Social choice and popular control

Sean Ingham
Department of Political Science, University of Georgia, USA

Abstract
In democracies citizens are supposed to have some control over the general direction of policy. According to a pretheoretical interpretation of this idea, the people have control if elections and other democratic institutions compel officials to do what the people want, or what the majority want. This interpretation of popular control fits uncomfortably with insights from social choice theory; some commentators—Riker, most famously—have argued that these insights should make us abandon the idea of popular rule as traditionally understood. This article presents a formal theory of popular control that responds to the challenge from social choice theory. It makes precise a sense in which majorities may be said to have control even if the majority preference relation has an empty core. And it presents a simple game-theoretic model to illustrate how majorities can exercise control in this specified sense, even when incumbents are engaged in purely re-distributive policymaking and the majority rule core is empty.

Keywords
Accountability; popular control; popular will; Riker; social choice theory

1. Introduction

In a democracy, citizens should have some control over how they are governed. They should have tools for holding their leaders accountable and influencing decisions on public policy. Policymakers and government officials are the citizens' public servants. They are entrusted with the power to set policy, but citizens should retain an indirect, controlling influence over the general direction of policy, if not the day-to-day decisions on particulars.

Corresponding author:
Sean Ingham, Department of Political Science, University of Georgia, 104 Baldwin Street, Athens, GA 30602, USA.
Email: ingham@uga.edu

This ideal of popular control over government and political elites is a familiar aspiration for democracy. What does it mean? One answer is that the people have control if they can compel public officials to do what they want. The people are like a principal who delegates responsibilities to an agent of its choosing, but whose power to confer rewards or penalties—such as re-election or removal from office—gives it control over the agent's behavior. Principal–agent models find frequent employment in studies of electoral accountability and control (Barro, 1973; Besley, 2006; Fearon, 1999; Ferejohn, 1986; Maskin and Tirole, 2004).

These models have produced valuable insights, but they have an important limitation. They typically assume away the problems of preference aggregation studied in social choice theory. In place of a multitude of citizens with diverse commitments and preferences, principal–agent models typically posit a representative voter, such as the median voter. But in some contexts there may not be any median voters; references to 'what the majority want' or 'what the people want' may not be meaningful. Consider, for example, incumbent officeholders who can distribute public revenues between themselves ('rents') and self-interested voters. No matter the choice of distribution, there will be an alternative that an overwhelming majority of citizens would prefer instead. The officeholder cannot help but act contrary to the wishes of some majority. None of these divisions is 'what the people want' or 'what the majority want'. Following Riker (1982), one might conclude from such examples, and the results in social choice theory that generalize them, that interpreting public officials as the agents of the people, who should be under their collective control, is not the right way to think about democracy.

The puzzle here is not simply about whether elections enable a heterogeneous citizenry to control its elected officeholders. In these cases it is not clear even how to define 'popular control', much less whether elections establish popular control. What does popular control mean if there is no popular will?

If we wish to defend popular control as an intelligible and feasible ideal for democracies, then we need an explanation of what it means for citizens to have popular control. The explanation should be compatible with findings in social choice theory; it should not evade the problems of preference aggregation with the stipulation of a representative voter or homogeneous citizenry. And it should facilitate theoretical and empirical inquiry into the conditions under which elections enable citizens to control public officials.

To that end the present article proposes a definition of 'popular control' and applies it to a model of elections. The definition makes precise a sense in which majorities can control policy even if policymakers inevitably act contrary to the wishes of some majority. Relative to this definition, social choice theory shows that majorities cannot always have *perfect* control. But majorities can always have imperfect control, even in the paradoxical situations in which every possible choice of policy runs counter to the wishes of some majority. Moreover, we can define a sense in which majorities might have *maximal* control, even if this level falls short of perfect control.

After introducing these definitions, I apply them to a simple principal–agent model of elections. I consider a case in which an incumbent officeholder chooses

how to divide a fixed surplus among self-interested voters and himself; thus, no matter his choice, a majority will prefer some other division. With the definitions, we can meaningfully ask whether citizens can use the threat of removing the incumbent from office as a mechanism for controlling his behavior. Under the assumptions of the model, the answer is that they can: there are equilibria in which majorities have maximal or nearly maximal control over the incumbent's choice, provided he places enough value on re-election. In these equilibria, they can compel him to divide all or almost all of the surplus equally among the citizenry.

In the next section I explain briefly why previous literature in normative and formal political theory does not settle the questions about popular control that this article addresses. Section 3 introduces the definitions of 'popular control' and *maximal majority control* and establishes two desirable implications of establishing popular control in the stipulated sense. Section 4 introduces the model of elections and describes the equilibria in which majorities have maximal or nearly maximal control over the incumbent. Section 5 concludes; all proofs are contained in the Appendix.

2. Background

When the policy choice is how to divide a fixed surplus among self-interested citizens, then every policy disappoints some majority. Social choice theory shows that this state of affairs arises under more general conditions; one cannot chalk it up to the narrow assumptions of self-interested citizens and a zero-sum policy choice. If policies differ continuously along multiple dimensions, then under weak assumptions about voters' preferences the 'core' of the majority preference relation will almost always be empty (Le Breton, 1987; McKelvey, 1979; McKelvey and Schofield, 1987; Plott, 1967; Saari, 1997).[1] When it is empty, it does not matter which policy the government enacts; there will be an alternative choice that a majority of citizens prefer. Aggregating the pairwise preferences of different majorities does not yield a well-defined 'majority will'.

Riker argued that social choice theory discredited the idea of a 'popular will' and therefore undermined any conception of democracy in which the purpose of elections is to ensure that public officials respect the will of the people. Elections, he argued, give citizens the power to remove leaders from office, but nothing more. In particular, it would be a mistake to think of public officials as agents of the people and elections as the device by which a principal disciplines its agents. The threat of being removed from office does not give officeholders incentives to do as the people want—the lesson of social choice theory is that expressions like 'what the people want' are empty. Elections do not make 'officials [...] responsive to some imaginary popular will' (Riker, 1982: 244–245). As Riker interprets democracy, there is no meaningful sense in which elections give citizens control over public officials.

Given the attention that normative and formal political theorists have given Riker's argument and social choice theory more generally, one might wonder what more there is to be said on the subject. But social choice theory raises multiple puzzles, and commentators have tended to focus on questions distinct from the

question about popular control that I address here. One finds discussions of questions about the stability of democratic decision-making; how elections and legislative institutions aggregate preferences into policy choices; the conditions for legitimate decision-making when there is no 'best' social choice (Patty and Penn, 2014); whether the problematic configurations of preferences—those that produce 'cycles' or an empty core—occur in actual politics (Mackie, 2003).[2] These studies address questions different from the question I address, even if the same body of formal results provides the impetus for all of these inquiries. Despite all of the discussion of Riker's argument among normative political theorists (Cohen, 1986; Coleman and Ferejohn, 1986; Dryzek and List, 2003; Mackie, 2003; Miller, 1992), no-one has responded with a definition of 'popular control' or 'popular rule' that retains a connection with individual preferences and renders the possibility of popular control robust to whether the core of the majority preference relation is empty.

Pettit's recent statement of a republican theory of democracy has an account of popular control at its center (Pettit, 2013). On his account, the people have control over government policy if their influence constrains the government to comply with commonly endorsed policymaking norms, that is, with norms that *all* citizens accept. One may worry that if a policymaking norm is so uncontroversial that it strikes *everyone* as a reasonable restriction on government, then it is likely not much of a restriction at all, and one may accordingly worry that this notion of popular control is too weak. But, while Pettit's position might be vulnerable to this criticism, it does not run into the aggregation problems studied in social choice theory.[3] The premise of these problems is, put loosely, that different individuals have different points of view—concerning policy or higher order questions like which norms should constrain the government. While a full defense of the conception of popular control I formalize below would require a more careful discussion of Pettit's important contribution, these brief remarks suffice, I hope, to indicate the main point of divergence.

List and Pettit (2011) present a theory of group agency that draws on the formal literature concerning judgment aggregation. Their theory is meant to explain how group agency, in general, is possible, and not how an electorate or citizenry, in particular, can acquire the status of a group agent with control over its government. Nonetheless, some might wonder whether the theory could be deployed for this purpose.

But in my view, a theory of popular control does not need a theory of group agency.[4] The account I present below is not an account of group agency, rival to List and Pettit's, but a description of one sense in which citizens may be said to control their government—even if they do not, collectively, have the status of an agent.

One might look for a definition of 'popular control' in formal political theory, in particular formal models that bill themselves as models of electoral control. But principal–agent models that address questions of electoral control and electoral accountability typically adopt the 'representative voter' approach described in the introduction (Barro, 1973; Besley, 2006; Fearon, 1999; Ferejohn, 1986; Maskin and

Tirole, 2004). Even when they admit heterogeneous voters, the inquiry still proceeds without an analysis of what it *means* for the people—or majorities—to control public officials in the event of an empty majority rule core.

Another place to look would be models of how elections aggregate citizens' preferences when the core is empty. I have in mind models that assume probabilistic voting or assume that voters respond not simply to candidates' policy promises but also to 'valence' characteristics such as perceived competence (Ansolabehere and Snyder, 2000; Coughlin, 1992; Schofield, 2003, 2004). Even though these models are not typically formulated in order to answer questions about popular control per se, they might nonetheless explain how popular control is consistent with an empty majority rule core.

These models give us ways of conceptualizing popular influence—how citizens' preferences causally influence policy outcomes. But it is useful to distinguish between *influence* and *control* (Pettit, 2013). If we take control as a particular form of influence—as I argue we should in the next section—then the patterns of popular influence that these models describe may or may not amount to what one should regard as popular control. We need a criterion for determining when popular influence counts as the kind of popular control that one might want democratic institutions to deliver.

In the next section I formalize an interpretation of popular control and give a definition of what it means for majorities to have maximal control over incumbents' behavior. With the definition in hand, we can meaningfully ask whether representative democracy gives citizens not only influence but popular control over elected officials and policy.

3. Definitions

A good definition of 'popular control' should harmonize with any firm intuitions we have about the concept, else it will be misleading to advertise it as a definition of 'popular control'. The one special case in which we have firm intuitions is the case where all citizens have identical preferences. In this case, 'what the people want' is unambiguous. Intuitively speaking, they have control if their influence over public officials leads to the outcome they want.

For example, suppose that an elected officeholder can divide some fixed surplus, $S > 0$, between himself and n citizens. But he is constrained to give citizens equal shares of whatever surplus remains after he appropriates a rent, $0 \leqslant r \leqslant S$. If citizens can influence the officeholder's choice, and their influence keeps him from extracting any positive rent, then their influence should register as popular control on any good definition.

Compelling officeholders to do exactly as citizens want should suffice for popular control, but it should not be necessary. If citizens' influence causes the officeholder to refrain from extracting more than a very small but positive amount of rent, then, intuitively speaking, this influence should again register as popular control, albeit less than perfect popular control.

An informal definition that fits with these intuitive judgments is that control is influence that leads to a *satisfactory* outcome. The citizens have control in this hypothetical scenario because their influence leads, if not to the perfect outcome of $r = 0$, then to an outcome $r > 0$ that is perfectly acceptable. If $r > 0$ is small enough, then it is plausibly described as 'acceptable' or 'satisfactory' in this sense: while there are alternative outcomes that all citizens would prefer, there is no alternative that each citizen would be willing to pay more than a small amount—specifically, r/n units of the surplus—in order to obtain.

Note that on this definition, control is not equivalent to influence; it is rather a particular kind of influence. It differs from other species of influence, including what one might describe as 'responsiveness' to citizens' preferences, often taken as a hallmark of democracy (Dahl, 1989). There are of course different ways to define 'responsiveness', but its ordinary meaning suggests influence in a particular direction. If we do not build into its definition some reference to the acceptability of the results of popular influence, then responsiveness—in contrast with control—will not be well-described as a key feature of democracy. To see why, suppose an office-holder chooses a policy in $[0, \infty)$, citizens all have Euclidean preferences with an identical ideal point of x_c, and the officeholder's incentives are such that if their ideal point is x_c, then he chooses $x(x_c) = \alpha x_c$, where $\alpha > 0$ is large. If citizens' common ideal point shifts from x_c to x'_c, then policy shifts from αx_c to $\alpha x'_c$. Thus, the officeholder's choice is positively responsive to citizens' ideal point. But this responsiveness is neither popular control—policy may differ drastically from what citizens want—nor an example of what democracy is meant to establish. In a democracy, citizens should be able to influence policy, and their influence should produce policies that they find acceptable.

Citizens do not in general agree on how to rank all of a public official's possible choices. How should we extend the definition to handle disagreement? One idea would be to define 'popular control' as popular influence that induces policy outcomes that *everyone* finds acceptable (Pettit, 2013). But if citizens strongly disagree about the merits of policies, then there may be no policy that proves acceptable to everyone.

This interpretation of popular control runs into an additional difficulty. Asking that a policy be acceptable to *everyone* means that it be acceptable to every bigot, eccentric crank, and ideological extremist, however repugnant, outlandish or self-serving their policy preferences may be.[5]

A more compelling and arguably more democratic aspiration is for a system of popular influence that ensures that policies are acceptable to majorities. Popular control means control by majorities; 'in all countries where the people reign, the majority rules in the name of the people' (de Tocqueville, 2000: 173). We can make this idea more precise with an extension of the informal definition given above.

Consider again the rent-seeking officeholder who chooses how to divide a surplus, S, between himself and the citizenry, but who is unconstrained in how he divides the citizens' portion among the citizens. His possible choices include every possible division of S into $n + 1$ shares (n citizens plus himself). If a majority can influence his choice and their influence induces him to allocate the entire surplus to

members of this majority, then their influence should register as control on any good definition. Being able to compel the officeholder to give them everything is sufficient for control.

But, intuitively speaking, it is not necessary. If the officeholder appropriates a slice of the surplus $r > 0$, but the majority's influence ensures that r is small and induces him to divide the remaining $S - r$ equally among members of this majority, then their influence should again register as control, albeit less than perfect control. They are able to secure an *acceptable* policy; there is no alternative that they all consider *significantly* better. If r is small enough, this outcome is plausibly described as 'acceptable' or 'satisfactory' to the majority in the same sense suggested above: while there are alternatives that each member of the majority would prefer, there is no alternative that each would be willing to pay more than a small amount—specifically, r/m units of the surplus, where m is size of the majority—in order to obtain.

To formalize and generalize this idea, let X be the set of policies. For $i = 1,2,\ldots n$, let $v_i : X \rightarrow \mathbf{R}$ represent citizen i's policy preferences, where differences in v_i are assumed to be measurable in terms of i's willingness to forego units of a numeraire good.[6] For example, $v_i(x) - v_i(y) = m$ means that i would be willing to pay up to m units of the numeraire in order to obtain x in place of y. Say that $x \in X$ is *unacceptable* to a coalition $L \subset N$ if, for some $y \in X$, and some 'largish' $\varepsilon > 0$

$$v_i(y) - v_i(x) > \varepsilon, \ \forall i \in L$$

and *acceptable* otherwise. In words, a policy x is unacceptable to a group if there is another policy y that they all consider significantly better, i.e. so much better that they would each gladly give up ε units of the numeraire in order to obtain it.

Of course, what counts as 'significantly better' than a given policy—whether ε is significantly large—is vague. Whether the policy is 'acceptable' is therefore also vague. A less vague, albeit more cumbersome way to put it would be to say that x is acceptable to L *relative to* $\varepsilon > 0$ if and only if there is no $y \in X$ such that $v_i(y) - v_i(x) > \varepsilon$ for all $i \in L$.

We can now put these definitions together with the informal definition of 'control' as influence that leads to an acceptable or satisfactory outcome. Say that L has *some control* if its influence leads to an outcome that is acceptable relative to some $\varepsilon > 0$; and the smaller the ε, relative to which the outcome of its influence is acceptable, the *more control* it has.

Relative to $\varepsilon = 0$, no policy will be acceptable to every majority coalition, unless the core of the majority preference relation is nonempty. But even with an empty core, there will be some $\delta > 0$ and some policy x such that x is acceptable (relative to δ) to every majority. If this policy results from the aggregated influence of the various majorities, then, as defined here, these majorities all have some control. The smaller the δ term, relative to which this condition holds, the more control they have.

Some additional notation will be handy in stating subsequent definitions and results. Define the binary relations $R_\varepsilon, P_\varepsilon \subset X \times X$ as follows: for $x, y \in X$, $(y, x) \in P_\varepsilon$ if and only if for some majority coalition $L \subset N$

$$v_i(y) - v_i(x) > \varepsilon, \ \forall i \in L$$

and $(x, y) \in R_\varepsilon$ if $(y, x) \notin P_\varepsilon$. Write $xR_\varepsilon y$ for $(x, y) \in R_\varepsilon$ and $xP_\varepsilon y$ for $(x, y) \in P_\varepsilon$. So $x \in X$ is acceptable (relative to ε) to every majority if and only if

$$x \in M(R_\varepsilon, X) := \{z \in X | zR_\varepsilon y, \ \forall y \in X\}$$

As a solution concept, the set $M(R_\varepsilon, X)$ resembles what is sometimes called the ε-core (Bräuninger, 2007; Tovey, 2010, 2011).[7] Feld and Grofman (1991) describe it as the 'benefit-of-the-doubt' set, since it has a natural interpretation in terms of the set of policies that an incumbent could safely choose if his constituents were prepared to give him some benefit of the doubt in contests with a challenger. The next section explores a similar interpretation with a simple game-theoretic model of elections.

If majorities have more control, the smaller the ε term relative to which the outcome of their influence is acceptable to them, then there is a particular value of ε that corresponds to what we might naturally define as 'maximal' majority control over an officeholder's policy choice.

Definition 1. *Say that majorities have* maximal control *over the incumbent's policy choice if their influence leads him to choose a policy in the set* $M(R_\varepsilon, X)$, *where*

$$\varepsilon := \min\{\tilde{\varepsilon} > 0 | M(R_{\tilde{\varepsilon}}, X) \neq \varnothing\}$$

and they have nearly maximal control *if their influence leads him to choose a policy in the set* $M(R_\delta, X)$, *where* $\delta > \varepsilon$ *is nearly equal to* ε.

Social choice theory reveals the lack of precision in the familiar idea that democracy involves rule by the majority, or respect for the 'will of the majority': there are multiple majorities, and adding up the pairwise preferences of different majorities need not produce a coherent 'will'. But we can now appreciate one way of meeting the challenge from social choice theory and patching up the logical holes in this idea. There is a sense in which majorities (plural) can all have control: their aggregated influence over the decision may ensure a decision that no majority finds unacceptable, i.e. a decision that some majority may consider inferior to an alternative, but no majority considers *greatly* inferior to any alternative. Instead of associating popular control with implementing the 'will of the majority', majoritarians can associate popular control with giving majorities as much control as possible, i.e. giving majorities 'maximal control'. Popular control is intelligible whether or not the core is nonempty.

One might fear the consequences of giving majorities maximal control in the stipulated sense, just as de Tocqueville feared the 'tyranny of the majority'. But this fear is not a reason for objecting to the definition of 'maximal majority control'. On the contrary, if the prospect of giving majorities maximal control did not elicit fear in liberal critics of majoritarianism, then it would be a sign that the definition strayed too far from its mark. Provided we have not distorted the concept beyond all recognition, critics of majoritarianism should find something to dislike in it.

I conclude this section by recording two implications of majorities having maximal control. The implications should render this interpretation of popular control

more appealing to majoritarians and their critics alike. The first implication (Proposition 1) supplies a partial defense against the criticism that maximal majority control would threaten the interests of minorities. In some settings there is reason to expect *equitable* policies when majorities have maximal control. If a policy *x* treats a minority badly, then there should be many alternatives that members of this minority greatly prefer to *x*. All else equal, that makes it easier to assemble a majority that includes this minority and whose members all strongly prefer some alternative *y* to *x*—in which case *x* is not the outcome that one would associate with maximal majority control. The next proposition confirms this intuition in the 'divide-the-dollar' setting, where each person has self-interested preferences over the possible divisions of a fixed surplus.

Proposition 1. *Let* $X = \left\{ \mathbf{x} \in \mathbf{R}^n_+ \mid \sum_i x_i = 1 \right\}$ *and* $v_i(\mathbf{x}) = x_i$ *for* $i = 1,2,\dots n$, *with* n *odd. Then*

$$\varepsilon := \frac{n-1}{n(n+1)} = \min\{\delta > 0 \mid M(R_\delta, X) \neq \varnothing\}$$

and $M(R_\varepsilon, X) = \{(1/n,\dots, 1/n)\}$.

Figure 1 illustrates the proposition for $n = 3$. The 'satisfaction threshold' is $\varepsilon = 1/6$. Relative to this threshold, the equal division $\mathbf{e} := (1/3, 1/3, 1/3)$ is acceptable to every majority coalition. There is no $\mathbf{z} \in X$ such that $\mathbf{z} P_\varepsilon \mathbf{e}$, because there is no way to divide the amount allocated to one voter between the other two such that both receive more than 1/2. As the figure indicates, for any threshold $\delta < 1/6$ the adjusted indifference sets would intersect and there would be some $\mathbf{z} \in X$ such that $\mathbf{z} P_\delta \mathbf{e}$.

A second desirable implication of majorities having maximal control in the stipulated sense is that the resulting policy is guaranteed to satisfy a weak Pareto efficiency criterion. Say that *y strongly Pareto dominates x* if $v_i(y) > v_i(x)$ for all $i \in N$.
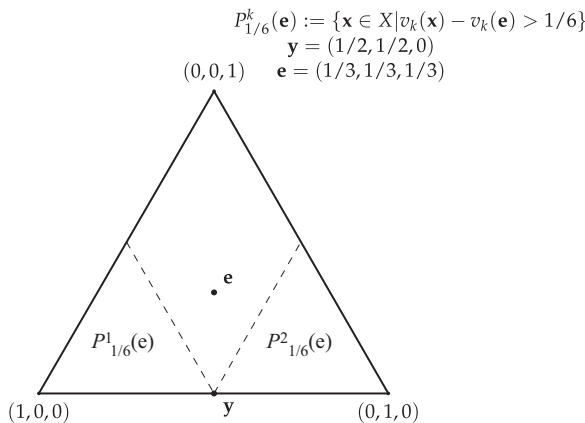

$$P^k_{1/6}(\mathbf{e}) := \{\mathbf{x} \in X \mid v_k(\mathbf{x}) - v_k(\mathbf{e}) > 1/6\}$$
$$\mathbf{y} = (1/2, 1/2, 0)$$
$(0,0,1)$ $\mathbf{e} = (1/3, 1/3, 1/3)$

$\mathbf{e}$

$P^1_{1/6}(e)$ $P^2_{1/6}(e)$

$(1,0,0)$ $\mathbf{y}$ $(0,1,0)$

Figure 1. The triangle and its interior are the set of points in $\mathbf{R}^3$ the components of which are nonnegative and sum to 1. Each vertex is an individual's ideal point; the dashed lines are 1's and 2's indifference sets for **y**. There is no division **z** such that $v_k(\mathbf{z}) > v_k(\mathbf{e}) + 1/6$ for each member k of a majority, thus $\mathbf{e} \in M(R_{1/6}, X)$.

Proposition 2. *If $x \in M(R_\varepsilon, X)$ and $\varepsilon = \min\{\delta > 0 | M(R_\delta, X) \neq \emptyset\}$, then there is no $y \in X$ that strongly Pareto dominates x.*

We cannot replace 'strongly Pareto dominates' with 'weakly Pareto dominates' in the statement,[8] as the next example demonstrates.

Example. Let $X = \{w, x, y, z\}$, $N = \{1, 2, 3\}$ and

$$
\begin{array}{lll}
v_1(x) = 3 & v_2(y) = 3 & v_3(z) = 3 \\
v_1(y) = 2 & v_2(w) = 2 & v_3(x) = 2 \\
v_1(w) = 2 & v_2(z) = 1 & v_3(y) = 1 \\
v_1(z) = 1 & v_2(x) = 0 & v_3(w) = 1
\end{array}
$$

Then, $M(R_\varepsilon, X) = \emptyset$ for all $\varepsilon < 1$ and $M(R_1, X) = \{w, x, y, z\}$, but y weakly Pareto dominates w.

A good definition should fit with any firm intuitions we have about popular control and bring consistency and guidance to our thinking in cases where we lack firm intuitions. As motivation for the definition, I began with a hard case in which elected officeholders face opportunities to divide surpluses any which way among self-interested citizens, and so, for any given division they might choose, an overwhelming majority will prefer an alternative division. What would it even mean for elections to give majorities control over public officials in such cases? If majorities had control, would we expect a minimal majority coalition to receive the lion's share or everyone to receive an equal share? Or is the idea of majority control simply unintelligible when the core of the majority preference relation is empty? I assume most of us lack firm intuitions about how to answer these questions. A formal definition is useful if for no other purpose than to ensure that our use of the concept of majority control is consistent and logically coherent. But having given the concept a rigorous definition, we get an unexpected implication as a bonus: in the divide-the-dollar setting, giving majorities maximal control would lead to the egalitarian division.

Now that we have a definition of 'popular control', we can meaningfully ask whether and under what conditions elections establish popular control over elected officials. The next section uses a simple principal–agent model of elections to address the question.

4. Elections and popular control: a model

One might worry that the definition of 'maximal majority control' merely picks out an ideal that, whatever its appeal on normative grounds, lacks any connection with the logic of actual elections. In this section I present a model that is intended as a 'proof of concept': an illustration of how elections might serve to give majorities maximal or nearly maximal control as defined, even when the core of the majority preference relation is empty.

There are different electoral mechanisms by which citizens might exercise a controlling influence over public officials. They might use elections as mechanisms for selecting good types of officials or as mechanisms for rewarding or punishing

incumbents' behavior (Fearon, 1999).[9] Here I present a simple model of the second mechanism. The model deliberately abstracts from many features of democracies that are consequential for the principal–agent relationship between citizens and elected officeholders. Its aim is to keep everything as simple as possible, but without relying on the one simplifying device that most principal–agent models depend upon, namely the assumption of a representative voter. Instead of a representative voter, the officeholder faces a collection of individual voters, who have common as well as conflicting interests. In the equilibria I describe, citizens condition their votes on assessments of the incumbent's performance, and majorities have maximal or nearly maximal control over the incumbent.

The players of the game are an incumbent, a challenger, and n voters, with n odd. Prior to any player's choice of action, the value of a random variable $\Theta \in [0,1]$ is drawn from a distribution F, which only the incumbent and challenger observe. The set of policies available to the incumbent is then $X_\theta := \left\{ \mathbf{x} \in \mathbb{R}_+^{n+1} \mid \sum_k x_k = \theta \right\}$. Voters then announce their voting rules, the incumbent chooses $\mathbf{x} \in X_\theta$, and each voter i observes her share x_i. The challenger then chooses $\mathbf{y} \in X_\theta$, which voters observe before casting their ballots for the incumbent or challenger.

Let $X := \bigcup_\theta X_\theta$. The order of play in the first period is as follows.

1. Nature chooses $\theta \in [0, 1]$, which only the incumbent and challenger observe.
2. Each voter announces her voting rule, $f_i : [0, 1] \times X \to \{0, 1\}$.
3. The incumbent chooses $\mathbf{x} \in X_\theta$, and each voter i observes x_i.
4. The challenger announces $\mathbf{y} \in X_\theta$, which all voters observe.
5. Each voter casts her vote, $a_i \in \{0, 1\}$, and the incumbent wins re-election if and only if $\sum_i a_i \geq (n + 1)/2$, otherwise the challenger holds office in Period 2.

Thus, if the voter observes x_i and $\mathbf{y} \in X$ and she votes as her announced rule prescribes, then $a_i = f_i(x_i, \mathbf{y})$.

Voters receive utility $u_i(\mathbf{x}) = x_i$. Let $r(\mathbf{x})$ refer to the $n + 1^{\text{th}}$ component of $\mathbf{x}$, which is interpreted as the rent captured by the incumbent. The incumbent receives a payoff of $r(\mathbf{x}) + \theta R$ if re-elected and $r(\mathbf{x})$ otherwise, and the challenger receives a payoff of θR if elected and 0 otherwise, where $R > 0$. Thus, the greater the size of the pie from which officeholders can take a cut for themselves, the greater the second-period payoff from holding office. If $R \geq 1$, then the incumbent is patient and willing to refrain from taking any first-period rent, if necessary, in order to win re-election; if $R < 1$, the incumbent is impatient and unwilling to do so.

The model is a variation on Ferejohn's (1986) model of electoral accountability. In this model, each citizen knows the impact of the incumbent's policy on her own welfare, but in the absence of any electoral competition—in the absence of a challenger who can point out the feasible alternatives to the incumbent's policy—the citizen would have no basis for determining whether the incumbent had 'misbehaved'. Elections create the possibility for popular control within the model because they allow the challenger to reveal to voters what the incumbent could

have done for them. Voters can thereby learn whether the incumbent's behavior was bad enough, relative to feasible alternatives, that it calls for punishment at the polls.

If each voter were prepared to abandon the incumbent whenever the challenger identified an alternative under which the voter would have been better off, then the incumbent could never win the election. For he could then win only if he adopted a policy in the core of the majority preference relation, but the core is empty in this setting. Thus, such voting rules would fail to give the incumbent incentives to refrain from seizing the maximum rent in the first period.

But if voters cut the incumbent some slack, then they can create incentives that work to give majorities maximal or nearly maximal control over the incumbent. Consider voting rules of the following form

$$f_i(x_i, \mathbf{y}) = \begin{cases} 1 & \text{if } y_i - x_i \leqslant \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

The ε term captures the amount of slack that the voter gives the incumbent. Provided the challenger cannot prove that a significantly more attractive alternative choice was available to the incumbent, the voter remains loyal to the incumbent. Let us refer to such rules as *incumbent-favoring voting rules*.

Equilibria in strategies using incumbent-favoring voting rules exist. With the appropriate choice of the incumbency bias, these strategies give majorities maximal or nearly maximal control over incumbents and the incumbent secures no or almost no rent.

Let $\mathbf{e} := (1/n, \ldots, 1/n, 0)$. For $\mathbf{y} \in X$, let $\sigma(\mathbf{y}) := \sum_k y_k$. Note that $\sigma(\mathbf{y}) = \theta$ for $\mathbf{y} \in X_\theta$

Proposition 3. *Let* $\varepsilon = (n-1)/[n(n+1)]$. *For all* $\delta < \theta\varepsilon$, $M(R_\delta, X_\theta) = \emptyset$, *and*

$$M(R_{\theta\varepsilon}, X_\theta) = \{\theta\mathbf{e}\}$$

With ε specified as in this proposition, define

$$f_i^*(x_i, \mathbf{y}) := \begin{cases} 1 & \text{if } y_i - x_i \leqslant \sigma(\mathbf{y})\varepsilon \\ 0 & \text{otherwise} \end{cases}$$

Since θ measures what the incumbent is capable of providing to voters in the aggregate, the voting rules allow voters to hold the incumbent accountable to a standard that is sensitive to this capacity constraint.

Proposition 4. *If* $R > 1$, *then there is a subgame perfect Nash equilibrium in which voters announce and use the voting rules* $f_1^*, \ldots, f_n^*$, *and in any such equilibrium, the incumbent chooses* $\theta\mathbf{e} \in X_\theta$ *after observing* $\theta \in [0,1]$, *and always wins re-election.*

Taken together, the propositions say that when voters use the incumbent-favoring voting rules defined above and incumbents are willing to forego all rents for the sake of winning re-election, i.e. $R > 1$, then majorities achieve maximal control over the incumbent's behavior.

If the incumbent is unwilling to forego all rents in the first period for the sake of re-election, then no voting rules give majorities maximal control. But if $1 - R$ is near 0, then there are incumbent-favoring voting rules that give majorities nearly maximal control. And the more strongly the incumbent values re-election, the greater the degree of control that majorities exercise over the incumbent's behavior, and the smaller the rent that the incumbent is able to extract in equilibrium.

Proposition 5. *Let* $\tilde{\mathbf{e}} = (R/n, \ldots, R/n, 1 - R)$ *and* $\gamma := \varepsilon + (1 - R)/n$, *with* ε *defined as in Proposition 3. For all* $\theta \in [0, 1]$,

1. $\theta\tilde{\mathbf{e}} \in M(R_{\theta\gamma}, X_{\theta})$,
2. $\theta\mathbf{x} \in M(R_{\theta\gamma}, X_{\theta})$ *and* $\mathbf{x} \neq \tilde{\mathbf{e}} \Rightarrow \theta x_{n+1} < \theta(1 - R)$.

Thus, if $1 - R$ is near zero and the influence of majorities ensures an outcome in $M(R_{\theta\gamma}, X_{\theta})$, then majorities have nearly maximal control, as defined above. With γ as specified, define

$$g_i^*(x_i, \mathbf{y}) := \begin{cases} 1 & \text{if } y_i - x_i \leqslant \sigma(\mathbf{y})\gamma, \\ 0 & \text{otherwise} \end{cases}$$

Like the previous voting rules, these rules allow voters to hold the incumbent accountable according to a standard that is relative to the capacity constraint, θ. But with $R < 1$, voters must now grant the incumbent additional slack if they wish to give him any incentives; without the additional slack his impatience will lead him to take the maximum rent in the first period.

Proposition 6. *If* $R < 1$, *then there is a subgame perfect Nash equilibrium in which voters announce and use the voting rules* $g_1^*, \ldots, g_n^*$, *and in any such equilibrium, the incumbent chooses* $\theta\tilde{\mathbf{e}} \in X_{\theta}$ *after observing* $\theta \in [0,1]$, *and always wins re-election.*

This proposition and part (1) of the preceding proposition imply that majorities achieve nearly maximal control when they use the specified voting rules.

This completes the proof of concept that this section was meant to furnish. The previous section offered an interpretation of popular control that reconciled it with insights from social choice theory. The proof of concept was meant to allay worries that popular control, as defined, might bear no connection to any of the usual mechanisms by which elections are thought to give citizens control. One such mechanism is retrospective voting when incumbents and challengers are identical with respect to their preferences and abilities. We have now seen that this mechanism is indeed capable of producing popular control in the stipulated sense.

5. Conclusion

Democracy is widely thought to provide citizens with control over their elected leaders, but findings in social choice theory raise puzzles about what their control could possibly consist of. Social choice theory shows that there are deep problems

in the naïve idea that popular control is achieved when elected leaders have incentives to 'do as the people want'. In some situations, no matter which policy is enacted, there will be some alternative that an overwhelming majority of citizens believe should have been enacted instead.

Riker's (1982) response was to give up on the idea that democratic institutions provide any significant popular control over policy. He argued for a vision of democracy, like Schumpeter's (1942), in which citizens have the power to remove leaders from office, but in which there is no expectation that this power gives leaders incentives to respect the will of the people—simply because the 'will of the people' is ill-defined.

In this article I have described one sense in which majorities may still be said to have control, even if the core of the majority preference relation is empty. This interpretation of popular control does not presuppose the existence of a 'popular will'. With the definition it becomes intelligible to ask, and with the help of models to study, the question of whether elections enable citizens to control officials, even when the core is empty. It tells us what would count as an answer to this question. And so it tells us what to look for when exploring principal–agent models like the one analyzed, in which officeholders face, not a 'representative voter' but separate individuals with their own distinct points of view.

Appendix

Let $N = \{1, \ldots, n\}$. Throughout I assume that $n \geqslant 3$, n is odd, and $v_i(\mathbf{x}) = x_i$, where x_i is the ith component of $\mathbf{x} \in \mathbf{R}_+^{n+1}$. The sets $X, X_\theta \subset \mathbf{R}_+^{n+1}$ are defined as in Section 4.

Lemma 1. *Let any $r \geqslant 0$ be given and set $\gamma := \frac{n-1}{n(n+1)} + \frac{r}{n}$. Define $\mathbf{e} \in X_1$ by $e_i := \frac{1-r}{n}$ for $i \in N$ and $e_{n+1} := r$.*

   (i) $\mathbf{e} \in M(R_\gamma, X_1)$.
   (ii)   *For any $\mathbf{x} \in X_1$ with $x_i > (1-r)/n$ for some $i \in N$ and $x_{n+1} \geqslant r$, there is a $\mathbf{w} \in X_1$ and a $M \subset N$ with $|M| \geqslant (n+1)/2$ and $i \notin M$, such that*

   $$w_j > x_j + \gamma, \ \forall j \in M \ \& \ w_{n+1} = 0$$

   (iii) $\mathbf{x} \in M(R_\gamma, X_1) \ \& \ \mathbf{x} \neq \mathbf{e} \Rightarrow x_{n+1} < r$.

Proof: (i) Let $m := (n+1)/2$. If $\mathbf{y} P_\gamma \mathbf{e}$ for some $\mathbf{y} \in X_1$, then for at least m members of N, $y_i > (1-r)/n + \gamma$. But that implies $\sum_{i \in N} y_i > 1$, a contradiction. Thus, (i).

(ii) Take any $\mathbf{x} \in X_1$ with $x_i > (1-r)/n$ for some $i \in N$. Choose $M \subset N \setminus \{i\}$ so that $|M| = m$ and for all $j \in M$ and $k \in N \setminus (M \cup \{i\})$, $x_j \leqslant x_k$. Then, since $x_i > (1-r)/n$ and $\sum_{k \in N} x_k = 1 - x_{n+1} \leqslant 1 - r$, we must have $\sum_{j \in M} x_j < \frac{m(1-r)}{n}$. Thus

$$\frac{1}{m}\left(x_{n+1} + \sum_{j\in N\setminus M} x_j\right) = \frac{1}{m}\left(1 - \sum_{j\in M} x_j\right)$$
$$> \frac{1}{m}\left(1 - \frac{m(1-r)}{n}\right)$$
$$= \gamma$$

Construct $\mathbf{w} \in X_1$ by setting

$$w_k = x_k + \frac{1}{m}\left(x_{n+1} + \sum_{j\in N\setminus M} x_j\right), \ \forall k \in M$$

and $w_k = 0$ for all $k \notin M$. Thus, $w_k > x_k + \gamma$ for all $k \in M$. Thus, (ii).

(iii) Assume $\mathbf{x} \neq \mathbf{e}$ and $\mathbf{x} \in M(R_\gamma, X_1)$. By (ii), we must have either $x_{n+1} < r$ or $x_i \leqslant (1-r)/n$ for $i = 1,2,\ldots,n$. If the former, we're done, so assume the latter. This implies $x_{n+1} > r$, since $\mathbf{x} \neq \mathbf{e}$. Now choose $M \subset N$ such that $|M| = m$ and $x_j \leqslant x_k$ for all $j \in M$ and $k \in N \setminus M$. We must have $\sum_{j\in M} x_j < \frac{m(1-r)}{n}$ because $x_j \leqslant (1-r)/n$ for all $j \in M$ and $x_{n+1} > r$. We can therefore construct $\mathbf{w}$, as in the proof of (ii), such that $\mathbf{w}P_\gamma \mathbf{x}$, a contradiction. Thus, we reject the supposition that $x_i \leqslant (1-r)/n$, $\forall i \in N$ and conclude $x_{n+1} < r$. $\qquad\square$

Proof of Proposition 1. Let $Y := \{\mathbf{x} \in \mathbf{R}_+^{n+1}| \sum x_i = 1, x_{n+1} = 0\}$ and $\mathbf{e} := (1/n,\ldots,1/n,0)$. Set $\varepsilon := \frac{n-1}{n(n+1)}$. Proposition 1 is then equivalent to the statement that (1) $M(R_\varepsilon, Y) = \{\mathbf{e}\}$ and (2) $M(R_\delta, Y) = \emptyset$ for all $\delta < \varepsilon$.

Part (i) of the lemma (setting $r = 0$) implies $\mathbf{e} \in M(R_\varepsilon, Y)$. Observe that the antecedent of part (ii) of the lemma (with $r = 0$) necessarily holds if $\mathbf{x} \in Y$ and $\mathbf{x} \neq \mathbf{e}$, and its consequent implies $\mathbf{w}P_\varepsilon \mathbf{x}$ for some $\mathbf{w} \in Y$. Hence, $M(R_\varepsilon, Y) = \{\mathbf{e}\}$.

To prove (2), let $\delta < \varepsilon$. Since $\mathbf{x} \in M(R_\delta, Y) \Rightarrow \mathbf{x} \in M(R_\varepsilon, Y)$ and, as just shown, $M(R_\varepsilon, Y) = \{\mathbf{e}\}$, if $M(R_\delta, Y) \neq \emptyset$ then $M(R_\delta, Y) = \{\mathbf{e}\}$. But there exists $\mathbf{x} \in Y$ with $x_i = 1/n + \varepsilon$ for $i = 1,2,\ldots,(n+1)/2$. Thus, $\mathbf{x}P_\delta \mathbf{e}$ and so $\mathbf{e} \notin M(R_\delta, Y)$. $\qquad\square$

Proof of Proposition 2. Assume the hypothesis, but suppose y strongly Pareto dominates x. Let

$$\hat{\varepsilon} = \min_{i\in N}\{v_i(y) - v_i(x)\}$$

Since $x \in M(R_\varepsilon, X)$, we must have $\varepsilon > \hat{\varepsilon}$. Note that

$$v_i(y) + \varepsilon - \hat{\varepsilon} \geqslant v_i(y) + \varepsilon - (v_i(y) - v_i(x)) = v_i(x) + \varepsilon, \ \forall i \in N$$

and, for every $z \in X$, there is a majority $M \subset N$ such that $v_i(x) + \varepsilon \geqslant v_i z$ for all $i \in M$. It follows that $y \in M(R_{\varepsilon-\hat{\varepsilon}}, X)$. But this contradicts the hypothesis that $\varepsilon = \min\{\delta > 0 | M(R_\delta, X) \neq \emptyset\}$. $\qquad\square$

A.1 The model of electoral control

Proofs of Propositions 3 and 5. Observe that (given the specifications of v_i) for any $\theta \in [0,1]$ and $\mathbf{x},\mathbf{y} \in X_1$, $\mathbf{y} P_\varepsilon \mathbf{x}$ if and only if $\theta\, \mathbf{y} P_{\theta\varepsilon} \theta\, \mathbf{x}$ and, therefore, $\mathbf{x} \in M(R_\varepsilon, X_1)$ if and only if $\theta\, \mathbf{x} \in M(R_{\theta\varepsilon}, X_\theta)$. Proposition 3 is thus a corollary of Proposition 1.

To prove Proposition 5, first set $\theta = 1$ and apply parts (i) and (iii) of the lemma, setting $r = 1 - R$. Then apply again the observation that $\mathbf{x} \in M(R_\varepsilon, X_1)$ if and only if $\theta\mathbf{x} \in M(R_{\theta\varepsilon}, X_\theta)$.

Proofs of Propositions 4 and 6. In both cases—$R < 1$ or $R \geqslant 1$—each voter is indifferent between the incumbent and challenger at the time of election, so voting according to the voting rule announced in stage (2) is a best response.
In any subgame perfect Nash equilibrium, the challenger's strategy $\mathbf{y}^*$ must satisfy

$$\mathbf{y}^*(\theta; f_1, \ldots, f_n; \mathbf{x}) \in \left\{ \mathbf{z} \in X_\theta \Big| \sum_{i=1}^{n} f_i(x_i, \mathbf{z}) < (n+1)/2 \right\}$$

whenever this set is nonempty. Let $\mathbf{y}^*$ be any such strategy.
Define

$$I(\theta; f_1, \ldots, f_n) := \left\{ \mathbf{z} \in X_\theta \Big| \sum_{i=1}^{n} f_i(z_i, \mathbf{y}) \geq (n+1)/2, \ \forall \mathbf{y} \in X_\theta \& r(\mathbf{z}) + \theta R \geq \theta \right\}$$

In words, this set contains the policy choices that guarantee re-election and give the incumbent enough of a rent that he prefers re-election under these terms to seizing the maximum rent in Period 1. Given the equilibrium behavior from the challenger, in any subgame perfect Nash equilibrium the incumbent's strategy must satisfy $\mathbf{x}^*(\theta; f_1, \ldots, f_n) \in I(\theta; f_1, \ldots, f_n)$ whenever this set is nonempty and $\mathbf{x}^*(\theta; f_1, \ldots, f_n) = (0, \ldots, 0, \theta)$ otherwise, for all $\theta \in [0,1]$. Let $\mathbf{x}^*$ be any strategy that satisfies this condition.

Proposition 4. Assume $R \geqslant 1$ and define $f_1^*, \ldots, f_n^*$ as in the main text above, with $\varepsilon = \frac{n-1}{n(n+1)}$.
Proposition 3 implies

$$I(\theta; f_1^*, \ldots, f_n^*) = \{\theta \mathbf{e}\}$$

so the incumbent's strategy $\mathbf{x}^*$ has him choosing $\theta \mathbf{e}$ whenever nature chooses $\Theta = \theta$ and voters announce $f_1^*, \ldots, f_n^*$.

It remains to prove that no voter i can deviate from the choice of f_i^* at stage (2). If voter i does have a profitable deviation, it must leave him strictly better off for some $\theta \in [0, 1]$.

Suppose that i deviates and this prompts the incumbent, after the observation of some $\theta \in [0,1]$, to choose a policy $\mathbf{x} \in X_\theta$ with $x_i > \theta/n$. Since $\frac{1}{\theta} x_i > 1/n$, we can invoke part (ii) of the lemma (with $r = 0$) and conclude that there is some majority $M \subset N$, which excludes i, and some $\mathbf{w} \in X_1$ such that $w_j > \frac{1}{\theta} x_j + \varepsilon$ for all $j \in M$. Hence, $\theta w_j > x_j + \theta \varepsilon$ for all $j \in M$. Note that $\sigma(\theta\, \mathbf{w}) = \theta$. Thus, each $j \in M$, voting according to f_j^*, will vote for the challenger when she chooses $\theta \mathbf{w} \in X_\theta$ after the

incumbent chooses $\mathbf{x} \in X_\theta$, but that contradicts the assumption of optimal play from the incumbent.   □

Proposition 6. Assume $R < 1$ and let $g_1^*, \ldots, g_n^*$ be the incumbent-favoring voting rules defined in the main text, with $\gamma = \varepsilon + (1 - R)/n$.

Proposition 5 implies

$$I(\theta; g_1^*, \ldots, g_n^*) = \{\theta\widetilde{\mathbf{e}}\}$$

and so the incumbent's strategy $\mathbf{x}^*$ has him choosing $\theta\widetilde{\mathbf{e}}$ whenever nature chooses $\Theta = \theta$ and voters announce $g_1^*, \ldots, g_n^*$.

It remains to prove that no voter can profitably deviate from the voting rules announced at stage (2). Suppose that i deviates and this prompts the incumbent, after the observation of some $\theta \in [0,1]$, to choose a policy $\mathbf{x} \in X_\theta$ with $x_i > \theta R/n$. Then $\frac{1}{\theta}x_i > R/n$, so we can invoke part (ii) of the lemma once more (with $r = 1 - R$) and conclude that for some $\mathbf{w} \in X_1$ and majority M excluding i, $w_j > \frac{1}{\theta}x_j + \gamma$ for all $j \in M$. Hence, $\theta w_j > x_j + \theta\gamma$ for all $j \in M$. Since each $j \in M$ votes according to g_j^*, the incumbent will lose the election when the challenger adopts $\theta\mathbf{w} \in X_\theta$, contradicting the assumption of optimal play from the incumbent.   □

Acknowledgements

Declaration of Conflicting Interests

Funding

Notes

1. The core of the majority preference relation is the set of policies 'at least as good' as all others, when comparisons are made according to the preferences of majorities.
2. For an overview of the puzzles that social choice theory raises for positive political theory, and the relevant literature, see Cox and Shepsle (2007). For an informative discussion of Mackie's take on social choice theory, see Patty and Penn (2014).
3. Pettit does discuss the implications of his and others' work on judgment aggregation for direct democracy; see pp. 191–194. See footnote 9 on p. 194 for Pettit's explanation of the focus on judgment aggregation instead of Arrovian preference aggregation.

4. Pettit (2013) holds the same position, as I understand it; see his remarks on the distinction between the constituting people—the citizens who influence and control the government but who are not, taken together, a group agent—and the constituted people—the state, which can be an agent under the right conditions (Pettit, 2013). That popular control is intelligible independently of a popular collective agent is also implied, I take it, by a passing comment in List and Pettit (2005): 'We think that [the people can operate as a personifiable entity] insofar as government can be constrained by public opinion or public reason to the point where its judgments […] are reasonably treated as the people's own' (List and Pettit, 2005: 389). If we check the people's status as an agent by checking whether government is subject to popular control—constraint by public opinion—then the latter must be intelligible independently of the former.

5. I take it that Pettit believes he can avoid this objection because he only defends the possibility of popular control for those 'who are willing to live on equal terms with others' (Pettit, 2013: 170). Another way of avoiding the objection, which Pettit rejects (footnote to p. 170) is to use a moralized definition of 'acceptability'—construing it as what each *ought* to be willing to accept—so that a person's morally objectionable preferences do not determine what counts as 'acceptable' to her. Moralized acceptability tests are the cornerstones of many contractualist moral and political theories (Beitz, 1989; Estlund, 2008; Rawls, 1971, 1993; Scanlon, 2000). I believe 'control' should be defined so that moral disagreements do not stand in the way of agreement about the concept's application, which is unlikely if we use a moralized acceptability test to define 'control'.

6. That is, we assume that each person's utility from a policy x and an amount m of the numeraire is given by a quasilinear function, such as $u_i(x, m) = v_i(x) + m$. In my view, the willingness-to-pay measure of how much a person values one suboptimal policy relative to better alternatives represents just one way to formalize the basic, intuitive idea behind the definition of 'popular control' that I present here.

7. These authors discuss it in the context of the spatial model. They interpret ε in terms of the metric on the policy space, instead of willingness to pay units of a numeraire good.

8. A policy y would weakly dominate x if $v_i(y) \geqslant v_i(x)$ for all $i \in N$ and $v_j(y) > v_j(x)$ for some $j \in N$.

9. If citizens use elections to select officials whose policy preferences align with their own, and these officials then act at their own discretion and according to their own policy preferences—like good 'trustees'—then citizens may not influence or control the behavior of any given elected official. But they may nonetheless influence and control the general direction of policy that ultimately results from their choices at the polls.

References

Ansolabehere S and Snyder J (2000) Valence politics and equilibrium in spatial election models. *Public Choice* 103: 327–336.

Barro RJ (1973) The control of politicians: an economic model. *Public Choice* 14: 19–42.

Beitz C (1989) *Political Equality*. Princeton: Princeton University Press.

Besley T (2006) *Principled Agents? The Political Economy of Good Government*. Oxford: Oxford University Press.

Bräuninger T (2007) Stability in spatial voting games with restricted preference maximizing. *Journal of Theoretical Politics* 19: 1173–1191.

Cohen J (1986) An epistemic conception of democracy. *Ethics* 97: 26–38.

Coleman J and Ferejohn J (1986) Democracy and social choice. *Ethics* 97: 6–25.

Coughlin P (1992) *Probabilistic Voting Theory*. Cambridge: Cambridge University Press.

Cox GW and Shepsle KA (2007) Majority cycling and agenda manipulation: Richard McKelvey's contributions and legacy. In: Aldrich JH, Alt JE and Lupia A (eds) *Positive Changes in Political Science*. Ann Arbor: University of Michigan Press.

Dahl R (1989) *Democracy and its Critics*. New Haven: Yale University Press.

de Tocqueville A (2000) *Democracy in America*. New York: Perennial.

Dryzek JS and List C (2003) Social choice theory and deliberative democracy: a reconciliation. *British Journal of Political Science* 33: 1–28.

Estlund DM (2008) *Democratic Authority*. Princeton: Princeton University Press.

Fearon J (1999) Electoral accountability and the control of politicians: selecting good types versus sanctioning poor performance. In: Przeworski A, Manin B and Stokes S (eds) *Democracy, Accountability, and Representation*. Cambridge: Cambridge University Press.

Feld SL and Grofman B (1991) Incumbency advantage, voter loyalty and the benefit of the doubt. *Journal of Theoretical Politics* 3: 115–137.

Ferejohn J (1986) Incumbent performance and electoral control. *Public Choice* 50: 5–26.

Le Breton M (1987) On the core of voting games. *Social Choice and Welfare* 4: 295–305.

List C and Pettit P (2005) On the many as one: a reply to Kornhauser and Sager. *Philosophy & Public Affairs* 33: 377–390.

List C and Pettit P (2011) *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.

McKelvey R (1979) General conditions for global intransitivities in formal voting models. *Econometrica* 47: 1085–1112.

McKelvey R and Schofield N (1987) Generalized symmetry conditions at a core point. *Econometrica* 55: 923–933.

Mackie G (2003) *Democracy Defended*. Cambridge: Cambridge University Press.

Maskin E and Tirole J (2004) The politician and the judge: accountability in government. *American Economic Review* 94: 1034–1054.

Miller D (1992) Deliberative democracy and social choice. *Political Studies* 40: 54–67.

Patty J and Penn EM (2014) *Social Choice and Legitimacy: The Possibilities of Impossibility*. Cambridge: Cambridge University Press.

Pettit P (2013) *On the People's Terms: A Republican Theory and Model of Democracy*. Cambridge: Cambridge University Press.

Plott CR (1967) A notion of equilibrium and its possibility under majority rule. *American Economic Review* 57: 787–806.

Rawls J (1971) *A Theory of Justice*. Cambridge: Harvard University Press.

Rawls J (1993) *Political Liberalism*. New York: Columbia University Press.

Riker W (1982) *Liberalism Against Populism*. New York: Waveland Press.

Saari D (1997) The generic existence of a core for *q*-rules. *Economic Theory* 9: 219–260.

Scanlon T (2000) *What We Owe to Each Other*. Cambridge: Belknap Press of Harvard University Press.

Schofield N (2003) Valence competition in the spatial stochastic model. *Journal of Theoretical Politics* 15: 371–383.

Schofield N (2004) Equilibrium in the spatial 'valence' model of politics. *Journal of Theoretical Politics* 16: 447–481.

Schumpeter J (1942) *Capitalism, Socialism, and Democracy*. New York: Harper & Brothers.

Tovey CA (2010) The instability of instability of centered distributions. *Mathematical Social Sciences* 59: 53–73.

Tovey CA (2011) The finagle point and the epsilon-core: a comment on Bräuninger's proof. *Journal of Theoretical Politics* 23: 135–139.