Defining ‘democracy’: Are we staying on topic?

Sean Ingham and David Wiens

April 15, 2020

Abstract

Political scientists’ failure to pay careful attention to the content (as opposed to the operationalization) of their chosen definition of ‘democracy’ can make them liable to draw invalid inferences from their empirical research. With this problem in mind, we argue for the following proposition: if one wishes to conduct empirical research that contributes to an existing conversation about democracy, then one must choose a definition of ‘democracy’ that picks out the topic of that conversation as opposed to some other (perhaps nearby) topic of conversation. We show that, as a practical matter, one of the most effective methods for preserving “topic continuity” is to choose a definition of ‘democracy’ that concurs with prevailing judgments about how to classify particular regimes, emphasizing the superiority (in this regard) of judgments about stylized hypothetical scenarios as opposed to judgments about the actual regimes we observe in our datasets.

Before we can study the causes and consequences of democracy, we must first choose how to define and measure the concept of democracy.[1] Political scientists recognize this, of course, as attested by the large literature that addresses these and related topics (among others, see Cheibub, Gandhi and Vreeland, 2010; Collier and Adcock, 1999; Coppedge et al., 2016; Goertz, 2006; Munck, 2009; Paxton, 2000). Existing discussions predominantly focus on measurement issues: whether to use a dichotomous or many-valued measure; the choice of measurement scale (e.g., nominal, ordinal, or interval); the choice of indicators to operationalize the underlying concept; the choice of mathematical operations to aggregate indicator scores. Even those contributions that discuss conceptual issues more directly tend to

[1] Strictly speaking, concepts are not defined; instead, words are defined so as to refer to one concept rather than another. So, instead of talking about “how to define and measure the concept of democracy”, it is more accurate to say “how to define the word ‘democracy’ — that is, how to choose which concept ‘democracy’ should refer to — and measure the concept referred to by that word”. But this formulation is unacceptably cumbersome. For the remainder of the paper, with some abuse of language, we use phrases like “define the concept of democracy” and “define [the word] ‘democracy’” interchangeably to mean “choose the concept to which the word ‘democracy’ should refer”. We use ‘democracy’ (in single quotation marks) to mark the word and, where appropriate, the concept picked out by the word.

use measurement-related criteria to assess choices about how to define 'democracy'. For example, Cheibub, Gandhi and Vreeland (2010), echoing Przeworski et al. (2000), defend their minimalist definition of *democracy as competitive elections* on the grounds that it can be operationalized by "clear and stark" coding rules, which results in a measure that conveys "precise information" about observations; additionally, these coding rules "involve[] no subjectivity", referring wholly to observable features of regimes, thus producing an "easily reproducible" measure (2010, p. 71).

In this paper, we set aside measurement concerns to emphasize issues that arise when we make choices about conceptual content. To motivate and focus our discussion, we start by showing how political scientists' failure to pay careful attention to the content (as opposed to the operationalization) of their chosen definition of 'democracy' can make them liable to draw invalid inferences from their empirical research. These threats to valid inference raise important questions about how to assess candidate definitions of 'democracy' with regard to their conceptual content. We show that attempts to answer these questions while retaining a heavy emphasis on measurement-related issues are implausible. We then introduce the idea of *topic continuity* (Cappelen, 2018) and demonstrate its importance for choosing among alternative definitions of 'democracy'. The basic idea is that definitions of 'democracy' determine *topics of conversation*; that is, how one understands the concept of democracy shapes what one believes one is thinking/speaking/writing/inquiring about when using the term 'democracy'. With this in mind, we argue for the following proposition, roughly stated for now: If one wishes to conduct empirical research that contributes to an existing conversation about democracy (scholarly or not), then one must choose a definition of 'democracy' that "picks out" the topic of *that* conversation as opposed to some other (perhaps nearby) topic of conversation. We further show that, as a practical matter, one of the most effective methods for preserving topic continuity is to choose a definition of 'democracy' that concurs with prevailing judgments about how to classify particular regimes.

This practical recommendation may raise a skeptical concern that anchoring choices regarding conceptual content to "subjective" judgments about particular cases is anti-scientific and arbitrary (cf. Cheibub, Gandhi and Vreeland, 2010). We address these concerns by showing that political scientists already implicitly (though unsystematically) appeal to judgments about particular cases as a way to craft a shared understanding of the topic under discussion. One way to characterize our aim, then, is to make explicit the criterion of topic continuity so that we can think systematically about its relevance, as well as craft best practices for its application in choosing among candidate definitions of 'democracy'. With respect to best practices, we address lingering skepticism by showing how the application of topic continuity to test and choose among candidate definitions of 'democracy' can be improved and made scientifically credible by appealing to judgments about stylized hypothetical scenarios as opposed to judgments about the actual observations we record in our empirical datasets.

Although we confine our attention to adjudicating among competing definitions of 'democracy', this is just one

example of a more general phenomenon in political science. Other central concepts in the discipline — for example, the concepts of power, accountability, political representation, rule of law, and so on — admit of multiple, conflicting interpretations. A scholar's choice of definition can be consequential, affecting the questions she poses, the hypotheses she entertains, and the framing of the research findings. When we face such choices, a natural question is whether some definitions are better than others, and by what criteria we might make such comparisons. Our reflections in this paper generalize to this broader context: whatever the concept one is using to articulate empirical hypotheses and explanatory theories, one must not only pay careful attention to the operationalization of that concept, but must also ensure that one's choices with respect to conceptual content produce a definition that corresponds to the topic of the conversation to which one wishes to contribute.

1 The consequences of neglecting conceptual content

Following Joseph Schumpeter's (1942) minimalist view of democracy, political scientists have largely converged on a *procedural* approach to defining 'democracy': democracy is roughly defined as an institutional arrangement for selecting political leaders through competitive elections. No doubt, there is dispute about which institutions and procedures constitute a democracy. All definitions accept that competitive elections are a necessary condition for democracy. Remaining disputes thus concern whether competitive elections are sufficient for democracy and, if not, which features or attributes must be added for a regime to count as democratic. Some argue that competitive elections are sufficient (Przeworski et al., 2000; Cheibub, Gandhi and Vreeland, 2010). Most, however, accept some sort of minimum suffrage or participation requirement as necessary for democracy (e.g., Boix, Miller and Rosato, 2012; Dahl, 1971; Paxton, 2000). Many also argue that effective guarantees for citizens' civil and political rights are a necessary condition (e.g., Collier and Levitsky, 1997; Dahl, 1971; Freedom House, 2020; Levitsky and Way, 2010), although it is not always clear whether such guarantees are treated merely as a way to operationalize the idea of competitive elections (because, e.g., civil rights guarantees tend to support genuinely competitive elections; see Levitsky and Way 2010) or whether they are treated as a defining feature of democracies and, thus, should be included in addition to the condition of competitive elections. Beyond these, a range of further conditions are proposed as defining features of democracy, for example: that elected governments must be able to rule effectively and without interference from unelected agencies (Adcock and Collier, 2001), or that the executive's decision-making authority is constrained by other government or civil society agencies (Marshall, Gurr and Jaggers, 2018).

Not every political scientist subscribes to the procedural view, to be sure; especially in the literature on democratic representation, scholars frequently cite Dahl's claim that a "key characteristic of a democracy is continuing respon-

siveness of the government to the preferences of its citizens, considered as political equals" (1971, p. 1). Dahl's claim can be seen as one interpretation of the traditional idea, rooted in the original meaning of the Greek *demokratia*, that democracy is a form of rule (*kratos*) by the people (*demos*). Besides popular rule, one could think of various related concepts with which 'democracy' has historically been closely identified, such as popular control, popular sovereignty, popular government, collective self-government, and so on. Advocates for the procedural view do not typically reject this traditional understanding of 'democracy' but instead propose the procedural view as a definition for a technical concept, one that "systematizes" (Adcock and Collier, 2001) the vague and intuitive connotations of the traditional definition and, thus, renders the traditional definition amenable to operationalization in terms of observable features of regimes (e.g., Coppedge et al., 2016; Levitsky and Way, 2010; Munck, 2009).[2] But these differences are generally considered relatively minor. We think it is more than fair to say that there is broad agreement among political scientists that a procedural definition of 'democracy' is best suited to the purposes of empirical political science.

One might be tempted to resist our claim of a broad agreement on a procedural definition by highlighting the wide range of democracy measures on offer and the relatively vigorous disputes about their respective merits as measurement instruments (e.g., Cheibub, Gandhi and Vreeland, 2010; Collier and Adcock, 1999; Coppedge et al., 2016; Elkins, 2000; Treier and Jackman, 2008). But disputes about conceptual content are distinct from disputes about measurement validity insofar as the latter presuppose agreement on conceptual content. Indeed, agreement on a procedural definition has allowed scholars to acknowledge this distinction only to set aside conceptual issues and focus instead on measurement validity (e.g., Adcock and Collier, 2001; Seawright and Collier, 2012). Measurement validity is obviously an important issue, given the problems that measurement error can pose for scientific inference about the causes and consequences of democracy.[3] However, this overwhelming focus on avoiding measurement error has obscured the ways in which inferential validity can depend not only on one's choice of measure but also on one's choice of conceptual content. Put differently, due to their fixation on finding a valid measure of the concept of democracy, political scientists have failed to see how choices related to conceptual content can threaten the validity of certain inferences from empirical research even if that research uses a valid measure of the chosen concept.

To get an intuition for how one's choice of conceptual content can threaten inferential validity, consider the following schematic scenario. Imagine a researcher wants to study the effect of democracy on some outcome of interest, say, economic growth, which is operationalized using a variable Y. Suppose the researcher adopts a procedural definition of 'democracy' and operationalizes this concept using a variable C ("competitive elections"), which we assume is a valid measure of the procedural definition. Using a research design that is appropriate for causal identification,

[2] Schumpeter (1942); Przeworski et al. (2000); Riker (1982) are examples of scholars who present a minimal procedural definition as a *replacement* for the traditional definition of 'democracy', rejecting popular rule and its cognates for one reason or another.

[3] For an example, see Paxton's (2000) discussion of how measurement error arising from neglect for women's suffrage threatens inferences about democratic transitions.

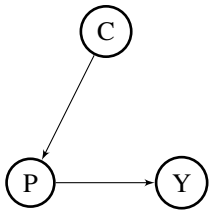

Figure 1: A regime's status as a competitive electoral regime (C) affects whether it is a regime that provides citizens with an adequate degree of control (P), which in turn affects economic growth (Y). Competitive elections affect growth only through their effects on citizens' degree of control. Evidence that they affect the growth is evidence that popular control affects growth.

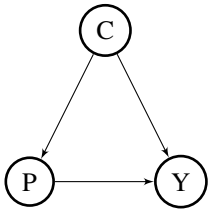

Figure 2: A regime's status as a competitive electoral regime (C) affects the extent to which citizens have control over political decisions (P), which in turn affects economic growth (Y). Competitive elections also directly affect growth (Y). Evidence regarding the effect of competitive elections on growth is not evidence for the effect of popular control on growth.

the researcher finds that C has a substantial and significant positive effect on Y. In summarizing these findings, the researcher writes: "Contrary to conventional wisdom, democracy increases economic growth." If the conventional wisdom that is allegedly overturned is articulated using a procedural definition of 'democracy', then we can validly conclude that the study's findings challenge the conventional wisdom. Since this case is straightforward, we set it aside.

Things are less straightforward when the conventional wisdom presupposes a different definition of 'democracy' — for example, "rule by the people" or one of its cognates. Can we validly infer that the empirical findings overturn the conventional wisdom in this case? Our answer to this question will turn on our beliefs about the true causal structure relating competitive elections, economic growth, and rule by the people (or whatever the definition presupposed by conventional wisdom picks out). To illustrate this point, let P be a (perhaps unobservable) variable that indicates citizens' degree of control over political outcomes (which corresponds to a popular control definition of 'democracy'). Let's stipulate that the "conventional wisdom" under investigation claims that increasing citizens' control over political outcomes reduces economic growth rates.

Suppose, to start, that the true causal relationships are as depicted in figure 1: having competitive elections causally affects citizens' control over political decisions and influences economic growth, if at all, only via its effect on popular control. Given this causal structure, if the empirical study shows that competitive elections cause an increase in growth, then we can validly conclude that the study's findings overturn the conventional wisdom.

Now suppose, instead, that the true causal structure is given by figure 2: competitive elections causally influence economic growth rates indirectly, via its effect on citizens' control, but also directly. Given this causal structure, the finding that competitive elections increase growth presents no evidence against the conventional wisdom. Indeed, the conventional wisdom could still be consistent with the hypothetical findings. This would be so if increasing popular

control reduces growth rates while competitive elections increase growth.

Clearly, in this example, one cannot validly conclude that the empirical study overturns the conventional wisdom until one persuasively argues that the true causal structure is one way rather than another. Even if C is a valid measure of a procedural definition of 'democracy,' and one's research design ensures valid inferences about the effects of C on Y, still one cannot draw conclusions about the effect of democracy *as the conventional wisdom conceptualizes it* without first establishing the causal structure relating the variables associated by the procedural and conventional definitions of 'democracy'. Observe, too, that, absent assumptions about the true causal structure, such conclusions would be invalid even if there is perfect correlation between C and P, the two measures of democracy corresponding to the procedural and conventional definitions of 'democracy', respectively. This points to a broader lesson: In contexts where the term 'democracy' is used to refer to several distinct concepts, these conceptual differences pose threats to inferential validity that cannot be addressed by establishing that one is using a valid measure of one's chosen definition of 'democracy' or demonstrating or that one's empirical research design ensures causal identification. Put simply, drawing valid inferences from our empirical research requires us to pay careful attention to conceptual content.[4]

This is not a merely hypothetical scenario. For a familiar example, Przeworski and colleagues (2000) endorse a minimalist procedural definition of 'democracy': "'democracy,' for us, is a regime in which those who govern are selected through contested elections" (p. 15). This definition is operationalized using a dichotomous measure of regime type, which is produced using four coding rules that focus on the presence or absence of competitive elections within a regime (see pp. 18–29 for details). Using this measure of contested elections, Przeworksi et al. subsequently study the empirical relationship between democracy and economic investment. Responding to a literature that "claim[s] that democracy undermines investment" and, in turn, economic growth, Przeworski et al. argue that this claim "finds no support in the evidence" (p. 146). On closer inspection, however, the existing literature they are addressing is not about the economic effects of elections *per se*, but about the effect of increasing citizens' influence over economic policy. Indeed, two of the articles they cite to motivate their countervailing analysis do not even mention elections or voters but instead focus on the investment effects of increasing workers' policy influence via unions and labor parties (see p. 142).[5] Przeworski et al. thus find themselves in a situation where the literature they address argues that increasing citizens' policy influence deters investment, while their empirical analysis shows that the presence or

[4]One might argue here that we can validly infer that the imagined empirical study overturns the conventional wisdom if we show that C is a valid instrument for drawing inferences about the effect of popular control on economic growth. We agree with this point, but it does nothing to blunt the force of the point we are making here: that valid inference requires paying attention to conceptual content in addition to measurement validity and empirical research design. To wit, to establish that C is a valid instrument, one must argue for claims about the true causal structure, an argumentative burden that is revealed by attending to potential differences in conceptual content. And, further, arguing that C is a valid instrument for popular control is only one approach to resolving these conceptual differences — one that leaves unresolved the questions about how to define 'democracy' that we focus on here.

[5]E.g., they quote from the following passage in de Schweinitz (1959): "Trade unions and labor parties, however, raise problems of a different order. If they are successful in securing a larger share of the national income and in limiting the freedom of action of entrepreneurs, they may have the effect of restricting the investment surplus so much that the rate of economic growth is inhibited" (p. 388).

absence of competitive elections makes no difference to rates of investment.

One might defend the relevance of their findings to the broader debate by arguing that their measure of competitive elections is a valid instrument for popular influence, as in Figure 1. Ironically, Przeworski et al. themselves express doubts about the link between competitive elections and popular influence when they defend their choice of a procedural definition of 'democracy' (p. 33), so they are not in a position to defend their research design in these terms. In any case, our claim is not that Przeworski et al.'s conclusions vis-à-vis the existing literature are in fact invalid. Our point is instead that the validity of their inferences vis-à-vis the existing literature depends on assumptions about the causal relationship between the variable associated with their concept and the variable associated with their interlocutors' distinct concept. Vindicating this assumption is a burden they fail to acknowledge, much less address.

As a second example, Acemoglu et al. (2019) claim to show that democracy causes economic growth. For their statistical analyses, they construct a novel dichotomous measure of democracy from several extant indices: Freedom House, Polity IV, Cheibub, Gandhi and Vreeland (2010) (which updates Przeworski et al.'s [2000] measure), and Boix, Miller and Rosato (2012) (see pp. 53–4 and their online appendix for details). This measure operationalizes a procedural definition of 'democracy': a democracy is a regime in which political leaders are selected via competitive elections, there are institutional checks on executive authority, and citizens enjoy minimal civil and political rights (appendix, p. A4). Using this measure, and a causal identification strategy whose assumptions we can grant for the sake of argument, Acemoglu et al. infer that democracy causes economic growth, and draw conclusions about a host of subsidiary results pertaining to mechanisms: that democracy "increas[es] investment, encourag[es] economic reforms, improv[es] the provision of schooling and health care, and reduc[es] social unrest" (p. 51). These results are said to challenge widespread "[s]kepticism about the [economic] performance of democracy," skepticism that is said to go back to Plato and Aristotle (p. 96). But if Plato and Aristotle are supposed to be representatives of this skeptical view, then it cannot be a view about the economic consequences of selecting leaders via competitive elections or of granting citizens minimal liberal rights—contemporary democratic procedures and institutions had no place in classical political theorists' theoretical framework (cf. Ober, 2017). Consider the quotes Acemoglu et al. include (p. 96): Plato's concern is with a regime in which "idle spendthrifts" "are almost the entire ruling power", while Aristotle's contention is that "it is not safe to trust them [the bulk of the people] with the first offices of the state" (original interpolation). More generally, classical skepticism about democracy up to the late eighteenth century focused on the consequences of granting ordinary citizens significant political power by allowing them to occupy political offices. We see, then, that Acemoglu et al. use a measure of a procedural view of democracy to address a skeptical view that defines 'democracy' in terms of more traditional notions such as popular rule or popular control. Thus, to validly infer that their empirical results challenge this skeptical view, it is incumbent on them to defend assumptions about how the

variable associated with their concept of democracy relates to the distinct variable associated with their interlocutors' concept of democracy. Because they ignore questions about conceptual content, Acemoglu et al. fail to acknowledge this burden.

These are not exotic examples drawn from the periphery of the discipline. These are some of the most influential social scientists drawing the kinds of inferences that are the bread-and-butter of empirical democracy research. Our point, to be clear, is not that the inferences we've highlighted are in fact invalid. It is instead to point out the ways in which failure to pay careful attention to the content of different definitions of 'democracy' leaves the validity of these kinds of inferences in doubt and open to challenges that few empirical social scientists appreciate.

2 Choosing definitions of democracy

Our discussion in the previous section raises pressing questions about how to choose among different definitions of 'democracy'. How should we settle on a definition of 'democracy', given the myriad possibilities? More specifically, how (if at all) should one's choice of definition be constrained by others' understanding of 'democracy'? To prepare the ground for a plausible answer to this question, let's consider some proposals that would prompt little to no change to the current measurement-focused practice.

Consider first the proposal that social scientists are free to stipulate just any definition of 'democracy'. A moment's reflection is enough to see that this view is far too crude. A good definition has to be logically consistent and must be deemed better, all else being equal, if its implications are clear and unambiguous. Moreover, if the goal is a definition that is fruitful for empirical inquiry, it is better, all else being equal, if it picks out a concept that can be measured, and so criteria for the choice of measures (e.g., Adcock and Collier, 2001; Collier and Adcock, 1999; Coppedge, 2012, chap. 2; Munck, 2009, chap. 2; Seawright and Collier, 2012) must indirectly constrain how one stipulates a definition of 'democracy'. Political scientists recognize this point; they err, not in neglecting it, but in focusing almost exclusively on measurement-related concerns when evaluating definitions.

Internal consistency and the requirements of good measurement cannot be the only constraints on stipulated definitions of democracy. To see why, suppose (what is absurd) that Przeworski et al. (2000) had stipulated that 'democracy' is to refer to political regimes in which the chief executive's astrological sign is Libra. This definition can be readily operationalized so as to meet their technical criteria for a good measure; in particular, it can be "formalize[d]…in terms of rules that can be decisively and reliably applied to the observable aspects of national histories" (p. 13). Nonetheless, the definition is absurd. One reason someone might give is that it violates a criterion that is at least implicitly acknowledged by most political scientists:

Explanatory significance. All else being equal, a good measure (and, in turn, the definition chosen to fit with the measure) picks out attributes (e.g., of regimes) that figure in explanations of otherwise puzzling facts (cf. Przeworski et al., 2000, p. 14; Cheibub, Gandhi and Vreeland, 2010, pp. 72–3).

Note that this criterion imposes a content-related (as opposed to a measurement-related) constraint on acceptable definitions of 'democracy'. Insofar as one accepts this criterion — and we cannot think of anyone who rejects it — one must reject the thought that operationalization issues are all that matter when choosing a definition of 'democracy'.

One might still persist with the thought that scholars are permitted to stipulate a definition of democracy, although within the bounds set by the constraints noted thus far. Even this more nuanced proposal runs into immediate problems. To illustrate, suppose we agree that a procedural minimalist definition (e.g., that given by Przeworski et al.) satisfies all relevant operationalization-related criteria as well as the criterion of explanatory significance. Imagine now that a scholar stipulates the

Kindergartener's definition: A regime is a democracy if (and only if) it is *not* a competitive electoral regime.

From the perspective of a procedural minimalist, one who stipulates the kindergartener's definition is like the child who decides to use words as if they meant the opposite of what they actually mean. We expect that most scholars would reject the kindergartner's definition. Note, however, that it fares no worse than existing minimalist definitions with respect to the criteria we've accepted thus far since it simply reverses the labels of the classification scheme produced by existing minimalist measures.

One obvious reason to reject the kindergartner's definition in favor of a procedural minimalist definition is that the former is liable to produce confusion whereas the latter is not. This is on the right track, but it's important to be clear about the kind of confusion that would ensue from using the kindergartner's definition. Notice, in particular, that whatever confusion it would create, it would do so despite satisfying the criteria of explanatory significance and operationalizability. It can be operationalized using the same rules for coding indicators and aggregating indicator scores as those used by existing minimalist measures, and if whether a regime *is* a competitive electoral regime has explanatory significance, then whether a regime *is not* a competitive electoral regime must also have such significance.

If the kindergartner's definition creates confusion, it is not because it fares poorly by the criteria noted thus far, but because it deviates from the concept of democracy as it is typically understood. This suggests that an acceptable definition of 'democracy' must be tethered to others' understanding of the concept's content. Exactly how is still unclear, but we can put to rest the idea that purely stipulative definitions of 'democracy', which ignore how one's interlocutors understand the concept of democracy, are acceptable.

How should existing understandings of the concept of democracy constrain one's definition? During its long his-

tory, people have apparently meant all kinds of things by 'democracy', often considered the paradigmatic example of an "essentially contested concept" (Gallie, 1956; cf. Collier, Hidalgo and Maciuceanu, 2006). Scholars routinely warn against the unproductive and hopeless exercise of trying to characterize what 'democracy' "really" means (Adcock and Collier, 2001). The jumble of vague and inconsistent connotations the term has acquired—what one might call the folk concept or "background" concept of democracy (Adcock and Collier, 2001) — may seem an unpromising source of constraints on a definition intended for social scientific inquiry. What social science needs is a technical (or "formalized" or "systematized") concept that refines our fuzzy background concept of democracy (Przeworski et al., 2000, p. 13; Adcock and Collier, 2001).

While that may be so, the example of the kindergartener's definition shows that a technical refinement of our fuzzy background concept still has to be tethered to the latter somehow. It needs to be recognizable as a refinement *of the background concept of democracy*, as opposed to a refinement of its opposite or something else entirely.[6] Moreover, the examples in the previous section underscore the dangers of failing to attend to the possibly different choices about conceptual content across different scholarly contexts (or by different scholars working within the same context). Scholars may think they are putting to the test widely held beliefs about democracy, or past scholars' or thinkers' claims about democracy, when they are in fact just talking past their intended interlocutors, unwittingly addressing a different topic altogether. If social scientists want to frame their research as part of a conversation about democracy, then they have to consider whether their technical concept of democracy is sufficiently anchored to the background concept that they are addressing the same topic of conversation as their interlocutors.

Despite its being vague or inconsistent, the background concept of democracy may still have enough content to anchor our choice of a technical definition. Multifarious uses of the concept could, despite their differences, share certain core denotations. To wit, scholars frequently acknowledge that, across a wide range of contexts, the various uses of 'democracy' more or less consistently denote the classical idea of "rule by the people" or some nearby notion such as "popular sovereignty" or "popular control" (e.g., Dunn, 2004). Indeed, numerous political scientists start from this classical idea and present their proposed measures as refinements of it, meant to render it precise and determinate enough to make measurement possible (e.g., Coppedge et al., 2011; Dahl, 1971). Even those who wish to detach the concept of democracy from its classical denotation recognize this venerable tradition (e.g.,Achen and Bartels, 2016; Przeworski, 2010; Riker, 1982; Schumpeter, 1942).

These remarks raise several questions. What does it mean for a background concept to "anchor" a technical definition of democracy? What does it mean for two people who use different definitions of 'democracy' (e.g., an

[6]Compare Carnap's notion of "explication", which aims to replace an inexact concept (the *explicandum*) with a more exact and scientifically fruitful concept (the *explicatum*). An explication is successful insofar as the explicatum is "similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used" (Carnap, 1950, p. 7, quoted in Cappelen, 2018). Carnap's emphasis on similarity of use makes clear why a technical definition of a concept must be tethered to others' use of that concept.

informal, ordinary concept of democracy and a technical concept of democracy) to discuss the same topic, and how do we establish the topic of their conversation? We take these questions up in the next section, where we propose *topic continuity* as a criterion for good technical definitions of 'democracy'.

In what follows, we will not argue for any specific technical definition of 'democracy', nor will we argue that acceptable technical definitions of 'democracy' must be tethered to any particular background or folk notion of democracy; we use specific definitions of 'democracy' only as examples to make our reasoning more determinate. Sorting out these issues is, of course, an important task. But it is most effectively carried out within a framework of shared criteria for assessing candidate answers to these questions. Our objective in this paper is contribute to the development of these criteria.

3 Topic continuity

We propose the following criterion, inspired by Cappelen (2018), for assessing when a technical definition of 'democracy' is acceptable as a refinement of a background concept:

Topic continuity. A candidate technical definition of 'democracy' is better, all else equal, if it allows one to inquire (think, speak, write) about the same topic as one's interlocutors.

Making this criterion maximally precise requires us to say something about how to individuate topics and what it would take for two or more people to be thinking, speaking, writing, and inquiring about the same topic when using the term 'democracy'. We can't offer a general analysis of these matters, but fortunately using topic continuity to assess candidate technical definitions of 'democracy' doesn't require maximal precision. It will be enough for our purposes to develop rules of thumb for applying the criterion in practice.

To warm up, consider a stylized conversation inspired by our examples in section 2. One speaker (who we'll call "Poppy") uses the term 'democracy' in a way that emphasizes the classical notions of popular rule and its cognates; the other (who we'll call "Minnie") is a proponent of a Schumpeterian minimalist definition of 'democracy'.

POPPY: Democracy is bad for economic performance. By giving ordinary citizens — workers and the poor in particular — a significant measure of influence over economic policy choices, resources are diverted away from investment in productive activities and toward consumption, which, in turn, decreases economic growth.

MINNIE: Our empirical study shows otherwise: societies that select their political leaders using competitive elections do not systematically differ in their investment rates from societies that do not.

POPPY: That's interesting, but let's make sure we're talking about the same thing. As I use the term 'democracy', classical Athens is an archetype of a democracy. Do you agree?

MINNIE: No. According to my technical definition, classical Athens is not a democracy.

POPPY: Really? That's surprising. Do you agree that Athenian citizens exercised significant control over political decisions?

MINNIE: Yes.

POPPY: And you still say that Athens was not a democracy?

MINNIE: Correct.

There's an intuitive sense in which Poppy could rightly complain at this point that Minnie has changed the topic of conversation by adopting a minimalist definition. But why, exactly? Can we make this intuition more precise?

The most salient feature of our imagined scenario is that Poppy and Minnie disagree on how to classify classical Athens. More than that, however, they disagree about how to classify a case that Poppy uses to anchor her understanding of democracy as a topic of conversation. This isn't to say that Poppy is using a precisely defined concept of democracy; it may be vague and in desperate need of refinement. Even still, in Poppy's mind, whatever we're talking about when we talk about democracy, we are talking about a set that includes classical Athens. In effect, Poppy treats classical Athens as a test case for acceptable definitions of 'democracy'. That's why she finds Minnie's exclusion of classical Athens from the set of democracies surprising and counterintuitive.

An important assumption about topics and definitions is that two people who use distinct definitions of 'democracy' can nonetheless be talking about the same topic (see Cappelen, 2018, chaps. 9 and 10 *passim*; also Sawyer, 2018). For example, early definitions of "the derivative" expressed in terms of infinitesimals diverge from later definitions expressed in terms of the modern concept of a limit; the former posit the existence of infinitesimally small numbers, while the latter do not. Nonetheless, mathematicians like Leibniz and Cauchy are accurately described as discussing the same topic — the derivative — despite relying on distinct definitions. Moreover, each may be described as discussing the same topic as an interlocutor who lacks a technical definition but who has asked a question about the rate at which a moving object's speed is increasing at a particular point in time. If in doubt, all the participants to this conversation could reassure each other that they are discussing the same topic — despite some lacking any technical definition of the key concept and some using alternative technical definitions — by referring to particular test cases. For example, one mathematician might ask another, "Whatever we're talking about when we talk about 'the derivative,' or 'instantaneous rate of change,' you agree that the derivative is zero at the peak of this parabolic trajectory describing the cannonball's vertical position as a function of time, yes?". A negative answer would be surprising; it would indicate that the interlocutor is confused about the topic that mathematicians are discussing when

they discuss the derivative; a positive answer would reassure the questioner that, however the respondent might define "the derivative," they are not talking past each other, but are discussing the same topic.

The history of the term "whale" is another example in which topic continuity has been maintained even as definitions change.[7] In pre-modern times, the term was defined to mean *large fish*, whereas, according to modern definitions, "whale" means *large sea mammal*. Nonetheless, we recognize pre-modern and modern users of the term as discussing the same topic, which can be readily confirmed by the fact that users of both definitions use the term to refer to paradigmatic instances of the kind (e.g., particular humpback whales, sperm whales, and so on).

Mathematicians can use the example of the cannonball's trajectory and zoologists can use examples of individual whales to assure themselves that they are discussing the same topic because these cases *anchor* definitions of "the derivative" and "whale". If one mathematician wants to define "the derivative" in a way that permits them to discuss the same topic as other mathematicians who have discussed the derivative, the chosen definition has to classify this case and similarly archetypical cases of the derivative in the same way their interlocutors would classify them; analogously for the case of "whale". We will refer to such cases, which serve to anchor a person's understanding of the concept whose definition is in question, as *anchor cases*. For the concept of democracy, anchor cases are those actual or hypothetical regimes that fix a person's understanding of the topic under discussion when they and others are discussing democracy. They are the cases that a person might insert into statements such as the following: "Whatever we're talking about when we talk about democracy, [anchor case] is a democracy", or "Whatever we're talking about when we talk about democracy, [anchor case] is not a democracy."

A judgment about an anchor case is "intuitive" in that it is pre-theoretical and not made on the basis of an explicit definition of democracy that one has already accepted for other reasons. Being intuitive, in this sense, does not mean it is a mere "gut reaction." What we have in mind are instead judgments that would continue to anchor one's evaluation of proposed definitions even as one continues reflecting on and scrutinizing one's beliefs about democracy and particular cases to which the concept is supposed to apply. If one scholar's technical definition of democracy classifies a particular case in a way that their interlocutor finds counterintuitive — that is, the interlocutor regards it as a misclassification of an anchor case — that is evidence that the two are talking past each other. It is evidence that, instead of disagreeing about which definition is best for discussing their shared topic of conversation, they are instead discussing different topics. It may be inconclusive evidence; it may be that, on further reflection, the interlocutor reconsiders their initial judgment about the case, or whether it should be regarded as an anchor case (more on this below). But as a rule of thumb, such conflicts between a scholar's technical definition and an interlocutor's judgments about putative anchor cases indicate violations of topic continuity.

[7] We take this example from Sawyer (2018).

First rule of thumb. We have good reason to suspect that the topic picked out by a technical definition of democracy diverges from the topic as understood by users of the background concept if the technical definition classifies anchor cases in a way that users of the background concept would find deeply counterintuitive.

Second rule of thumb. We have good reason to believe that a technical definition of democracy picks out the same topic as that picked out by the background concept if the technical definition classifies anchor cases as users of the background concept would classify those cases.

These rules focus on the classification of anchor cases in particular, not the classification of all possible cases. This is for good reason. It would be much too strong to argue that a technical definition preserves topic continuity only if its extension — i.e., the set of regimes classified as democracies according to the technical concept — is identical to the extension of the background concept. To begin with, such a requirement would effectively undermine the usefulness of technical concepts, which are introduced to make the background concept more precise; it would be surprising if such refinements left the two concepts with identical extensions. Moreover, some cases may strike people as borderline cases of democracy, and borderline cases make for poor tests of topic continuity. To see this, suppose that, when pushed, Poppy would classify post-independence Botswana as a democracy but that she doesn't have strong convictions about this. Suppose that Minnie also believes Botswana to be a borderline case but her technical definition excludes it from the set of democracies (cf. Przeworski et al., 2000, pp. 23–28). This disagreement would provide little (if any) evidence that the two are talking past each other when they use the term "democracy". It's not uncommon, after all, for people who share an understanding of a concept to nonetheless disagree about its application to borderline or obscure cases. Poppy's complaint that Minnie has changed the topic is plausible not because the extensions of the technical and background concepts diverge, but rather because Minnie's technical concept misclassifies what she regards as an *archetypical* case of democracy.

Insofar as one cares to preserve topic continuity, judgments about how to classify anchor cases should constrain choices about how to define a technical concept of democracy. Some political scientists may resist at this point, arguing that a science of democracy should seek to eliminate such apparently "subjective" elements from political scientific practice (e.g., Green and Gerber, 2003; but see Schedler, 2012). For one thing, intuitive judgments are often vague and opaque whereas proper technical definitions must be precise and transparent; it's thus unclear how intuitions about cases could provide reasons to favor one technical definition over another. A key advantage of procedural definitions as a technical concept is supposed to be their focus on procedures with publicly observable attributes, which means that it can be applied "strictly based on objective judgment and observational criteria" (Cheibub, Gandhi and Vreeland, 2010, p. 74). While some political scientists are willing to admit the relevance of intuitive judgments, those who do tend

to discount their significance. To wit, "face validity", which gauges the extent to which a measure fits with intuitive judgments about conceptual content, is typically thought to be "the weakest way to try to demonstrate construct validity" (Trochim and Donnelly, 2006, p. 130; cf. Adcock and Collier, 2001, pp. 538–40). This general skepticism about "subjective" or intuitive judgments plausibly explains why political scientists have overwhelmingly emphasized measurement-related criteria for choosing among candidate technical definitions.

To deflect some of this resistance, let's observe that using intuitive judgments about cases to anchor one's technical definition is a familiar, if unnoticed, feature of existing political scientific practice. To take a typical example, Przeworski and colleagues write,

> Most people think that Argentina under President Arturo Illia (1963–66) was democratic, even though the largest party in the country was prohibited from competing in the elections of July 1963. In turn, most agree that Mexico is not democratic, even though no party is legally banned from contesting elections. The reason is that Illia won narrowly, with 26.2 percent of votes cast, and he could have lost. In contrast, in Mexico it was certain that the Partido Revolucionario Institucional (PRI) would win. (2000, p. 17)

The authors are here appealing to generally shared judgments about the democratic status of Argentina under President Illia and the undemocratic status of Mexico; later, they point out how their minimalist definition of "democracy," which includes "ex ante uncertainty" about electoral outcomes as a necessary condition, concurs with each of these classificatory judgments. They register this concurrence as a reason to favor their definition.

For another example, consider the opening passage from Levitsky and Way's study of "competitive authoritarian regimes":

> ...Unlike single-party or military dictatorships, post-Cold War regimes in Cambodia, Kenya, Malaysia, Mexico, Nigeria, Peru, Russia, Serbia, Taiwan, Ukraine, Zimbabwe, and elsewhere were competitive in that opposition forces used democratic institutions to contest vigorously — and, on occasion, successfully — for power. *Nevertheless, they were not democratic.* Electoral manipulation, unfair media access, abuse of state resources, and varying degrees of harassment and violence skewed the playing field in favor of incumbents. In other words, competition was real but unfair. (2010, p. 1, our emphasis)

Minimal procedural definitions classify these cases as democratic in virtue of the fact that they hold competitive elections in which incumbents sometimes lose (e.g., Boix, Miller and Rosato, 2012; Przeworski et al., 2000). In the authors' view — which they seem to expect will be widely shared — minimal definitions misclassify these cases: they should be classified as undemocratic despite the presence of real electoral competition. These intuitive judgments are then used to motivate and justify a refinement of the minimal definition to include "fair competition" as a necessary

condition for democracy (see also p. 12–13).[8]

Some readers may wonder whether there can be a systematic and scientific approach to incorporating judgments about anchor cases. Given disagreement about how to classify cases, whose judgments should we take as anchors for our technical definition? Users of the background concept have a vague understanding of democracy as a topic. Should we then avoid using their (vague) intuitions about particular cases to anchor our technical definition? Should we be concerned that a practice of fitting a technical definition to judgments about particular cases gives researchers one too many degrees of freedom? Won't such a practice enable researchers to "over-fit" the definition of democracy in pursuit of a desired empirical result?

The answer to the first question is that the relevant judgments for anchoring a technical definition are the shared judgments of the participants in the conversation to which a scholar wishes to contribute. If a scholar is uninterested in contributing to, say, discussions of democracy among members of the Chinese Communist Party, then maintaining continuity with *their* topic of conversation is not a priority, and a conflict between the scholar's definition of democracy and Chinese communists' judgments about the democratic status of particular regimes is no reason at all to modify or abandon the definition. The relevant judgments are those of one's interlocutors, who include not just one's contemporaries, but anyone whom one takes to be addressing the same topic and whose claims about this topic one might wish to dispute or put to the test.

For example, when Acemoglu et al. (2019) claim that their results challenge skepticism about democracy that goes back to Plato and Aristotle, they are treating Plato and Aristotle as interlocutors in the sense that is relevant here: they are taking themselves to be advancing a claim *about democracy* that conflicts with skeptical claims *about democracy* that Plato and Aristotle defended. But there is only a conflict if all parties to this conversation are addressing the same topic. To figure out whether there is continuity of topic, Acemoglu et al. should ask whether their definition of democracy implies that regimes that Plato and Aristotle would classify as *archetypical*, *anchoring* cases of democracy are in fact not democracies. If so, that is evidence that Acemoglu et al. are in fact discussing a different topic altogether, and are thus not actually disagreeing with or putting to an empirical test the claims that Plato and Aristotle defended. Indeed, this would appear to be the case. Like other modern political scientists, Acemoglu et al. use a definition according to which a regime is democratic only if it selects its highest executive and legislative officeholders through competitive elections. On the most straightforward interpretation of this definition, classical Athens was not a democracy, as the members of the assembly and the executive council (*boule*) were not elected (the assembly was open to all citizens and the council's members were selected by lot). Yet classical Athens was obviously *the paradigm case*

[8]This is an example of what Collier and Levitsky (1997) call "precising", a "strategy of conceptual innovation" that seeks to improve a technical definition so as to resolve a "mismatch" between intuitive judgments about particular cases and an existing technical definition (p. 42). Adcock and Collier (2001) endorse a practice of refining a technical definition in response to unexpected implications of a definition for particular cases.

of democracy for ancient Greeks, one that would have anchored their understanding of the topic they were addressing with their claims about democracy.

Let us emphasize that none of this implies there is one "true" definition of democracy, or that we need to identify one unique set of classificatory judgments that should constrain all efforts to develop a technical definition of democracy. Our point is simply that if one aims to empirically study democracy in a way that can contribute to an existing conversation about democracy, then one's technical definition should be constrained by participants' judgments about how to classify the cases that anchor their shared understanding of the topic of conversation.

To the second question: It is too quick to move from the thought that users have a vague understanding of the topic to the thought that their judgments about how to classify anchor cases are vague. To say that one has a vague understanding of the background concept is to say that one is unable to precisely define the contours of the topic picked out by that concept. But that is consistent with having clear and well-considered convictions about how to classify certain central cases. Consider an illustration from a different domain. Many of us have, at best, a hazy understanding of core emotion concepts, such as love or fear; this is just to say: many of us are unable to give precise definitions of these concepts or enumerate all behavioral responses that fall under these concepts. At the same time, most of us have a perfectly clear conviction about how to classify certain particular cases: a mother feeding and gently stroking her infant is a clear manifestation of love; a child cowering alone behind a tree to hide from a wild bear is a clear manifestation of fear. Such judgments anchor our understanding of what we're talking about when we talk about love or fear. To take another example, most of us have only a vague understanding of the concept of a living organism; there are certain entities we may be unsure how to classify, such as a virus. All the same, we have no reservations judging that a cat, but not a stone, is an example of a living organism; these are anchor cases that help to fix the topic of conversation. Analogously, participants in a conversation about democracy can have similarly clear judgments about how to classify particular regimes without having a precise understanding of the concept.

What of the remaining concerns about researcher degrees of freedom and "over-fitting"? Note, first of all, that we are proposing an *additional constraint* on top of the rather minimal set of constraints on the choice of definition that scholars already recognize. Scholars would have *fewer* degrees of freedom in choosing their definitions if they recognized the criterion of topic continuity as a constraint, in addition to the constraints of logical coherence and operationalizability that they already recognize.

Nonetheless, the reader may worry that to endorse topic continuity as a criterion, with the associated practice of checking definitions against judgments about anchor cases, is just to sanction a dubious practice of choosing one's definitions so as to arrive at desired conclusions. For example, a scholar may wish (consciously or not) to arrive at the conclusion that democratization raises average income. If they then choose a definition based in part on whether it

classifies certain regimes as democracies, is there not a danger that the scholar, after observing which countries have low average incomes, might form the "anchoring judgment" that regimes like Liberia, Malawi, and Niger are not really democracies, and that "democracy" and associated terms like "competitive elections" need to be defined in such a way that they exclude these regimes?

While this is a serious concern, we think it can be largely circumvented with a simple change to existing political scientific practice: namely, that we anchor our technical definitions using judgments about how to classify *stylized hypothetical scenarios* rather than judgments about actual cases. Notice how the concern in the previous paragraph arises because the researchers' judgment that Liberia is not a democracy is influenced not only by their observations of Liberia's political institutions and procedures but also of its level of economic development and level of civil violence, dimensions that are plausibly irrelevant for defining the concept of democracy. This will be true of our judgments about how to classify *any actual case*, due to the simple fact that we cannot isolate our observations along conceptually relevant dimensions from our observations along conceptually irrelevant dimensions. As a result, judgments about how to classify actual cases provide an inefficient tool for investigating one's understanding of the concept of democracy. To wit, many political scientists (plausibly enough) treat the United States as an anchor case for our understanding of the concept of democracy: "Whatever we're talking about when we talk about democracy, it must imply that the United States has been a democracy since [insert preferred year here]". But what should we take away from such a report for the purposes of crafting a technical definition? Which features of the US are the ones that the speakers have in mind when they report this judgment? Is it the fact that the US selects its political leaders via competitive elections? Is it a belief that US citizens have an adequate measure of control over political decisions? Is it the fact that the US has a bicameral legislature, or that it is a federal state, or that it has a politically independent judiciary, or that is has a politically independent civil service? To what extent do observations along potentially extraneous dimensions influence this judgment? Since actual cases always come as a bundle of intertwined attributes, only some of which are conceptually relevant, using judgments about how to classify such cases is an ineffective way to go about defining a technical concept of democracy.

Judgments about how to classify stylized hypothetical scenarios offer are a better tool for testing and refining candidate definitions of democracy because the attributes of hypothetical scenarios are a matter of stipulation. One can specify the attributes of a hypothetical regime along those dimensions that are (perhaps only provisionally) taken to be conceptually relevant while ignoring all other dimensions. If a hypothetical scenario elicits strong convictions about its classification among participants in a conversation, then such a case is useful for anchoring one's understanding of how to define 'democracy' as it is being used in the conversation.

To make this point concrete, consider how one might use hypothetical scenarios to test whether minimalist tech-

nical definitions refine the background concept in a way that preserves topic continuity. We start by specifying some hypothetical cases; for each case, ask yourself whether you think it should be classified as a democracy and then spend a moment checking whether you would retain this judgment after careful reflection or, what amounts to the same thing, whether you can readily anticipate a line of argument that could persuade you to overturn your initial reaction.

Citizen assemblies. Suppose a small country were to replace its elected legislative bodies with legislative assemblies in which all adult citizens can participate without restriction — any citizen is permitted to cast a vote for or against any piece of legislation. The results of assembly votes are binding and enforced by the government.

Selection by lot. Suppose a country were to replace its elected legislative bodies with legislative assemblies whose members were periodically selected by lottery to serve a fixed term. All adult citizens are entered into the lottery and eligible to serve in the legislative assembly. The results of assembly votes are binding and enforced by the government.

Franchise for rich men. Suppose a country selects its political leaders by competitive elections (fill in "competitive" as you like) but only the wealthiest 50% of male citizens are permitted to vote in these elections.

Public officials for sale. Suppose a country selects is political leaders by competitive elections (again, fill in "competitive" as you like) and let it be the case that all adult citizens are permitted to vote in these elections. But suppose that it is legal in this country for candidates for office and elected officials to enter into binding *quid pro quo* contracts with private citizens and corporations, whereby they commit to taking certain actions while in office (e.g., proposing certain legislation) in exchange for monetary payment. And suppose that in this country, a private association, which receives donations from industries and wealthy private citizens, enters into contracts with all elected officials, who then act according to the association's instructions. Neither the association nor the elected officials it contracts with interfere in elections, which continue to be contested and free and fair. Incumbents are regularly removed from power and replaced with challengers, but they, too, then enter into contracts with the private association.

Keeping in mind your considered judgments about how to classify these cases,[9] observe now that most minimalist definitions of 'democracy' (e.g., Boix, Miller and Rosato, 2012; Przeworski et al., 2000; Schumpeter, 1942) classify the first two cases as nondemocracies while the third and fourth are classified as democracies. We submit that most typical users of the background concept — including many political scientists — will find these implications of mini-

[9]One possible reaction is that we haven't given enough information for you to make a confident judgment. In addition, perhaps you'd want to know, for example, whether minorities' rights are protected. Such a reaction demonstrates how using hypothetical scenarios can be useful for not only identifying which observations along the relevant dimensions qualify a regime as democratic, but also for figuring out which aspects of a regime are conceptually relevant.

malist definitions deeply counterintuitive. Given our two rules of thumb, these counterintuitive implications are strong evidence that the topic picked out by minimalist definitions is not the same as the topic many people have in mind when they are thinking, speaking, writing, and inquiring about democracy. This isn't to say that minimalist definitions are "false" or "incorrect", whatever that might mean. It's just to say that they likely fail to pick out what many people — including many scholars — have in mind when they use the term 'democracy'.[10]

Let's take stock of the advantages of using hypothetical scenarios as test cases. To begin with, notice that hypothetical scenarios focus our attention on attributes that are put forward as conceptually relevant. For example, most minimalist definitions propose that competitive elections and perhaps a minimally extensive franchise are individually necessary and jointly sufficient to qualify as a democracy (note that Przeworski et al. and Schumpeter do not put forward any suffrage requirement). The first two scenarios directly test the claim that competitive elections are necessary; the third and fourth cases directly test the joint sufficiency claim. Moreover, our hypothetical scenarios isolate these attributes for our consideration and thereby avoid conflating our judgments about democracy with our judgments about potentially irrelevant attributes such as economic development or civil violence. Finally, because these hypothetical scenarios are not drawn from any dataset we might use to empirically test hypotheses about democracy, we need not worry that using judgments about these scenarios to anchor our technical definition will produce empirical measures that are "over-fit" to the data and thus biased in favor of finding results that confirm antecedently-held convictions.

Some political scientists might object here that these counterfactual cases are irrelevant given their scientific aims. "Empirical political science aims to study the political world as we observe it," they might reply; "these scenarios are beyond empirical observation and, thus, inappropriate anchors for a technical definition." It's true, of course, that we wish to study the countries of the real world and not imagined utopias. But one thing that our growing sophistication about causal inference has taught us is that is that our efforts to explain our observations of actual countries depends on our ability to describe what would happen, counterfactually, if they were different in certain respects. Would North Korea grow richer if it became a democracy? Would the United States grow poorer if it became an autocracy? Political scientists wish to answer questions about the counterfactual versions of actual countries both because, like scientists generally, they are curious about the world and seek satisfying causal explanations of what they observe in it, and because the answers to causal questions have implications for policy and other important practical consequences. Given the centrality of causal inference in political science, how a definition of 'democracy' and its operationalized measure direct us to classify regimes in counterfactual scenarios is no less important than the classifications they yield for actual regimes.

[10] We recognize that defenders of the minimalist approach might just as well respond by rejecting topic continuity as a desideratum for technical definitions. We will address this response at the end of the paper.

4 Concluding remarks: Should we change the topic?

Democracy is a central topic of public political discourse. Many political scientists conduct empirical research on democracy as a means to make informed contributions to these conversations. Against this background, we have argued for the following modest proposition: If a scholar wishes to conduct research that contributes to ongoing conversations about democracy, then she must adopt a definition of 'democracy' that picks out the topic of those conversations as understood by participants in those conversations. For practical purposes, we have proposed that scholars test candidate definitions against the criterion of topic continuity by considering how well they concur with participants' judgments about how to classify particular regimes, emphasizing judgments about stylized hypothetical scenarios as opposed to the regimes we observe in the actual world. We recognize that our argument will raise more questions than it answers but that's the point: to show that conceptual content matters when choosing a definition of 'democracy' and to initiate a discussion about the criteria we might use for assessing these choices.

Our proposal faces a skeptical objection that we have thus far left unaddressed. Our discussion of topic continuity and its application to political scientific practice presupposes that political scientists should aim to preserve topic continuity. But, after seeing the potential implications of accepting this criterion, one might reject topic continuity as a desideratum rather than adjust political scientific practice to meet its demands. Moreover, one might do so with good reason: "Typical conversations about democracy are vague and confused. Indeed, it is rare to find a conversation about democracy that has a well-defined topic with which to maintain continuity. In view of this confusion, political scientists should focus on steering conversations about democracy toward conversations with well-defined topics, even if doing so means changing the topic of conversation."[11]

We think there's something important in this objection. More specifically, we agree that it can be appropriate for scholars to engage in some "conceptual engineering" (Cappelen, 2018) with the aim of facilitating more productive conversations. Alas, we can't provide any general principles for determining when a change of topic is preferable to preserving topic continuity. Instead, we register three notes of caution for those who are tempted to dismiss topic continuity out of hand in favor of topic change.

First, we should be aware that, once we have changed the topic, we may no longer be in a position to contribute to existing conversations on the basis of our empirical research, for the reasons we highlighted in section 2. This might be obvious, but it's worth emphasizing that there might well be a trade-off between, on the one hand, conducting scientifically credible research on democracy and, on the other hand, contributing (perhaps as a corrective) to broader conversations about democracy. We must be alert to the realistic possibility that we cannot do both.

This naturally suggests a second caution. There is good reason to think that changing the topic should be a last

[11] Przeworski (2010, esp. chaps. 1–2) provides a forceful expression of this thought.

resort and not the default starting point. As we have noted several times above, we should hesitate to conclude that the topic of an ongoing conversation about democracy is confused, equivocal, or otherwise ill-defined just because participants in that conversation do not have a precise definition of 'democracy' at hand. We should thus hesitate to change the topic of a conversation before we have made a sympathetic effort to reconstruct it. This means starting from the assumption that the conversation we wish to join has a well-defined topic, abandoning that assumption only once it has become clear that the conversation is genuinely confused and, as a result, unproductive.

Third, if a change of topic is required, we can probably do better than to hijack a term that is already in wide circulation. Continuing to use the term 'democracy' while giving it a definition that deviates from ordinary usage is, at a minimum, an odd strategy for gaining clarity. Why not instead choose a more precise label for one's technical concept, such as "polyarchy" (Dahl, 1971) or "competitive electoral regime"? If one insists on using the label 'democracy' for a technical concept that diverges sharply from the background concept, this suggests that one wishes to be seen as discussing the same topic as others who use the term 'democracy', which implies some concern for preserving topic continuity.

References

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

Achen, Christopher H. and Larry M. Bartels. 2016. *Democracy for Realists: Why Elections Do Not Produce Responsive Government*. Princeton and Oxford: Princeton University Press.

Adcock, Robert and David Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3):529–46.

Boix, Carles, Michael Miller and Sebastian Rosato. 2012. "A Complete Data Set of Political Regimes, 1800–2007." *Comparative Political Studies* 46(12):1523–54.

Cappelen, Herman. 2018. *Fixing Language: An Essay on Conceptual Engineering*. Oxford: Oxford University Press.

Carnap, Rudolph. 1950. *The Logical Foundations of Probability*. Chicago: University of Chicago Press.

Cheibub, José Antonio, Jennifer Gandhi and James Raymond Vreeland. 2010. "Democracy and Dicatatorship Revisited." *Public Choice* 143:67–101.

Collier, David, Fernando Daniel Hidalgo and Andra Olivia Maciuceanu. 2006. "Essentially Contested Concepts: Debates and Applications." *Journal of Political Ideologies* 11(3):211–46.

Collier, David and Robert Adcock. 1999. "Democracy and Dichotomies: A Pragmatic Approach to Choices About Concepts." *Annual Review of Political Science* 2:537–65.

Collier, David and Steven Levitsky. 1997. "Democracy With Adjectives: Conceptual Innovation in Comparative Research." *World Politics* 49(3):430–51.

Coppedge, Michael. 2012. *Democratization and Research Methods*. New York: Cambridge University Press.

Coppedge, Michael, John Gerring, David Altman, Michael Bernhard, Steven Fish, Allen Hicken, Matthew Koenig, Staffan Lindberg, Kelly McMann, Pamela Paxton, Holli A. Semetko, Svend-Erik Skaaning, Jeffrey Staton and Jan Teorell. 2011. "Conceptualizing and Measuring Democracy: A New Approach." *Perspectives on Politics* 9(2):247–67.

Coppedge, Michael, Staffan Lindberg, Svend-Erik Skaaning and Jan Teorell. 2016. "Measuring High Level Democratic Principles Using the V-Dem Data." *International Political Science Review* 37(5):580–93.

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

de Schweinitz, Karl. 1959. "Industrialization, Labor Controls, and Democracy." *Economic Development and Cultural Change* 7(4):385–404.

Dunn, John. 2004. *Democracy: A History*. New York: Atlantic Monthly Press.

Elkins, Zachary. 2000. "Gradations of Democracy? Empirical Tests of Alternative Conceptualizations." *American Journal of Political Science* 44(2):287–94.

Freedom House. 2020. "Freedom in the World." https://freedomhouse.org/report/freedom-world.

Gallie, W. B. 1956. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56:167–98.

Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton and Oxford: Princeton University Press.

Green, Donald P. and Alan Gerber. 2003. "The Underprovision of Experiments in Political Science." *The ANNALS of the American Academy of Political and Social Sciences* 589(1):94–112.

Levitsky, Steven and Lucan Way. 2010. *Competitive Authoritarianism: Hybrid Regimes After the Cold War*. New York: Cambridge University Press.

Marshall, Monty, Ted Robert Gurr and Keith Jaggers. 2018. *Polity IV Project: Dataset Users Manual*. Center for Systemic Peace.

Munck, Gerardo. 2009. *Measuring Democracy: A Bridge Between Scholarship and Politics*. Baltimore: The Johns Hopkins University Press.

Ober, Josiah. 2017. *Demopolis: Democracy Before Liberalism in Theory and Practice*. Cambridge: Cambridge University Press.

Paxton, Pamela. 2000. "Women's Suffrage in the Measurement of Democracy: Problems of Operationalization." *Studies in Comparative International Development* 35(3):92–111.

Przeworski, Adam. 2010. *Democracy and the Limits of Self-Government*. Cambridge: Cambridge University Press.

Przeworski, Adam, Michael E. Alvarez, José Antonio Cheibub and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. Cambridge: Cambridge University Press.

Riker, William H. 1982. *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice*. Long Grove, IL: Waveland Press.

Sawyer, Sarah. 2018. "The Importance of Concepts." *Proceedings of the Aristotelian Society* 118(2):127–47.

Schedler, Andreas. 2012. "Judgment and Measurement in Political Science." *Perspectives on Politics* 10(1):21–36.

Schumpeter, Joseph. 1942 [2008]. *Capitalism, Socialism, and Democracy*. 3rd ed. New York: Harper.

Seawright, Jason and David Collier. 2012. "Rival Strategies of Validation: Tools for Evaluating Measures of Democracy." *Comparative Political Studies* 47(1):111–38.

Treier, Shawn and Simon Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–217.

Trochim, William and James P. Donnelly. 2006. *The Research Methods Knowledge Base*. Cengage Learning.