DOI: 10.1111/jopp.12307

ORIGINAL ARTICLE

Is One More Powerful with Numbers on One's Side?

Sean Ingham[1] | **Niko Kolodny**[2]

[1]Political Science, University of California, San Diego

[2]Philosophy, University of California, Berkeley

Correspondence: University of California San Diego, Department of Political Science, Social Sciences Building 301, 9500 Gilman Drive, #0521, La Jolla, CA 92093-0521, USA
Email: singham@ucsd.edu

Suppose that in a jurisdiction there are 2,000,001 white voters and 1,000,000 black voters, all of whom enjoy equally weighted votes. The question of white supremacy is routinely put to a majority-rule plebiscite. In each such plebiscite, all white voters vote *yes* for white supremacy and all black voters vote *no*. This has been going on as long as anyone can remember, and it will continue for as long as anyone can foresee. This is a paradigm of a persistent minority, to which, intuitively, each black voter has an objection.

What is their objection? One answer is that black voters don't get their preferences satisfied. Another answer is that black voters are oppressed by the eventuating policies of white supremacy. Yet another answer considers the white majority *as a group*. As a group, they have greater power to determine the outcome than have the blacks as a group. Indeed, the white majority as a group is always decisive. In the last plebiscite, all the whites voted for white supremacy and it passed; and if all the whites had voted against, it would have failed. By contrast, the black minority as a group is never decisive. They voted against and it passed; and if any assemblage of them had voted *yes*, it would still have passed.

It seems more difficult, however, to say that members of the minority as individuals have less voting power or power to determine the outcome than members of the majority as individuals. Accordingly, some say they may be unlucky, but are no less powerful than members of the majority, as power is not the same as prospects for success.[1] Recently, however, Arash Abizadeh has argued ingeniously that members

[1]Brian Barry, "Is it better to be powerful or lucky? Part 2," *Political Studies*, 28 (1980), 338–52; Keith Dowding, *Rational Choice and Political Power: The Original Text with Two New Chapters* (Bristol: Bristol University Press, 2019), p. 62; Keith Dowding, "Resources, power and systematic luck: a response to Barry," *Politics, Philosophy and Economics*, 2 (2003), 305–22. See also Peter Morriss, *Power: A Philosophical Analysis*, 2nd edn (Manchester: Manchester University Press, 2002), pp. xxxv–ix; and Annick Laruelle and Federico Valenciano, "Assessing success and decisiveness in voting situations," *Social Choice and Welfare*, 24 (2005), 171–97, at p. 77.

of persistent minorities are in fact less powerful than members of persistent majorities, provided further background conditions are met.[2] Abizadeh argues that the bad luck diagnosis ignores "the power of numbers":

> This is the agential social power enjoyed by one who finds herself among others disposed to act in concert with her to effect the realization of her intentions, preferences, or objectives, even if she has no power over them. It is the power enjoyed by members of a persistent majority under majoritarian decision-making: if over a broad range of issues over time, conditional on a voter supporting an alternative, a majority would also support it, then he has greater decision-making power than a voter for whom there is no such correlation of voting intentions—even if he exercises no power over the persistent majority.[3]

In this article, we raise a number of doubts about Abizadeh's suggestion that the power-of-numbers thesis can vindicate the thought that members of the minority as individuals have less voting power and thereby account for their objection to belonging to a persistent minority. Perhaps the most serious doubt is that while Abizadeh correctly holds that voting power must be assessed in part by counterfactualizing on votes—by asking what would have happened if a voter had voted otherwise than he in fact did—he does not counterfactualize in the right way.

I | A POSTERIORI VOTING POWER

We cannot vindicate the thought that members of the minority have less voting power if we assume an a priori measure of voting power that abstracts from information about the distribution of political preferences and its causes. An example is the Banzhaf (or Penrose–Banzhaf) measure of voting power, according to which a voter's power is her probability of casting a decisive vote if all other voters vote independently and with equal probability for either alternative.[4] (A voter's vote in favor of (or against) a measure is *decisive* if the measure passes (fails), but would have failed

---

[2] Arash Abizadeh, "The power of numbers: on agential power-with-others without power-over-others," *Philosophy and Public Affairs*, 49 (2021), 290–318; Arash Abizadeh, "Counter-majoritarian democracy: persistent minorities, federalism, and the power of numbers," *American Political Science Review*, 115 (2021), 742–56.

[3] Abizadeh, "The power of numbers." p. 291.

[4] L. S. Penrose, "The elementary statistics of majority voting," *Journal of the Royal Statistical Society*, 109 (1946), 53–7; John F. Banzhaf, "Weighted voting doesn't work: a mathematical analysis," *Rutgers Law Review*, 19 (1965), 317–43; John F. Banzhaf, "Multi-member electoral districts. do they violate the 'one man, one vote' principle?", *Yale Law Journal*, 75 (1966), 1309–38; Harry Brighouse, "Egalitarianism and equal availability of political influence," *Journal of Political Philosophy*, 4 (1996), 118–41, at pp. 119, 139; Morriss, *Power*, pp. 154–66; Niko Kolodny, "Rule over none ii: social equality and the justification of democracy," *Philosophy and Public Affairs*, 42 (2014), 287–336, at pp. 320–2. See Dan Felsenthal and Moshé Machover, *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes* (Cheltenham: Edward Elgar, 1998), for technical discussion of the Penrose–Banzhaf measure and other measures of voting power.

(passed) if the voter had instead voted against (in favor).) An a priori measure of voting power will not register any differences between members of the persistent minority and anyone else, because it ignores, by construction, the facts about the distribution of preferences and its causes, in virtue of which some voters qualify as members of a persistent minority.

Abizadeh therefore seeks to vindicate the power-of-numbers thesis with an a posteriori measure: power is calculated on the basis of, rather than in abstraction from, information about how social structure influences the distribution of political preferences. This measure of power also differs from the Penrose–Banzhaf measure in making voting power a function of the degree to which the voter can expect her actions to be “efficacious,” where even non-decisive, redundant votes for the winning alternative count as partially efficacious.

So let us consider voting power in the a posteriori context, in which we take into account information about how voters are likely to vote. The information on the basis of which power is measured may be more or less predictive of voting behavior. At the limit, knowledge of a person's structural position, together with knowledge of how they will vote, fully predicts how everyone else will vote. For the sake of simplicity, we start by assuming that such fully predictive information is available and is the appropriate basis for measuring voting power. This assumption is obviously unrealistic, but it simplifies the discussion by relieving us of the need to calculate probabilities, and it is an innocent simplification because it does not prejudice the assessment of the power-of-numbers thesis. It would be bizarre to argue that a member of a persistent minority has an objection only when there is genuine uncertainty about the future distribution of votes, but not when the future distribution, and her status as a political minority, is certain. We will, in any case, relax this assumption in due course.

We assume, in particular, that the relevant information perfectly predicts the scenario described above: in every plebiscite, the 2,000,001 white voters are certain to vote in favor of white supremacy and the 1,000,000 black voters are certain to vote against. Our question is whether a given white voter can be said to have significantly more power than a given black voter when voting power is measured on the basis of this information.

II | THE POWER OF NUMBERS

In this situation, everyone's probability of casting a decisive vote is zero. Thus, if a voter's power corresponds to the probability of casting a decisive vote (an assumption behind the Penrose–Banzhaf measure), then every voter, white or black, has the same amount of power.

What if, following Abizadeh, we grant that a member of the winning side can enjoy partial efficacy, even if they are not decisive? When everyone casts equally weighted votes, as will be true in all our examples, we can assume the degree of partial efficacy is a function of the size of the winning coalition (relative to the size of the electorate). At one end of the spectrum is the case of someone who votes for the winning side along with every other member of the electorate; their vote has some partial efficacy, but not as much as it would have if the relative size of the winning coalition were smaller. At the other end is the case of the fully decisive voter who votes for the winning side and is part of a minimal decisive coalition (for example, a bare majority under simple majority rule).

In our example, when a white voter votes with the winning coalition, in favor of white supremacy, then his vote has some partial efficacy, namely the degree of partial efficacy that corresponds to voting with 2,000,000 other voters, in an electorate of 3,000,001 voters, for the winning alternative. If the white voter were to vote against white supremacy, the efficacy of his vote would be zero. If a black voter were to vote with the winning coalition, in favor of white supremacy, then her vote would have some partial efficacy, namely the degree of partial efficacy that corresponds to voting with 2,000,001 other voters, in an electorate of 3,000,001 voters, for the winning alternative. When the black voter votes against white supremacy, the efficacy of her vote would be zero. Their partial efficacy scores are virtually identical.

These claims rest on the tacit assumption that if a voter were to vote differently from how they will actually vote, everyone else would still vote the same as they are going to vote in actual fact. One must reject this assumption if one wishes to argue for the power-of-numbers thesis.

Some might object to the tacit assumption on the following grounds. Given our stipulations about the case, all white voters are *certain* to vote in the same way. So, it follows that if a given white voter were to vote *against* white supremacy rather than *for*, then all other white voters would also vote against, too. But that does *not* follow. Suppose newspapers A and B always report the same events; when the scandal breaks, they are each certain to report it. It does not follow that if A were to refrain from reporting the scandal, then B would refrain as well. Probabilistic dependence does not imply counterfactual dependence.[5]

One would reject the tacit assumption with good reason if one voter's decision about how to vote causally influenced other voters' decisions. But let us assume that no one's vote has any causal influence over anyone else's vote. If you like, assume everyone votes secretly and simultaneously. The power-of-numbers thesis is not supposed to rest on voters' abilities to influence each other.

A different reason one might reject the assumption is if one thought that measurements of voting power ought to reflect counterfactualizing not only the voter's action, but also the past events that causally influence the voter's action as well as other voters' behavior. One might reason as follows. The example assumes that all white voters always vote as a block and all black voters always vote as a block. This

[5] Put more abstractly, $Pr(A|B) = p$ and $Pr(A|\neg B) = q$ does not imply that if B were to occur, then A would occur with probability p, and if B were not occur, then A would occur with probability q.

pattern of correlation can only hold in virtue of underlying social-structural causes that have one (fully determining) impact on white voters' preferences and an opposite (fully determining) effect on black voters' preferences. If a given white voter were to vote against the proposed ballot measure, it would have to be because the underlying social-structural influences caused him and all other white voters to be opposed. Thus if a given white voter were to vote against, they would find themselves on the winning side and would enjoy some partial efficacy, just as they enjoy some partial efficacy when they vote in favor; by contrast, if a given black voter were to vote in favor, she would still be on the losing side and her vote would still be inefficacious, just as her actual vote against is inefficacious. Thus the white voter enjoys greater power than the black voter.

This reasoning involves "backtracking" counterfactuals: the assumption is that when one counterfactualizes over a white voter's action and looks for the closest possible world in which he voted no, one should include worlds with *different histories up until the time of that vote*.[6] This allows one to say that the white voter would be in the majority even if he were to vote against the proposal, provided that the closest possible world in which he votes *no* is a world with a different history, one in which some common cause led all the other whites to prefer *no* and so to vote *no* as well.

In the next section, we will argue that the counterfactuals that enter into the measurement of power should not be interpreted in a way that permits backtracking.[7] But even if we allow backtracking counterfactuals, it may not help the power-of-numbers thesis. Suppose that the closest world in which a given white voter votes no is a world in which the normal link between social-structural causes and political preferences is broken for him alone, but not for other white voters (he regularly converses about racial justice with a colleague, and the closest possible world in which he votes against white supremacy is one in which these conversations induce a moral epiphany, severing the causal link that continues to make other white voters' political preferences a deterministic function of their position in the racial hierarchy). Then the crucial counterfactual would be false: if he were to vote against white supremacy, the other whites would still vote in favor, and his situation is not appreciably different from the situation of any black voter. Nor would it help much if the closest world is one in which he and, say, a thousand others had racial justice epiphanies.

One worry about the power-of-numbers thesis is that, offhand, it's not clear why the closest possible world in which a given white voter votes no (even among those worlds with different histories) is one in which the social-structural causes

[6] David Lewis, "Counterfactual dependence and time's arrow," *Nous*, 13 (1979), 455–76.

[7] For discussions of the counterfactuals that enter into the analysis of power, see Alvin Goldman, "Toward a theory of social power," *Philosophical Studies*, 23 (1972), 221–68, at p. 232; and Morriss, *Power*, pp. 71–9, 216–21. To our knowledge, no previous authors take up questions about backtracking counterfactuals in the analysis of power, although Morriss's discussion of counterfactuals is relevant.

that influence all other white voters' preferences are different. So even if we permit backtracking, it is not clear why, in our example, it would be true that if a given white voter were to vote against white supremacy, then all other white voters (or at least a subset that constitutes a majority) would also vote *no*. Another worry about the power-of-numbers thesis, as a diagnosis of our intuitive objection to a paradigm case of a persistent minority, is that the intuitive objection does not seem to depend on which of these backtracking counterfactuals is correct, but the power-of-numbers thesis does.

III | REJECTING BACKTRACKING COUNTERFACTUALS

In any event, this backtracking counterfactual isn't the relevant conditional for assessing agential power generally or voting power in particular. If we are measuring the power of a given white voter on the day of the election, we should ask what would happen if he were to vote *no*, *holding fixed the actual history of the world up until the time he votes*. The following cases illustrate the importance of excluding backtracking from the analysis of agential power.

Jones sometimes prays for rain after eating breakfast. To be precise, he prays on all and only those mornings when, while eating breakfast and reading the newspaper, he notices that the weather forecast predicts rain. Thus, conditional on praying for rain in the morning, the probability that it will rain in the afternoon is high. And conditional on not praying, the probability that it will rain is low (because, conditional on not praying, there was no rain in the forecast, and forecasting the weather in his area of southern California is not hard). Jones does not have the power to make it rain, of course. We can explain this fact with normal, non-backtracking counterfactuals, but not with backtracking counterfactuals. On any given morning when Jones prays, the probability of rain is high, but would remain high even if he were not to pray for rain; and on any given morning when he does not pray, the probability of rain is low, but would remain low even if he were to pray. If agential power is measured with backtracking counterfactuals, however, Jones does have the power to make it rain. With backtracking, we have: on any given morning, if Jones were to pray, then (it would be because rain was in the forecast and so) it would probably rain, and if he were not to pray, then (it would be because rain was not in the forecast and so) it would probably not rain.

Backtracking counterfactuals are no more plausible a basis for ascriptions of power when we move from the single-agent example to a voting situation. A boss threatens to fire anyone who votes for a proposal to unionize. As a result, each worker knows that if she votes for the proposal to unionize and it fails, she will lose her job; otherwise she will keep it. Needless to say, no one sticks her neck out for the union. It seems false to say, and a cruel joke in the mouth of the boss, that each worker has the power to effect, with the others, the passage of the unionization proposal.[8] And yet it is plausibly true if we assess a worker's power using backtracking counterfactuals. The closest world in which she votes for the union might be one in which the boss never made the threat and no one faces the prospect of losing their job for voting for it, in which case everyone else would also vote for the union and the unionization proposal would pass. Thus if she were to vote in favor of the unionization proposal, then—backtracking and counterfactualizing the boss's past behavior—it would pass, in which case she would have effected, with others, the passage of the proposal.

Or suppose an American football coach is deciding whether to go for a touchdown on fourth down or instead have his kicker kick a field goal. One consideration when he makes such decisions is the wind: if the wind is too strong, the kicker will not have the power to make the field goal, and so in such circumstances the coach decides to go for the touchdown. Suppose in actual fact the winds are strong, so he decides to go for the touchdown. After they fail, his critics say: "Coach made the wrong call. The kicker *did* have the power to make a field goal. In the closest possible world in which he kicked, he kicked because coach told him to, and coach told him to because the wind conditions were favorable, and since they were favorable, kicker made the field goal." We have taken Monday-morning quarterbacking to a whole new level.

IV | PARTIALLY PREDICTIVE INFORMATION

We have so far considered an example in which the measurement of voting power is made on the basis of information that is fully predictive of everyone's voting behavior. Under this assumption, the differences in voting power between an individual member of the persistent majority and an individual member of the persistent minority were negligible. If one relaxes this assumption and measures power instead on the basis of information that is only partially predictive of the distribution of votes, there are cases in which members of a structural minority have *more* power in expectation than members of a structural majority—provided, again, that one does not measure power with backtracking counterfactuals.

Voter 1 belongs to a structural minority, voters 2, 3, 4, and 5 to a structural majority. For each plebiscite t, there is a social-structural variable S_t that influences the voting behavior of all five voters. Either $S_t = 0$ or $S_t = 1$. If $S_t = 1$, then the probability that voters 1, 2, and 3 will support the measure and voters 4 and 5 will oppose it

[8]We are *not* claiming here that the individual worker does not really enjoy the power to vote for the unionization proposal because the boss can retaliate, nor are we denying that the workers *as a group* have the power to effect the passage of the proposal. We are denying that an *individual* worker has the power to effect, with the others, the passage of the unionization proposal, and pointing out that this false claim would be true if backtracking counterfactuals were used in the analysis of voting power.

is p; and the probability that voter 1 supports it, but voters 2, 3, 4, and 5 oppose it is $1-p$. If $S_t = 0$, then the probability that voters 1, 2, and 3 will oppose the measure and voters 4 and 5 will support it is p; and the probability that voter 1 opposes it, but voters 2, 3, 4, and 5 support it is $1-p$.[9] Assume $0 < p < 1/2$. Most of the time (a fraction $1-p > 1/2$ of the time), voter 1 finds herself in the minority and voters 2, 3, 4, and 5 find themselves in the majority.

We now show that if $p > 3/23$, then in any given plebiscite, voter 1 has more power than voter 5, despite the fact that voter 5 has numbers on his side. This conclusion holds for every plebiscite t, whatever the history preceding the plebiscite and, in particular, whatever the realized value of S_t. The argument for the conclusion abstracts from all facts about the voters' situation other than (1) the voting rule and (2) the social structure and probabilistic dependence that it induces in voting behavior.

Our argument uses the measure of partial efficacy proposed by Abizadeh.[10] According to this measure, in an electorate of five voters, where everyone casts equally weighted votes, a vote for the winning alternative has an efficacy score of 1 if it is one of three votes cast for the winning alternative, a score of 3/4 if it is one of four votes cast for the winning alternative, and an efficacy score of 3/5 if it is one of five votes cast for the winning alternative. A vote cast for the losing alternative has no efficacy. Perhaps a different measure of efficacy could be proposed. But so long as the measure holds that a successful voter's efficacy is a function $f(x)$ of the fraction x of the electorate voting for the winning alternative, and (in the 5-voter case) the function satisfies $0 < f(5/5) < f(4/5) < f(3/5)$, one can construct an example similar to the one below (see Appendix).

Choose any plebiscite t. Either voter 1 supports the measure or opposes it. Assume, without loss of generality, that voter 1 supports the measure. This implies $S_t = 1$. Thus the probability that voters 2 and 3 also support it, while voters 4 and 5 oppose it is p, and the probability that voters 2, 3, 4, and 5 all oppose it is $1-p$. Thus,

(1) If voter 1 were to vote for the measure, her vote would be fully decisive with probability p and inefficacious with probability $1-p$. Expected efficacy: $1p + 0(1-p) = p$.

[9]The case is described with deliberate abstraction so as to avoid the impression that our conclusion depends on anything other than the assumed voting rule and the social-structural influences on voters' preferences. But if one would like a more concrete story to accompany it, one can imagine that the structural categories are racial categories (voter 1 is a racial minority), and one can think of S_t as a description of whether *passing* the measure on the ballot would materially benefit voter 1 at the expense of voters 4 and 5 ($S_t = 1$), or whether *rejecting* the measure would materially benefit voter 1 at the expense of voters 4 and 5 ($S_t = 0$). Voters 4 and 5 always have preferences opposed to voter 1's, while voters 2 and 3 usually side with voters 4 and 5 when, as usually happens (with probability $1-p$), the policy question has been framed in racialized terms, but sometimes (with probability p) side with voter 1 if it has been framed in non-racialized terms.

[10]Arash Abizadeh, "A recursive measure of voting power with partial decisiveness or efficacy," *Journal of Politics*, 84 (2022), 1652–66.

(2) If voter 1 were to vote against the measure, her vote would be fully decisive with probability p and would have a partial efficacy score of 3/5 (a member of a five-voter supermajority) with probability $1-p$. Expected efficacy: $1p+3/5(1-p)=3/5+2/5p$.

The critical assumption here is that each counterfactual must be assessed without backtracking. Conditional on its being the case that voter 1 supports the measure, voters 4 and 5 are sure to vote against, and voters 2 and 3 will vote for the measure with probability p and vote against it with probability $1-p$. This describes the way in which the other voters' behavior is probabilistically dependent on voter 1's behavior. But it does not imply counterfactual dependence (unless one allows for backtracking counterfactuals). Their voting behavior is counterfactually independent of what voter 1 chooses to do, so even if voter 1 were to vote against the measure, it would still be the case that voters 4 and 5 are sure to vote against and voters 2 and 3 vote in favor with probability p.[11]

Now turn to voter 5. As we assumed (without loss of generality) that voter 1 supports the measure, it follows that voters 4 and 5 oppose it, and, with probability $1-p$, voters 2 and 3 also oppose it, and with probability p, voters 2 and 3 support it. Thus,

(3) If voter 5 were to vote for the measure, his vote would have a partial efficacy score of 3/4 (he would belong to the four-voter supermajority comprising 1, 2, 3, and himself) with probability p, and his vote would be inefficacious with probability $1-p$. Expected efficacy: $3/4p+0(1-p)=3/4p$.
(4) If voter 5 were to vote against the measure, his vote would have a partial efficacy score of 3/4 (he would belong to the four-voter supermajority comprising 2, 3, 4, and himself) with probability $1-p$, and his vote would be inefficacious with probability p. Expected efficacy: $0p+3/4(1-p)=3/4(1-p)$.

Again, the critical assumption is that the behavior of the other voters is counterfactually independent of what voter 5 chooses to do.

Now notice that for any nonzero value of p, voter 1's action of voting *yes* has a greater expected efficacy than voter 5's action of voting *yes*. And voter 1's action of voting *no* has a greater expected efficacy than voter 5's action of voting *no* so long as $3/5+2/5p>3/4(1-p)$, which is equivalent to $p>3/23$. Thus, provided $p>3/23$, each of voter 1's actions is more efficacious in expectation than the

[11] See n. 5, above. That $Pr(A|B)=0$ and $Pr(A|\neg B)=1$ does *not* imply that if B were not to occur, then A would occur (here $A=$"voters 4 and 5 vote in favor" and $B=$"voter 1 votes in favor"). The probability that Jones has diabetes, conditional on its being the case that he injects himself with insulin every day, is high; the probability that he has diabetes, conditional on its not being the case that he injects himself with insulin every day, is low. It does *not* follow that if Jones were, counterfactually, to refrain from injecting himself with insulin, then the probability that he would have diabetes would be low. He has diabetes, and he will have it whether he injects himself with insulin or not. Plug in $A=$"it rains" and $B=$"Jones does not pray for rain" for another intuition primer.

corresponding action of voter 5, so voter 1 must be deemed more powerful than voter 5. But voter 1 belongs to a structural minority and voter 5 to a structural majority. Voter 1 is in the minority, and voter 5 is in the majority, a fraction $1 - p > 1/2$ of the time, due to their positions in the social structure and its influence on voting behavior.

We assumed without loss of generality that we were dealing with a plebiscite in which voter 1 supports the measure. But we would have reached the same conclusion if we had instead assumed that voter 1 opposed it. Thus, our conclusion does not depend on the actual history leading up to the plebiscite (in particular on the actual realization of the social-structural variable S_t). It relied only on the assumption that plebiscites are decided by simple majority rule and that the social structural variable S_t induces probabilistic dependence in voting behavior in the manner described.

The intuition for the result is simple. Voter 1 can have a higher probability of being in the minority than voter 5, even though she also has a higher probability of casting a decisive vote. If the latter probability is sufficiently high (in our example, if $p > 3/23$), the conclusion will be that voter 1 is more powerful. The critical assumption is that we assess counterfactual suppositions about voter 1's actions without counterfactualizing the features of her history (the realized value of S_t) that fix the probability distribution over other voters' behavior and thereby fix the probability that her vote will be decisive.

V | COMPENSATING POLITICAL ECCENTRICS

In several places, Abizadeh cautions us that we cannot define persistent majorities or minorities in terms of the voters' policy preferences:

> The problem is that appealing directly to the content of other voters' preferences, but not the power-assessee's, is insufficient for explaining the inequality faced by a persistent-minority member: if the persistent-majority and persistent-minority groups are defined in terms of members' substantive preferences, then counterfactualizing over the power-assessee's preferences implies abstracting from her membership of the persistent minority. Her power inequality would consequently fail to register.
>
> … A persistent minority must therefore be defined and picked out in terms of its members' position in other social structures—such as language, ethnicity, race, class, or geography—which position induces a correlation in preferences.[12]

[12]Abizadeh, "Counter-majoritarian democracy," p. 750.

This claim about how persistent minorities must be defined is important, because it is all that Abizadeh invokes to avoid the result that the power of numbers thesis "absurdly requires compensating political eccentrics on democratic-equality grounds." It is his explanation for why "If libertarians or Bolsheviks are persistently outnumbered, democratic equality does not call for formal-procedural inequalities to compensate them."[13]

What exactly is the problem with defining "persistent minority" in terms of actual preferences? The worry seems to be that, so defined, there is something incoherent about asking "what if the persistent minority had different preferences?", just as there might seem to be something incoherent about asking, "what if every student in the class got a grade better than the median grade?". But neither question is incoherent provided we take "the persistent minority" and "the median student" to be "rigid designators," rather than terms whose referents vary across possible worlds. If, in the actual world, the median grade is a B, then the second question is asking what if every student in the class got better than a B. And if you are the one person who, in actual fact, belongs to the minority in every election, then the first question is asking what would happen if *you* had different preferences and voted differently.

The justification for defining "persistent minority" in terms of structural position cannot be that this is the only definition that permits a coherent formulation of counterfactual statements about the preferences and actions of the persistent minority. We can intelligibly ask what would happen if a libertarian or Bolshevik, who in *actual* fact persistently finds himself in the minority, were counterfactually to vote differently from how he will actually vote, and compare that with what would happen if someone who in actual fact is in the mainstream were to vote differently from how she will actually vote. If the answer—in virtue of backtracking counterfactuals—is that the Bolshevik would still be in the minority, and the mainstream voter would still be in the majority, then the power-of-numbers thesis implies there is an inequality in power.

If one rejects the use of backtracking counterfactuals to answer this question, then the libertarian and the Bolshevik will have no less power than anyone else, at least not in virtue of being persistently outnumbered. But the argument for the power-of-numbers thesis needs backtracking counterfactuals, so the question is whether there is any justification for excluding them when the question concerns libertarians and Bolsheviks, but including them when it concerns structural minorities.

Backtracking counterfactuals need not be any less plausible in these other "non-structural" cases. Suppose there are 1,000,000 voters who consistently vote in whatever way the Bolshevik party leaders advocate, and 2,000,001 voters who consistently vote against whatever the Bolshevik party leaders advocate. Take any election, and

[13] Ibid., p. 754.

assume (without loss of generality) that the pro-Bolshevik voters are all going to vote *yes*, the anti-Bolshevik 2,000,001 all *no*. If a pro-Bolshevik were to vote *no*, it could only be because the party vanguard had instructed them to vote *no*, in which case all the anti-Bolshevik voters would vote *yes*, and our voter would still be in the minority. That judgment is no less reasonable than the judgment that if a black voter were to vote *no*, in a society polarized along racial lines, it could only be because of social-structural conditions that induce all white voters to vote *yes*, such that the black voter would still be in the minority. We have argued that backtracking counterfactuals should be excluded from analyses of power in all cases, but if there were any justification for including them in the case of structural minorities, it is unclear why the justification would not carry over to non-structural minorities. And then one gets the absurd conclusion that the power-of-numbers thesis, if it justifies compensating the power deficits of racial or religious minorities, also justifies compensating libertarians and Bolsheviks.

It is also perhaps worth noting here that once backtracking counterfactuals are admitted, the majority or minority need not be *persistent* in order for there to be power-of-numbers inequalities. Imagine a one-hit-wonder at the momentary apex of their popularity. Wonder's fans, for a day, form a majority. A vote is taken on that day whether to declare the day Wonder-appreciation-day. Wonder's fans will vote for whatever Wonder decides. If Wonder says *yes*, then the fans will vote *yes*; if Wonder, in modesty, says *no*, then the fans will vote *no*. Each of the fans in the majority enjoys power of numbers that each nonfan does not. But none of this will persist.

VI | ABSTRACTING VERSUS COUNTERFACTUALIZING

Even if an agent's power should not *in general* be assessed on the basis of backtracking counterfactuals, could there nonetheless be a justification for incorporating them into a specific kind of voting power measure? We anticipate an erroneous argument for doing so that proceeds from the reasonable assumption that the desired measure of voting power ought to *abstract* from features of a voter's actual situation, such as the actual history preceding the vote, and draws the errant conclusion that voting power ought to be assessed by counterfactualizing these features of the voter's situation.

The assumption that a measure of voting power ought to abstract from certain features of a voter's situation is reasonable. The Penrose–Banzhaf measure, for example, only purports to register the power a voter enjoys in virtue of the voting rule.[14] That is why it abstracts from all other features of a voter's situation (such as how the

[14] Dan Felsenthal and Moshé Machover, "Myths and meaning of voting power: comments on a symposium," *Journal of Theoretical Politics*, 13 (2001), 81–97, at p. 90; Felsenthal and Machover, *The Measurement of Voting Power*, pp. 20–1; Laruelle and Valenciano, "Assessing success and decisiveness in voting situations," p. 183. Dan Felsenthal, Dennis Leech, Christian List, and Moshé Machover, "In defence of voting power analysis: responses to Albert," *European Union Politics*, 4 (2003), 473–97, at p. 476.

voter can expect other voters to behave) that are plausibly relevant to a *comprehensive* assessment of agential power. Analogously, if a measure only aims to capture the power a voter enjoys in virtue of (1) the voting rule and (2) their position in the social structure, then it ought to abstract from the voter's actual circumstances.

We anticipate an erroneous argument that illicitly infers from this reasonable assumption the mistaken conclusion that the measure of voting power must "counterfactualize" the voter's circumstances. The mistaken reasoning might go something like the following:

> One cannot say what would happen if the voter were to vote *yes*, unless one makes some assumption or other about her circumstances, but one's conclusion should not be sensitive to what her circumstances actually are. So one must ask, "if she were to vote *yes*, what would her circumstances probably be, in virtue of the history that would have preceded her decision to vote *yes*?". One then considers all these probable circumstances, her efficacy in each, and weights these scores by the probability with which such circumstances would have occurred in virtue of the history that would have preceded a decision to vote *yes*. That is, one uses backtracking counterfactuals to compute the probabilities of different configurations of votes for the other voters.

The second proposition does not follow from the first, however. There is a logical gap between the reasonable supposition that a measure of voting power ought to abstract from a voter's actual circumstances and the conclusion that it ought to do so by asking backtracking counterfactual questions. To see that the conclusion does not follow, it suffices to establish the existence of a measure that abstracts from the voter's actual circumstances, but eschews backtracking counterfactuals. Consider the following approach.

For each of the possible circumstances the voter might find herself in, compute the power she would have in those circumstances with normal, non-backtracking counterfactuals in just the way we did for the paradigm case and the example in Section IV. For example, there is the scenario in which a worker is voting on whether to unionize, and her boss has issued a threat to her and all the other workers; and the scenario in which they are voting, but the boss has not issued a threat. In the example from Section IV, there is the scenario in which $S_t = 1$ and another scenario in which $S_t = 0$. Compute the voter's power in each of these scenarios (just as we have already done for the first of these two scenarios in Section IV).[15] Now average over the measurements of the voter's power taken in all these

[15]Because we want the measure of power to make its determinations on the basis of (1) the voting rule and (2) the underlying social structure and the probabilistic dependence it induces among voters' preferences, we want to hold these two features of the voter's situation constant as we consider the different possible circumstances she is in. That is, each of the possible scenarios that one considers as part of this computation should be a scenario in which the underlying social structure and voting rule are the same; the scenarios are individuated along other dimensions (e.g., the content of the proposal they are voting on and whether the boss has issued a threat, or the value of the variable S_t).

possible circumstances, weighting according to the probability of each circumstance. The result does not depend on which of the possible scenarios is the actual one. The power-of-numbers thesis will still fail to hold if one considers it a thesis about "aggregate" voting power in this sense, rather than a thesis about voting power in a particular circumstance. Since no voter derives special power from having numbers on her side in any particular circumstance, no special aggregate power will emerge when one computes a weighted average of her power across the possible circumstances.

Abstracting is not the same as counterfactualizing. One can measure power in a way that abstracts from a voter's actual circumstances and their history—in a way that makes measurements of power independent of a voter's actual circumstances and history—without counterfactualizing them. Because the use of backtracking counterfactuals in the measurement of voting power does not follow from the (reasonable) assumption that a voting power measure should abstract from the voter's circumstances, a proponent of the power-of-numbers thesis must find some other argument for it. We are not sure what it could be.

VII | IS A POSTERIORI POWER THE CURRENCY OF POLITICAL EQUALITY?

Provided that a posteriori power is measured using standard, non-backtracking counterfactuals, facts about who has more or less of it will not closely track membership in structural minorities and majorities. In our paradigm case, members of structural minorities and majorities have all but equal a posteriori power. In the case from Section IV, members of structural minorities have greater a posteriori power than members of majorities.

In so far as this article's argument underscores the vagaries of a posteriori power distributions, it may make one skeptical that democracy could require equality in this dimension, or even an approximation to equality. Once we move from a priori to a posteriori measures of voting power, there is no reason to expect an equal scheme of basic political rights to produce an equal distribution of (a posteriori) voting power. Just the opposite. Almost all political institutions, except those that have been calibrated to the actual social-structural conditions just so, will produce asymmetries in a posteriori voting power. And when the social-structural conditions underlying the distribution of preferences change, political institutions will also need to change if they are to preserve equality.

Moreover, the kind of changes in social-structural conditions that could undermine equal a posteriori voting power are not normally considered threats to democratic equality. So-called swing voters will enjoy more a posteriori voting power than voters who are reliably on one side of a partisan divide. But even when this voting behavior reflects structural causes—the ways in which race, class, and other social-structural identity categories pull some citizens towards one party, some towards the other, and leave some conflicted and open to appeals from both parties—the fact that the rigidly partisan voter has less a posteriori voting power than the swing voter is not plausibly considered a breach of democratic principles.

One weird implication of the view is that some citizens could violate political equality, as though they were denying someone an equally weighted vote, merely by "freeing" themselves of social-structural influences and thereby altering the probability distribution of preferences in a way that reduces their efficacy. As a simple illustration, suppose an electorate of three voters, 1, 2, and 3, faces a series of binary votes, and due to social-structural causes they always vote as a block and with equal probability for either alternative. Whether one uses backtracking counterfactuals or not, everyone has equal a posteriori power.[16]

But now imagine that voter 1 "frees" himself of the causal influence of the background social structure and comes to vote independently of voters 2 and 3, who still vote as a block, and everyone still votes with equal probability for either alternative. Now voter 1 has less power than either of the other two. Each of his actions has an expected efficacy of 1/6: if he were to vote *yes*, then with probability 1/2, the block would vote *yes* and his efficacy would be 1/3; and with probability 1/2, the block would vote *no* and his efficacy would be zero. The expected efficacy of voting *no* is calculated analogously and also comes out to 1/6.[17] But each of voter 2's actions has an expected efficacy of 2/3, assuming we use backtracking counterfactuals in the calculation:[18] if she were to vote *yes*, then with probability 1/2, voters 1 and 3 would each vote *yes*, in which case her efficacy is 1/3; and with probability 1/2, voter 1 would vote *no*, but 3 would vote *yes*, in which case her efficacy is 1. Thus her expected efficacy of voting *yes* is $1/6 + 1/2 = 2/3$. The expected efficacy of voting *no* is calculated analogously and also comes out to 2/3.

Some might find this result counterintuitive, because they do not think one can become less powerful in virtue of increasing one's ability to form political opinions independently of social-structural influences. But that is not our point. Our point is

[16]With backtracking counterfactuals, we would say that for each voter, each of his actions would have a partial efficacy of 1/3, with probability 1—because if he were to vote *yes*, the other two voters would also vote *yes*, and if he were to vote *no*, they would also vote *no*—so, for each voter, each action has an expected partial efficacy of 1/3. Without backtracking counterfactuals, we would say that for each voter, one of his actions (whichever one is the action he will actually choose) will have an efficacy of 1/3 with probability 1, thus an expected efficacy of 1/3, while the other would have zero efficacy (because if he voted differently from the other two, his action would be inefficacious). Either way we go with the counterfactuals, no power inequalities exist.

[17] These calculations come out the same with or without backtracking, because the assumption of voter 1's probabilistic independence from voters 2 and 3 renders the point moot. The calculations for voter 2 in the next sentence assume backtracking, but if we excluded backtracking, we would still get the conclusion that voter 2 has more power. Without backtracking, we would say that whichever action voter 2 will actually choose has an expected efficacy of 2/3 (calculated as in the main text). The other, counterfactual action, which differs from voter 3's with certainty, has an expected efficacy of 1/2: with probability 1/2, it agrees with voter 1's, but differs from 3's, and is therefore decisive; and with probability 1/2, it differs from both 1's and 3's and has no efficacy. Thus each of her actions has greater expected efficacy (either 2/3 or 1/2) than either of voter 1's (1/6).

[18] See n. 17 above.

rather that if one has reasons to try to preserve political equality, but political equality is understood in terms of equal a posteriori voting power, then the value of political equality gives voter 1 reasons not to cultivate an ability to form political opinions independently of social-structural influences. *That* implication seems strange.

These considerations make us doubt that a posteriori voting power is the currency of democratic equality.

VIII | CONCLUSION

The power-of-numbers thesis is supposed to vindicate the intuition that a member of a persistent minority has a complaint against majoritarian democracy. If the thesis is true, then majoritarian democracy does not in general respect a principle of democratic equality, and the member of the persistent minority can ground their objection by appeal to this principle. We have noted several grounds for doubts about the power-of-numbers thesis, most importantly its reliance on backtracking counterfactuals.

Supposing one accepts our critique of the power-of-numbers thesis, how else might one try to ground the objection? We noted several alternatives at the outset. One is that the real objection is not to an inequality in voting power, but instead to an inequality in the satisfaction of preferences: namely, that black voters' preferences go unsatisfied vote after vote, whereas white voters' preferences are reliably satisfied. Another alternative is that, again, the real objection is not to an inequality in voting power, but instead to the substantive inequality that results from the policies voted into effect: that black voters are subordinated by a regime of white supremacy. A final alternative recognizes an inequality in voting power, but at the level of groups rather than individuals. Even if no individual member of the structural minority has appreciably less power than any individual member of the structural majority, the latter *group* has more power than the former. Perhaps the distribution of power across social groups takes on normative significance—above and beyond the distribution of power across individuals—in societies with histories of group-based injustice.[19] We do not have space in this article, however, to develop these suggestions and evaluate their merits. Those are tasks for other work.

ACKNOWLEDGMENTS
The authors would like to thank Zeynep Pamuk, two anonymous referees, and especially Arash Abizadeh for helpful comments on previous drafts of the article.

FUNDING INFORMATION
None relevant.

CONFLICT OF INTEREST STATEMENT
There are no potential conflicts of interest relevant to this article.

[19]See Niko Kolodny, *The Pecking Order* (Cambridge, MA: Harvard University Press, 2023), pp. 155–6.

DATA AVAILABILITY STATEMENT

All relevant data are included in the article.

ETHICS STATEMENT

The authors declare human ethics approval was not needed for this study.

How to cite this article: Ingham, Sean and Niko Kolodny. 2023. "Is One More Powerful With Numbers on One's Side?." *Journal of Political Philosophy* 00, 1–18. https://doi.org/10.1111/jopp.12307.

APPENDIX

Here we explain why the argument from Section IV generalizes to any reasonable measure of partial efficacy. We restrict attention to the five-voter example for the sake of simplicity.

Let f_k refer to the efficacy of a voter who votes for the winning alternative when a total of k voters do so, in an electorate of five voters. In the main text, we adopted Abizadeh's proposed measure of partial efficacy, according to which $f_5 = \frac{3}{5}, f_4 = \frac{4}{5}$, and $f_3 = 1$. We now relax this assumption and assume merely that $0 < f_5 < f_4 < f_3$. The argument from Section IV goes through so long as the measure of partial efficacy satisfies these inequalities. Substituting these terms for particular numerical values that were assumed in Section IV, and reasoning about the example in the same way (in particular, with normal, non-backtracking counterfactuals), we have:

- If voter 1 were to vote for the measure, her efficacy would be f_3 with probability p and inefficacious with probability $1 - p$. Expected efficacy: pf_3.
- If voter 1 were to vote against the measure, her vote would be fully decisive with probability p and would have a partial efficacy score of f_5 (a member of a five-voter supermajority) with probability $1 - p$. Expected efficacy: $pf_3 + (1 - p)f_5$.
- If voter 5 were to vote for the measure, his vote would have a partial efficacy score of f_4 (he would belong to the four-voter supermajority comprising 1, 2, 3, and himself) with probability p, and his vote would be inefficacious with probability $1 - p$. Expected efficacy: pf_4.
- If voter 5 were to vote against the measure, his vote would have a partial efficacy score of f_4 (he would belong to the four-voter supermajority comprising 2, 3, 4, and himself) with probability $1 - p$, and his vote would be inefficacious with probability p. Expected efficacy: $(1 - p)f_4$.

Thus voter 1's action of voting *yes* is more efficacious than voter 5's action of voting *yes* for any nonzero value of p, because $f_3 > f_4$ by assumption. Voter 1's action of voting *no* is more efficacious than voter 5's action of voting *no* if and only if

$$pf_3 + (1-p)f_5 > (1-p)f_4,$$

which is equivalent to

$$p > \frac{f_4 - f_5}{f_3 + f_4 - f_5}$$

Our assumptions ensure that the right-hand side is a number strictly greater than 0 and strictly less than 1/2. So if p is greater than this threshold, but less than 1/2, voter 1 is more powerful than voter 5—each of voter 1's actions has greater expected efficacy than the corresponding action for voter 5—even though voter 1 is more often than not in the minority and voter 5 is more often than not in the majority, due to the structural causal influences on their and other voters' preferences.

DOI: 10.1111/jopp.12307DOI: 10.1111/jopp.12307