscnslnl = na.spline(v2pscnslnl, na.rm = F)) %>%
  # Ungrouping before SVM regression
  ungroup() %>%
  # Removing still missing data
  drop_na()

# Running SVM function created for the purpose of this contribution
SVM_index(svm_dataframe = svm_df, id_vars = id_vars, support_vars =
support_vars,
          svm_data_vars = svm_data_vars, priming_selec = priming_selection,
          save_path = "result/result_SVM_extern_emb4.RDS")

# Comparison

# Reading individual SVM dataset created in this section
result_int_emb <- readRDS("result/result_SVM_intern_emb4.RDS")$Output

result_elec_reg <- readRDS("result/result_SVM_elec_reg4.RDS")$Output

result_full_emb <- readRDS("result/result_SVM_extern_emb4.RDS")$Output

# Reading SVMDI
SVMDI <- read_excel("data/Gruendler/ML indices (1).xlsx") %>%
  mutate(iso_year = paste(ISO, Year, sep = "_"))

# Joining all datasets
final_df <- result_elec_reg[,c(id_vars, "med")] %>%
  rename(`Elec. Reg.` = med) %>%
  full_join(
    result_int_emb[,c(id_vars, "med")] %>%
      rename(`Intern. Emb.` = med),
    by = id_vars
  ) %>%
  full_join(
    result_full_emb[,c(id_vars, "med")] %>%
      rename(`Fully Emb.` = med),
    by = id_vars
  ) %>%
  full_join(
    SVMDI[,c("iso_year","C_MLDI")] %>%
      rename(SVMDI = C_MLDI),
    by = "iso_year"
  ) %>%
  pivot_longer(cols = c("Elec. Reg.","Intern. Emb.","Fully Emb.","SVMDI"))
```

```
  # Figures

    ## Figure 7

agg_plot_7 <- final_df %>%
  filter(iso3c == "DEU") %>%
  ggplot(aes(x = year, y = value, color = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index"))

agg_plot_7

png("img/figure7.png", res = 120, width = 700, height = 700)
agg_plot_7
dev.off()

    ## Figure 8

agg_plot_8 <- final_df %>%
  filter(iso3c == "HUN" & year > 2000) %>%
  ggplot(aes(x = year, y = value, color = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index"))

agg_plot_8

png("img/figure8.png", res = 120, width = 700, height = 700)
agg_plot_8
dev.off()

    ## Figure 9

agg_plot_9 <- final_df %>%
  filter(iso3c == "POL" & year > 2000) %>%
  ggplot(aes(x = year, y = value, color = name)) +
  geom_line(size = 1.2, alpha = 0.6) +
  labs(y = "Score", x = "Year") +
  guides(color = guide_legend(title = "Index"))

agg_plot_9

png("img/figure9.png", res = 120, width = 700, height = 700)
agg_plot_9
dev.off()
```

Declaration on Honor

I hereby declare that the paper submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the paper in digital form can be examined for the use of unauthorized aid and in order to determine whether the thesis as a whole or parts incorporated in it may be deemed a plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future theses submitted. Further rights of reproduction and usage, however, are not granted here. This paper was not previously presented to another examination board and has not been published.